2 Equations, from linear to nonlinear

2a Introduction

> Born: *I should like to put to Herr Einstein a question, namely, how quickly the action of gravitation is propagated in your theory. . .*
>
> Einstein: *It is extremely simple to write down the equations for the case when the perturbations that one introduces in the field are infinitely small. . . . The perturbations then propagate with the same velocity as light.*
>
> Born: *But for great perturbations things are surely very complicated?*
>
> Einstein: *Yes, it is a mathematically complicated problem. It is especially difficult to find solutions of the equations, as the equations are nonlinear.* — Discussion after lecture by Einstein in 1913.
>
> . . .
>
> The hardest part of differential calculus is determining when replacing a nonlinear object by a linear one is justified.[1]

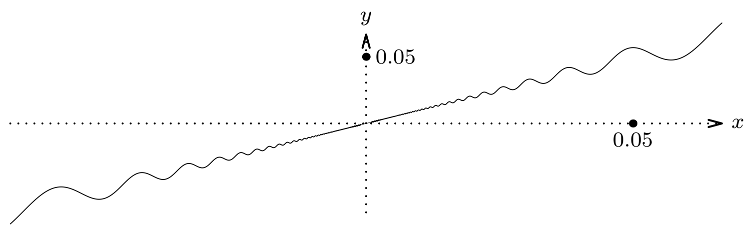
In other words, we want to know, when the linear approximation

$$f(x_0 + h) \approx f(x_0) + (Df)_{x_0} h$$

may be trusted near x_0.

2a1 Example. [2]

$f : \mathbb{R} \to \mathbb{R}$, $f(x) = x + 3x^2 \sin \frac{1}{x}$ for $x \neq 0$, $f(0) = 0$, $x_0 = 0$.



[1] Quoted from: Hubbard, Sect. 1.7, pp. 125–126.

[2] Hubbard, Example 1.9.4 on p. 157; Shifrin, Sect. 6.2, Example 1 on pp. 251–252.

This function is differentiable everywhere; its linear approximation near 0, $f(x) \approx x$, is one-to-one. Nevertheless, f fails to be one-to-one near 0 (and the equation $f(x) = y$ has more than one solution).[1] The linear approximation cheats. In fact,

$$\liminf_{x\to 0} f'(x) = -2\,, \qquad \limsup_{x\to 0} f'(x) = 4$$

(think, why); f is differentiable, but not continuously.

This is why throughout this section we require f to be *continuously* differentiable near x_0.

LINEAR ALGEBRA

2a2 Example. The matrix $A = \begin{pmatrix} 3 & 2 & 0 & 1 \\ 0 & 1 & 3 & -1 \\ 3 & 1 & -3 & 2 \end{pmatrix}$ is of rank 2, it has (at least one) non-zero minor 2×2, but not 3×3, since the first row is the sum of the two other rows. Treated as a linear operator $A : \mathbb{R}^4 \to \mathbb{R}^3$ it maps $\mathbb{R}^4$ onto a 2-dimensional subspace of $\mathbb{R}^3$, the image of A: $A(\mathbb{R}^4) = \{(z_1, z_2, z_3) : z_1 - z_2 - z_3 = 0\}$. The kernel of A, $A^{-1}(\{0\}) = \{u \in \mathbb{R}^4 : Au = 0\}$ is a 2-dimensional subspace of $\mathbb{R}^4$ spanned (for instance) by two vectors $(-1, 1, 0, 1)$ and $(0, -1, 1, 2)$, according to two linear dependencies of the columns: $-\begin{pmatrix} 3 \\ 0 \\ 3 \end{pmatrix} + \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} = 0$, $-\begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ 3 \\ -3 \end{pmatrix} + 2\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} = 0$.

It is convenient to denote a point of $\mathbb{R}^4$ by (x_1, x_2, y_1, y_2); that is, $x \in \mathbb{R}^2$, $y \in \mathbb{R}^2$, $(x, y) \in \mathbb{R}^4$. The equation $A\begin{pmatrix} x \\ y \end{pmatrix} = 0$ becomes $\begin{pmatrix} 3 & 2 \\ 0 & 1 \end{pmatrix} x + \begin{pmatrix} 0 & 1 \\ 3 & -1 \end{pmatrix} y = 0$ (the third row is redundant, being a linear combination of other rows); $y = -\begin{pmatrix} 0 & 1 \\ 3 & -1 \end{pmatrix}^{-1} \begin{pmatrix} 3 & 2 \\ 0 & 1 \end{pmatrix} x = -\begin{pmatrix} 1 & 1 \\ 3 & 2 \end{pmatrix} x$, since $\begin{pmatrix} 0 & 1 \\ 3 & -1 \end{pmatrix}^{-1} = \frac13 \begin{pmatrix} 1 & 1 \\ 3 & 0 \end{pmatrix}$. Not unexpectedly, $\begin{pmatrix} 0 \\ 1 \end{pmatrix} = -\begin{pmatrix} 1 & 1 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 2 \end{pmatrix} = -\begin{pmatrix} 1 & 1 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} 0 \\ -1 \end{pmatrix}$. The more general equation $A(x, y) = z$ for a given $z \in A(\mathbb{R}^4)$ may be solved similarly; $\begin{pmatrix} 3 & 2 \\ 0 & 1 \end{pmatrix} x + \begin{pmatrix} 0 & 1 \\ 3 & -1 \end{pmatrix} y = \tilde z$ (where $\tilde z \in \mathbb{R}^2$ is (z_1, z_2) for $z = (z_1, z_2, z_3)$); $y = \begin{pmatrix} 0 & 1 \\ 3 & -1 \end{pmatrix}^{-1} \Big(\tilde z - \begin{pmatrix} 3 & 2 \\ 0 & 1 \end{pmatrix} x\Big) = \frac13 \begin{pmatrix} 1 & 1 \\ 3 & 0 \end{pmatrix} \tilde z - \begin{pmatrix} 1 & 1 \\ 3 & 2 \end{pmatrix} x$.

