Chapter 3
Foundations of Differentiability: Functions of several variables

In this section we shall consider finite dimensional normed vector spaces V and W. We recall that it will be no essential restriction of generality if we often assume $V = \mathbb{K}^n$ and $W = \mathbb{K}^m$.

In earlier chapters we have frequently used the concept of an *inner* or *interior* point of a subset X of a metric space Y; since X was almost always an interval in $Y = \mathbb{R}$, all points with the possible exception of end pints were inner points. We recall the definition:

Definition 3.1. Let (Y, d) be a metric space. A point $a \in X \subseteq Y$ is called an *inner point* or an *interior point* of X if and only if a neighborhood of a in Y is contained in X. (We recall that a subset U is a neighborhood of a in Y if there is a number $r > 0$ such that $U_r(a) = \{x \in Y \mid d(x, a) < r\}$ is contained in U.) The set of all inner points is called the *interior* of X. □

We observe that X is open if and only if every point of X is an inner point of X, i.e. if X agrees with its interior.

In the main definition for this chapter we formulate the differentiability of functions $f: X \to M$ for $X \subseteq V$ and $M \subseteq W$; it follows exactly the lead of the first definition of Chapter 4 of Analysis I. But the more familiar definition of differentiability via condition 4.7(ii) in the case of one variable must fail here in the case of several variable because we cannot form a quotient of two vectors.

DEFINITION OF DIFFERENTIABILITY

Definition 3.2. Let V and W be two finite dimensional normed vector spaces. A function $f: X \to M$, $X \subseteq V$, $M \subseteq W$ is called *differentiable in an inner point* a of X if there is a linear function $L: V \to W$ and a function $r: X \to W$ such that the following conditions hold:

(i) $f(x) = f(a) + L(x - a) + r(x)$ and

(ii) $\lim_{\substack{x \to a \\ x \neq a}} \|x - a\|^{-1} \cdot r(x) = 0$

This is equivalent with the existence of a linear function L such that

(iii) $\lim_{\substack{x \to a \\ x \neq a}} \frac{1}{\|x-a\|} \cdot \|f(x) - f(a) - L(x - a)\| = 0$.

We say that f is *differentiable*, if X is open and f is differentiable in all points of X. □

Notice that $A(x) = f(a)+L(x-a) = Lx+(f(a)-La)$ defines an affine function, and that $f(x) = A(x)$ plus a remainder function $f - A$ which is very small near a if f is differentiable at a. These circumstances should allow us to reduce various local properties of f to those of A. For instance it is not unreasonable to surmise that, locally, f is invertible if A is invertible and that is the case if L is invertible; and for this we have a very effective test via the determinant $\det L$.

If $f\colon I \to W$ is a curve in the finite dimensional normed vector space (see 2.1) then Definitions 2.2 and 3.1 are easily seen to be compatible in view of the fact that each linear map $L\colon \mathbb{R} \to W$ is given by a unique vector $v \in W$ via $L(t) = t{\cdot}v$ so that $v \mapsto L\colon W \to \mathrm{Hom}(\mathbb{R}, W)$ is a linear bijection.

Recall that a Banach space is a normed vector space in which every Cuachy sequence converges. We notice that Definition 3.2 can be straightforwardly generalized to the case of two Banach spaces V and W with the only proviso that the linear map L is *postulated to be continuous*, which in the infinite dimensional case is not automatic.

Remark 3.3. If f is differentiable in a, then the linear map L is uniquely determined.

Proof. We indicated the proof as a variant of the proof of 4.2. Indeed, for a nonzero vector v, set $\mathrm{rad}\, v = \|v\|^{-1}{\cdot}v$; then $\mathrm{rad}\, v$ is a unit vector, and $v = \|v\|{\cdot}\,\mathrm{rad}\, v$. Assume now that $f(x) = f(a)+L_1(x-a)+r_1(x) = f(a)+L_2(x-a)+r_2(x)$ such that $\|x-a\|^{-1}{\cdot}r_1(x) \to 0$ and $\|x-a\|^{-1}{\cdot}r_2(x) \to 0$ for $x \to a$. We conclude that

$$(*)\qquad \begin{aligned}(L_1-L_2)\big(\mathrm{rad}(x-a)\big) &= (L_1-L_2)\big(\|x-a\|^{-1}{\cdot}(x-a)\big)\\ &= \|x-a\|^{-1}{\cdot}(r_2-r_1)(x) \to 0 \quad \text{for} \quad x \to a.\end{aligned}$$

Let e be an arbitrary unit vector in V; since a is an inner point of X, there is a $\delta > 0$ such that $0 < t < \delta$ implies $x = a + t{\cdot}e \in X$. Then $\mathrm{rad}(x-a) = e$ and $\|x-a\|^{-1}{\cdot}(r_2-r_1)(x) = t^{-1}(r_2-r_1)(a+t{\cdot}e) \to 0$ for $t \to 0$. Therefore $(*)$ implies $(L_1-L_2)(e) = 0$ for all unit vectors e. Thus $L_1 = L_2$. □

The proof works even if a is a boundary point of X, provided there is a basis $e_1, \dots, e_n$ of unit vectors of V and a $\delta > 0$ such that $a + t{\cdot}e_k \in X$ for $k = 1, \dots, n$ and $0 < t < \delta$. But a condition close to this one will be necessary because of the following example: Let $X = \{(x,y) \in \mathbb{R}^2 : x < 1 \Rightarrow y = 0\}$ and $a = (0,0)$. Then the zero function $f\colon X \to \mathbb{R}$ is differentiable in a and each of the infinitely many linear functions L with $L(x,y) = cy$, $c \in \mathbb{R}$ satisfy the condition of differentiability of f in a with $r = 0$ even though a is an accumulation point of X.

Definition 3.4. The uniquely determined linear map L of Definition 3.2 is called the *derivative* of f in a and is denoted by df_a, or $df(a)$, or D_af, or $f'(a)$.

It is of paramount importance to remember always that D_af (or $f'(a)$) is a linear map $V \to W$, that is, an element of $\mathrm{Hom}(V, W)$.

Notation. Assume that V and W are finite dimensional normed vector spaces. Let X be an open subset of V. Then a function $\omega: X \to \mathrm{Hom}(V,W)$ is called a W-valued *differential form.* If $W = \mathbb{R}$, one omits the adjective "W-valued" and speaks of a *differential form* or also as a *Pfaffian* form. If now $f: X \to Y \subseteq W$ is a differentiable function, then $df: X \to \mathrm{Hom}(V,W)$ (recall $df(x) = f'(x)$) is in fact a W-valued differential form. It is in this context that the notation $df(x)$ is the preferred notation for the derivative of f at $x \in X$.

If $V = \mathbb{K}^n$ and $W = \mathbb{K}^m$ then the elements of $\mathrm{Hom}(V,W)$ may be identified canonically with $m \times n$ matrices. After such an identification the derivative $D_a f$ is a matrix, and we will have to determine its coefficients.

Even in the case of the more special situation $V = W = \mathbb{K}$, the derivative D_f is a linear map, but in the case of one dimension one is not consciously aware of this fact since the linear maps $\mathbb{K} \to \mathbb{K}$ may be identified with 1×1-matrices, that is, with elements of $\mathbb{K}$.

Students tend to be confused about the significance of the derivative $f'(a)$ of f at a. This is a function $f'(a): V \to W$. That is, if $v \in V$ then $f'(a)(v) \in W$. If X is open and f differentiable on X, then f' is a function $f': X \to \mathrm{Hom}(V,W)$ which associates with each point $a \in X$ a linear map $f'(a): V \to W$. In terms of the terminology of differential forms, f' is in fact a differential form.

Remark 3.5. If a function f is differentiable in a, then it is continuous in a.

Proof . Exercise. □

Exercise E3.1. Prove Remark 3.5. □

The converse implication already fails in one variable.

Proposition 3.6. *An affine function $x \mapsto Lx + v: V \to W$ is differentiable and has the derivative L.*

Proof . Exercise. □

Exercise E3.2. Prove 3.5.

Rules of differentiation: The Sum Rule

We hasten to secure the rules of differentiation which we know from the one dimensional case.

Proposition 3.6. (Rules for sums and scalar products) *If $f, g: X \to \mathbb{R}^m$, $X \subseteq \mathbb{R}^n$ are functions and if a is an inner point of X, and if $r \in \mathbb{R}$ is a number, then the differentiability of f and g in a implies that of $f + g$ and $r{\cdot}f$. Moreover, $D_a(f+g) = D_a f + D_a g$ and $D_a(r{\cdot}f) = r{\cdot}D_a f_a$. (Equivalent formulation:*

$$(f+g)'(a) = f'(a) + g'(a) \quad \textit{and} \quad (r{\cdot}f)'(a) = r{\cdot}f'(a).).$$

Proof . Exercise. □

Exercise E3.3. Prove Proposition 3.6. □

As a consequence, the set $\mathcal{D}_a(X) \subseteq W^X$ of a functions $X \to W$ which are differentiable in a form a vector space, and the function $D_a: \mathcal{D}_a \to \mathrm{Hom}(V, W)$ is a linear map.

Rules of differentiation: The Chain Rule

The Chain Rule remains the most important single differentiation rule in the general context.

Chain Rule

Theorem 3.7. *Assume that U, V and W are finite dimensional vector spaces, $X \subseteq U$, $Y \subseteq V$ and assume that $g: X \to Y$ and $f: Y \to W$ functions such that $D_a g$ exists in the inner point $a \in X$ and that $D_b f$ exists in the inner point $b \stackrel{\text{def}}{=} g(a) \in Y$. Then the composition $f \circ g: X \to M$ is differentiable in a and*

$$D_a(f \circ g) = (D_b f)(D_a g) = (D_{g(a)} f)(D_a g). \tag{1}$$

Equivalent notation:

$$(f \circ g)'(a) = f'\big(g(a)\big)g'(a). \tag{2}$$

Proof . The proof of Theorem 4.15 was deliberately organized in such a fashion that it painlessly applies to the present situation. □

Exercise E3.4. Rewrite the proof of 4.15 In Analysis I, being conscious of the present context, properly replacing absolute values by the appropriate norms.

If $X \subseteq U$ and $Y \subseteq V$ and $\varphi: X \to \mathrm{Hom}(V, W)$ and $\psi: X \to \mathrm{Hom}(U, V)$, let us write $(\varphi\psi)(a) = \varphi(a) \circ \psi(a): U \to W$.

Corollary. *If $X \subseteq U$ and $Y \subseteq V$ are open subsets and f and g are differentiable, then, using the notation we just introduced, we can summarize the Chain Rule also in the form*

$$(f \circ g)' = (f' \circ g)g'. \tag{$*$}$$ □

Exercise E3.5. Verify that $(*)$ is an acceptable abbreviation. Note

$$\begin{aligned} (f \circ g)': X &\to \mathrm{Hom}(U, W), \\ g': X &\to \mathrm{Hom}(U, V), \\ f' \circ g: X &\to \mathrm{Hom}(V, W). \end{aligned}$$

For an understanding of the Chain Rule we alert the student again to the fact that $D_a g$ and $D_b f$, $b = g(a)$, are linear maps and that the juxtaposition $(D_b f)(D_a g)$ of linear maps denotes their composition which could have been denoted by $(D_b f) \circ (D_a g)$.

In a computational vein, if $U = \mathbb{K}^p$, $V = \mathbb{K}^n$, and $W = \mathbb{K}^m$ then the linear maps $D_a g$ and $D_b f$ "are" $n \times p$–, respectively, $m \times n$–matrices so that $(D_b f)(D_a g)$ "is" an $m \times p$ matrix product.

Or, again reformulated in other words:

i) the affine approximation of g near a is $x \mapsto g(a) + (D_a g)(x - a)$,

ii) the affine approximation of f near $b = g(a)$ is $y \mapsto f(b) + (D_b f)(y - b) = f\big(g(a)\big) + \big(D_{g(a)} f\big)\big(y - g(a)\big)$,

iii) the affine approximation of $f \circ g$ near a is $f\big(g(a)\big) + D_a(f \circ g)(x - a)$.

iv) The composition of the affine maps in i) and ii) is

$$\begin{aligned} x \mapsto & f\big(g(a)\big) + (D_{g(a)} f)\big(g(a) + (D_a g)(x - a) - g(a)\big) \\ = & f\big(g(a)\big) + (D_{g(a)} f)(D_a g)(x - a). \end{aligned}$$

Therefore we have the following reformulation of the Chain Rule:
The affine approximation of a composition of differentiable functions is the composition of their affine approximations.

The General Mean Value Theorem

As a first simple application of the Chain Rule we formulate the final version of the Mean Value Theorem for vector valued functions of several variables. The decisive work was done in Chapter 2 leading us to the key Lemma 2.10. In particular we recall from 2.8ff. the concept of geodesic distance $d(x, y)$ of two points x, y of a connected open subset of a normed vector space.

MEAN VALUE THEOREM

Theorem 3.8. *Let X be a connected open subset of finite dimensional normed vector space V. Let $f: X \to W$ be a differentiable function with values in a finite dimensional normed vector space W and assume that the function $x \mapsto f'(x): X \to$ $\mathrm{Hom}(V, W)$ is bounded so that $\|f'\| = \sup\{\|f'(x)\| : x \in X\}$ is well defined. Then*

$$(*) \qquad (\forall x,\, y \in X)\ \|f(x) - f(y)\| \le \|f'\| \cdot d(x, y).$$

If x and y are connected in X by a straight line segment, then

$$(**) \qquad \|f(x) - f(y)\| \le \|f'\| \cdot \|x - y\|.$$

Proof. Let $\gamma: [a, b] \to X$ be a piecewise differentiable curve. Then from the Chain Rule we compute that $(f \circ \gamma)'(t) = f'(x)\big(\gamma'(t)\big)$. In particular, for all such γ and all t in the domain of γ we get $\|(f \circ \gamma)'(t)\| \le \|f'\big(\gamma'(t)\big)\| \cdot \|\gamma'(t)\| \le \|f'\| \cdot \|\gamma'(t)\|$. Now Lemma 2.10 applies and immediately yields the assertion of the theorem. □

We should remark, that this theorem hold for not necessarily finite dimensional Banach spaces V and W as well.

Theorem 3.8 allows us at once to conclude:

If two (vector valued) functions (of several variables) have the same derivatives on an open connected set then they differ on this set by at most a constant.

