LECTURE 7. LINEARITY, LINEARIZABILITY, LINEARIZATION

ROTHSCHILD–CAESARIA COURSE, 2011/2

Continuous functions (and maps) were defined as maps which admit a good local approximation by constant maps sending everything to a single point (value). The resulting notion of continuity prompted us to study properties of sets (subsets of $\mathbb{R}^n$), such as connectedness and compactness, which are preserved by continuous maps.

Now we make one step further and consider maps (again between subsets of the Euclidean spaces $\mathbb{R}^n$) which admit a good local approximation by *linear* (more precisely, affine) maps. This will bring us to the realm of Analysis proper.

1. Linear and affine maps: a crash course

In this section we recall the most basic terms and constructions from the *Linear Algebra*. The subject itself is much more elaborate and deep, and its infinite-dimensional version, called *Functional Analysis*, is one of the most powerful tools of the 20th century Mathematics.

1.1. **Linear space.** A linear space (over a number field, which will almost always be $\mathbb{R}$ in our case) is a set of objects V (called *vectors*) with a special element 0 and two operations, addition $V \ni v, w \mapsto v + w \in V$ and multiplication by constants, $\lambda \in \mathbb{R},\ v \in V \mapsto \lambda \cdot V \in V$ (sometimes the result is written as λv for brevity). These two operations are expected to satisfy all natural properties:

$$\begin{aligned} v + w &= w + v, & \lambda \cdot (v + w) &= \lambda \cdot v + \lambda \cdot w, \\ (u + (v + w)) &= ((u + v) + w), & (\lambda + \mu) \cdot v &= \lambda \cdot v + \mu \cdot w, \\ v + 0 &= v, & (\lambda\mu) \cdot v &= \lambda \cdot (\mu \cdot v), \\ 0 \cdot v &= 0, & 1 \cdot v &= v. \end{aligned} \tag{1}$$

Note that *no multiplication between vectors is defined*!

Problem 1. Prove that for any $v \in V$ there exists $v' \in V$ such that $v + v' = 0$ (the zero vector). This means that V is a commutative group with respect to the operation $+$.

Date: January 17, 2012.

Example 1. The easiest mathematically, but somewhat difficult psychologically is the field $\mathbb{R}$ itself. Indeed, all above axioms are clearly satisfied if u, v, w, λ, μ are real numbers.

The psychological difficulty is to separate between $\mathbb{R}$ as a linear space (whose elements should not be multiplied between themselves) and $\mathbb{R}$ as the field of constants which allow for any arithmetic operations. To stress this difference, we will use the symbol $\mathbb{R}^1$ when $\mathbb{R}$ is considered as a linear space, so that the law of multiplication by constants will be the application $\mathbb{R} \times \mathbb{R}^1 \times \mathbb{R}^1$, $(\lambda, v) \mapsto \lambda \cdot v$.

Example 2. If V, W are two linear spaces over the same field, then their Cartesian product $V \times W$ has the *natural* structure of the linear space again. The naturality means that there is only one reasonable way one can define the sum of two pairs $(v_1, w_1), (v_2, w_2) \in V \times W$ and the result of multiplication of the pair $(v, w) \in V \times W$ by a constant $\lambda \in \mathbb{R}$, using the linear structures on V and W respectively.

Problem 2. Write the corresponding formulas and prove that all the axioms (1) will be satisfied.

Problem 3. Prove that the set $\mathbb{R}^n = \{(x_1, \dots, x_n) \colon x_i \in \mathbb{R}^1\}$ is a linear space. Why did we choose $\mathbb{R}^1$ rather than $\mathbb{R}$, as before, in this formulation?

The number n is (not surprisingly) called the *dimension* of the space $\mathbb{R}^n$. The initial chapters of Linear algebra explain, how the notion of dimension can be generalized from this rather specific example to the general case. The key idea is to introduce the notion of linear dependence between vectors, existence of a *linear combination* $\lambda_1 v_1 + \cdots + \lambda_n v_n$ which is equal to the zero vector 0, while not all coefficients $\lambda_1, \dots, \lambda_n$ vanish simultaneously. We will not pursue this subject, since will be mainly interested with the cases $n = 1, 2$ where the definitions will be completely transparent. In some sense, there are no other examples of finite-dimensional linear spaces apart from $\mathbb{R}^n$.

Problem 4. Prove that the following sets form linear spaces over $\mathbb{R}$:
(1) polynomials in one variable $\mathbb{R}[x]$;
(2) polynomials of degree $\leqslant n - 1$ with real coefficients;
(3) functions on any subset $A \subset \mathbb{R}^n$;
(4) continuous functions on any subset $A \subseteq \mathbb{R}^n$;
(5) polynomials vanishing at all points of a given set $A \subseteq \mathbb{R}$;

Which of these spaces are finite-dimensional in your opinion?

1.2. **Linear maps.** A map between two linear spaces is called linear, if it "respects" both operations.

Definition 1. A map $F \colon V \to W$ is called *linear*, if $F(v_1 + v_2) = F(v_1) + F(v_2)$ and $F(\lambda v) = \lambda F(v)$ for any $v, v_1, v_2 \in V$ and $\lambda \in \mathbb{R}$. If $W = \mathbb{R}^1 \simeq \mathbb{R}$, such map is called *linear function* or *linear functional.*

Proposition 1. *Composition of two linear maps is again a linear map. If a linear map is invertible, then the inverse is also linear.*

Sum of two linear functionals is a linear functional. The constant multiple of a linear functional is a linear functional.

Note that in general we cannot define the sum of two linear maps: if $F\colon V \to W$ and $G\colon V \to W'$ are two linear maps with different ranges, then addition of elements $w + w'$ from two different spaces is in general not defined. However, if both maps have the same domain and target spaces, then one can well define their sum $F + G$. It will again be a linear map.