In general, a matrix $A : \mathbb{R}^n \to \mathbb{R}^m$ has some rank $r \le \min(m, n)$. The image is r-dimensional, the kernel is $(n - r)$-dimensional. We rearrange rows and columns (if needed) such that the upper right $r \times r$ minor is not 0, denote a point of $\mathbb{R}^n$ by (x, y) where $x \in \mathbb{R}^{n-r}$, $y \in \mathbb{R}^r$, then the equation $A(x, y) = 0$ becomes $Bx + Cy = 0$, $B : \mathbb{R}^{n-r} \to \mathbb{R}^r$, $C : \mathbb{R}^r \to \mathbb{R}^r$, $\det C \ne 0$;

$$A \;=\; \begin{array}{c|c|c|} \multicolumn{1}{c}{} & \multicolumn{1}{c}{n-r} & \multicolumn{1}{c}{r} \\ \cline{2-3} r & B & C \\ \cline{2-3} m-r & \multicolumn{2}{c|}{} \\ \cline{2-3} \end{array}$$

[1] Bad news... But here are good news: all solutions of the equation $f(x) = y$ are close (to each other and y), namely, $x = y + \mathcal{O}(y^2)$.

the solution is $y = -C^{-1}Bx$. More generally, the equation $A(x, y) = z$ for a given $z \in A(\mathbb{R}^n)$ becomes $Bx + Cy = z$; the solution: $y = C^{-1}(z - Bx)$.

Note existence of an r-dimensional subspace $E \subset \mathbb{R}^n$ such that the restriction $A|_E$ is an invertible mapping from E onto $A(\mathbb{R}^n)$.[1]

Special cases:

$*$ $r = m \le n \iff A(\mathbb{R}^n) = \mathbb{R}^m$ (onto);

$*$ $r = n \le m \iff A^{-1}(\{0\}) = \{0\}$ (one-to-one); no x variables;

$*$ $r = m = n \iff A$ is invertible (bijection).

Note that A is onto if and only if its rows are linearly independent.

ANALYSIS

We turn to the equation $f(x) = y$ where $f : \mathbb{R}^n \to \mathbb{R}^m$ is continuously differentiable near x_0, and introduce $A = (Df)_{x_0}$. We want to compare two mappings, f (nonlinear) and $x \mapsto f(x_0) + A(x - x_0)$ (linear),[2] near x_0. Or, equivalently, of $h \mapsto Ah$ (linear) and $h \mapsto f(x_0 + h) - f(x_0)$ (nonlinear), near 0 (that is, for small h). The relevant properties of f, including its derivative A, are insensitive to a change of the origin[3] in $\mathbb{R}^n$ and $\mathbb{R}^m$, and therefore we may assume WLOG[4] that $x_0 = 0$ and $f(x_0) = 0$.

Also, all relevant properties of f are insensitive to a change of basis, both or $\mathbb{R}^n$ and $\mathbb{R}^m$; this argument will be used later.

Also, values of f outside a neighborhood of 0 are irrelevant; we consider f near 0 only. This is why I often write just "$f : \mathbb{R}^n \to \mathbb{R}^m$" rather than "$f : U \to \mathbb{R}^m$ where $U \subset \mathbb{R}^n$ is a neighborhood of 0".

The linear algebra gives us properties of $A : \mathbb{R}^n \to \mathbb{R}^m$, and we want to prove the corresponding *local* properties of $f : \mathbb{R}^n \to \mathbb{R}^m$ near 0. Here are some relevant definitions, global and local; the local definitions are formulated for the case $x_0 = 0$, $f(0) = 0$; you can easily generalize them to arbitrary $x_0 \in \mathbb{R}^n$ and $y_0 = f(x_0) \in \mathbb{R}^m$.

2a3 Definition. (a) Let $U, V \subset \mathbb{R}^n$ be open sets. A mapping $f : U \to V$ is a *homeomorphism,* if it is bijective, continuous, and $f^{-1} : V \to U$ is also continuous.

(b) $f : \mathbb{R}^n \to \mathbb{R}^n$ is a *local homeomorphism,* if there exist open sets $U, V \subset \mathbb{R}^n$ such that $0 \in U$, $0 \in V$, and f is a homeomorphism $U \to V$.

---

[1] Another proof: take a basis $\beta_1, \dots, \beta_r$ of $A(\mathbb{R}^n)$; choose $\alpha_1, \dots, \alpha_r \in \mathbb{R}^n$ such that $A\alpha_1 = \beta_1, \dots, A\alpha_r = \beta_r$; consider E spanned by $\alpha_1, \dots, \alpha_r$.

[2] More exactly: affine.

[3] It means, all *points* are changed (shifted), but *vectors* remain intact; x and y are points, while h and Ah are vectors.

[4] Without Loss Of Generality.