Rules of differentiation: The Product Rule

The product rule plays a distinctly smaller role here than it does in the one-variable situation. We discuss it anyhow to maintain the parallelity of our proceeding with the elementary situation. In a first reading, this subsection may be skipped. The first thing we have to realize is that we do not have a given multiplication of vectors. This requires a systematic approach to multiplications as we see if we analyze the general situation as follows:

Assume that we are given functions $f{:}\,X \to W_1$, $X \subseteq V_1$ and $g{:}\,Y \to W_2$, $Y \subseteq V_1$, both of which are differentiable in the inner points a of X, respectively, b of Y. Moreover, let $B{:}\,W_1 \times W_2 \to U$ be a bilinear function, that is, a function which is linear in each of its arguments if the other one is held fixed. We would like to make statement about the differentiability of the function

$$(x, y) \mapsto B\big(f(x), g(y)\big) : X \times Y \to U \text{ in } (a, b) \in X \times Y.$$

This function is the composition of the functions $f \times g : X \times Y \to W_1 \times W_2$, $(f \times g)(x, y) = \big(f(x), g(y)\big)$, and, secondly, the function B. The function $f \times g$ is differentiable in (a, b) and has the derivative $D_a f \times D_b g$ where $(D_a f \times D_b g)(u, v) = \big(D_a f(u), D_b g(v)\big)$. By the Chain Rule, the function $B \circ (f \times g)$ now is differentiable in (a, b) if B is differentiable in $\big(f(a), g(b)\big)$. Therefore we have to investigate the differentiabilityof bilinear functions. We begin by a Lemma, that belongs largely to linear algebra:

Lemma. *Let W_1, W_2 and U be finite dimensional normed vector spaces and $B{:}\,W_1 \times W_2 \to U$ be a bilinear map. Then there is a a unique smallest real number $\|B\|$ such that*

(3) $$(\forall w_1 \in W_1,\ w_2 \in W_2)\ \|B(w_1, w_2)\|_U \le \|B\|{\cdot}\|w_1\|_{W_1}{\cdot}\|w_2\|_{W_2}.$$

Proof . Exercise. □

Exercise E3.6. Prove the preceding lemma.
[Hint. Pick a basis $e_1, \ldots, e_m$ of W_1 and a basis $f_1, \ldots, f_n n$ of W_2 and define the vectors $b_{j,k} \in U$ by $b_{j,k} = B(e_j, f_k)$, $j = 1, \ldots, m$, $k = 1, \ldots, n$. Write $w_1 = \sum_{j=1}^m x_j{\cdot}e_j$ and $w_2 = \sum_{k=1}^n y_k{\cdot}f_k$. As usual set $\|x\|_\infty = \max\{|x_1|, \ldots, |x_m|\}$, similarly for $\|y\|_\infty = \max\{|y_1|, \ldots, |y_n|\}$, and define $\beta = \|\sum_{\substack{j=1,\ldots,m\\k=1,\ldots,n}} b_{jk}\|_U$. Then $\|B(w_1, w_2)\|_U = \|\sum_{\substack{j=1,\ldots,m\\k=1,\ldots,n}} x_j y_k{\cdot}b_{jk}\|_U \le \|x\|_\infty{\cdot}\|y\|_\infty\|$. By 1.27,there are numbers c_1 and c_2 such that $\|w_p\|_\infty c_p\|w_p\|_{W_p}$, $p = 1, 2$. Thus $\|B(w_1, w_2)\|_U \le C{\cdot}\|w_1\|_{W_1}{\cdot}\|w_2\|_{W_2}$. Set $\|B\| = \sup\{\|B(w_1, w_2)\|_U : \|w_1\|_{W_1} \le 1, \|w_2\|_{W_2} \le 1$. Complete the proof.]

Theorem 3.9. (i) *A bilinear map $B{:}\,W_1 \times W_2 \to U$ is differentiable in every point (x, y) and has the derivative $D_{(x,y)}B$ given by $D_{(x,y)}B(u, v) = B(x, v) + B(u, y)$.*

(ii) For differentiabl functions f and g as in the discussion preceding the theorem, the composite function $F = B \circ (f \times g)$ is differentiable in (a, b) and has the derivative $D_{(a,b)}f = (D_{\big(f(a),g(b)\big)}B)(D_af \times D_bg)$ which is given by

$$(D_{(a,b)}F)(x,y) = B\big(D_af(x), g(b)\big) + B\big(f(a), D_bg(y)\big).$$

Proof . (i) Set $L(u,v) = B(u,y) + B(x,v)$. Then $B(x+u, y+v) = B(x,y) + B(x,v) + B(u,y) + B(u,v) = B(x,y) + L(u,v) + B(u,v)$. By the preceding lemma,

$$(*) \qquad \|B(u,v)\|_U \le \|B\|{\cdot}\|u\|_{W_1}{\cdot}\|v\|_{W_2}.$$

Now assume that $(u,v) \ne (0,0)$, say $u \ne 0$, the case $v \ne 0$ is similar. Consider the projection $P: W_1 \times W_2 \to W_1$, $P(u,v) = u$. Then $\|u\| = \|P(u,v)\| \le \|P\|{\cdot}\|(u,v)\|_{W_1\times W_2}$ with the operator norm $\|P\|$ of P (see 6.33) and with the product norm $\|(u,v)\|_{W_1\times W_2} = \max\{\|u\|_{W_1},\ \|v\|_{W_2}\}$ (cf. 3.49(iii) where this definition was used for a product of metric spaces). Since P is not the zero operator, $\|P\| \ne 0$ and thus

$$(**) \qquad \|(u,v)\|_{W_1\times W_2} \ge \|P\|^{-1}{\cdot}\|u\|_{W_1}.$$

Setting $\|B\|{\cdot}\|P\| = C$, from $(*)$ and $(**)$ we obtain

$$(\dagger) \qquad \frac{\|B(u,v)\|_U}{\|(u,v)\|_{W_1\times W_2}} \le \frac{\|B\|{\cdot}\|u\|_{W_1}{\cdot}\|v\|_{W_2}}{\|P\|^{-1}{\cdot}\|u\|_{W_1}} = C{\cdot}\|v\|_{W_1} \le C{\cdot}\|(u,v)\|_{W_1\times W_2}.$$

Since this tends to 0 for $(u,v) \to (0,0)$ the theorem follows from Definition 3.2.

(ii) This is an immediate consequence of (i) and the Chain Rule. □

If we take $W_1 = W_2 = U = \mathbb{K}$ and consider functions of one variable combined with the bilinear map $B: \mathbb{K} \times \mathbb{K} \to \mathbb{K}$ defined by $B(x,y) = xy$ then we find $F(x,y) = f(x)g(y)$ and $D_{(a,b)}F(x,y) = f'(a)g(b)x + f(a)g'(b)y$. The old product rule follows if we invoke the $\varphi: \mathbb{K} \to \mathbb{K}$, $\varphi(x) = F(x,x)$. Then $\varphi = F \circ \delta$ where $\delta: \mathbb{K} \to \mathbb{K}\times\mathbb{K}$ is the diagonal map defined by $\delta(x) = (x,x)$, a linear function. Thus $\varphi' = F' \circ \delta$, and we finally get $\varphi'(a) = (fg)'(a) = f'(a)g(a) + f(a)g'(a)$. This same argument can be carried out more generally if $m = n$ and $X = Y$. Again we set $\delta(x) = (x,x)$ and observe $\delta' = \delta$. This yields the consequence

Corollary 3.10. *Assume that $f: X \to W_1$ and $g: X \to W_2$ are functions on an open set X of some finite dimensional vector space into finite dimensional vector spaces and assume that f and g are differentiable in the inner point a of X; assume further that $B: \mathbb{R}^p \times \mathbb{R}^q \to U$ is a bilinear map into some finite dimensional vector space, then the function $x \mapsto B\big(f(x), g(x)\big)$ is differentiable and has the derivative $x \mapsto B\big((D_af)(x), g(a)\big) + B\big(f(a), (D_ag)(x)\big)$.* □

Directional derivatives and partial derivatives

A simple application of the chain rule arises from a specialisation of the general case. Let $f: X \to M \subseteq W$, $X \subseteq V$ be a function and let a be an inner point a of X. Let $e \in V$ be an arbitrary vector. Since a is an inner point, there is a $\delta > 0$ such that $|t| \le \delta$ implies $a + t{\cdot}e \in X$. We consider the curve $\gamma: [-\delta, \delta] \to W$, $\gamma(t) = f(a + t{\cdot}e)$ in W.

Definition 3.11. If γ is differentiable at 0, then the derivative $D_0\gamma$ of γ at 0 is a vector $\partial_{a;e}f$ in W defined by

$$\partial_{a;e}f = \lim_{\substack{t\to 0\\ t\neq 0}} \frac{1}{t}\big(f(a+t\cdot e) - f(a)\big) = \frac{d}{dt}f(a+t\cdot e)\bigg|_{t=0} = D_0\gamma.$$

The vector $\partial_{a;e}$ is called the *directional derivative* of f at a in the direction of $e \in V$. □

The symbol ∂ is spoken *"partial"*, and we shall presently see why; a German abbreviation is „del", apparently a transmogrification of "delta".

If a notation is used it must indicate the place a at which it is taken and the direction e into which "it points." The directional derivative is defined also for the "zero direction" $e = 0$ but it is the the zero vector and is not particularly interesting. Most often V is a normed space and e is a unit vector, i.e. $\|e\| = 1$.

A special situation arises if $V = \mathbb{R}^n$. Then we let e be one of the standard basis vectors

$$\begin{aligned} e_1 &= (1,0,\dots,0,\dots,0),\\ e_2 &= (0,1,\dots,0,\dots,0),\\ &\vdots\\ e_k &= (0,0,\dots,1,\dots,0),\\ &\vdots\\ e_n &= (0,0,\dots,0,\dots,1); \end{aligned}$$

where in the row e_k the element 1 is in the k-th position. These vectors yield n directional derivatives $\partial_{a;e_k}f$ which we abbreviate $(\partial_k f)(a)$; thus $(\partial_k f)(a) =$

$$\lim_{\substack{t\to 0\\ t\neq 0}} \frac{1}{t}\big(f(a_1,\dots,a_{k-1},a_k+t,a_{k+1},\dots,a_n) - f(a_1,\dots,a_{k-1},a_k,a_{k+1},\dots,a_n)\big),$$

$k = 1,\dots,n$. Each of these is a vector in W.

Definition 3.12. If $W = \mathbb{R}$, then the real number $(\partial_k f)(a)$, $k = 1,\dots,n$ is called the k-th *partial derivative* of the function $f\colon X \to \mathbb{R}$ at the inner point $a \in X$. Various notations are used:

$$\frac{\partial f}{\partial x_k}\bigg|_{x=a} \quad\text{or}\quad \frac{\partial f(x)}{\partial x_k}\bigg|_{x=a} \quad\text{or}\quad \frac{\partial f}{\partial x_k}(a) \quad\text{or}\quad (\partial_k f)(a).$$ □

If $V = \mathbb{R}^n$ and $W = \mathbb{R}^m$, then for each $x = (x_1,\dots,x_n) \in X$ the vector $f(x) \in W$ is of the form

$$f(x) = \begin{pmatrix} f_1(x_1,\dots,x_n)\\ f_2(x_1,\dots,x_n)\\ \vdots\\ f_m(x_1,\dots,x_n) \end{pmatrix},$$

and thus f is really an m-tuple of scalar valued functions $f_j\colon X \to \mathbb{R}$. If f has all directional derivatives $(\partial_k f)(a)$, $k = 1,\dots,n$, then we obtain a full $m \times n$ matrix of mn partial derivatives

$$\big(\partial_k f_j)(a)\big)_{\substack{j=1,\dots,m\\k=1,\dots,n}} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}\Big|_{x=a} & \cdots & \frac{\partial f_1}{\partial x_n}\Big|_{x=a} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}\Big|_{x=a} & \cdots & \frac{\partial f_m}{\partial x_n}\Big|_{x=a} \end{pmatrix}$$

What do all these partial derivatives have to do with the possible differentiability of f at a in the sense of Definition 3.2?

Proposition 3.13. (i) *Assume that the function $f: X \to M \subseteq W$, $X \subseteq V$ for finite dimensional normed vector spaces V and W is differentiable at the inner point a of X, and that e is any vector in V. Then the directional derivative of f at a in the direction of e exists and is equal to*

$$\partial_{a;e} f = (D_a f)(e) = f'(a)(e). \tag{4}$$

(ii) *Now let $V = \mathbb{K}^n$ and $W = \mathbb{K}^m$. Then the linear map $D_a f = f'(a): \mathbb{K}^n \to \mathbb{K}^m$ has the matrix*

$$(a_{jk})_{\substack{j=1,\dots,m\\k=1,\dots,n}}, \quad a_{jk} = (\partial_k f_j)(a) = \frac{\partial f_j}{\partial x_k}\Big|_{x=a}. \tag{5}$$

Proof. (i) We define $\gamma(t) = f(a + t{\cdot}e)$ for all $t \in [-\delta, \delta]$ for a sufficiently small δ and have $\partial_{a;e} f = \frac{d\gamma}{dt}\Big|_{t=0}$. Set $g(t) = a + t{\cdot}e$, $g: [-\delta, \delta] \to W$; then $g'(0)(t) = t{\cdot}e$ for all $t \in \mathbb{R}$ and $\gamma = f \circ g$. By the Chain Rule 3.7, for all $t \in \mathbb{R}$ we have $t{\cdot}\partial_{a;e} f = D_0(f \circ g)(t) = D_{g(0)} f \circ D_0 g(t) = D_a f\big(g'(0)(t)\big) = D_a f(t{\cdot}e) = t{\cdot}D_a f(e)$. It follows that $\partial_{a;e} f = D_a f(e)$.

(ii) Here we take $e = e_k$ and find $(\partial_k f)(a) = D_a(e_k)$ for all $k = 1, \dots, n$. We know from the definition of the matrix of $D_a f$ that its k-th column is precisely the image $(D_a f)(e_k)$ written as a column. But $(\partial_k f)(a) =$

$$\lim_{\substack{t\to 0\\t\neq 0}} \frac{1}{t}\big(f(a + t{\cdot}e_k) - f(a)\big) = \lim_{\substack{t\to 0\\t\neq 0}} \frac{1}{t}\cdot \begin{pmatrix} f_1(a + t{\cdot}e_k) - f_1(a) \\ \vdots \\ f_m(a + t{\cdot}e_k) - f_m(a) \end{pmatrix} = \begin{pmatrix} (\partial_k f_1)(a) \\ \vdots \\ (\partial_k f_m)(a) \end{pmatrix}.$$

Taken together, these two observations prove the claim. □

Let us stress this point again: In order to compute a partial derivative of a function $f: X \to \mathbb{R}$, $X \subseteq \mathbb{K}^n$ with respect to x_k at the point $(a_1, \dots, a_n)$ one fixes all coordinates except for the k-th, putting them equal to $a_1, \dots, a_{k-1}, a_{k+1}, \dots, a_n$, and one considers the one variable function $x \mapsto f(a_1, \dots, a_{k-1}, x, a_{k+1}, \dots, a_n)$; now one differentiates this one-variable function it at the point a_k as in Chapter 4 of Analysis I; the result is $(\partial_k f)(a)$. Once more: If we set $\varphi(t) = f(a_1, \dots, t, \dots, a_n)$ then $(\partial_k f)(a) = \varphi'(a_k)$.

Now the derivative $D_a f = f'(a)$ of a function f is computationally accessible, since we can compute the matrix coefficients immediately as partial derivatives of the coefficient functions f_j, and as we have observed, partial derivatives are computed as derivatives of one variable functions.