The "product" (not a composition!) of two linear maps is in general not defined even if both maps have the same target space. The only exception appears when one of these target spaces is the field $\mathbb{R}$ itself, i.e., when one of the maps is a linear function. Then the product will be well-defined map, however, it will not be linear (give an example!).

Proof. All assertions except the inverse map are obvious consequences of the definition. To prove it, assume that $G = F^{-1}\colon W \to V$ is the inverse, and consider $G(w_1 + w_2)$. By invertibility, $w_i = F(v_i)$, $i = 1, 2$, $v_i \in V$, that is, $G(w_i) = v_i$. By the linearity of F, we have

$$G(w_1+w_2) = G\big(F(v_1)+F(v_2)\big) = G\big(F(v_1+v_2)\big) = v_1+v_2 = G(w_1)+G(w_2).$$

The second check is even easier. □

Problem 5. Find necessary and sufficient conditions for a map $\mathbb{R}^1 \to \mathbb{R}^1$, $x \mapsto ax + b$, be linear. Note the difference with the "high school" definition of a linear function.

Example 3. The linear maps from $\mathbb{R}^1$ to $\mathbb{R}^1$ are obvious: all of them have the form $x \mapsto ax$, $a \in \mathbb{R}$. The constant a is called the *multiplicator* of a linear 1D-map.

The multiplicator of the composition of two linear 1D-maps is the product of the multiplicators. Since the product in $\mathbb{R}$ is commutative, the composition of 1D-linear maps is commutative: $F \circ G = G \circ F$. This is not the case in higher dimensions!

A linear map is invertible if and only if its multiplicator is non-zero. The constant (identically zero) map $x \mapsto 0x \equiv 0$ is linear, but not invertible.

There is a distinction between linear 1D-maps with $a > 0$ and $a < 0$. The linear maps with positive multiplicator preserve the order: for any two points $x, y \in \mathbb{R}^1$ if $x > y$, then $F(x) > F(y)$. Usually such functions are called *strictly* monotonously increasing. The maps with negative multiplicator reverse the order (are strictly decreasing).

Example 4. Linear maps from $\mathbb{R}^1$ to $\mathbb{R}^2$ are pairs of two linear maps $t \mapsto x$ and $t \mapsto y$ and hence are defined by two multiplicators:

$$t \mapsto \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} at \\ bt \end{pmatrix}, \quad a, b \in \mathbb{R}.$$

If $a \neq 0$, then the equation $x = at$ can be resolved with respect to t, $t = \frac{1}{a}x$, and the result substituted into the second equation, yielding $y = \frac{b}{a}x$. Thus the range of a linear map is a line passing through the origin, horizontal if $b = 0$. The same can be done if $b \neq 0$. Thus in all cases except for $a = b = 0$ the image is a line. In the exceptional case the map is constant and its image is the origin $x = 0$, $y = 0$. Clearly, maps from $\mathbb{R}^1$ to $\mathbb{R}^2$ cannot be invertible.

Example 5. Linear functionals from $\mathbb{R}^2$ to $\mathbb{R}^1$ are also described by two numbers, but these two numbers play completely different role. Note that any vector in $\mathbb{R}^2$ can be written as $x \cdot (1, 0) + y \cdot (0, 1) = x \cdot e_1 + y \cdot e_2$, where $x, y \in \mathbb{R}$ are *real numbers* (coordinates of the vector) and $e_1 = (1, 0)$, $e_2 = (0, 1)$ are two coordinate unit vectors along the horizontal and vertical axes (in the standard drawing). By linearity, any linear functional can be written as

$$\begin{aligned} F(x, y) &= F\big(x \cdot e_1 + y \cdot e_2\big) = xF(e_1) + yF(e_2) \\ &= ax + by, \qquad a = F(e_1),\ b = F(e_2) \in \mathbb{R}. \end{aligned}$$

The linear functional is constant if and only if $a = b = 0$, otherwise it is surjective but never injective. The preimage of any point $(p, q) \in \mathbb{R}^2$ is (assuming the nonconstant case) a line, passing through the origin if $(p, q) = (0, 0)$ is the origin, and orthogonal (perpendicular) to the vector with the coordinates (a, b) (prove that!).

The world of linear maps from $\mathbb{R}^2$ to $\mathbb{R}^2$, geometrically visualized as the transformations of the plane, is much more rich than the previous examples.

Example 6. A linear map $\mathbb{R}^2 \to \mathbb{R}^2$ is determined by 4 real numbers: each of the two coordinate function is a linear functional defined by 2 real numbers. These four real numbers are conveniently organized into the 2×2-table (called the *matrix*). For those familiar with the matrix multiplication law, the image of the vector (x, y) considered as a column 1×2-matrix is the matrix product:

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix}$$

Invertibility of the map depends on the *determinant* of the matrix, the number $\Delta = ad - bc \in \mathbb{R}$. If $\Delta \neq 0$, then the system of equations

$$\begin{aligned} ax + by &= u, \\ cx + dy &= v, \end{aligned}$$

can be solved for any $(u, v) \in \mathbb{R}^2$ with respect to (x, y) (write down the explicit formulas for the solution!), and this solution is unique. In particular, from this uniqueness it follows that $F^{-1}(0, 0) = (0, 0)$, that is, for $\Delta \neq 0$ the linear map is both injective and surjective, i.e., invertible. By Proposition 1 the inverse map is also linear (write the coefficients of the matrix for it). Such linear maps are also called *nondegenerate*.

The linear maps with $\Delta = 0$ can be of two types: if not all 4 numbers are zeros, then the image is a line passing through the origin in $\mathbb{R}^2$, and the preimage of the origin is also a line (compute these lines). If all numbers are zero, the linear map is constant (everything is sent to the origin).