2a4 Definition. (a) Let $U, V \subset \mathbb{R}^n$ be open sets. A mapping $f : U \to V$ is a *diffeomorphism,*[1] if it is bijective, continuously differentiable, and $f^{-1} : V \to U$ is also continuously differentiable.

(b) $f : \mathbb{R}^n \to \mathbb{R}^n$ is a *local diffeomorphism,* if there exist open sets $U, V \subset \mathbb{R}^n$ such that $0 \in U$, $0 \in V$, and f is a diffeomorphism $U \to V$.

2a5 Exercise. For a linear $A : \mathbb{R}^n \to \mathbb{R}^m$ prove that the following conditions are equivalent:

(a) A is invertible;
(b) A is a homeomorphism;
(c) A is a local homeomorphism;
(d) A is a diffeomorphism;
(e) A is a local diffeomorphism.

2a6 Definition. (a) Let $U \subset \mathbb{R}^n$ be an open set. A mapping $f : U \to \mathbb{R}^m$ is *open,* if for every open subset $U_1 \subset U$ its image $f(U_1) \subset \mathbb{R}^m$ is open.

(b) $f : \mathbb{R}^n \to \mathbb{R}^m$ is *open at* 0, if for every neighborhood $U \subset \mathbb{R}^n$ of 0 there exists a neighborhood $V \subset \mathbb{R}^m$ of 0 such that $f(U) \supset V$.[2,3]

2a7 Exercise. (a) Prove that f is open at 0 if and only if for every sequence $y_1, y_2, \dots \in \mathbb{R}^m$ such that $y_k \to 0$ there exists a sequence $x_1, x_2, \dots \in \mathbb{R}^n$ such that $x_k \to 0$ and $f(x_k) = y_k$ for all k large enough;

(b) generalize 2a6(b) to arbitrary x_0 and $y_0 = f(x_0)$;

(c) prove that $f : U \to \mathbb{R}^m$ is open if and only if f is open at x for every $x \in U$.

2a8 Exercise. Prove or disprove: a continuous function $\mathbb{R} \to \mathbb{R}$ is open if and only if it is strictly monotone.

2a9 Exercise. For a linear $A : \mathbb{R}^n \to \mathbb{R}^m$ prove that the following conditions are equivalent:

(a) $A(\mathbb{R}^n) = \mathbb{R}^m$ ("onto");
(b) A is open at 0;[4]
(c) A is open.

---

[1]This is a C^1 diffeomorphism (most important for this course); C^0 diffeomorphism is just a homeomorphism, and C^k diffeomorphism must be continuously differentiable k times (both f and f^{-1}).

[2]This notion is seldom used; but see for instance Sect. 2.8 in Basic Complex Analysis by G. De Marco.

[3]In this form, we may interpret the phrase "U is a neighborhood of 0" as "0 is an interior point of U" or, equally well, as "$\exists \varepsilon > 0 \; U = \{x : |x| < \varepsilon\}$".

[4]Hint: (b) use subspace E such that $A|_E$ is an invertible mapping from E onto $A(\mathbb{R}^n)$.

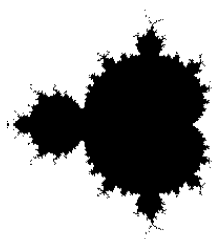
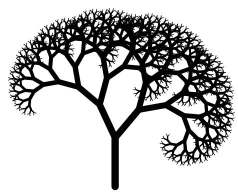
2a10 Exercise. Consider the mapping $f : U \to \mathbb{R}^2$, where $U = (1,2) \times (-T,T) \subset \mathbb{R}^2$ (for a given $T \in (0,\infty)$) and $f(r,\theta) = (r\cos\theta, r\sin\theta)$. Denote $V = f(U)$. For each of the following conditions (separately) find all T such that the condition is satisfied:[1]

(a) V is open;
(b) f is continuous;
(c) f is uniformly continuous;
(d) f is continuously differentiable;
(e) $f : U \to V$ is bijective;
(f) $f : U \to V$ is a homeomorphism;
(g) $f : U \to V$ is a homeomorphism and $f^{-1} : V \to U$ is uniformly continuous;
(h) $f : U \to V$ is a diffeomorphism;
(i) f is a local homeomorphism near each point of U;
(j) f is a local diffeomorphism near each point of U;
(k) f is an open mapping.

In the linear case we may ignore the last $m - r$ (redundant) equations. In the nonlinear case we cannot.

2a11 Example. Consider $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x_1,x_2) = \big(x_1, x_1 + c(x_1^2 + x_2^2)\big)$ for a given c. The linear approximation: $f(x_1,x_2) \approx (x_1,x_1)$; $A = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$. The equation $Ax = 0$ is satisfied by all $x = (0,x_2)$. However, the equation $f(x) = 0$ is satisfied by $x = (0,0)$ only (unless $c = 0$). The linear approximation cheats.

In fact, for *every* closed set $F \subset \mathbb{R}^n$ containing 0 there exists a continuously differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ such that $F = \{x : f(x) = 0\}$ and $f(0) = 0$, $(Df)_0 = 0$.[2]

In the linear approximation, $f(x) \approx 0$; $A = (0,\dots,0)$; $Ax = 0$ for all x. However, $f(x) = 0$ for $x \in F$ only.

[1] Hint: (h) use arccos and arcsin (you really need both); (i), (j): generalize definitions 2a3(b), 2a4(b) to arbitrary x_0 and $y_0 = f(x_0)$.