We have concluded, in particular, that the existence of a derivative $D_a f$ implies the existence of all directional derivatives and, in particular, all partial derivatives $(\partial_k f_j)(a)$. The converse however, is false; we want to illustrate that by constructung an example of a function $f\colon \mathbb{R}^2 \to \mathbb{R}$, $a = (0,0)$ which is not differentiable at $(0,0)$ but all directional derivatives exist at $(0,0)$, so certainly the two partial derivatives exist.

In order to understand better the construction of such examples let us return to the polar coordinate function $P\colon \mathbb{R} \times [0,\infty[\to \mathbb{R}^2$, $P(t,r) = (r\cos t, r\sin t)$ of 5.40(42) whose domain and codomain we have now extended with the result that P is no longer bijective. Now assume that a function $f\colon \mathbb{R}^2 \to \mathbb{R}$ is given to us. We obtain a new function $F = f \circ P : \mathbb{R} \times [0,\infty[\to \mathbb{R}$ with the property $F(t,0) = F(0,0)$ for all $t \in \mathbb{R}$ and $F(t,r) = F(t + 2\pi n, r)$, $n \in \mathbb{Z}$, $t \in \mathbb{R}$, $0 \le r$. Conversely, every function F with these properties can be written in the form $F = f \circ P$ with a uniquely determined function $f\colon \mathbb{R}^2 \to \mathbb{R}$.

Exercise E3.7. Prove the existence of f as asserted.
[Hint. Set $f(0,0) = F(0,0)$, and recall 5.40(42) which yields a function $P^{-1}\colon \mathbb{R}^2 \setminus \{(0,0)\} \to]-\pi,\pi] \times]0,\infty[$. Then use the given properties of F.]

In this way we wish to construct the function of our example we have announced. Indeed, we take functions $g\colon \mathbb{R} \to \mathbb{R}$ and $h\colon [0,\infty[\to \mathbb{R}$ such that $-g(t) = g(t+\pi)$ and $h(0) = 0$ for $t \in \mathbb{R}$. This implies, in particular, $g(t+2\pi) = g(t)$. The we define $F(t,r) = h(r)g(t)$ and obtain a uniquely determined function $f\colon \mathbb{R}^2 \to \mathbb{R}$ such that $f(r\cos t, r\sin t) = h(r)g(t)$ and $f(0,0) = 0$. We assume that the one-sided derivative $h'(0)$ of h in 0 exists and is different from 0. For $e = (\cos t, \sin t)$ we set $\varphi_e(r) = f(r{\cdot}e)$ and assert that $\varphi_e'(0) = \frac{d}{dr} f(r{\cdot}e)\big|_{r=0} = h'(0)g(t)$, because

$$f(r{\cdot}e) = \begin{cases} h(r)g(t) & \text{for } r > 0, \\ h(-r)g(t+\pi) & \text{for } r < 0. \end{cases}$$

The derivative at $r = 0$ on the right side of the function $r \mapsto f(r{\cdot}e)$ is $h'(0)g(t)$, and its derivative on the left side is $(-1)h'(0)g(t+\pi) = h'(0)g(t)$ because of $-g(t) = g(t+\pi)$. Thus the asserted directional derivative exists and equals $h'(0)g(t)$. In particular, $(\partial f/\partial x)_{(x,y)=(0,0)} = h'(0)g(0)$ and $(\partial f/\partial y)_{(x,y)=(0,0)} = h'(0)g(\pi/2)$. The function f is continuous at $(0,0)$, since $h'(0)$ exists. If it is also differentiable in $(0,0)$, then the derivative $D_{(0,0)}f$ has to be equal to $\big(h'(0)g(0), h'(0)g(\pi/2)\big) = h'(0)(g(0), g(\pi/2))$. Now one has great freedom in the selection of g. For instance, we can choose g so that $g(0) = g(\pi/2) = 0$, but that g is not identically 0. Then the function f has directional derivatives in all directions, but is *not* differentiable in $(0,0)$.

Exercise E3.8. (i) Consider the following function

$$f(x,y) = \begin{cases} xy^2/(x^2+y^2) & \text{for } (x,y) \neq (0,0), \\ 0 & \text{otherwise.} \end{cases}$$

Show that this is a continuous function having all directional derivatives everywhere and that it is not differentiable in $(0,0)$.

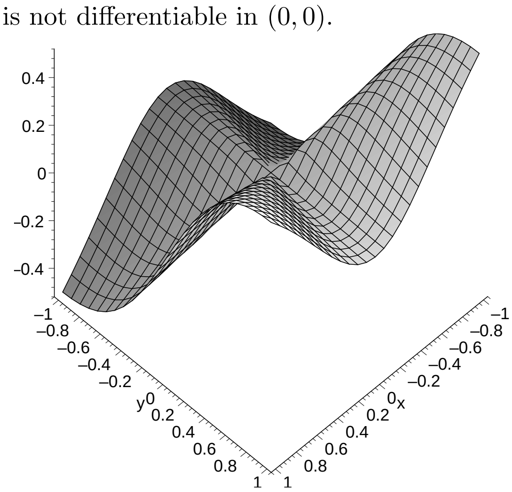

Figure 3.1

(ii) Define a closed subset $S \subseteq \mathbb{R}^2$ of the plane by

$$S = \{(x,y) \in \mathbb{R}^2 : y \leq 0 \text{ or } y \geq x^2\},$$

and let $f:\mathbb{R}^2 \to R$ denote the characteristic function of S, that is, $f(x,y) = 1$ if $(x,y) \in S$ and $= 0$ elswhere. Show that this function has all directional derivatives in $(0,0)$ but is discontinuous at $(0,0)$. The directional derivatives all vanish at $(0,0)$.

[Hint. (i) Apply our previous discussion with $h(r) = r$, $g(t) = (\cos t)(\sin^2 t)$. (ii) Let e be a unit vector in $\mathbb{R}^2$. Show that there is a $\delta > 0$ (depending on e) such that $|t| < \delta$ implies $t{\cdot}e \in S$ and thus $f(t{\cdot}e) = 0$.]

This situation may look a bit awkward. But it changes as soon as the partial derivatives exist in an entire neighborhood of the point a and are continuous in a. Indeed we have the following theorem:

Theorem 3.14. *Let $f: X \to \mathbb{R}^m$, $X \subseteq \mathbb{R}^n$ be a function and a an inner point of X. Assume that all partial derivatives exist on a neighborhood $U_r(a)$ of a and are continuous in a. Then f is differentiable in a and $D_a f$ has the matrix $\big((\partial_k f_j)(a)\big)_{\substack{j=1,\dots,m\\k=1,\dots,n}}$.*

Proof . First we note that f is differentiable in a as soon as all the coefficient functions $f_j: X \to \mathbb{R}$ are differentiable in a: Indeed let e_j, $j = 1,\dots,m$ denote the standard basis vectors of the range space $\mathbb{R}^m$; then $f(x) = f_1(x){\cdot}e_1 + \cdots +$

$f_m(x){\cdot}e_m = \big(f_1(x), \ldots, f_n(x)\big)$. We may therefore assume without restricting the generality that $m = 1$. We shall do this in the following. We also fix a norm, well aware of the fact that it is immaterial which one we fix.

We consider $x \in U_r(a)$ and notice

$$\begin{aligned} f(x) - f(a) &= f(x_1, a_2, \ldots, a_n) - f(a_1, \ldots, a_n) \\ &\quad + f(x_1, x_2, \ldots, a_n) - f(x_1, a_2, \ldots, a_n) \\ &\quad \vdots \\ &\quad + f(x_1, x_2, \ldots, x_n) - f(x_1, x_2, \ldots, x_{n-1}, a_n). \end{aligned}$$

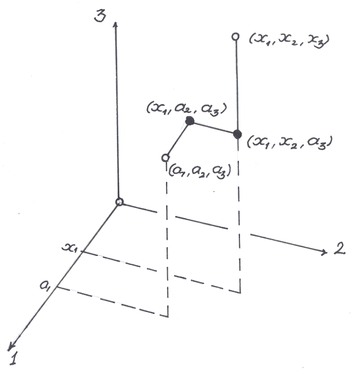

Figure 3.2

Since the partial derivatives p_j, $p_j(u) \stackrel{\text{def}}{=} \partial f/\partial x_j|_{x=u}$ of f exist on $U_r(a)$, the Mean Value Theorem 4.29 yields numbers t_j between a_j and x_j such that

$$\begin{aligned} &f(x_1, \ldots, x_{k-1}, x_k, a_{k+1}, \ldots, a_n) - f(x_1, \ldots, x_{k-1}, a_k, a_{k+1}, \ldots, a_n) = \\ &p_j(x_1, \ldots, x_{k-1}, t_k, a_{k+1}, \ldots, a_n)(x_k - a_k), \quad k = 0, \ldots, n. \end{aligned}$$

Then

$$f(x) = f(a) + \sum_{j=1}^{n} p_j(a)(x_j - a_j) + r(x)$$

where

$$r(x) = \sum_{j=1}^{n} \big(p_j(a^{(k)}) - p_j(a)\big)(x_j - a_j),$$

and

$$a^{(k)} = (x_1, \ldots, x_{k-1}, t_k, a_{k+1}, \ldots, a_n), \quad a = (a_1, \ldots, a_n).$$

We will finish the proof by showing that $\|x - a\|^{-1} r(x) \to 0$ for $x \to a$ with $x \neq a$ in $U_r(a)$. Now we notice that $a^{(k)}$ tends to a if $x = (x_1, \ldots, x_n)$ tends to a, because t_k is between a_k and x_k. Since the partial derivatives p_j are continuous at a, the function $p_j(a^{(k)}) - p_j(a)$ tends to 0 as x tends to a. Since all norms on

$\mathbb{R}^n$ are equivalent, there is a number $C > 0$ such that $C \cdot \|x - a\| \geq \|x - a\|_\infty = \max_\ell |x_\ell - a_\ell|$. Now we have $\|x - a\|^{-1}|x_j - a_j| \leq C\|x - a\|_\infty^{-1} \cdot |x_j - a_j| \leq C$ for all $j = 1, \ldots, n$, and then it follows that $\|x - a\|^{-1} r(x) \to 0$ for $x \to a$ with $x \neq a$ in $U_r(a)$. But this is what we had to show. □

It is worth emphasizing that from modest assumptions on partial derivatives which amounts to information on n directions only (if the domain is contained in n-space) we derive the strongest possible differentiability property, namely, differentiability itself.

Scalar valued functions on higher dimensional domains

In the general theory we considered functions from (open subsets of) $\mathbb{R}^n$ to $\mathbb{R}^m$. Several special situations arise:

(a) $n = 1$ and m arbitrary. This the case of curves which we considered in Chapter 2 above.

(b) $m = 1$ and n arbitrary. We encountered this in form of the coefficient functions f_j above; this special case captures most of the general features of the theory.

(c) $m = n$. This arises whenever we consider, for instance, self-maps of some open domain in $\mathbb{R}^n$. Issues of (local) invertibility of functions take place in this setting as we shall see below.

But now we turn the special case $m = 1$, n arbitrary which is opposite to that of curves. In other words, we consider functions $f: X \to \mathbb{R}$ with $X \subseteq E$, $E = \mathbb{R}^n$. The graph $G = \{(x, f(x)) \in E \times \mathbb{R} : x \in X\}$ of such a function may be visualized as a surface in $n + 1$ dimensional space $E \times \mathbb{R}$ projecting onto the base surface $X \times \{0\} \cong X$ with $f(x) \in \mathbb{R}$ denoting the "elevation," "height," or "level" of the point $(x, f(x))$ above the base plane. Therefore, such a function is sometimes called a *level function* or, in German, a *Höhenfunktion*. (See Figures 3.1 above and 3.3 below.) If a is an inner point of X, then f is differentiable in a, if there is a linear map $D_a f = df_a = f'(a): E \to \mathbb{R}$ and a function $r: X \to \mathbb{R}$ with $|r(x)|/\|x - a\| \to 0$ for $x \to a$, $x \neq a$ such that

$$f(x) = f(a) + f'(a)(x - a) + r(x).$$

By Proposition 3.13(5) the matrix of the linear map $D_a f$ is

$$\begin{aligned}\text{matrix of } D_a f &= \big((\partial_1 f)(a), \ldots, (\partial_n f)(a)\big) \\ &= \left(\frac{\partial f(x)}{\partial x_1}\bigg|_{x=a}, \ldots, \frac{\partial f(x)}{\partial x_n}\bigg|_{x=a} \right).\end{aligned}$$

The derivative $df(a) = D_a f: E \to \mathbb{R}$ is a linear form. However if we consider E as a real Hilbert space with the inner product $(x \mid u) = \sum_{j=1}^n x_j u_j$ there we have a unique vector $g \in E$ such that $(D_a f)(v) = (v|g) = (g|v)$. This calls for a name.

Definition 3.15. The unique vector g in the Hilbert space E for which $(\forall v \in E)\, df(a)(v) = (D_a f)(v) = (g|v)$ is called the *gradient* of f at a and is written $\operatorname{grad}_a f$, or $(\operatorname{grad} f)(a)$, or $\nabla_a f$, or $\nabla f|_{x=a}$. □

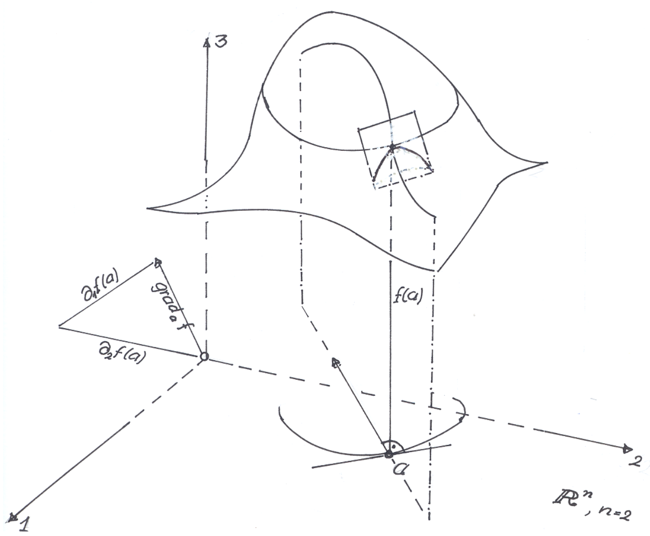

Figure 3.3

Gradient and directional derivative

In the n-tuple space $E = \mathbb{R}^n$ we have

$$\operatorname{grad}_a f = \big((\partial_1 f)(a), \ldots, (\partial_n f)(a)\big) \in \mathbb{R}^n.$$

With this notation, for a function f which is differentiable at a, we get the representation

$$f(x) = f(a) + \big(\operatorname{grad}_a f | (x-a)\big) + r(x) \quad \text{such that} \quad \lim_{\substack{x\to a\\ x\neq a}} |r(x)|/|x-a| = 0.$$

The affine approximation $x \mapsto f(a) + (\operatorname{grad}_a f \mid x - a)$ describes the behavior of f up to a very small error near a. In particular, we recall from 1.23(6) that

$$(\operatorname{grad}_a f|x-a) = \| \operatorname{grad}_a f\| \cdot \|x-a\| \cos w(\operatorname{grad}_a f, x-a)),$$
$$w(\operatorname{grad}_a f, x-a) = \text{nonoriented angle between } \operatorname{grad}_a f \text{ and } x-a.$$

(Cf. 1.21. The angle is undefined if $\operatorname{grad}_a f = 0$!)