The difference between the cases of positive and negative (nonzero) Δ can be best understood from the two examples:

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \\ -y \end{pmatrix}, \qquad \text{and} \qquad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} -x \\ -y \end{pmatrix}.$$

The first map is the reflection in the x-axis, and it "changes the orientation". The second map is the reflection in the origin (the central symmetry) and it preserves the orientation: the result can be considered as the rotation by π. The determinant is negative (-1) for the first map and positive ($+1$) for the second.

Orientation. The notion of orientation is rather subtle, despite its "apparent" geometrical meaning. It is determined by an *ordered* pair v_1, v_2 of non-collinear vectors on $\mathbb{R}^2$. Such a pair forms two angles, one smaller than π, another larger (in the absolute value). We say that the orientation of the pair is *positive*, if the rotation from v_1 to v_2 in the "shorter" direction is counterclockwise, otherwise the orientation of the pair is called *negative*. The standard coordinate vectors $e_1 = (1, 0)$ and $e_2 = (0, 1)$ form a positive pair (this is the way to remember the sign). We say that a nondegenerate linear map F preserves orientation, if the orientation of any pair of images $F(v_1), F(v_2) \in \mathbb{R}^2$ is the same as the orientation of the initial pair $v_1, v_2 \in \mathbb{R}^2$.

1.3. **Continuity and the norm of a linear operator.** For a linear map $L\colon \mathbb{R}^m \to \mathbb{R}^n$ each component of the image is a linear combination of the coordinates $x_1, \dots, x_m$ with constant coefficients. Therefore the map is continuous: for any $\varepsilon > 0$ there exists a positive $\delta > 0$ such that the cube $\delta \boldsymbol{I}^m$ is mapped inside the cube $\varepsilon \boldsymbol{I}^n$, $L(\delta \boldsymbol{I}^m) \subseteq \varepsilon \boldsymbol{I}^n$.

For general continuous maps the dependence of δ on ε can be arbitrary: what is important is the condition that $\delta \to 0^+$ as $\varepsilon \to 0^+$. For linear operators this dependence can be made much more explicit.

Proposition 2. *For any linear operator $L\colon \mathbb{R}^m \to \mathbb{R}^n$ there exists a finite constant $C < +\infty$ such that*

$$\forall v \in \mathbb{R}^m \quad \|L(v)\| \leqslant C \cdot \|v\|. \tag{2}$$

Proof. Consider the ratio $f(v) = \frac{\|L(v)\|}{\|v\|}$ defined on $\mathbb{R}^m \smallsetminus \{0\}$. This is a positive real-valued function continuous on its domain. By linearity, $f(\lambda v) = |\lambda|\, f(v)$ for any $\lambda \neq 0$. Thus $\sup_{v \neq 0} f(v) = \sup_{\|v\|=1} f(v)$. But the "unit sphere" $S = \{v : \|v\| = 1\}$ is a compact subset, so by continuity the right hand side is a finite positive number C. □

Corollary 3. *In the definition of the continuity, one can choose $\delta = \frac{1}{C}\varepsilon$.*

Remark 1. The maximum

$$\|L\| = \sup_{v \neq 0} \frac{\|L(v)\|}{\|v\|}$$

is called the *norm* of the linear operator L. As follows from Proposition 2, this norm is always finite.

In the coordinates $(x_1, \dots, x_m)$ on $\mathbb{R}^m$ and $y_1, \dots, y_n$ on $\mathbb{R}^n$, the map L can be defined by the formulas

$$\begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \mapsto \begin{pmatrix} y_1 \\ x_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} c_{11}x_1 + \cdots + c_{1m}x_m \\ c_{21}x_1 + \cdots + c_{2m}x_m \\ \vdots \\ c_{n1}x_1 + \cdots + c_{nm}x_m \end{pmatrix}, \tag{3}$$

where $\{c_{kl}\}_{k=1,\dots,n,\ l=1,\dots,m}$ are nm real numbers forming the matrix of L. One can easily verify that

$$\|L(x)\| = \max_{k=1,\dots,n} |c_{k1}x_1 + \cdots + c_{km}x_m| \leqslant Cm \max_{l=1,\dots,m} |x_l| = Cm\|x\|,$$

where $C = \max |c_{kl}|$ is the maximal absolute value of the matrix elements of L. Thus $\|L\| \leqslant \max_{k,l} |c_{kl}|$.

1.4. **Affine maps.** The notion of linearity is very natural algebraically, but geometrically it is not broad enough to provide a sufficient supply for approximations. Indeed, any linear map always maps origin (the zero vector) to the origin, $F(0) = 0$ (note that the same symbol 0 denotes two different zero vectors in different spaces!), so even constant maps cannot be approximated unless they are identically zero. However, this can be easily corrected.

Definition 2. An affine map between two linear spaces is any map which is the sum of a linear map and a constant map between these spaces:

$$A(x) = L(v) + C(v), \qquad L\colon V \to W \quad \text{linear}, \ C\colon V \to W \quad \text{constant}.$$

The linear map $L\colon V \to W$ is called the *linear part* or *differential* of the affine map A.