[2] Hint: cover the complement with a sequence of open balls and take the sum of an appropriate series of functions positive inside these balls and vanishing outside.

The case $r < m$ is intractable. This is why we restrict ourselves to the cases

$r = m < n$; $A(\mathbb{R}^n) = \mathbb{R}^m$ (onto, not one-to-one);
$r = m = n$; A is invertible (bijection).

2b Main results formulated and discussed

First, the case $r = m = n$. Here is a theorem called "the inverse function theorem" [1] or "inverse mapping theorem".[2]

2b1 Theorem. Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be continuously differentiable near 0, $f(0) = 0$, and $(Df)_0 = A : \mathbb{R}^n \to \mathbb{R}^n$ be invertible.[3] Then f is a local diffeomorphism, and $\big(D(f^{-1})\big)_0 = A^{-1}$.

2b2 Remark. The relation $\big(D(f^{-1})\big)_0 = A^{-1}$ is included for completeness. It follows easily from the chain rule: $f^{-1} \circ f = \mathrm{id}$, therefore $\big(D(f^{-1})\big)_0 (Df)_0 = \mathrm{id}$. (However, differentiability of f^{-1} does not follow from the chain rule!) Similarly, $\big(D(f^{-1})\big)_{f(x)} = \big((Df)_x\big)^{-1}$ for all x near 0.

Second, the case $r = m < n$. Here is the "implicit function theorem".[4]

2b3 Theorem. Let $f : \mathbb{R}^{n-m} \times \mathbb{R}^m \to \mathbb{R}^m$ be continuously differentiable[5] near $(0, 0)$, $f(0, 0) = 0$, and $(Df)_{(0,0)} = A = (B \big| C)$, $B : \mathbb{R}^{n-m} \to \mathbb{R}^m$, $C : \mathbb{R}^m \to \mathbb{R}^m$, with C invertible. Then there exists $g : \mathbb{R}^{n-m} \to \mathbb{R}^m$, continuously differentiable near 0, such that the two relations $f(x, y) = 0$ and $y = g(x)$ are equivalent for (x, y) near $(0, 0)$; and $(Dg)_0 = -C^{-1}B$.

Clearly, $g(0) = 0$ (since $f(0, 0) = 0$).

2b4 Exercise. Deduce from 2b3 existence of $\varepsilon > 0$, $\delta > 0$ such that for every x satisfying $|x| < \delta$ there exists one and only one y satisfying $|y| < \varepsilon$ and $f(x, y) = 0$; namely, $y = g(x)$.

Clearly, $f\big(x, g(x)\big) = 0$ for x near 0.

2b5 Remark. The relation $(Dg)_0 = -C^{-1}B$ is included for completeness. It follows easily from the chain rule: $f\big(x, g(x)\big) = 0$, that is, $f \circ \varphi = 0$ where

---

[1] Fleming, Hubbard, Shifrin, Shurman, Zorich.

[2] Lang.

[3] Recall 2a5.

[4] "It is without question one of the most important theorems in higher mathematics" (Shifrin p. 255).

[5] As a mapping $\mathbb{R}^n \to \mathbb{R}^m$.

$\varphi : x \mapsto \binom{x}{g(x)}$; we have $\varphi(0) = (0,0)$ and $(D\varphi)_0 = \Big(\dfrac{\mathrm{id}}{(Dg)_0}\Big)$; thus, $0 = D(f \circ \varphi)_0 = (Df)_{(0,0)}(D\varphi)_0 = \big(B \big| C\big)\Big(\dfrac{\mathrm{id}}{(Dg)_0}\Big) = B + C(Dg)_0$. (However, differentiability of g does not follow from the chain rule!) Similarly, for all x near 0 holds $(Dg)_x = -C_x^{-1}B_x$ where $\big(B_x \big| C_x\big) = A_x = (Df)_{(x,g(x))}$.

In dimension $1 + 1 = 2$ we have

$$g'(x) = -\frac{(D_1 f)_{(x,y)}}{(D_2 f)_{(x,y)}} \quad \text{where } y = g(x)\,.$$

Less formally, $\frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{\partial g/\partial x}{\partial g/\partial y}$ since $\frac{\partial g}{\partial x}\mathrm{d}x + \frac{\partial g}{\partial y}\mathrm{d}y = \mathrm{d}g(x,y) = 0$.

2b6 Exercise. Given $k \in \{1, 2, 3, \dots\}$, we define $f : \mathbb{R}^2 \to \mathbb{R}$ by $f(x,y) = \operatorname{Im}\big((x + \mathrm{i}y)^k\big)$ (where $\mathrm{i}^2 = -1$ and $\operatorname{Im}(a + \mathrm{i}b) = b$).

(a) Find all k such that f satisfies the *assumptions* of Theorem 2b3.

(b) Find all k such that f satisfies the *conclusions* of Theorem 2b3 (except for the last equality).

2b7 Exercise. Let f satisfy the assumptions of Theorem 2b3. Show that f^2 (pointwise square) violates the assumptions of Theorem 2b3 but still satisfies its conclusions (except for the last equality).

It is not easy to prove these two theorems, but it is easy to derive one of them from the other. First, 2b3$\Longrightarrow$2b1. The idea is simple: x is implicitly a function of y according to the equation $\varphi(y,x) = f(x) - y = 0$.