We continue to consider a function $f\colon X \to \mathbb{R}$ which is differentiable at the inner point $a \in X \subseteq E$. From Definition 3.11 we recall the concept of the directional derivative $\partial_{a;e} f = \frac{df(a+t\cdot e)}{dt}\Big|_{t=0}$ of f at a in the direction of $e \in E$. In Proposition 3.13(4) we observed that $\partial_{a;e} f = (D_a f)(e)$. In the present situation this means that the directional derivative can be computed with the gradient:

Remark 3.16. For all $e \in E$ we have

$$\partial_{a;e} f = (\operatorname{grad}_a f|e).$$

In particular, if e is a unit vector then, assuming that the gradient does not vanish and using the nonoriented angle $w(u,v)$ between two nonzero vectors (see Definition 1.21), we can also write

$$\partial_{a;e} f = \|\operatorname{grad}_a f\| \cos w(\operatorname{grad}_a f, e). \tag{6}$$
□

If $e = e_k$ is the k-th standard basis vector of $\mathbb{R}^n$, then the directional derivative in the direction e_k is exactly $(\operatorname{grad}_a f | e_k) = \frac{\partial f(x)}{\partial x_k}\Big|_{x=a}$ the k-th partial derivative of f.

From (6) is is clear that (in case $\operatorname{grad}_a f \neq 0$) the directional derivative in the direction of e is maximal iff $\cos w(\operatorname{grad}_a f, e) = 1$ iff $w(\operatorname{grad}_a f, e) = 0$ iff $e = \|\operatorname{grad}_a f\|^{-1} \cdot \operatorname{grad}_a f$. In other words, the vector $\operatorname{grad}_a f$ points into the direction of the largest ascent of the function f, and its length is the directional derivative in that direction, that is, the rate of change in this direction. If on the other hand we select a unit vector which is perpendicular to $\operatorname{grad}_a f$, then $(\operatorname{grad}_a f \mid e) = 0$, that is, the directional derivative in the direction of e is zero. Hence the function $t \mapsto f(a + t \cdot e)$ is stationary at $t = 0$.

Level sets

In order to get an intuitive idea of the function $f\colon X \to \mathbb{R}$ we consider, for each $y \in \mathbb{R}$ the inverse images $f^{-1}(y) = \{x \in X : f(x) = y\}$. In our present context one speaks of *level sets* in X. For example, if $n = 2$, then the graph of f in $X \times \mathbb{R}$ is a surface in three-space lying above the planar region X, and the level sets are, as a rule, at least locally, the range of a curve, and are called *level lines* (German *Höhenlinien*) or *level curves* known from geographic maps.

The point a itself lies on the level set $f^{-1}\big(f(a)\big) = \{x \in X : f(x) = f(a)\}$. Then a point $f(x)$ is on this level set iff $(\operatorname{grad}_a f | x - a) = -r(x)$. Let us define the affine function $\alpha\colon E \to \mathbb{R}$ by $\alpha(x) = (\operatorname{grad}_a f | x - a)$. It is now plausible that in the case that $\operatorname{grad}_a f \neq 0$ the level set $f^{-1}\big(f(a)\big)$ is approximated near a by the level set $\alpha^{-1}\big(\alpha(0)\big)$, that is by the hyperplane

$$\{x \in X : (\operatorname{grad}_a f | x) = (\operatorname{grad}_a f | a)\}$$

which is perpendicular to the gradient $\operatorname{grad}_a f$ and which passes through a. This makes it also plausible that the level set may be described near a as the graph of a function; we shall later see the so-called "Implicit Function Theorem" which will allow us to prove rigorously our plausibility arguments. If $\operatorname{grad}_a f \neq 0$, then we can form the unit vector $e \stackrel{\text{def}}{=} \|\operatorname{grad}_a f\|^{-1} \cdot \operatorname{grad}_a f$. Thus the directional derivative $\partial_{a;e} f$ of f at a in the direction of e is precisely $(\operatorname{grad}_a f | e) = \|\operatorname{grad}_a f\| > 0$, and the function $t \mapsto f(a + te)$, defined for all sufficiently small t as a is an inner point of the domain X, has the derivative

$$\frac{df(a + t \cdot e)}{dt}\bigg|_{t=0} = \partial_{a;e} f = (\operatorname{grad}_a f | e) = \|\operatorname{grad}_a f\|.$$

According to Theorem 4.25 of Analysis I and since a is an inner point of X, there is a $\delta > 0$ such that $0 < t < \delta$ implies $f(a + t \cdot e) > f(a)$ and that $-\delta < t < 0$

implies $f(a+t\cdot) < f(a)$. We conclude that f cannot attain a local extremal value (Definition 4.26 of Analysis I) at a. We reformulate this as follows:

Proposition 3.17. *Let $X \subset \mathbb{R}^n$ and assume that a is an inner point of X. If the function $f: X \to \mathbb{R}^n$ is differentiable at a and attains in this point a local extremal value, then* $\operatorname{grad}_a f = 0$. □

Thus local extremal values are to be found at most in the stationary point*s* of f, that is the points at which the gradient vanishes. Sometimes these points are also called *critical* points. However, the function $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = x^2 - y^2$ has at $(0,0)$ a stationary point, but attains at $(0,0)$ neither a local minimum nor a local maximum. In fact, the function $t \mapsto f(t,0)$ attains in 0 a local minimum, the function $t \mapsto f(0,t)$, however, a local maximum. The two functions $t \mapsto f(t, \pm t)$ are constant. A stationary point with such properties is called a *saddle point.* The level lines of this function are the hyperbolas $\{(x,y) : x^2 - y^2 = r\}$, provided that $r \neq 0$. We will defer a more thorough analysis of the local behavior of a level function in a stationary point until we discuss higher derivatives.

Exercise E3.9. Prove the following assertion.
Proposition. *Let $f: X \to \mathbb{R}$, $X \subseteq E = \mathbb{R}^n$ be differentiable in the inner point a of X. Then the graph of the affine approximation $x \mapsto f(a) + (\operatorname{grad}_a f | x - a)$ of the function f at a is the tangent hyperplane T to the graph of the function f at the point $(a, f(a))$. If we identify the vector spaces $\mathbb{R}^n \times \mathbb{R}$ and $\mathbb{R}^{n+1}$, then the vector $(\operatorname{grad}_a f, -1) \in \mathbb{R}^n \times \mathbb{R} = \mathbb{R}^{n+1}$ is perpendicular to T.*
[Hint. The difference $r(x) = f(x) - \big(f(a) + (\operatorname{grad}_a f | x - a)\big)$ satisfies $\lim_{x\to a} \frac{|r(x)|}{\|x-a\|} = 0$. This justifies the term "tangent hyperplane." The graph of the affine approximation is $\{(x,y) \in \mathbb{R}^n \times \mathbb{R} = \mathbb{R}^{n+1} : y = (\operatorname{grad}_a f | x) + q\}$, $q = f(a) - (\operatorname{grad}_a f | a)$. The hyperplane through the orgin which is parallel to it has the equation $\big((\operatorname{grad}_a f, -1) | (x,y)\big) = 0$ with the inner product on $\mathbb{R}^{n+1}$.]

The Implicit Function Theorem

In this section we consider differentiable functions $f: U_1 \to U_2$ where U_1 and U_2 are open subsets of finite dimensional normed vector spaces V and W of the same dimension. (In fact, all arguments apply to the case of Banach spaces V and W if we take it for granted that a derivative is defined as a *continuous* linear map, which is not guaranteed in infinite dimensions.)

In many of the preceding sections we noted the importance of inverse functions, notably in the context of differentiable functions of one variable (cf. Proposition 4.18 through Exercise E4.8 in Analysis I). In the context of functions $f: I \to J$ between intervals f has an inverse $g: J \to I$ iff for all $y \in J$ the equation $y = f(x)$ has precisely one solution $x \in I$. In the case of a continuous f this property is equivalent with strict monotonicity. A sufficient condition was that f had everywhere a positive derivative. The investigation of inverses (at least locally) in higher dimensions therefore must concentrate on several questions:

(i) Assume that a differentiable function $f\colon U_1 \to U_2$ has an inverse function, is the inverse function differentiable and what is its derivative?

(ii) When does a given differentiable function f have an inverse function?

These issues are certainly no less important for functions in several variables than they are for functions of one variable. Indeed at stake is the solvability of entire *systems* of equations

$$\begin{array}{rcl} y_1 & = & f_1(x_1,\dots,x_n), \\ \vdots & \vdots & \vdots \\ y_m & = & f_m(x_1,\dots,x_n), \end{array}$$

where the y_k are given and we have to solve for the x_j.

As first order of business we deal with the differentiability of the inverse function of a differentiable function in case it does have an inverse.

Thus let $U_1 \subseteq V$ and $U_2 \subseteq W$ open subsets of finite dimensional normed vector spaces. Let us now assume that two functions $f\colon U_1 \to U_2$ und $g\colon U_2 \to U_1$ are inverse functions of each other. Further assume that f is differentiable in $a \in U$. Then $f(x) = f(a)+L(x-a)+r(x)$ with $L = D_af\colon V \to W$ and the usual remainder function r. We set $b = f(a)$ and thus $a = g(b)$.

As a first step we shall show that L has to be invertible *if* g is differentiable in b. If this is shown then $\dim V = \dim W$ follows. We have $g \circ f = \mathrm{id}_{U_1}$ and $f \circ g = \mathrm{id}_{U_2}$. By the Chain Rule 2.22. this implies $\mathrm{id}_V = \mathrm{id}'_{U_1}(a) = D_bg \circ D_af$. Since D_bg and D_af are linear maps between finite dimensional vector space, this suffices for $D_bg = (D_af)^{-1}$.

Now we assume, conversely, that $L = D_af$ is invertible. We set $f(x) = f(a) + L(x-a) + \|x-a\|{\cdot}R(x)$ with $R(x) \to 0$ for $x \to a$. We set $y = f(x)$ and derive $y - b = L(g(y)-g(b)) + \|g(y)-g(b)\|R\big(g(y)\big)$, that is

$$g(y) = g(b) + L^{-1}(y-b) - \|g(y)-g(b)\|{\cdot}L^{-1}R\big(g(y)\big).$$

Now we have

$$\|g(y)-g(b)\| = \|y-b\|{\cdot}\frac{\|g(y)-g(b)\|}{\|y-b\|} = \|y-b\|{\cdot}\frac{\|x-a\|}{\|f(x)-f(a)\|}.$$

Set $c = \min\{\|Lu\| : \|u\| = 1\}$; since L is invertible, this number is well defined. Let us consider x so closed to a that $\|R(x)\| < \frac{c}{2}$. Then for these x we have

$$\frac{\|f(x)-f(a)\|}{\|x-a\|} \ge \left|\|R(x)\| - \left\|L\left(\frac{x-a}{\|x-a\|}\right)\right\|\right| > \frac{c}{2}.$$

Therefore $\frac{\|x-a\|}{\|f(x)-f(a)\|}$ stays bounded for $x \to a$. If g is assumed to be continuous at b, then $y \to b$ and $x \to a$ are equivalent. Therefore $\|g(y)-g(b)\| = \|y-b\|{\cdot}B(y)$ with a function $B\colon U_2 \to \mathbb{R}$ which stays bounded for $y \to b$. Thus

$$g(y) = g(b) + L^{-1}(y-b) + \|y-b\|{\cdot}\big[B(y){\cdot}L^{-1}R\big(g(y)\big)\big].$$

It follows that g is differentiable in b and has the derivative L^{-1}. Thus we have proved the following result:

Proposition 3.18. *Assume that $U_1 \subseteq V$ und $U_2 \subseteq W$ are open sets in finite dimensional normed vector spaces and that $f: U_1 \to U_2$ and $g: U_2 \to U_1$ are inverse functions of each other. Further assume that f is differentiable in a and g is continuous in $b = f(a)$. Then the following two statements are equivalent.*

(i) *g is differentiable in b.*

(ii) *$D_a f$ is invertible.*

If these statements hold, then $D_b g = D_{f(a)} g = (D_a f)^{-1}$, and the vector spaces V and W are necessarily isomorphic. □

Let us recall that statement (ii) above is equivalent to

(iii) $\det D_a f \neq 0$.

Exercise E3.10. Compare Proposition 3.18 with the discussions of the one dimensional case in Proposition 4.18 in Analysis I.

The question of the *existence* of a (local) inverse function is harder, but more informative and much more fascinating. The result is a standard tool in analysis in all of its branches.

So let U and V be open subsets of $\mathbb{R}^n$, say, and consider a continuous function $f: U \to V$. We noted in 3.18 that, in the context of locally invertible differentiable functions it would be absurd to investigate real vector spaces of different dimensions. We will show that a suitable strong condition of differentiability of f at a point $a \in U$ with an invertible derivative implies that f maps an open neighborhood of a bijectively and continuously invertibly onto a neighborhood of $f(a)$. Recall that, conversely, the condition that f maps a sufficiently small open neighborhood $U_r(a)$ of a point $a \in U$ bijectively onto a neighborhood of $f(a)$ in V does not imply that f is differentiable in a. Indeed $\sqrt[3]{\ }: \mathbb{R} \to \mathbb{R}$ is continuous and bijective but fails to be differentiable at 0. Furthermore, its inverse function $x \mapsto x^3$ is smooth and bijective, but its derivative at 0 is not invertible.

For our purpose we consider an interesting variation of the concept of differentiability of a function at a point.

Definition 3.19. Let V and W be finite dimensional normed spaces. A function $f: X \to W$ with $X \subseteq V$ is called *strongly differentiable* at an inner point a of X, if there is a linear map $L: V \to W$ and a function $R: X \times X \to W$ so that

$$(7) \qquad f(u) - f(v) = L(u-v) + \|u-v\| \cdot R(u,v) \quad \text{and} \quad \lim_{(u,v)\to(a,a)} R(u,v) = 0. \quad \square$$

Taking $v = a$ we see at once that a function which is strongly differentiable at a is differentiable. The converse may fail. (See E2.21 below.)

The following theorem clarifies the situation.

Theorem 3.20. *Let V and W be finite dimensional normed vector spaces and X an open subset of V; assume that $f\colon X \to W$ is a differentiable function. Then for a point $a \in X$ the following statements are equivalent:*

(i) f' is continuous at a.
(ii) f is strongly differentiable at a.