The constant map C is completely characterized by the image point $p = F(0) \in V$, thus knowledge of linear maps allows to write down general formulas for affine maps:

(1) $A\colon \mathbb{R}^1 \to \mathbb{R}^1$ has the form $A(x) = ax + b$, $a, b \in \mathbb{R}$;
(2) $A\colon \mathbb{R}^2 \to \mathbb{R}^1$ has the form $A(x, y) = ax + by + c$, $a, b, c \in \mathbb{R}$;
(3) $A\colon \mathbb{R}^1 \to \mathbb{R}^2$ has the form $A(t) = (a_1t + b_1, a_2t + b_2)$, $a_i, b_i \in \mathbb{R}$, $i = 1, 2$;
(4) $A\colon \mathbb{R}^2 \to \mathbb{R}^2$ has the form

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} ax + by + u \\ cx + dy + v \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} u \\ v \end{pmatrix}, \qquad a, b, c, d, u, v \in \mathbb{R}.$$

Problem 6. Prove that an affine map is invertible if and only if its linear part is invertible. Show that the inverse map is also affine.

1.5. **Translations. Affine maps as compositions of linear maps with translations.** The representation $A(v) = L(v) + b$, $b \equiv C(v)$ means that the affine map A is the *post*composition $T_b \circ L$ of the linear map $L\colon V \to W$ with the *parallel translation map*, or *shift* $T_b\colon W \to W$, $T_v(w) = w+b$ which is a self-map of W. Note that the translation T_b by the vector b *is not* a linear map (the image of the zero vector is nonzero), neither it is constant (translates of two distinct points are always distinct).

The translations by different vectors obviously are closed by the operation of composition, moreover, they commute between themselves:

$$\forall b, c \in W \quad T_b \circ T_c = T_c \circ T_b = T_{b+c}, \quad T_b \circ T_{-b} = \mathrm{id} \quad \text{(the identity map)}.$$

Note that the translation maps always have the same domain and target. Thus in general it makes no sense to say that translations commute with linear maps unless $V = W$. However, instead of representing an affine map as a *post*composition with a translation, one can represent it as a *pre*composition and write

$$A = L \circ T_a, \qquad T_a\colon V \to V, \quad T_a(v) = v + a,$$

with a suitable shift vector a. In is instructive to compare two representations for *the same* affine map:

$$L \circ T_a = T_b \circ L \iff L(v+a) = L(v) + b \iff L(a) = b,$$

i.e., for a nondegenerate linear map one can freely choose one of the two compositions.

Problem 7. Prove that the symmetry in the x-axis commutes with any shift parallel to this axis.

If L is not invertible, then sometimes the post-translation cannot be replaced by pre-translation. However, any affine map can always (and in a unique way) defined by its linear part and the image of any point in the domain.

Proposition 4. *For any two points $p \in V$, $q \in W$ and any linear map $L : V \to W$ there exists a unique affine map $A\colon V \to W$ which has the linear part L and sends p to q, $A(p) = q$.*

Proof. We define the map A explicitly:

$$A(v) = L(v-p) + q, \qquad \text{i.e.,} \quad L = T_q \circ L \circ T_{-p}. \tag{4}$$

To show the uniqueness, note that for any two affine maps A and B with the same linear part L, their difference $F(v) = B(v) - A(v)$ has zero linear part and is a translation. By assumption, $F(0) = B(p) - A(p) = q - q = 0$, but the only translation that keeps the origin at its place, is T_0, the identity (translation by the zero vector). □

Problem 8. Find an affine map $\mathbb{R}^1 \to \mathbb{R}^1$ whose graph passes through the given point $(p, q) \in \mathbb{R}^2$ and has the slope $a \in \mathbb{R}$. How this is related to the above proposition?

Problem 9. Let $A\colon \mathbb{R}^1 \to \mathbb{R}^1$ be a function whose graph (subset of $\mathbb{R}^2$) is a straight line. Prove that A is an affine function and find (in any way) its linear part.

Problem 10. Let $v_i = (p_i, q_i) \in \mathbb{R}^2$, $i = 1, 2$, are two non-collinear vectors. Prove that there exists a unique *linear* map $L\colon \mathbb{R}^2 \to \mathbb{R}^2$ such that L sends the unit coordinate vectors $e_1 = (1, 0)$, $e_2 = (0, 1)$ to v_1 and v_2 respectively. Find this map.

Problem 11. Let $p_1 \neq p_2 \in \mathbb{R}^1$ are two distinct points and $q_1, q_2 \in \mathbb{R}^1$ any other points (eventually, coinciding). Prove that there exists a unique affine map $A\colon \mathbb{R}^1 \to \mathbb{R}^1$ satisfying the assumptions $A(p_i) = q_i$, $i = 1, 2$. Can one always place 3 different points at 3 arbitrary positions by an affine map of the real line $\mathbb{R}^1$?

Problem 12. Let v_1, v_2, v_3 and w_1, w_2, w_3 be two non-collinear triplets of points on $\mathbb{R}^2$. Prove that there exists a unique *affine* map $A\colon \mathbb{R}^2 \to \mathbb{R}^2$ such that $A(v_i) = w_i$, $i = 1, 2, 3$.

Hint. Assume that $v_3 = v_3 = 0$ and then show how one can treat the general case by pre/post/compositions with shifts.

Proposition 5. *Composition (in any admissible order) of affine maps is again an affine map. The linear part of the composition is the composition of the linear parts (in the same order).*

Problem 13. Prove this obvious statement.

2. Linearizability at a point

In what follows we will be dealing with maps between subsets of linear spaces and their *local behavior* near certain points. To reduce the number of letters used for various notations and bury some $\exists$-quantifiers inside, we will use the following mnemonic rule.

2.1. **What is a linear approximation?** If V is a linear space and $p \in V$ a point in this space, then the notation (V, p) will be used for *some open neighborhood of the point p in the domain V*. The size (and specific shape) of this neighborhood will be not important for our purposes (unless specifically mentioned). If F is defined in some such neighborhood (V, p) and is continuous, then by definition the image of this neighborhood (if sufficiently small) will belong to any specified neighborhood (W, q) in the target space, with $q = F(p) \in W$. Thus the notation

$$F\colon (V, p) \to (W, q), \qquad p \in V,\ q \in W, \tag{5}$$

means that F is defined in some open neighborhood of p in V, maps p into q and takes values in some small neighborhood (W, q) around the image $q = F(p)$. Mathematicians say that we are given only the *germ* of a map.