***Proof** of the implication 2b3 $\Longrightarrow$2b1.* Given $f : \mathbb{R}^n \to \mathbb{R}^n$ as in 2b1, we define $\varphi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ by $\varphi(y,x) = f(x) - y$ and check the conditions of 2b3 for $2n, n, \varphi$ in place of n, m, f. We have $(Df)_{(0,0)} = \big(-\mathrm{id}\big|(Df)_0\big)$; $C = (Df)_0$ is invertible. Theorem 2b3 gives $g : \mathbb{R}^n \to \mathbb{R}^n$, continuously differentiable near $(0,0)$, such that $f(x) - y = 0 \iff x = g(y)$ near $(0,0)$.

We take $\varepsilon > 0$ such that both f and g are continuously differentiable on $\{x : |x| < \varepsilon\}$, and $f(x) = y \iff x = g(y)$ whenever $|x| < \varepsilon, |y| < \varepsilon$. We define $U = \{x : |x| < \varepsilon, |f(x)| < \varepsilon\}$ and $V = \{y : |y| < \varepsilon, |g(y)| < \varepsilon\}$. Both U and V are open and contain 0.

If $x \in U$ and $y = f(x)$, then $|y| < \varepsilon$ and $x = g(y)$, therefore $y \in V$. We see that $f(U) \subset V$ and $g \circ f = \mathrm{id}$ on U. Similarly, $g(V) \subset U$ and $f \circ g = \mathrm{id}$ on V. It means that $f|_U : U \to V$ and $g|_V : V \to U$ are mutually inverse; thus, f is a local diffeomorphism. □

Second, 2b1 $\Longrightarrow$2b3. The idea is the diffeomorphism $\varphi : (x,y) \mapsto \big(x, f(x,y)\big)$ and its inverse.

***Proof** of the implication 2b1 $\Longrightarrow$2b3.* Given $f : \mathbb{R}^{n-m} \times \mathbb{R}^m \to \mathbb{R}^m$ as in 2b3, we define $\varphi : \mathbb{R}^{n-m} \times \mathbb{R}^m \to \mathbb{R}^{n-m} \times \mathbb{R}^m$ by $\varphi(x,y) = \big(x, f(x,y)\big)$ and check the conditions of 2b1 for n, φ in place of n, f. We have

$$(D\varphi)_{(0,0)} = \left(\begin{array}{c|c} \mathrm{id} & 0 \\ \hline B & C \end{array}\right);$$

an invertible matrix, since its determinant is $\det(\mathrm{id})\det(C) = \det(C) \ne 0$. [1] By Theorem 2b1, φ is a local diffeomorphism. Its inverse $\psi = \varphi^{-1}$ is also a local diffeomorphism. Both φ and ψ do not change the first component x.

We define $g : \mathbb{R}^{n-m} \to \mathbb{R}^m$ by $\big(x, g(x)\big) = \psi(x,0)$ and note that g is continuously differentiable near 0, since it is the composition of three mappings

$$x \overset{\text{linear}}{\longmapsto} \begin{pmatrix} x \\ 0 \end{pmatrix} \overset{\psi}{\longmapsto} \psi\begin{pmatrix} x \\ 0 \end{pmatrix} = \begin{pmatrix} x \\ g(x) \end{pmatrix} \overset{\text{linear}}{\longmapsto} g(x)\,.$$

Finally, $f(x,y) = 0 \iff \varphi(x,y) = (x,0) \iff \psi(x,0) = (x,y) \iff y = g(x)$ for (x,y) near $(0,0)$. $\square$

Having 2b1 $\Longleftrightarrow$ 2b3, we need to prove only one of the two theorems. Which one? Both options are in use. Some authors[2] prove 2b3 by induction in dimension, and then derive 2b1. Others[3] prove 2b1 and then derive 2b3; we do it this way, too.

2c Proof, the easy part

Given $f : \mathbb{R}^n \to \mathbb{R}^n$ as in 2b1, we may choose at will a pair of bases (as explained in Sect. 1f and mentioned in Sect. 2a, Item "analysis"). We choose bases $(\alpha_1, \dots, \alpha_n)$ and $(\beta_1, \dots, \beta_n)$ such that $A\alpha_1 = \beta_1, \dots, A\alpha_n = \beta_n$ (here $A = (Df)_0$, as before), then A becomes id. That is, we may (and will) assume WLOG that $A = \mathrm{id}$. [4]

2c1 Lemma. f is one-to-one near 0.

Proof. We take $\delta > 0$ such that on the set $U = \{x : |x| < \delta\} \subset \mathbb{R}^n$ the function f is continuously differentiable and $\|Df - \mathrm{id}\| \le \frac{1}{2}$. We have

[1] Alternatively, invertibility is easy to check with no determinant. The equation $\left(\begin{array}{c|c} \mathrm{id} & 0 \\ \hline B & C \end{array}\right)\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} u \\ v \end{pmatrix}$ for given u, v becomes $x = u$, $Bx + Cy = v$, and clearly has one and only one solution $x = u$, $y = C^{-1}(v - Bu)$.

[2] Curant, Zorich.

[3] Fleming, Hubbard, Lang, Shifrin, Shurman.

[4] See also: Fleming, Sect. XVIII.3, p. 515.