Proof . In place of the function f we consider the function $F\colon X \to W$ defined by $F(x) = f(x) - f(a) - f'(a)(x-a)$. If we prove the equivalence of (i) and (ii) for F, then it is also secured for f, since the two functions differ only by an affine function. For the function F, however, we have

$$F(a) = 0 \qquad \text{and} \qquad F'(a) = 0.$$

We will now show that (i) implies the strong differentiability of F and that (ii) entails the continuity of F' at 0.

(i)⇒(ii): By the continuity of F' at a and in view of $F'(a) = 0$, for a given $\varepsilon > 0$ we find a $\delta > 0$ so that $w \in U_\delta(a)$ implies $\|F'(w)\| \le \varepsilon$ with the operator norm (s. 1.32, 1.33). The set $U_\delta(0)$ is a ball and thus is convex, i.e. for two points $u, v \in U_\delta(0)$ and $t \in [0,1]$ the point $w = (1-t).u + t{\cdot}v$ on the straight line segment S between them satisfies $\|w\| \le (1-t)\|u\| + t\|v\| < (1-t)\delta + t\delta = \delta$ also belongs to $U_\delta(0)$. By the Mean Value Theorem 2.10(∗∗) we have

$$\|F(v) - F(u)\| \le \varepsilon\|v-u\|$$

for $u, v \in U_\delta(a)$, and this shows, that F is strongly differentiable in a.

(ii)⇒(i): Since F is differentiable on X, for any $b \in X$ and $h \in V$ with $b+h \in X$ we know $F(b+h) - F(b) = F'(b)(h) + \|h\|{\cdot}r_b(h)$ such that

$$(\forall b \in X)(\forall \varepsilon)\big(\exists \delta = \delta(b,\varepsilon)\big)\|h\| < \delta \Rightarrow \|r_b(h)\| < \varepsilon.$$

By (ii), in view of $F'(a) = 0$ we have

$$F(b+h) - F(b) = F'(a)(h) + \|h\|{\cdot}R_a(b,h) = \|h\|{\cdot}R_a(b,h)$$

with $\lim_{(b,h)\to(a,0)} R_a(b,h) = 0$. We conclude

$$F'(b)(h) = \|h\|{\cdot}\big(R_a(b,h) - r_b(h)\big).$$

Now let $\varepsilon > 0$ be given. We select $\delta > 0$ so, that u, $u+h \in U_\delta(a)$ implies $\|R_a(u,h)\| < \varepsilon/4$. Let $b \in U_{\delta/2}(a)$. Then we determine a $\delta' = \delta'(b,\varepsilon)$ with $0 < \delta' < \delta/2$ so that $\|h\| < \delta'$ entails $\|r_b(h)\| < \varepsilon/4$. Then b and $b+h$ are still contained in $U_\delta(a)$. Hence we have $\|R_a(b,h)\| < \varepsilon/4$. Now let v be an arbitrary element of V. Set $h = \frac{\delta'}{\|v\|+1}{\cdot}v$. Then $\|h\| < \delta'$ and thus

$$\|F'(b)(h)\| \le \|h\|{\cdot}\|R_a(u,h) - r_b(h)\| \le \|h\|(\varepsilon/4 + \varepsilon/4) = \|h\|\varepsilon/2.$$

Hence $\|F'(b)(v)\| \le \|v\|\varepsilon/2$ and thus $|F'(b)\| \le \varepsilon/2 < \varepsilon$ for all $b \in U_{\delta/2}(a)$. This shows that F' is continuous at a. □

This theorem illustrates the significance of strong differentiability. If differentiability is secured on a neighborhood of a, then strong differentiability is a consequence of the continuity of the derivative; the theorem makes precise, to which extent the converse is true. Strong differentiability therefore is a concept applying to functions whose properties are known at one point a only, and it is being "continuously differentiable at this point". In particular the theorem implies the equivalence of the following two conditions Then the following statements are equivalent:
(i) $f' : X \to \mathrm{Hom}(V, W)$ *is a continuous differential form.*
(ii) f *is strongly differentiable on* X.

After these preparations we turn to the local invertibility of a function F which is strongly differentiable in an inner point a of its domain X.

We reduce the problem to a special, more managable case. For a function $f : U \to V$, where U and V are open subsets of a finite dimensional normed vector space E assume that f is differentiable at $a \in U$ such that $D_a f$ is an invertible vector space endomorphism of E. Now we define

$$F : U - a \to (D_a f)^{-1} V - f(a), \quad F(x) = (D_a f)^{-1}(f(x + a) - f(a)).$$

Then $F(0) = 0$, and $D_0 F$ exists and equals the identity map $\mathbf{1}_E$ of E. Also $f(u) = f(a) + (D_a f)\big(F(u - a)\big)$. Moreover, if f is strongly differentiable at a, then F is strongly differentiable at 0. We claim that if F is locally invertible near 0 then f is locally invertible near a: Indeed let $GF(x) = x$ and $FG(y) = y$ for all x, y near 0 then we set $g(y) = a + G\big[(D_a f)^{-1}\big(y - f(a)\big)\big]$ and quickly verify $g\big(f(x)\big) = x$ and $f\big(g(y)\big) = y$ for all x sufficiently close to a and all y sufficiently close to $b = f(a)$.

Thus we shall now assume that f is strongly differentiable at 0 i.e. there is a function $R : U \times U \to E$ such that

$$f(u) - f(v) = u - v + \|u - v\| \cdot R(u, v) \text{ such that } \lim_{(u,v)\to(0,0)} R(u, v) = 0$$

for $u, v \in U$. In particular, this implies

$$f(x) = x + \|x\| \cdot R(x, 0) \text{ such that } \lim_{x\to 0} R(x, 0) = 0$$

for $x \in U$.

Now we observe that for a sufficiently small numbers $r > 0$ we have
(a) $U_r(0) \subseteq U$,
(b) $\|R(u, v)\| \le 1/2$ holds for $\|u\|$, $\|v\| < r$.

We fix an element y with $\|y\| \le \frac{r}{2}$ and set $K(x) = x - f(x) + y$. Then $y = f(x)$ iff $K(x) = x$. Moreover,

$$\|K(u) - K(v)\| = \| - f(u) + f(v) + (u - v)\| \le \|u - v\| \cdot \|R(u, v)\| \le \frac{1}{2}\|u - v\|$$

for u, $v \in U$. We note that then
(c) $\|K(x)\| \le \|x\| \cdot \|R(x, 0)\| + \|y\| < \frac{r}{2} + \frac{r}{2} = r$.

We summarize that with the $r > 0$ so determined, we have

$$\text{(C)}\qquad \begin{aligned} \big(\forall x \in U_r(0)\big)\ \|K(x)\| &< r,\\ \big(\forall u,\, v \in U_r(0)\big)\ \|K(u) - K(v)\| &\le \frac{1}{2}\|u - v\|. \end{aligned}$$

Thus we have the following information on K: For each y with $\|y\| \le \frac{r}{2}$ the function K maps $U_r(0)$ into itself and properly contracts distance in the sense that $d(K(u), K(v)) < \frac{1}{2}d(u, v)$. This situation calls for an interlude on metric spaces.

The Banach Contraction Principle

Recall that a metric space is called *complete* if every Cauchy sequence converges (cf. Definition 6.5).

Definition 3.21. A self-map $K: X \to X$ of a metric space X is called a *proper contraction*, if there is a number c with $c < 1$ such that $d\big(K(x), K(y)\big) \le c{\cdot}d(x, y)$.
A point x satisfying $K(x) = x$ is called a *fixed point.* □

The function $K: U_r(0) \to U_r(0)$ introduced above as $K(x) = x - f(x) + y$ is a proper contraction.

For self-maps of metric space we have the following extremely useful result. ,

THE BANACH FIXED POINT THEOREM

Theorem 3.22. *A proper contraction of a complete metric space has a unique fixed point.*

Proof . Uniqueness: If $K(x) = x$ and $K(y) = y$ then $d(x, y) \le d\big(K(x), K(y)\big) \le c{\cdot}d(x, y)$, that is, $0 \le (1 - c)d(x, y) \le 0$; this implies $d(x, y) = 0$ and thus $x = y$ as asserted.

Existence. First note via induction, that for any two points x and y in X we have $d\big(K^n(x), K^n(y)\big) \le c^n d(x, y)$. Now let $x_0 \in X$ be completely arbitrary. Set $x_n = K^n(x_0)$. I.e., $x_{n+1} = K(x_n)$ for $n = 0, 1 \ldots$.. Then, by induction, we get

$$\begin{aligned} d(x_{n+k}, x_n) &\le d(x_n, x_{n+1}) + \cdots + d(x_{n+k-1}, x_{n+k}),\ \text{and thus}\\ d(x_{n+k}, x_n) &= d\big(K^{n+1}x_0, K^n x_0\big) + \cdots + d\big(K^{n+k}x_0, K^{n+k-1}x_0\big)\\ &\le (c^n + \cdots + c^{n+k-1}){\cdot}d\big(K(x_0), x_0\big) \end{aligned}$$

Looking at the geometric series $1 + c + c^2 + \cdots$ (see 4.7) we first note that $c^n + \cdots c^{n+k-1} \le \frac{c^n}{1-c}$, whence

$$\text{(8)}\qquad (\forall n, k = 0, 1, \ldots)\ d(x_{n+k}, x_n) \le c^n{\cdot}\frac{d\big(K(x_0), x_0\big)}{1 - c}.$$

This implies at once that $(x_n)_n$ is a Cauchy sequence; since X is complete, $x \stackrel{\text{def}}{=} \lim_n x_n = \lim_n K^n(x_0)$ exists. Since every proper contraction is clearly continuous we obtain $K(x) = K(\lim_n x_n) = \lim_n K(x_n) = \lim_n x_{n+1} = x$. Thus x is the required fixed point. □

We observe quickly that the proof actually yielded a more precise estimate how close the n-th iteration of $K^n(x_0)$ is to the limit:

A PRIORI ESTIMATE FOR A BANACH CONTRACTION

Corollary 3.23. *Let $d\big(K(x), K(y)\big) \leq c{\cdot}d(x, y)$ for all x, y in a complete metric space and set $x_n = K^n(x_0)$. Let $x = \lim x_n$ be the unique fixed point of K according to* 2.40*. Then*

$$(\forall n = 0, 1, \ldots)\, d(x, x_n) \leq c^n{\cdot}\frac{d\big(K(x_0), x_0\big)}{1 - c}. \tag{9}$$

Proof . This follows at once from (8) by letting x_{n+k} tend to x, $k \to \infty$. □

If we have any ideal where the fixed point might be located, then we would naturally pick x_0 near the likely location of x and thus make $d(K(x_0), x_0)$ small, as $d(K(y), y) \to 0$ when $y \to x$. But in the estimate (9), the factor $d(K(x_0), x_0)$ does not play a very significant role, whereas c^n does: This factor decreases exponentially to 0.

It is most remarkable how elementary these proofs are that yield such powerful results. The Banach Contraction Principle is of the utmost importance for applications in many branches of pure and paplied mathematics. It is constructive in the sense that it does not only prove the existence of a fixed point but actually allows us to *construct* a sequence of "iterates" starting from an arbitrarily selected initial element x_0 which takes us very quickly near the unique fixed point. The recursively defined sequence $x_{n+1} = K(x_n)$ is ideally set up for being programmed.

The Banach Fixed Point Theorem requires a *proper* contraction. It generally fails for self-maps f satisfying $d\big(f(x), f(y)\big) < d(x, y)$.

Back to Local Inverses

With the help of the Banach Contraction Principle we now derive very quickly the following intermediate result:

Lemma 3.24. *Let $f{:}\,U \to E \cong \mathbb{K}^n$ be defined on an open set $U \subseteq E$ containing* 0 *and assume that $f(0) = 0$ and that f is strongly differentiable in* 0 *with derivative $\mathbf{1}_E$. Then* (i) *there is an $r > 0$ such that for all $r' \in\,]0, r]$ it follows that $U_{r'}(0) \subseteq U$ and that for each $y \in \overline{U_{r'/2}(0)}$ there is a unique $x \overset{\text{def}}{=} g(y) \in U_{r'}(0)$ such that $f(x) = y$.*

(ii) *In particular, $\overline{U_{r'/2}(0)} \subseteq f\big(U_{r'}(0)\big)$, and thus f maps* every *neighborhood of* 0 *in U* onto *a neighborhood of* 0.

(iii) *Moreover, for $x_1, x_2 \in U_r(0)$ the equation $f(x_2) - f(x_1) = x_2 - x_1 + \|x_2 - x_1\|{\cdot}R(x_1, x_2)$ holds with $\|R(x_1, x_2)\| < \frac{1}{2}$ and $R(x_1, x_2) \to 0$ for $(x_1, x_2) \to (0, 0)$.*

Proof. We continue the notation we have introduced in the discussion leading us to (**C**); in particular, we choose $r > 0$ as we did in that discussion. Whatever holds for r is also true for every smaller positive number r' in place of r. Assume $\|y\| \leq r/2$. Now we apply the Banach Contraction Principle 3.22 to $K: U_r(0) \to U_r(0)$ as given by $K(x) = x - f(x) + y$ and find that there is a unique fixed point $x \in U_r(0)$; thus $K(x) = x$ and hence $f(x) = y$. This proves (i). Since every neighborhood of 0 contains a neighborhood $U_{r'}(0)$ for some $r' < r$, (ii) follows readily. Finally, (iii) follows from the choice of r, yielding condition (b) that precedes (**C**). □

Lemma 3.25. *In the circumstances of* Lemma 3.24*, the function* $g: U_{r/2}(0) \to U_r(0) \subseteq E$ *is strongly differentiable in* 0.