Note that this shortcut already conceals quite a lot of information about F (in particular, continuity and the value at the given point). Manipulation with this notation assumes that the reader verifies that the usage is indeed legitimate.

Consider first the case of maps which "preserve" the origin, $F(0) = 0$, that is, the maps $F\colon (V,0) \to (W,0)$ in our notation. We need to introduce the notion of "accurate approximation" of F by a linear operator $L\colon V \to W$ in such a way that:

(i) there is at most one "accurate approximation" operator L;
(ii) reasonably large and interesting class of nonlinear maps would admit such approximation.

The idea is to consider the "relative error function"

$$E(x) = \frac{1}{\|x\|} \cdot \bigl(F(x) - L(x)\bigr) \in \mathbb{R}^n, \qquad 0 \neq x \in (\mathbb{R}^m, 0). \tag{6}$$

Since the norm $\|x\|$ is nonzero for $x \neq 0$, this map is is well defined near the origin. However, the ratio apriori need not be continuous at $x = 0$, so one cannot replace $\mathbb{R}^n$ by $(\mathbb{R}^n, 0)$.

Example 7. If $F(v) \equiv 0$ and L is a nontrivial (i.e., not identically zero) linear map, then the relative error function of the corresponding "approximation"

$$E(x) = -\frac{L(x)}{\|x\|}, \qquad x \neq 0,$$

is *zero order homogeneous*, $E(x) = E(\lambda x)$ for any scalar (number) $\lambda \neq 0$. This means that the image of the small cube $\delta \overline{\boldsymbol{I}}^m$ coincides with the image of the unit cube $\overline{\boldsymbol{I}}^n$ and even its boundary, the "sphere" $S^m = \{x \in \mathbb{R}^m \colon \|x\| = 1\}$.

Note that the sphere is compact (why?), so its image $L(S^m)$ is compact. If L is a nontrival operator, this image does not reduce to a single point (the origin), thus contains a nonzero vector $0 \neq v$ and its antipode $-v$, and hence $E(x)$ cannot be continuous at the origin.

Definition 3. A continuous map $F\colon (V,0) \to (W,0)$ is called *differentiable* (or *linearizable*) at the origin $0 \in V$, if it admits a linear map $L\colon V \to W$ which approximates F with the relative error which tends to zero as $x \to 0$:

$$\lim_{x \to 0} E(x) = 0 \in \mathbb{R}^m, \qquad E(x) = E_{F,L}(x) = \tfrac{1}{\|x\|} \cdot \bigl(F(x) - L(x)\bigr). \tag{7}$$

The map L is called *linearization*, or *differential* of F at the origin.

Problem 14. Prove that if the linearization exists, it must be unique.

Solution. Assume that F is differentiable and L its linearization: $F(x) = L(x) + \|x\|\, E(x)$ with $\lim_{x\to 0} E(x) = 0$. For any other linearization L' the corresponding relative error function will be

$$E'(x) = \frac{1}{\|x\|}(F(x) - L'(x)) = \frac{1}{\|x\|}(L(x) - L'(x)) + E(x).$$

The second term converges to the origin, while the first term has no limit unless the operator $L - L'$ is not identically zero by Example 7. □

Definition 4. A continuous map $F\colon (V,p) \to (W,q)$ is linearizable at the point $p \in V$, if the "translated" map $T_{-q} \circ F \circ T_p\colon (V,0) \to (W,0)$ is linearizable in the sense of Definition 3. The corresponding linear operator L is called the *differential* of F at p.

In other words, for general maps instead of the linear approximation we need to use the affine approximation. If L is the linear map approximating $T_{-q} \circ F \circ T_p$, then by construction the affine map $A = T_q \circ L \circ T_{-p}$ approximates F in the sense that the "modified" error function

$$E(x) = \frac{1}{\|x-p\|}\bigl(F(x) - q - L(x-p)\bigr)$$

is continuous and vanishing at $x = p$, $\lim_{x\to p} \|E(x)\| = 0$. We stress here that although the approximation is given by an affine map, the differential is linear rather than affine.

2.2. **Notation.** The construction of the linearization has too many parameters: the initial map F itself, the point $p \in V$ at which the linearization is computed, besides, the linear approximation is itself not a single number, but rather a linear map $L\colon V \to W$. The notation must incorporate all these dependencies without too much clutter.

The first convention consists in removing the braces around the argument of a linear map and identify the linear operator L with the corresponding $n \times m$-matrix. The image of a vector $v \in V$ will be denoted then by the usual product notation, with or without the dot: $L\colon v \mapsto Lv$ or $v \mapsto L \cdot v$.

The notation for the differential (the linear map) approximating a map $F\colon (V,p) \to (W,q)$ will be $\mathrm{d}F(p)$:

$$\begin{gathered} F(x) = F(p) + \bigl(\mathrm{d}F(p)\bigr)\cdot(x-p) + \|x-p\|E(x), \\ \mathrm{d}F(p)\colon V \to W \text{ linear map, the differential,} \qquad \lim_{x\to p} E(x) = 0. \end{gathered} \tag{8}$$

We stress again that *the differential has "two arguments"*: one is the "point" p, at which the approximation is computed, the other the "vector attached to the point p", which is to be fed into the linear operator to obtain the approximation:

$$F(p+v) = F(q) + \mathrm{d}F(p)\cdot v + \cdots$$

with the dots denoting the "negligible error".