$\|D(f - \mathrm{id})\| \le \frac{1}{2}$ on the convex set U; by (1f31), $|\big(f(b)-b\big) - \big(f(a)-a\big)| \le \frac12|b-a|$ for all $a,b \in U$. Thus, $|\big(f(b)-f(a)\big) - (b-a)| \le \frac12|b-a|$; $|f(b)-f(a)| \ge |b-a| - \frac12|b-a| = \frac12|b-a|$;[1] $a \ne b \implies f(a) \ne f(b)$ whenever $|a| < \delta$, $|b| < \delta$. □

Taking U as above, we introduce $V = f(U)$ and $f^{-1} : V \to U$ (really, this is $(f|_U)^{-1}$). The inequality $|f(b)-f(a)| \ge \frac12|b-a|$ for $a,b \in U$ becomes $|f^{-1}(y) - f^{-1}(z)| \le 2|y-z|$ for $y,z \in V$, which shows that f^{-1} is continuous on V. Also, taking $z = 0$ we get $|f^{-1}(y)| \le 2|y|$ for all $y \in V$.

2c2 Lemma. $f^{-1}(y) = y + o(y)$ for $y \in V$, $y \to 0$.

Proof. We use δ and U from the proof of 2c1, and generalize that argument as follows. For every $\varepsilon \in (0, \frac12]$ there exists $\delta_\varepsilon \in (0,\delta]$ such that $\|Df - \mathrm{id}\| \le \varepsilon$ on the subset $U_\varepsilon = \{x : |x| < \delta_\varepsilon\}$ of U, which implies (as before) $|\big(f(b)-f(a)\big) - (b-a)| \le \varepsilon|b-a|$ for all $a,b \in U_\varepsilon$. We need only the special case (for $a = 0$): $|f(b) - b| \le \varepsilon|b|$ for $b \in U_\varepsilon$. It is sufficient to check that

$$|f^{-1}(y) - y| \le 2\varepsilon|y| \quad \text{whenever } y \in V\,, |y| < \frac12\delta_\varepsilon\,.$$

Given such y, we consider $x = f^{-1}(y) \in U$, note that $x \in U_\varepsilon$ (since $|x| \le 2|y| < \delta_\varepsilon$), thus $|f(x) - x| \le \varepsilon|x|$, which gives $|y - f^{-1}(y)| \le 2\varepsilon|y|$. □

Did we prove differentiability of f^{-1} at 0? Not yet. Is f^{-1} defined near 0? That is, is 0 an interior point of V?

2c3 Theorem. Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be continuously differentiable near 0, $f(0) = 0$, and $(Df)_0 = A : \mathbb{R}^n \to \mathbb{R}^n$ be invertible. Then f is open at 0.

The proof is postponed to Sect. 2d.

2c4 Lemma. Theorem 2c3 implies Theorem 2b1.

Proof. Given f as in 2b1, we take U as in the proof of 2c1; now 2c3 gives $\varepsilon > 0$ such that the set $V_1 = \{y : |y| < \varepsilon\}$ satisfies $V_1 \subset f(U)$. The set $U_1 = \{x \in U : f(x) \in V_1\}$ is open (since f is continuous on U), and $f(U_1) = V_1$. Taking into account that f^{-1} is continuous on $f(U)$ we see that $f|_{U_1}$ is a homeomorphism between open sets U_1 and V_1; thus, f is a local homeomorphism. By 2c2, f^{-1} is differentiable at 0.

The same holds near every point of U_1 (the assumption $x_0 = 0$ was not a loss of generality, see page 20). Thus, f^{-1} is differentiable on V_1, and $\big(D(f^{-1})\big)_y = \big((Df)_{f^{-1}(y)}\big)^{-1}$ (as explained in 2b2). Continuity of $D(f^{-1})$ follows by 1f18(c), and therefore f is a local diffeomorphism. □

---

[1] The triangle inequality $|x+y| \le |x| + |y|$ implies $|x| \ge |x+y| - |y|$, that is, $|u+v| \ge |u| - |v|$.

2d Proof, the hard part

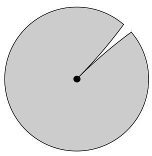
What is the problem

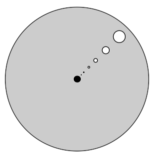
We want to prove Theorem 2c3. As before, we may assume WLOG that $A = \mathrm{id}$. We cannot use the theorems, but can use Lemma 2c1. We know that $f : U \to V$ is a homeomorphism, where $U = \{x : |x| < \delta\} \subset \mathbb{R}^n$ is an open ball, and δ is small enough; but we are not sure that V is open. Clearly, $0 \in V$ (since $f(0) = 0$). How to prove that 0 is an *interior* point of V?

In dimension 1 this is easy: $f(0) = 0$, $f'(0) = 1$; $f(\delta) > 0$, $f(-\delta) < 0$ (for δ small enough); 0 is an interior point of the interval $\big(f(-\delta), f(\delta)\big)$, and V contains this interval.

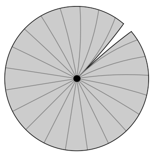
How do we know that V contains this interval? Being homeomorphic to the interval $U = (-\delta, \delta)$, V must be an interval. It is connected. Any hole inside the interval would disconnect it.

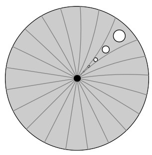
In dimension 2, V is homeomorphic to the disk $U = \{x : |x| < \delta\} \subset \mathbb{R}^2$, therefore, connected. So what? Could V be like these?





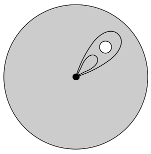
True, V must contain paths through 0 in all directions (images of rays). So what? This condition is also satisfied by these counterexamples.





Should we consider circles rather than rays? And what about higher dimension?