Proof. By 3.24(iii), for $x_1, x_2 \in U_r(0)$ we have

$$\begin{aligned} \|f(x_2) - f(x_1)\| =&\|x_2 - x_1 + \|x_2 - x_1\|{\cdot}R(x_1, x_2)\| \\ \geq&\big|\|x_2 - x_1\|| - \|x_2 - x_1\|{\cdot}\|R(x_2, x_1)\|\big| \\ =&\|x_2 - x_1\|\big|1 - \|R(x_2, x_1)\|\big| \geq \frac{1}{2}\|x_2 - x_1\|. \end{aligned} \tag{$*$}$$

Now let $y_1, y_2 \in U_{r/2}(0)$ and set $x_1 = g(y_1)$, $x_2 = g(y_2)$. Then $x_1, x_2 \in U_r(0)$ and $y_1 = f(x_1)$ and $y_2 = f(x_2)$. Accordingly, by 3.24(iii),

$$\|y_2 - y_1\| \geq \frac{1}{2}\|g(y_2) - g(y_1)\|, \text{ and}$$

$$\begin{aligned} g(y_2) - g(y_1) =&x_2 - x_1 = f(x_2) - f(x_1) - \|x_2 - x_1\|{\cdot}R(x_2, x_1) \\ =&y_2 - y_1 - \|g(y_2) - g(y_1)\|{\cdot}R\big(g(y_1), g(y_2)\big) \\ =&y_2 - y_1 + \|y_2 - y_1\|{\cdot}\rho(y_1, y_2), \end{aligned} \tag{$**$}$$

where

$$\rho(y_1, y_2) = \begin{cases} -\frac{\|g(y_2) - g(y_1)\|}{\|y_2 - y_1\|}{\cdot}\|R\big(g(y_1), g(y_2)\big) & \text{for } y_1 \neq y_2, \\ 0 & \text{for } y_1 = y_2. \end{cases}$$

Then by $(*)$, $\|\rho(y_1, y_2)\| \leq 2{\cdot}\big\|R\big(g(y_1), g(y_2)\big)\big\|$. Since $\big\|R\big(g(y_1), g(y_2)\big)\big| \leq \frac{1}{2}$ by $(*)$, after $(**)$ we have

$$\|g(y_2) - g(y_1) - y_2 - y_1\| \leq \|y_2 - y_1\|;$$

letting $y_1 = 0$ and $y_2 = y$ we see $\|g(y) - y\| \leq \|y\|$ which shows that g is continuous at 0. Hence $(y_1, y_2) \mapsto R\big(g(y_1), g(y_2)\big)$ tends to 0 as (y_1, y_2) tends to $(0, 0)$ and thus $\rho(y_1, y_2) \to 0$ as $(y_1, y_2) \to (0, 0)$. Then $(**)$ shows that g is strongly differentiable at 0 with derivative $\mathbf{1}$. □

Let us now collect the information we have on $f: U \to E$ under the given hypotheses that $f(0) = 0$, and that f is strongly differentiable at 0 with $D - 0f = \mathbf{1}_E$:

For all sufficiently small $r > 0$ we have $U_r(0) \subseteq U$ and there is a function $g: U_{r/2}(0) \to U_r(0)$ such that $f\big(g(y)\big) = y$ for all $y \in U_{r/2}(0)$ and that g is strongly differentiable at 0. Then Lemma 3.24 applies to g in place of f and shows that every neighborhood of 0 in $U_{r/2}(0)$ is mapped under g *onto* a neighborhood of 0. Thus $g\big(U_{r/2}\big)$ contains $U_s(0)$ for some $s \in\,]0, r]$. Then by definition of s, for each $x \in U_s(0)$ there is a $y \in U_{r/2}(0)$ such that $x = g(y)$. Hence $f(x) = f\big(g(y)\big) = y$ and then $g\big(f(x)\big) = g(y) = x$. Thus, if we set $W = f\big(U_s(0)\big) \subseteq U_{r/2}(0)$ then, firstly, W is a neighborhood of 0 and, secondly, the functions $\varphi: U_s(0) \to W$ and $\psi: W \to U_s(0)$, defined by $\varphi(v) = f(v)$, $\psi(w) = g(w)$ are inverse functions of each other.

We are ready for the major theorem in this context:

Inverse Function Theorem

Theorem 3.26. *Let X be an open set of a finite dimensional real Banach space E and $f: X \to E$ a function which is strongly differentiable at $a \in X$ such that $D_a f$ is invertible.*

Then for every sufficiently small neighborhood M of a in X there exists a neighborhood N of $b \stackrel{\text{def}}{=} f(a)$ in E contained in $f(X)$ such that there is a function $g: N \to M$, for which $f\big(g(n)\big) = n$ for all $n \in N$, and $g\big(f(m)\big) = m$ for all $m \in M$.

Moreover, g is strongly differentiable in b, and $D_b g = (D_a f)^{-1}$.

Proof. We define

$$F: U - a \to (D_a f)^{-1} V - f(a), \quad F(x) = (D_a f)^{-1}(f(x+a) - f(a)).$$

Our preceding discussion applies to F and yields neighborhoods M' and N' of 0 and a function $G: N' \to M'$ such that $F\big(G(m)\big) = m$ for all $m \in M'$, and $G\big(G(n)\big) = n$ for all $n \in N'$. We set $M = M' + a$ and $N = (D_a f)N' + b$. Then M is a neighborhood of a and N is a neighborhood of b.

Now $f(x) = f(a) + (D_a f)\big(F(x-a)\big)$; thus f maps M to $f(a) + (D_a f)^{-1} F M' = N$. We defined $g: M \to N$ by $g(y) = a + G\big[(D_a f)^{-1}\big(y - f(a)\big)\big]$. Then g is correctly defined and maps N to M. The required properties of f and g, as we said at the beginning of our discussion are quickly verified.

Since G is strongly differentiable at 0 with derivative $\mathbf{1}_E$, then g is strongly differentiable at b and $D_b g = (D_a f)^{-1}$. □

The proof remains intact verbatim if E is replaced by an arbitrary, not necessarily finite dimensional Banach space; we just have to make sure that in the definition of strong differentiability in 3.19 the linear map $L: E \to E$ is continuous.

It is a fact which is proved in Banach space theory that the inverse of a continuous linear self-map of a Banach space, it if exists, is also continuous.

Even in one dimension we had an example presented in Exercise E4.8 in Analysis I that showed that the differentiability of f in 0 in which we constructed a bijective function $f: f \to f$ with $f(0) = 0$ and $D_0 f = 1$ such that the inverse function g was discontinuous at g—let alone differentiable. This illustrates the fact that differentiability in one point alone is not sufficient—even if we know bijectivity beforehand. If one feels the impulse to relax the hypotheses of the preceding Theorem 3.26 one should recall this example and also the following one:

Exercise E3.11. Consider the function $f\colon \mathbb{R} \to \mathbb{R}$ given by

$$f(x) = \begin{cases} 2x^2 \cos \frac{1}{x} + x & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

Show that f is differentiable everywhere and strongly differentiable in each point of $\mathbb{R} \setminus \{0\}$, but fails to be strongly differentiable and to be locally invertible at 0.

Show that $f'(0) = 1$ but fails to have a local inverse at 0 by showing that f is injective on no interval $]-\delta, \delta[$, $\delta > 0$.
[Hint. Compute

$$f'(x) = \begin{cases} 4x \cos \frac{1}{x} + 2 \sin \frac{1}{x} + 1 & \text{if } x \neq 0, \\ 1 & \text{if } x = 0. \end{cases}$$

Note that f' is discontinuous at 0 and is continuous on $\mathbb{R} \setminus \{0\}$. Invoke Theorem 3.26 to see that f is strongly differentiable in all points of $\mathbb{R} \setminus \{0\}$ but fails to be strongly continuous at 0. (Alternatively, consider the sequences $u_n = \frac{1}{2n\pi}$ and $v_n = \frac{1}{(2n+1)\pi}$ and compute $\big(f(u_n) - f(v_n)\big)(u_n - v_n)^{-1}$.) Shows that every interval $]-\delta, \delta[$ contains an x such that $f'(x) < 0$. Conclude that f is not monotone on any such interval.]

In the proof of the Inverse Function Theorem 3.26 one encounters a relatively typical situation: The problem is reduced to finding a fixed point of a self-map, and then a restriction to a small domain is picked so that the hypotheses of the Banach Fixed Point Theorem apply: See (**C**) above.

Existence of Local Inverse Functions

Corollary 3.27. *Let X be an open set of a finite dimensional real Banach space E and $f\colon X \to E$ a continuously differentiable function, that is, $f'\colon X \to \mathrm{Hom}(E, E)$ is continuous. Assume that $f'(x)$ is invertible for all $x \in X$. Then $f(U)$ is open for every open subset of U; in particular the set $Y = f(X)$ is open in E. Moreover, there is a family $\mathcal{U}$ of open subsets of X such that the following conditions are satisfied:*

(i) $X = \bigcup \mathcal{U} = \bigcup_{U \in \mathcal{U}} U$, *and*
(ii) *for each $U \in \mathcal{U}$ there is a continuously differentiable inverse function denoted by $g_U\colon f(U) \to U$ of the restriction $f|U\colon U \to f(U)$, and $g_U'(v) = f'\big(g(v)\big)^{-1}$ for all $v \in f(U)$.*

Proof. By Theorem 3.20, f is strongly differentiable at every point. Thus Theorem 3.25 applies to every point $a \in X$. Hence if U is open in X and $u \in U$, then by 3.25, there is an open neighborhood M of u such that $f(M) \subseteq f(U)$ is a neighborhood of $f(u)$ in E. Hence each $f(u)$ is an inner point of $f(U)$; that is, $f(U)$ is open. This says that f maps every open set onto an open set.

Furthermore, for each $x \in X$, Theorem 3.25 gives us open sets M_x and $N_x \stackrel{\text{def}}{=} f(M_x)$ in E such that $x \in M_x$, and there is a function $g_x : N_x \to M_x$ inverting $f|M_x : M_x \to N_x$ and being strongly differentiable at $f(x)$. Let $m \in M_x$. Then by 3.25 there is an $s > 0$ so small that $U_s(m) \subseteq M_x$ and that $f\big(U_s(m)\big) \subseteq N_x$ is a neighborhood of $f(m)$ for which there is an inverse $g_* : f\big(U_s(m)\big) \to U_s(m)$ of $f|U_s(m)$ such that g_* is strongly differntiable at $f(m)$. For $m' \in U_s(m)$ we note $g_*\big(f(m')\big) = m' = g_x\big(f(m')\big)$. Thus g_* and g_x agree on $f\big(U_s(m)\big)$. Hence g_x is strongly differentiable at $f(m)$ for all $m \in M_x$. Thus by Theorem 3.26 the function $g_x : N_x \to M_x$ is continuously differentiable. Set $\mathcal{U} \stackrel{\text{def}}{=} \{M_x : x \in X\}$. Then (i) and (ii) are satisfied. □

One should not believe that under the circumstance of Corollary 3.27, f is invertible, and one should not believe that for $U_1, U_2 \in \mathcal{U}$, one could guarantee that $g_{U_1}(y) = g_{U_2}(y)$ for $y \in f(U_1) \cap f_2(U)$. See E3.12(b) below.

Exercise E3.12. (Complex functions considered as real functions).

Fill in the details in the following discourse and explain, how Sections (a) and (b) below are related.

(a) Let U be an open subset of $\mathbb{C}$ and $f : U \to \mathbb{C}$ a holomorphic function (cf. Definition 4.1 of Analysis I); this means that there are functions $f' : U \to \mathbb{C}$ and $r : U \times U \to \mathbb{C}$ such that $r(z, a) \to 0$ for $z \to a$ and

$$(*) \qquad f(z) = f(a) + f'(z)(z - a) + |z - a| r(z, a).$$

Define an open set $V \subseteq \mathbb{R}^2$ by $V = \{(x, y) : x + iy \in U\}$. A function $F : V \to \mathbb{R}^2$, $F(x, y) = \big(u(x, y), v(x, y)\big)$ is differentiable on V (cf. Definition 2.16) if there are functions $F' : V \to \operatorname{Hom}(\mathbb{R}^2, \mathbb{R}^2)$ and $R_{(a_1,a_2)} : V \to \mathbb{R}^2$ such that $R_{(a_1,a_2)}\big((x, y)\big) \to 0$ for $(x, y) \to (a_1, a_2)$ and

$(**)$

$$F(x, y) = F(a_1, a_2) + F'(x, y)(x - a_1, y - a_2) + \|(x - a_1, y - a_2)\| R_{(a_1,a_2)}\big((x, y)\big)),$$

where $\|(\xi, \eta)\| = \sqrt{\xi^2 + \eta^2}$ denotes the euclidean norm, and where

$$(\dagger) \quad F'(x, y)(\xi, \eta) = \big((\partial_1 u)(x, y)\xi + (\partial_2 u)(x, y)(\eta), (\partial_1 v)(x, y)\xi + (\partial_2 v)(x, y)(\eta)\big).$$

Now return to f and write $u(x, y) = \operatorname{Re} f(x + yi)$ and $v(x, y) = \operatorname{Im} f(x + yi)$, further $\varphi(x, y) = \operatorname{Re} f'(x + yi)$ and $\psi(x, y) = \operatorname{Im} f'(x + yi)$. Then

$$(\ddagger) \qquad f'(x + iy)(\xi + \eta i) = (\varphi(x, y)\xi - \psi(x, y)\eta) + (\varphi(x, y)\eta + \psi(x, y)\xi)i.$$

If we now set $F(x, y) = \big(u(x, y), v(x, y)\big)$, then the holomorphy of f via comparison of (8) and $(**)$ causes F to be differentiable, and via $(\dagger)$ and $(\ddagger)$ yields

$$\text{matrix of } D_{(x,y)}F = \begin{pmatrix} (\partial_1 u)(x, y) & (\partial_2 u)(x, y) \\ (\partial_1 v)(x, y) & (\partial_2 v)(x, y) \end{pmatrix} = \begin{pmatrix} \varphi(x, y) & -\psi(x, y) \\ \psi(x, y) & \varphi(x, y) \end{pmatrix}.$$

In particular,

$$\text{(CR)}\qquad \begin{aligned}(\forall x+iy\in U)\,(\partial_1 u)(x,y) &= \ (\partial_2 v)(x,y),\\ (\partial_1 v)(x,y) &= -\,(\partial_2 u)(x,y).\end{aligned}$$

These equations are called the *Cauchy-Riemann partial differential equations.* The derivative $D_{(x,y)}F$ is invertible iff $\det D_{(x,y)}F \neq 0$ iff $|f'(z)|^2 = \varphi(x,y)^2 + \psi(x,y)^2 \neq 0$ iff $f'(z)\neq 0$. the the function F and hence the function f have local inverses at all points (x,y), respectively $z = x+yi$ for which $f'(z)\neq 0$.

In particular, *any holomorphic function $f\colon U\to\mathbb{C}$ for which f' vanishes nowhere on U gives a function $F\colon V\to\mathbb{R}^2$ with local inverses.* But there are many instances where the function $z\mapsto f(z) : U\to f(U)$ is not invertible, e.g. $z\mapsto z^2\colon \mathbb{C}\setminus\{0\}\to\mathbb{C}\setminus\{0\}$ or $\exp\colon C\to\mathbb{C}\setminus\{0\}$.