2.3. **Particular cases: derivatives, partial derivatives, Jacobian matrix.** Now after we know that differentiability is a nontrivial property, we need to be sure that there is a large enough class of differentiable maps (functions).

Problem 15. Any affine map $A\colon V \to W$ with the linear part L is differentiable at any point $p \in V$, and the differential is the same L for all points. Prove that.

Example 8 (Differentiability and derivative for functions of one variable)**.** Let $m = m = 1$ and $F\colon (\mathbb{R}^1, p) \to (\mathbb{R}^1, q)$ is a continuous function defined in some neighborhood of the point p with $q = F(p)$. If this function is differentiable, this means that there exists an affine map $x \mapsto a(x-p)+q$ which approximates F with the relative error tending to zero as $x \to a$:

$$F(x) - q - a(x-p) = |x-p| \cdot E(x), \qquad \lim_{x\to p} E(x) = 0.$$

In this case the multiplicator (the only "matrix element") of the differential $\mathrm{d}F(p)\colon \mathbb{R}^1 \to \mathbb{R}^1$, $v \mapsto av$ can be expressed as follows:

$$a = \frac{F(x)-q}{x-p} + |E(x)|, \qquad x \in (\mathbb{R}^1, p),\ x \neq p. \tag{9}$$

Note that the second term in this formula has zero limit as $x \to p$. Thus we conclude that

$$a = \lim_{x\to p} \frac{F(x)-F(p)}{x-p}. \tag{10}$$

Conversely, if the limit in (10), called the *derivative* (of F at p) exists, then we see that the relative error in (9) tends to zero and the function F turns out to be differentiable.

Thus the usual notion of the derivative as the limit for the ratio (10) for functions of one variable is indeed a particular case of the differentiability of maps.

Remark 2. The notation for the derivative is one of the most variable in the entire Mathematics. The limit in (10) can be denoted by one of the following ways:

$$F'(p), \quad \frac{\mathrm{d}F}{\mathrm{d}x}(p), \quad \left.\frac{\mathrm{d}}{\mathrm{d}x}\right|_{x=p} F, \quad \frac{\mathrm{d}F(p)}{\mathrm{d}x}, \quad F_x(p), \quad \dots$$

For the differential (recall, this is a linear map $L\colon V \to W$, approximating F at a point p) the spectrum of notations is also quite impressive:

$$\mathrm{d}F(p), \qquad F_{*,p}, \qquad \left(\frac{\partial F}{\partial x}\right)_p, \qquad \dots$$

Note that in all cases the differential maps 0 to 0, but it usually depends on the point p at which it is computed.

Remark 3. The expression $\mathrm{d}x$ can be considered as the linear map approximating the identical map $F\colon \mathbb{R}^1 \to \mathbb{R}^1$, $F(x) = x$. The approximated map (function) is itself already affine (even linear), so $\mathrm{d}x(p)\colon \mathbb{R}^1 \to \mathbb{R}^1$ is also the identical map, $v \mapsto v$, which does not depend on p (and hence the indication for p can be omitted).

This tautology may be accepted as the answer for the sacramental question "what is $\mathrm{d}x$" in the notation for the derivative as a fraction and the integral: for a function of one variable $\mathrm{d}F(p)$ and $\mathrm{d}x(p)$ are two linear maps, $v \to av$ and $v \to v$ respectively. The *ratio of these linear functions* is the *number*, $\frac{av}{v} = a \in \mathbb{R}$.

Example 9 (Partial derivatives for functions of several variables)**.** Consider a scalar function $F(x, y)$ of two variables defined in a neighborhood of a point on $\mathbb{R}^2$. For simplicity we assume that this point is at the origin, and $F(0, 0) = 0$.

Differentiability of F at the origin means that there exist a linear map $L:(x,y)\mapsto ax+by$ with the two "matrix elements" $a,b\in\mathbb{R}$, which approximates F with a relative error tending to zero:

$$F(x,y)=ax+by+\max\{|x|,|y|\}\cdot E(x,y),\qquad \lim_{(x,y)\to(0,0)} E(x,y)=0.$$

If F is differentiable, then the coefficients a,b can be recovered as the limits of the appropriate differences,

$$a=\lim_{x\to 0}\frac{F(x,0)-F(0,0)}{x-0},\qquad b=\lim_{y\to 0}\frac{F(0,y)-F(0,0)}{y-0}. \tag{11}$$

Adding/subtracting the zero constants 0 and $F(0,0)=0$ was done on purpose, to suggest the form of the answer for the general case where the function is defined near a point $(p,q)\in\mathbb{R}^2$ and takes a nonzero value $c=F(p,q)$ there (write these formulas in the general case).

The limits in (11) are called *partial derivatives* and denoted by many different ways,

$$\frac{\partial F}{\partial x},\ \frac{\partial F}{\partial y},\qquad F_x,\ F_y,\qquad \dots$$

with the point (p,q) explicitly indicated when necessary.

However, unlike for univariate functions, the mere existence of the two limits (11) does not guarantee that the function admits a nice linear approximation.

Example 10. Indeed, consider the example

$$F(x,y)=\sqrt{|x|}\cdot\sqrt{|y|},\qquad (x,y)\in(\mathbb{R}^2,0).$$

This function has the following properties, immediately following from its definition:

$$F(x,0)\equiv 0,\qquad F(0,y)\equiv 0,\qquad F(z,z)=|z|.$$

The first two identities imply that if F were differentiable, the linear approximation must be zero, $L(v)=0$. Yet the relative error $E(x)$ computed along the line $y=x$, $x\neq 0$, is equal to 1 and does not tend to zero. Thus the function cannot be differentiable despite continuity and existence of both partial derivatives.