You see, n-dimensional topology is much more complicated than 1-dimensional. A hole disconnects a line, but not a plane. Rather, a hole on a plane disconnects the space of loops!



These two loops belong to different connected components in the space of loops.

In $\mathbb{R}^3$ a hole does not disconnect the space of loops; rather, it disconnects the space of… loops in the space of loops! And so on. Algebraic topology, a long and hard way…

2d1 Remark. In fact, for every open $U \subset \mathbb{R}^n$, every continuous one-to-one mapping $U \to \mathbb{R}^n$ is open (and therefore a homeomorphism between open sets U and $f(U)$). This is a well-known topological result, "the Brouwer invariance of domain theorem".[1]

2d2 Exercise. Prove invariance of domain in dimension one.[2]

Topology assumes only continuity of mappings, not differentiability. We wonder, are differentiable mappings more tractable than (just) continuous mappings?

Yes, fortunately, they are. Two analytical (rather than topological) proofs of Theorem 2c3 are well-known. Some authors[3] consider the minimizer x_y of the function $x \mapsto |f(x) - y|^2$ (for a given y near 0) and prove that $f(x_y)$ cannot differ from y, using invertibility of the operator $(Df)_{x_y}$. Others[4] use iteration (in other words, successive approximations), that is, construct a sequence of approximate solutions $x_1, x_2, \dots$ of the equation $f(x) = y$ (for a given y near 0) and prove that the sequence converges, and its limit is a solution; we do it this way, too.

Here is the idea. First, the linear approximation $f(x) \approx x$ for f leads to the same linear approximation $f^{-1}(y) \approx y$ for f^{-1} (see 2c2); thus, given y, we consider $x_1 = y$ as the first approximation to the (hoped for) solution of the equation $f(x) = y$. Alas, $y_1 = f(x_1)$ differs from y, and we seek a better approximation x_2. The inequality $|(f(b) - f(a)) - (b - a)| \le \varepsilon |b - a|$ (seen in the proof of 2c2) suggests that $f(b) - f(a) \approx b - a$, and in particular, $f(x_2) - f(x_1) \approx x_2 - x_1$. Seeking $f(x_2) \approx y$, that is, $f(x_2) - f(x_1) \approx y - y_1$, we take $x_2 = x_1 + y - y_1$. And so on: $y_2 = f(x_2)$; $x_3 = x_2 + y - y_2$; ...

It appears that every constant $\varepsilon \in (0, 1)$ ensures convergence $x_k \to x$, and then $f(x) = y$. Thus, we'll use just $\varepsilon = \frac{1}{2}$ (as in the proof of 2c1).

Proof *of Theorem 2c3.* Given f as in 2c3, we assume WLOG that $A = \mathrm{id}$ (as before) and take $U = \{x : |x| < \delta\} \subset \mathbb{R}^n$ as in the proof of 2c1; we know that $f|_U$ is a homeomorphism $U \to V = f(U)$, and

$$|(f(b) - f(a)) - (b - a)| \le \tfrac{1}{2} |b - a| \qquad \text{for all } a, b \in U \,. \tag{2d3}$$

First, we'll prove that 0 is an interior point of V (and afterwards we'll prove that the mapping f is open at 0). To this end it is sufficient to prove that $y \in V$ for all $y \in \mathbb{R}^n$ such that $|y| < \frac{1}{2}\delta$.

[1] By the way, it follows from the Brouwer invariance of domain theorem that an open set in $\mathbb{R}^{n+1}$ cannot be homeomorphic to any set in $\mathbb{R}^n$ (unless it is empty). Think, why.

[2] Hint: recall 2a8.

[3] Shurman; Zorich (alternative proof in Sect. 8.5.5, Exer. 4f).

[4] Fleming, Hubbard, Lang, Shifrin; Curant (alternative proof in Sect. 3.3g).

Given such y, we rewrite the equation $f(x) = y$ as

$$\varphi(x) = x \qquad \text{“fixed point”}$$

where $\varphi : U \to \mathbb{R}^n$ is defined by

$$\varphi(x) = y + x - f(x)\,.$$

In order to prove that $y \in V$ we need existence of a fixed point.

By (2d3),

$$|\varphi(b) - \varphi(a)| \le \tfrac{1}{2}|b - a| \qquad \text{for all } a, b \in U\,.$$

If $|x| < \delta$, then $|\varphi(x)| \le |\varphi(x) - \varphi(0)| + |\varphi(0)| \le \frac{1}{2}|x - 0| + |y| < \frac{1}{2}\delta + \frac{1}{2}\delta = \delta$. We take $x_1 = y$, $x_2 = \varphi(x_1)$, $x_3 = \varphi(x_2)$ and so on; then $|x_k| < \delta$, and

$$|x_2 - x_1| = |\varphi(y) - \varphi(0)| \le \tfrac{1}{2}|y|\,;$$
$$|x_3 - x_2| = |\varphi(x_2) - \varphi(x_1)| \le \tfrac{1}{2}|x_2 - x_1| \le \tfrac{1}{4}|y|\,.$$

And so on;[1] $x_k \in U$, $x_{k+1} = \varphi(x_k)$, $|x_{k+1} - x_k| \le 2^{-k}|y|$. The series $\sum_k |x_{k+1} - x_k|$ converges, thus x_k are a Cauchy sequence, therefore convergent: $x_k \to x$; and $|x| \le 2|y| < \delta$ (since $|x_k| \le 2|y|$ for all k), thus $x \in U$. By continuity of φ, $\varphi(x) = \lim_k \varphi(x_k) = \lim_k x_{k+1} = \lim_k x_k = x$; a fixed point is found, and so, 0 is an interior point of V.