(b) Let $X = \,]1,\log 2[\,\times\,]-\frac{3\pi}{2},\frac{3\pi}{2}[\,\subseteq\mathbb{R}^2$ and $Y = \{(x,y)\in\mathbb{R}^2 : 1 < x^2+y^2 < 4\}$. Define $f\colon X\to Y$ by $f(x,y) = \big(\operatorname{Re}\big(\exp(x+iy)\big),\operatorname{Im}\big(\exp(x+iy)\big)\big) = e^x{\cdot}(\cos y,\sin y)$. Then all partial derivatives exist and are continuous, yielding for the linear map $f'(x,y)\colon\mathbb{R}^2\to\mathbb{R}^2$ the matrix

$$\begin{pmatrix} e^x\cos y & -e^x\sin y\\ e^x\sin y & e^x\cos y\end{pmatrix} = e^x{\cdot}\begin{pmatrix}\cos y & -\sin y\\ \sin y & \cos y\end{pmatrix}.$$

Since $\det\begin{vmatrix}\cos y & -\sin y\\ \sin y & \cos y\end{vmatrix} = 1$ and $e^x\neq 0$ for all x, the derivative is invertible at all (x,y). But f is not injective, hence not invertible, because $f(\log\frac{3}{2},-\pi) = (-\frac{3}{2},0) = f(\log\frac{3}{2},\pi)$. We may take $\mathcal{U} = \{U_1,U_2\}$ with

$$\begin{aligned} U_1 &=]1,\log 2[\times I_1, \quad I_1 = \left]-\frac{3\pi}{2},\frac{\pi}{2}\right[,\\ U_2 &=]1,\log 2[\times I_2, \quad I_2 = \left]-\frac{\pi}{2},\frac{3\pi}{2}\right[.\end{aligned}$$

Write $\mathbb{S}^1 = \{(x,y)\in\mathbb{R}^2 : x^2+y^2 = 1\}$. For $j = 1,2$, the functions $t\mapsto(\cos t,\sin t)\colon I_j\to\mathbb{S}^1\setminus\{p_j\}$, $p_j = (0,(-1)^{j-1})$ have inverse functions α_j. Then we have two slit annular regions $V_1 = f(U_1)$, $V_2 = f(U_2)$ and inverse functions $g_j\colon V_j\to U_j$, $j = 1,2$, $g_j(u,v) = (\log\sqrt{u^2+v^2},\alpha_j\big(\|(u,v)\|^{-1}(u,v)\big)$. But for $\xi = (-\frac{3}{2},0)\in V_1\cap V_2$ we have

$$g_1(\xi) = (\log\frac{3}{2},-\pi)\text{ and } g_2(\xi) = (\log\frac{3}{2},\pi). \qquad \square$$

Prove the following result complementing 3.11(a):

Theorem. *Let V be open in $\mathbb{R}^2$ and consider a function $F\colon V\to\mathbb{R}^2$, $F(x,y) = \big(u(x,y),v(x,y)\big)$. Assume that all partial derivatives of u and v exist and are continuous on V. Set $U = \{x+yi : (x,y)\in V\}$ and $f(x+yi) = u(x,y)+v(x,y)i$. Then $f\colon U\to\mathbb{C}$ is holomorphic iff the Cauchy-Riemann partial differential equations* (CR) *hold.* □

[Hint. We saw that the (CR)-equations are necessary in 6.81. For the converse, recall that the continuity of the partial derivatives by 3.14 implies the representation 2.46(∗∗). This representation and (CR) imply the representation 2.46(∗).]

From Exercise E3.12(b) we learned that local invertibility does not imply invertibility. The local invertibility itself, however, has remarkable consequences. An easy one first!

Corollary 3.28. *Let X be an open set of a nonzero finite dimensional real Banach space E and $f: X \to E$ a continuously differentiable function such that $f'(x)$ is invertible for all $x \in X$. Then the function $x \mapsto \|f(x)\|$ attains no local maximum.*

Proof. Since $\|\cdot\|: E \to [0, \infty[$ maps open sets to open sets (Exercise!), by Corollary 3.26 $x \mapsto \|x\| : X \to [0, \infty[$ maps open sets to open sets. Assume that this function takes a local maximum at a. Then there is an open ball U around a such that $\|f(a)\| \in \|f(U)\| \subseteq]\|f(a)\| - \delta, \|f(a)\|]$ and thus $\|f(U)\|$ would not be open. □

Exercise E3.13. (i) Prove that for every nonzero normed vector space E, the function $x \mapsto \|x\|: E \to [0, \infty[$ maps open sets to open sets.

(ii) Prove elementarily the following statement

Let X be an open set in a finite dimensional Banach space E and $f: X \to E$ is differentiable in $a \in X$ such that $D_a f$ is invertible. Then $x \mapsto \|f(x)\|$ does not attain a local maximum at a.

[Hint. First assume $f(a) \neq 0$, set $v = (D_a f)^{-1} \cdot f(a) \neq 0$ and $e = \|v\|^{-1} \cdot v$ is a unit vector such that $\partial_{a;e} f = (D_a f)(e) = \|v\|^{-1} \cdot f(a)$. Then $\varphi(t) \stackrel{\text{def}}{=} \|f(a + t \cdot e) = \|(1 + \frac{t}{\|v\|}) \cdot f(a) + t \cdot r(t)\|$ with $r(t) \to 0$ for $t \to 0$. Conclude that φ does not attain a local maximum at $t = 0$. If $f(a) = 0$ and $\|f(\cdot)\|$ attains a local maximum at a then f is locally constant equal to 0 at a; this entails $D_a f = 0$.]

The Implicit Function Theorem

From the Inverse Function Theorem we derive a theorem that has many applications; it will turn out that it is equivalent to the Inverse Function Theorem.

We shall consider open sets $X \subseteq \mathbb{R}^p$ and $Y \subseteq \mathbb{R}^q$, and a function $F: X \times Y \to \mathbb{R}^m$. The simplest example is a level function $F: \mathbb{R}^2 \to \mathbb{R}$. Let $(a, b) \in X \times Y$; we raise the question whether there are neighborhoods U and V of a and b in X, respectively, Y, and a function $f: U \to V$ with $f(a) = b$ such that $F\big(x, f(x)\big) = F(a, b)$ for all $x \in U$. Such a function f will be called an *implicitly defined function.* As an example let us look at $F: \mathbb{R}^2 \to \mathbb{R}$, $F(x, y) = x^2 + y^2$, and $F(a, b) = 1$. If $-1 < a < 1$, then the equation $x^2 + f(x)^2 = 1$ can be solved for $f(x)$, yielding $f(x) = \sqrt{1 - x^2}$ for $-1 < x < 1$. If $a = -1$ or $a = 1$, such a solution function does not exist.

We assume that F is strongly differentiable at (a, b). We define functions $F_1: X \to \mathbb{R}^m$ and $F_2: Y \to \mathbb{R}^m$, given by $F_1(x) = F(x, b)$ and $F_2(y) = F(a, y)$. If $I_1(x) = (x, b)$ and $I_2(y) = (a, y)$, then $F_1 = F \circ I_1$ and and $F_2 = F \circ I_2$. Thus F_1 and F_2 are differentiable at a, respectively b. We denote their derivatives by

$\partial_1 F(a,b)$ and $\partial_2 F(a,b)$, respectively. Then, by the chain rule, $[\partial_1 F(a,b)](u) = (D_{(a,b)}F \circ D_a I_1)(u) = D_{(a,b)}F(u,0)$, similarly and $[\partial_2 F(a,b)](v) = D_{(a,b)}F(0,v)$. Thus

$$D_{(a,b)}F(u,v) = \partial_1 F(a,b)(u) + \partial_2 F(a,b)(v).$$

We shall call $\partial_1 F(a,b)$ and $\partial_2 F(a,b)$ the first, respectively, second *partial derivatives* of F at (a,b).

Now we assume that the second partial derivative $\partial_2 F(a,b)\colon \mathbb{R}^q \to \mathbb{R}^m$ is invertible. This means, in particular, that we assume $q = m$. The matrix of $D_{(a,b)}F$ is an $m \times (p+q)$-matrix $(m = q)$ which consists of an $m \times p$-block, the matrix of $\partial_1 F(a,b)$, and of an $m \times q$-block, the matrix of $\partial_2 F(a,b)$, and we notice that this latter one is quadratic.

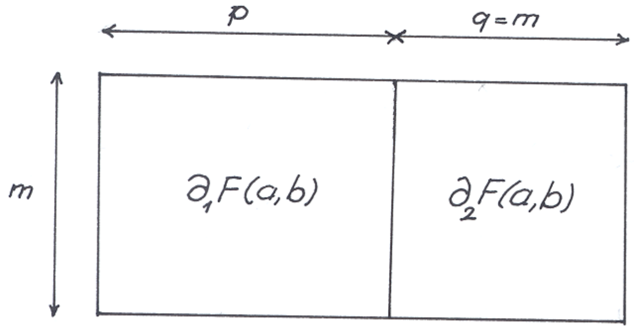

Figure 3.4

We now introduce a new function $G\colon X \times Y \to \mathbb{R}^p \times \mathbb{R}^m$ defined by $G(x,y) = \bigl(x, F(x,y)\bigr)$. Then G is strongly differentiable in (a,b) and its derivative is given by

$$D_{(a,b)}G(u,v) = \bigl(u, \partial_1 F(a,b)(u) + \partial_2 F(a,b)(v)\bigr).$$

For a given vector $(s,t) \in \mathbb{R}^p \times \mathbb{R}^q$, the equation $(s,t) = (u, \partial_1 F(a,b)(u) + \partial_2 F(a,b)(v))$ is easily solved for (u,v) by $u = s$ and $v = \bigl(\partial_2 F(a,b)\bigr)^{-1}\bigl(t - \partial_1 F(a,b)(s)\bigr)$. Therefore $D_{(a,b)}G$ is invertible and has an inverse given by

$$(10) \qquad (D_{(a,b)}G)^{-1}(s,t) = \Bigl(s, \bigl(\partial_2 F(a,b)\bigr)^{-1}\bigl(t - \partial_1 F(a,b)(s)\bigr)\Bigr).$$

Since F is strongly differentiable at (a,b) so is G. Hence the Theorem of the Existence of Local Inverse Functions 3.25 applies. Thus for all sufficiently small neighborhoods $U \times V$ of (a,b) in $X \times Y$ there exists a neighborhood W of $G(a,b) = \bigl(a, F(a,b)\bigr)$ in $\mathbb{R}^p \times \mathbb{R}^m$ and a unique function $H\colon W \to U \times V$ inverting $G|(U \times V) : U \times V \to W$. Since G is of the form $G(x,y) = \bigl(x, F(x,y)\bigr)$, the function H is of the form $H(u,v) = \bigl(u, h(u,v)\bigr)$ for $(u,v) \in W$ with a function $h\colon W \to \mathbb{R}^m$ which is strongly differentiable in $\bigl(a, F(a,b)\bigr)$. Then $(u,v) = GH(u,v) = G\bigl(u, h(u,v)\bigr) =$

$\big(u, F\big(u, h(u,v)\big)\big)$, that is

$$(\forall (u,v) \in W)\, v = F\big(u, h(u,v)\big). \tag{11}$$

Furthermore,

$$D_{(a,F(a,b))}H(s,t) = (D_{(a,b)}G)^{-1}(s,t) = \Big(s, \big(\partial_2 F(a,b)\big)^{-1}\big(t - \partial_1 F(a,b)(s)\big)\Big)$$

by (10). Hence $D_{(a,b)}h(s,t) = \big(\partial_2 F(a,b)\big)^{-1}\big(t - \partial_1 F(a,b)(s)\big)$.

There is an open set $W' \subseteq W$ containing $\big(a, F(a,b)\big)$. Then the set $\{x \in U : \big(x, F(a,b)\big) \in W'\}$ is an open neighborhood U_a of a, contained in U. Finally, we set $f(x) = h\big(x, F(a,b)\big)$ for $u \in U_a$. Then $F\big(x, f(x)\big) = F\big(x, h\big(x, F(a,b)\big)\big) = F(a,b)$ by (11) and f satisfies the requirement. Moreover, $D_a f(s) = (D_{(a,F(a,b))}h)(s,0) = \big(\partial_2 F(a,b)\big)^{-1}\big(0 - \partial_1 F(a,b)(s)\big) = -\big(\partial_2 F(a,b)\big)^{-1}\partial_1 F(a,b)(s)$, and thus $D_a f = -\big(\partial_2 F(a,b)\big)^{-1}\partial_1 F(a,b)(s)$. Finally, if $F(x,y) = F(a,b)$ for $(x,y) \in U_a \times V$, then $G(x,y) = \big(x, F(x,y)\big) = \big(x, F(a,b)\big)$ and thus $(x,y) = H\big(x, F(a,b)\big) = \big(x, h(x, F(a,b)\big) = \big(x, f(x)\big)$. Hence $f(x)$ is the unique solution of the equation $F(x,y) = F(a,b)$ with $x \in U_a$ and $y \in V$.

IMPLICIT FUNCTION THEOREM

Theorem 3.29. *Let $X \subseteq \mathbb{R}^p$ and $Y \subseteq \mathbb{R}^m$ be open sets and $F\colon X \times Y \to \mathbb{R}^m$ a function which is strongly differentiable at $(a,b) \in X \times Y$. We assume that the second partial derivative $\partial_2 F(a,b)\colon \mathbb{R}^m \to \mathbb{R}^m$ of F is invertibale. Then there are an open neighborhood U_a of a in X, an open neighborhood V of b in Y, and a function $f\colon U_a \to Y$ with $f(a) = b$ such that $F\big(x, f(x)\big) = F(a,b)$ holds for all $x \in U_a$, and $f(x)$ is the unique solution of the equation $F(x,y) = F(a,b)$ with $x \in U_a$ and $y \in V$.*

Moreover, f is strongly differentiable in a and

$$D_a f = -\big(\partial_2 F(a,b)\big)^{-1}\partial_1 F(a,b). \tag{12}$$ □

There is simple way to memorize (12). Once one has the function f and its properties, one differentiates both sides of the equation $F\big(x, f(x)\big) = 0$ and finds $\partial_1 F(a,b) + \partial_2 F(a,b) \circ D_a f = 0$ from which we obtain (12).

Level sets revisited

With the Implicit Function Theorem we can make our discussion of the level lines of a level function $f\colon X \to \mathbb{R}$, X open in $\mathbb{R}^n$ precise. The level set of level $f(a)$ is the set $H = \{x \in X : f(x) = f(a)\}$ where the graph of f attains the "height" $f(a)$. If f is continuously differentiable, then it is strongly differentiable in each point x of X by 3.20.

Consider a point $a \in X$ for which $\frac{\partial f}{\partial x_1}\Big|_{x=a} \neq 0$; as long as $\mathrm{grad}_a f \neq 0$, there is at least one partial derivative $\frac{\partial f}{\partial x_j}\Big|_{x=a} \neq 0$, and we can carry out the following argument with j in place of 1. We write $\mathbb{R}^n$ as $\mathbb{R} \times \mathbb{R}^{n-1}$ and consider a neighborhood $U\times V \subseteq X$, $a_1 \in U \subseteq \mathbb{R}$ and $a' = (a_2, \ldots, a_n) \in V \subseteq \mathbb{R}^{n-1}$. We look at the the restriction of the function f to $U \times V$. Since the first partial derivative $\partial_1 f(a_1, a') = \frac{\partial f}{\partial x_1}\Big|_{x=a}$ is nonzero, hence invertible in the sense of the Implicit Function Theorem, Theorem 3.29 implies and yields an open neighborhood $U_{a'}$ of a' in $\mathbb{R}^{n-1}$ and a function $\varphi\colon U_{a'} \to \mathbb{R}$ such that for $x' = (x_2, \ldots, x_n) \in U_{a'}$ we have $(s(x'), x') \in U \times V \subseteq X$ and

$$f(s(x_2, \ldots, x_n), x_2, \ldots, x_n) = f(a).$$

This means that the set

$$\{(s(x_2, \ldots, x_n), x_2, \ldots, x_n) : (x_2, \ldots, x_n \in U_{a'})\}$$

is firstly a neighborhood of $f(a)$ in the level set H and secondly is a piece of a hypersurface (of "dimension" $n-1$) in $\mathbb{R}^n$. Thus we are justified to call the level set H, whenever $\operatorname{grad}_a f$ does not vanish, a *level surface.* If $n = 2$, the expression level line is quite appropriate. We will again identify $\mathbb{R}^n$ and $\mathbb{R} \times \mathbb{R}^{n-1}$ and represent $x = (x_1, x_2, \ldots, x_n)$ in the form (x, x') with $x' = (x_2, \ldots, x_n)$.