Example 11 (Notation for the differential)**.** Consider two (linear) coordinate functions from $\mathbb{R}^2$ to $\mathbb{R}^1$:

$$x\colon (x,y)\mapsto x,\qquad y\colon (x,y)\mapsto y.$$

The differentials of these two functions are linear maps, denoted $\mathrm{d}x$ and $\mathrm{d}y$ respectively: since they do not depend on the point at which the approximation is computed, we can drop the point from the notation:

$$\mathrm{d}x:\mathbb{R}^2\to\mathbb{R}^1,\qquad \mathrm{d}y\colon\mathbb{R}^2\to\mathbb{R}^1.$$

Having fixed notation for these two linear maps, any other linear map $(x, y) \mapsto ax + by$, $a, b \in \mathbb{R}$, can be written as the sum $a\,\mathrm{d}x + b\,\mathrm{d}y$. In particular, we have the identity

$$\mathrm{d}F(p,q) = a\,\mathrm{d}x + b\,\mathrm{d}y,$$
$$a = a(p,q) = \frac{\partial F}{\partial x}(p,q) = \lim_{x\to p} \frac{F(x,q) - F(p,q)}{x - p},$$
$$b = b(p,q) = \frac{\partial F}{\partial y}(p,q) = \lim_{y\to q} \frac{F(p,y) - F(p,q)}{y - q}.$$

Dropping the indication of the point (p, q) at which the differential is computed, we obtain the formula

$$\mathrm{d}F = \frac{\partial F}{\partial x}\,\mathrm{d}x + \frac{\partial F}{\partial y}\,\mathrm{d}y. \tag{12}$$

You may have seen it before, now you know what every symbol in this formula stands for.

2.4. **Differential and derivative of the product. The Leibniz formula.** If F, G are two functions defined in a common neighborhood (V, p) and with the same range W, then one can define their sum $F + G$ pointwise:

$$(F + G)(x) = F(x) + G(x), \qquad x \in (V, p).$$

This definition makes sense because the target space is linear and the sum of two vectors is well defined.

Clearly, if both these maps are differentiable and $L = \mathrm{d}F(p)$ and $M = \mathrm{d}G(p)$ are the two corresponding linear approximations (differentials), then the linear map $L + M$, defined by the obvious rule $(L + M) \cdot v = Lv + Mv$ (recall our convention about arguments of the linear maps), is the approximation for the sum.

Indeed, in this case the relative error of the sum will be the sum of respective relative errors E_F and E_G:

$$(F + G)(x) - (F + G)(p) - (L + M) \cdot (x - p) = \|x - p\| \cdot (E_F(x) + E_G(x)),$$
$$E_F(x) + E_G(x) \to 0 \quad \text{as } x \to p,$$

by linearity.

However, the product of two functions is in general not defined, unless one (or both) functions have the one-dimensional space $\mathbb{R} = \mathbb{R}^1$ as the range. But even in this case the product of linear approximations will be non-linear anymore (quadratic).

Example 12. Assume $F\colon (\mathbb{R}^1, 0) \to (\mathbb{R}^1, c)$ and $G\colon (\mathbb{R}^1, 0) \to (\mathbb{R}^1, d)$ be two differentiable maps in a neighborhood of the point $x = 0$. Denote by $v \mapsto av$ and $v \mapsto bv$ their respective linearizations:

$$F(x) = c + ax + |x|E_F(x), \qquad G(x) = d + bx + |x|E_G(x),$$
$$E_F(x), E_G(x) \to 0 \text{ as } x \to 0.$$

Then the product FG can be represented as

$$\begin{aligned}(FG)(x) &= (c + ax + |x|E_F(x))(d + bx + |x|E_G(x)) \\ &= cd + (ad + bc)x + (ab)x^2 + |x|(b\,E_F(x) + \cdots + c\,E_G(x)) \\ &= cd + (ad + bc)x + |x|E_{FG}(x),\end{aligned}$$

where the terms collected and denoted by $E_{FG}(x)$ tend to zero as $x \to 0$. Thus we conclude that the affine approximation for FG at the origin is the affine function $ac + (ad + bc)x$.

Expressing the values $c = F(0)$, $c = G(0)$, $a = \frac{\mathrm{d}F}{\mathrm{d}x}(0)$, $a = \frac{\mathrm{d}G}{\mathrm{d}x}(0)$, we conclude that at the point $x = 0$, the derivative of the product has the form

$$\frac{\mathrm{d}(FG)}{\mathrm{d}x}(0) = F\,\frac{\mathrm{d}G}{\mathrm{d}x}(0) + G\,\frac{\mathrm{d}F}{\mathrm{d}x}(0). \tag{13}$$

This is the familiar Leibniz rule from the calculus. If the derivative is denoted by the "tag" it takes the form

$$(F \cdot G)' = F' \cdot G + F \cdot G' \tag{14}$$

valid at every point where both F and F are defined and differentiable.

Example 13. Consider the function of two variables $P\colon \mathbb{R}^2 \to \mathbb{R}^1$, $P(x, y) = xy$ ("the product function"). In the way completely similar to the one explained in the previous example, we see that at any point $(p, q) \in \mathbb{R}^2$,

$$\mathrm{d}P(p, q) = p\,\mathrm{d}y + q\,\mathrm{d}x.$$

Indeed, to see this one has to multiply the approximation identities

$$P(x, y) = (p + \mathrm{d}x + \cdots)(q + \mathrm{d}y + \cdots) = pq + (p\,\mathrm{d}y + q\,\mathrm{d}x) + \cdots,$$

where by the dots we denote the terms relatively smaller than the ones explicitly listed. Removing the explicit dependence of the differentials on the point, we arrive at the formula for the differential of the product:

$$\mathrm{d}(xy) = x\,\mathrm{d}y + y\,\mathrm{d}x. \tag{15}$$

Problem 16. Prove that for $x \neq 0$,

$$\mathrm{d}\frac{1}{x} = -\frac{1}{x^2}\mathrm{d}x.$$

Solution.