By 2a7(a) it remains to find, for given $y_k \to 0$, some $x_k \to 0$ such that $f(x_k) = y_k$ for all k large enough. This is immediate: we note that $y_k \in V$ (for large k), take $x_k = (f|_U)^{-1}(y_k)$ and use continuity of $(f|_U)^{-1}$. □

All theorems are proved, since 2c3 implies 2b1 by 2c4, and 2b1 implies 2b3 as shown in Sect. 2b.

2d4 Exercise. Assume that $f : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable near the origin, and $(D_1 f)_0 \ne 0, \dots, (D_n f)_0 \ne 0$. Then the equation $f(x_1, \dots, x_n) = 0$ locally defines n functions $x_1(x_2, \dots, x_n)$, $x_2(x_1, x_3, \dots, x_n)$, $\dots$, $x_n(x_1, \dots, x_{n-1})$. Find the product

$$\frac{\partial x_1}{\partial x_2}\frac{\partial x_2}{\partial x_3}\dots\frac{\partial x_{n-1}}{\partial x_n}\frac{\partial x_n}{\partial x_1}$$

at the origin.[2]

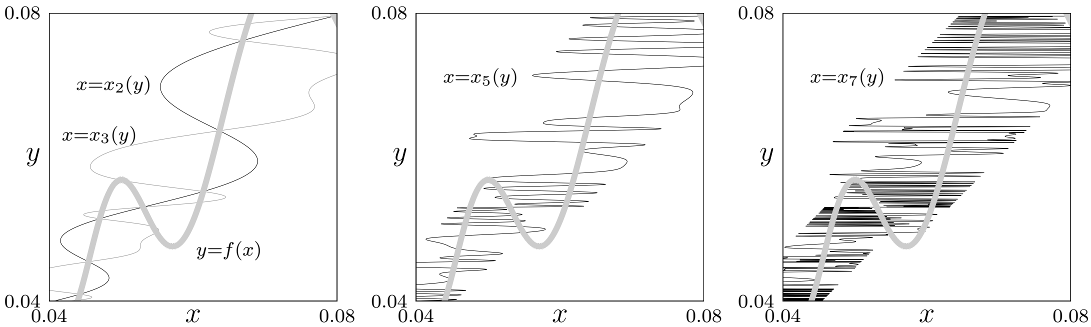
---

[1]More formally, we prove by induction in k existence of $x_1, \dots, x_k \in U$ such that $x_1 = y$, $x_2 = \varphi(x_1)$, $\dots$, $x_k = \varphi(x_{k-1})$, $|x_k - x_{k-1}| \le 2^{-k} \cdot 2|y|$, and $|x_k| \le \bigl(1 - 2^{-k}\bigr) \cdot 2|y|$.

[2]Hint: first, consider a linear f.

Using iteration in the proof only we need not bother about rate of convergence. However, iteration is quite useful in computation. For fast convergence, the transition from x_k to x_{k+1} is made via $A_k = (Df)_{x_k}$ rather than $A = (Df)_0$.[1]

On the other hand, using only $A = (Df)_0$ we could hope for convergence assuming just differentiability of f near 0 (rather than continuous differentiability). Let us try it for f of Example 2a1. Some x_n, shown here as functions of y, are discouraging.



This is instructive. Never forget the word "continuously" in "continuously differentiable"![2]

True, the mean value theorem, and the finite increment theorem (1f31), assume just differentiability. But this is a rare exception.

In the linear case, according to Sect. 2a (Item "linear algebra"), not only $A(x,y) = 0 \iff y = -C^{-1}Bx$, but also $A(x,y) = z \iff y = C^{-1}(z-Bx)$. In the nonlinear case the situation is similar.

2d5 Theorem. Let $f : \mathbb{R}^{n-m} \times \mathbb{R}^m \to \mathbb{R}^m$ and A, B, C be as in Th. 2b3. Then there exists $g : \mathbb{R}^{n-m} \times \mathbb{R}^m \to \mathbb{R}^m$, continuously differentiable near $(0,0)$, such that the two relations $f(x,y) = z$ and $y = g(x,z)$ are equivalent for (x,y,z) near $(0,0,0)$; and $(Dg)_{(0,0)} = \big(-C^{-1}B \,\big|\, C^{-1}\big)$.

Proof. Similarly to the proof of the implication 2b1 $\Longrightarrow$ 2b3 (in Sect. 2b) we introduce the local diffeomorphism φ, its inverse ψ, define g by $\big(x, g(x,z)\big) = \psi(x,z)$, note that $\binom{x}{z} \stackrel{\psi}{\longmapsto} \psi\binom{x}{z} = \binom{x}{g(x,z)} \stackrel{\text{linear}}{\longmapsto} g(x,z)$, and finally, $f(x,y) = z \iff \varphi(x,y) = (x,z) \iff \psi(x,z) = (x,y) \iff y = g(x,z)$ for (x,y,z) near $(0,0,0)$. $\square$

[1] If interested, see Hubbard, Sect. 2.7 "Newton's method" and 2.8 "Superconvergence".

[2] Differentiable functions are generally monstrous! In particular, such a function can be nowhere monotone. Did you know? Can you imagine it? See, for example, Sect. 9c of my advanced course "Measure and category".

Index