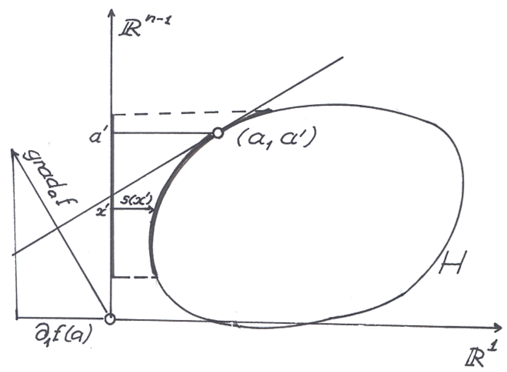

Figure 3.5

By Theorem 3.29, the function $s\colon U_{a'} \to \mathbb{R}$ is a differentiable level function (in one dimension lower), and the formula (12) for the derivative yields

$$(12') \qquad s'(a') = \operatorname{grad}_{a'} s = -(\partial_1 f)(a)^{-1}\cdot((\partial_2 f)(a), \ldots, (\partial_n f)(a)),$$

and thus also

$$(13) \qquad \begin{aligned} f'(a) = \operatorname{grad}_a f &= ((\partial_1 f)(a), -(\partial_1 f)(x)\operatorname{grad}_{a'} s) \\ &= (\partial_1 f)(a)\cdot(1, -\operatorname{grad}_{a'} s). \end{aligned}$$

The equation of the tangent hyperplane to the set $\{(s(x'), x') \in \mathbb{R} \times \mathbb{R}^{n-1} : x' \in U_{a'}\}$ at the point $(s(a'), a')$ is $x_1 = s(a') + \big(s'(a') \mid x'\big)$ since the affine approximation of s at a' is $x' \mapsto s(a') + (s'(a') \mid x')$, $x' = (x_2, \ldots, x_n)$. This equation is equivalent to

$$(14) \qquad \big((1, -s'(a')) \mid x\big) = s(a') \quad \text{with} \quad x = (x_1, \ldots, x_n).$$

In (13) and (14) we finally proved facts which, in our earlier discussion, we were able to illustrate only in an intuitive fashion. Indeed we have now shown:

Theorem 3.30. *Let $f: X \to \mathbb{R}$ be a continuously differentiable level function on an open set X of $\mathbb{R}^n$. If at $a \in X$ we have $\partial_1 f(a) \neq 0$ then the level set $H \stackrel{\text{def}}{=} \{x \in X : f(x) = f(a)\}$ is locally near a the graph of a differentiable function $s: U_{a'} \to \mathbb{R}$, where $U_{a'}$ is an open neighborhood of $a' = (a_2, \ldots, a_n)$ in $\mathbb{R}^{n-1}$:*

$$(\forall (x_2, \ldots, x_n) \in U_a)\, f\big(s(x_2, \ldots, x_n), x_2, \ldots, x_n\big) = f(a_1, \ldots, a_n).$$

The tangent hyperplane to H at $a = (a_1, \ldots, a_n)$ in $\mathbb{R}^n$ is perpendicular to the nonzero vector $\operatorname{grad}_a f$ *in $\mathbb{R}^n$.* □

Optimizing functions under constraints

The Implicit Function Theorem permits us to solve a classical problem of finding maxima and minima of a given real valued function $f: X \to \mathbb{R}$ on some subset X of $\mathbb{R}^n$ subject to extra conditions, called *constraints*, given frequently in the form of equations $g_1(x) = \cdots = g_m(x) = 0$ with level functions $g_j: X \to \mathbb{R}$. We combine these to a function $g: X \to \mathbb{R}^m$, $g(x) = \big(g_1(x), \ldots, g_m(x)\big)$. Let us make this precise in the following definition:

Definition 3.31. Let $f: X \to \mathbb{R}$ and $g: X \to \mathbb{R}^m$, $X \subseteq \mathbb{R}^n$ be functions. If for $a \in X$ there is a neighborhood U of a in X such that

$$f(a) = \max\{f(x) : x \in U \text{ and } g(x) = 0\},$$

then we say that f *attains in a a local maximum subject to the constraint* $g(x) = 0$. An analogous definition applies with "min" in place of "max". Also we say that f attains a *strict* local maximum in a subject to the constraint $g(x) = 0$ if U can be chosen so that

$$(\forall x \in U \setminus \{a\})\, g(x) = 0 \Rightarrow f(x) < f(a).$$

Similarly for "min" in place of "max". □

We are looking for (at least necessary) conditions for f to attain local extrema subject to constraints.

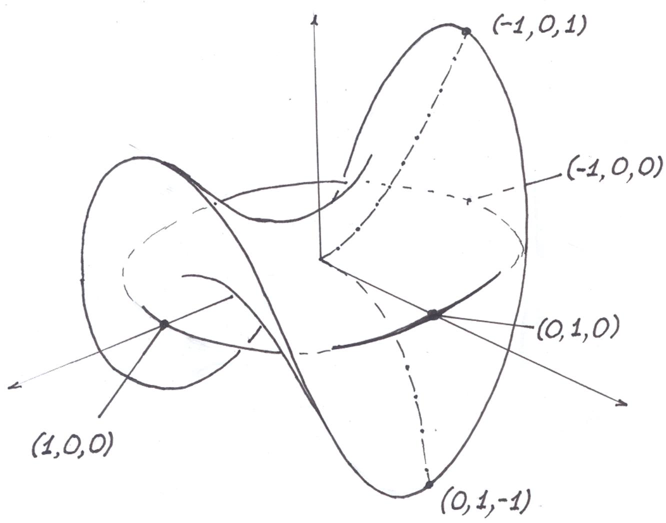

Figure 3.6

Figure 3..6 represents the optimisation problem: "find the local maxima and minima of the function $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = x^2 - y^2$ subject to the constraint $g(x,y) = 0$ $g(x,y) = x^2 + y^2 - 1$." This example shows that the function f may not have any local extrema at all, while under constraints there may be maxima and minima.

Now let a be an inner point of X with $g(a) = 0$. In finding necessary conditions for f to attain a local extremum a under constraints $g_1(x) = \cdots = g_m(x) = 0$ we expect to have fewer constraint equations than free variables $x_1, \ldots, x_n$; that is we assume $m < n$ and write $\mathbb{R}^n$ in the form $\mathbb{R}^{n-m} \times \mathbb{R}^m$ and each $x = (x_1, \ldots, x_n) \in X$ in the form $x = (x', x'')$ with $x' = (x_1, \ldots, x_{n-m})$ and $x'' = (x_{n+1-m}, \ldots, x_n)$. Then $f(x) = f(x', x'')$, and since a is an inner point of X there are open neighborhoods U of a' and V ov a'' such that $U \times V \subseteq X$. Now we assume that g is strongly differentiable at $a = (a', a'')$ and that the second partial derivative $\partial_2 g(a)$ is invertible. Then by the Implicit Function Theorem there is an open neighborhood $U_{a'}$ of a' in $\mathbb{R}^{n-m}$ and a function $h: U_{a'} \to \mathbb{R}^m$ such that $\big(x', h(x')\big) \in U \times V$ for all $x' \in U_{a'}$ and that

$$(\forall x' \in U_{a'})\, g\big(x', h(x')\big) = g(a', a'') = g(a) = 0.$$

Here $h(x')$ is the unique solution x'' of the equation $g(x', x'') = 0$ for $x' \in U_{a'}$. Now we know that the function $\varphi: U_{a'} \to \mathbb{R}$, $\varphi(x') = f\big(x', h(x')\big)$ attains a local maximum in a'. Hence by 3.17, $D_{a'}\varphi = 0$. Let $\theta: U_{a'} \to U \times V$ be given by $\theta(x') = \big(x', h(x')\big)$, then $D_{a'}\theta(u') = \big(u, D_{a'}h(u')\big)$. By the Chain Rule, $D_{a'}\varphi(u') = (D_{a',a''}f \circ D_{a'}\theta)(u') = D_a f\big(u', D_{a'}h(u')\big) = \partial_1 f(a)(u) + \partial_2 f(a)\big(D_{a'}h(u)\big)$. Thus we have

$$\partial_1 f(a) = -\partial_2 f(a) \circ D_{a'}h. \tag{15}$$

If we set $\psi\colon U_{a'} \to U \times V$, $\psi(x') = g\big(x', h(x')\big)$, then ψ is the zero function, and from differentiating it we obtain, in a fashion completely paralell to the one that yielded (61)

$$D_{a'}h = -\big(\partial_2 g(a)\big)^{-1}\partial_1 g(a). \tag{16}$$

Therefore, if we define the linear form L on $\mathbb{R}^m$ by $L = \partial_2 f(a) \circ \big(\partial_2 g(a)\big)^{-1}\colon \mathbb{R}^m \to \mathbb{R}$, that is $\partial_2 f(a) = L \circ \partial_2 g(a)$, then from (15) and (16) we obtain $\partial_1 f(a) = L \circ \partial_1 g(a)$, and thus altogether

$$D_a f = L \circ D_a g. \tag{17}$$

We have proved the following result

Theorem 3.32. *Let $X \subseteq \mathbb{R}^n$ have an inner point a and let $f\colon X \to \mathbb{R}$ and $g\colon X \to \mathbb{R}^m$ be functions such that f is differentiable in a and g is strongly differentiable in a. Assume that $m < n$, allowing us to write $\mathbb{R}^n$ as $\mathbb{R}^{n-m} \times \mathbb{R}^n$ and each $x \in X$ as $x = (x', x'')$ with $x' \in \mathbb{R}^{n-m}$ and $x'' \in \mathbb{R}^m$. We further assume that*
(i) *f attains a local extremum at $a = (a', a'')$ subject to the constraint $g(x) = 0$, and*
(ii) *$\partial_2 g(a)$ is invertible, i.e. the function $x'' \mapsto g(a', x'')$ has an invertible derivative. Then there exists a linear form $L\colon \mathbb{R}^m \to \mathbb{R}$ such that*

$$D_a f = L \circ D_a g. \tag{18}$$ □

Corollary 3.33. *Let $X \subseteq \mathbb{R}^n$ be open let $f, g_j\colon X \to \mathbb{R}$, $j = 1, \ldots, m < n$ functions with be continuous partial derivatives. We further assume that*
(i) *f attains a local extremum at a subject to the constraints $g_j(x) = 0$, $j = 1, \ldots, m$ and*
(ii) *We have*

$$\det\left(\frac{\partial g_j}{\partial x_{n-m+k}}\bigg|_{x=a}\right) \neq 0.$$

Then there exist numbers $\lambda_1, \ldots, \lambda_m$ such that

$$\frac{\partial f}{\partial x_k}\bigg|_{x=a} = \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_k}\bigg|_{x=a}, \quad k = 1, \ldots, n. \tag{18'}$$

Proof . We set $g = (g_1, \ldots, g_m)$. Since all partial derivatives of the g_j and of f are continuously differentiable, by 3.14 and 3.20, both f and g are strongly differentiable. The hypotheses (i) and (ii) correspond precisely to the hypotheses (i) and (ii) of 3.32. Hence 3.15 applies, and if we set $L = (\lambda_1, \ldots, \lambda_n)$ then (18) is exactly (18′). □

Equation (18′) is a system of m linear inhomogeneous equations in the unknowns $\lambda_1, \ldots, \lambda_n$, and the conclusion of Corollary 3.33 says that this system has a solution.

The actual algorithm for finding the numbers $a_1, \ldots, a_n$ is complicated by the task to find, in addition, the numbers $\lambda_1, \ldots, \lambda_m$; however, we must not forget that we have, in addition to the equations (18′), the m equations $g_j(x) = 0$, $j = 1, \ldots, m$. Indeed we can formulate the following

Algorithm 3.34. *If the data in* Corollary 3.33 *are given, define a new function* $F: X \times \mathbb{R}^m \to \mathbb{R}$ *by*

$$F(x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_m) = f(x_1, \ldots, x_n) - \lambda_1 g_1(x_1, \ldots, x_n) - \cdots - \lambda_m g_m(x_1, \ldots, x_n),$$

and find the local extrema of this function.

Proof . A necessary condition for F to attain a local extremum at

$$(a_1, \ldots, a_n, \lambda_1, \ldots, \lambda_m)$$

is $D_{(a_1,\ldots,a_n,\lambda_1,\ldots,\lambda_m)}F = 0$, and this is equivalent to the n equations of (18') and the m equations $g_j(x) = 0$, $j = 1, \ldots, m$. □

The numbers $\lambda_1, \ldots, \lambda_m$ are called *Lagrange multipliers.*

As is usual with optimization problems, if the algorithm yields one or several solutions

$$(a_1, \ldots, a_n, \lambda_1, \ldots, \lambda_m),$$

one has to determine on the basis of other information whether, in this critical point, f attains a local maximum, a local minimum, or neither on the set defined by the constraints. If the functions f and g are defined on a set X that is not open, the algorithm applies to the interior of X, i.e. the set of innner points of X in $\mathbb{R}^n$. Under these circumstances local minima or maxima may occur at boundary points which then have to be tested with other methods. As a guiding principle the one dimensional case gives an idea what happens: see Corollary 4.27 in Analysis I.

It seems as if, in our discussion of 3.32, 3.33, and 3.34, the last m coordinates $x_{n-m+1}, \ldots, x_n$ played a distinguished role in so far as we assumed 3.32(ii) or 3.33(ii). However, what is actually essential for the argument is that

- *the derivative* $D_a g: \mathbb{R}^n \to \mathbb{R}^m$ *is surjective, i.e. has rank* m.

If this condition is satisfied, then the theory of linear maps and matrices shows that there is a subset $I \subseteq \{1, \ldots, n\}$ such that the restriction of $D_a g$ to $R^I \stackrel{\text{def}}{=} \prod_{j=1}^n \mathbb{R}^{\varepsilon_j} \subseteq \mathbb{R}^n$, $\varepsilon_j = 1$ if $j \in I$ and $= 0$ otherwise, is bijective. This is tantamount to saying that the $m \times n$-matrix of $D_a g$ has an $m \times m$-submatrix A consisting of m columns with column indices $j \in I$, such that $\det A \neq 0$. We then select the m coordinates with the indices from I in the place of the last m coordinates and carry out the discussion for these.