$$\lim_{x\to p} \frac{\frac{1}{x} - \frac{1}{p}}{x - p} = \lim_{x\to p} \frac{p - x}{x - p} \cdot \frac{1}{xp} = -\frac{1}{p^2}.$$

□

Problem 17. Prove that

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

2.5. **Velocity of a motion.** If $\mathbf{F}\colon(\mathbb{R}^1,\tau)\to(\mathbb{R}^2,\mathbf{p})$ is a differentiable map of the real line (interpreted as the time) to the real plane, then the linear approximation admits a dynamic interpretation as the velocity. In this section we will stress by the boldface the points (vectors) on the plane, while keeping the usual font for the scalars (numbers).

Indeed, any linear map $\mathbf{L}\colon\mathbb{R}^1\to\mathbb{R}^2$ is completely determined by its value on the unit vector $\mathbf{L}\cdot 1=\mathbf{a}=(a_1,a_2)\in\mathbb{R}^2$. Any other vector t will be mapped then into $\mathbf{L}\cdot t=t\cdot\mathbf{a}\in\mathbb{R}^2$. The such linear map corresponds to the planar motion with the constant velocity vector $\mathbf{a}$, which at the moment $t=0$ started from the origin $\mathbf{0}=(0,0)\in\mathbb{R}^2$.

The general case corresponds to the differentiable motion which at the moment $\tau\in\mathbb{R}$ passes through the point $\mathbf{p}$ on the plane. The linearization (affine) map takes the form

$$\mathbf{F}(\tau+t)=\mathbf{F}(\tau)+t\mathbf{a}+|t|\cdot\mathbf{E}(t),\qquad \lim_{t\to\tau}\mathbf{E}(t)=0.$$

The differential of the map is the product $\mathrm{d}\mathbf{F}=\mathrm{d}t\cdot\mathbf{a}$: its value on the vector $v\in\mathbb{R}^1$ is the product of the number $\mathrm{d}t\cdot v$ and the vector $\mathbf{a}\in\mathbb{R}^2$. If the map F is defined by its coordinate functions, $\mathbf{F}(t)=(f_1(t),f_2(t))$, then the components of the velocity vector $\mathbf{a}=(a_1,a_2)$ at the moment t are two derivatives, $a_i=f_i'(t)$.

Problem 18. Let $f\colon(\mathbb{R}^1,0)\to\mathbb{R}^1$ be a scalar map (function of one variable) and $\mathbf{F}\colon(\mathbb{R}^1,0)\to\mathbb{R}^2$ a parameterized curve, both differentiable at $t=0$. Prove the Leibniz formula:

$$\frac{\mathrm{d}(f\cdot\mathbf{F})}{\mathrm{d}t}=\frac{\mathrm{d}f}{\mathrm{d}t}\cdot\mathbf{F}+f\cdot\frac{\mathrm{d}\mathbf{F}}{\mathrm{d}t}.$$

3. The chain rule of the derivation

Not very surprisingly, if each map admits a nice approximation by a linear map, then their composition is nicely approximated by the composition of the linear approximations.

Theorem 6 (Chain Rule of Derivation)**.** *If $F\colon(V,p)\to(W,q)$ and $G\colon(W,q)\to(Z,r)$ are two maps between linear spaces V,W and Z respectively, differentiable at the points p and $q=F(p)$, then the composition $H=G\circ F\colon(V,p)\to(Z,r)$ is also differentiable, and its differential $\mathrm{d}H(p)$ is the composition $\mathrm{d}G(q)\circ\mathrm{d}F(p)$.*

The proof of this theorem is one line long, yet it implies numerous corollaries.

Corollary 7. *The product of two differentiable functions $f,g\colon(V,p)\to\mathbb{R}$ of any number of variables is again a differentiable function and $\mathrm{d}(fg)(p)=f(p)\,\mathrm{d}g(p)+g(p)\,\mathrm{d}f(p)$.*

Indeed, this follows from the differentiability of the map $(V,p)\to\mathbb{R}^2$, $x\mapsto\big(f(x),g(x)\big)$, and the map $:\mathbb{R}^2\to\mathbb{R}$ from Example 13. In particular

$f, g\colon (\mathbb{R}^1, p) \to \mathbb{R}$ be two functions of one variable, differentiable at a point p, then the derivative of their product can be computed by the Leibniz formula $(fg)' = f'g + fg'$ (both parts should be evaluated at the point $x = p$).

Corollary 8. *If $f\colon (\mathbb{R}^1, p) \to (\mathbb{R}^1, q)$ and $g\colon (\mathbb{R}^1, q) \to (\mathbb{R}^1, r)$ are two differentiable functions with $q = f(p)$, then their composition is differentiable, and*

$$(g \circ f)'(p) = g'(q) \cdot f'(p). \tag{16}$$

Indeed, for functions of the real line into itself, the linearizations are linear (affine) maps whose differentials have the form $v \mapsto av$, where the multiplicator $a \in \mathbb{R}$ is the corresponding derivative. If a, b are the derivatives of f at p and g at q respectively, then the composition of the two linear maps takes the form $v \mapsto av \mapsto bav$ and has the multiplicator $ba = ab$.

Example 14. The derivative of the translated function $T_q \circ f$ at a point $p \in \mathbb{R}$ is equal to the derivative of f at the point p times the derivative of the translation, which is 1 at all points. Indeed, $T_q \circ f(x) = f(x) + q$, and the derivation eliminates the constant term.

On the other hand, the composition $f \circ T_p$ in the inverse order after the derivation yields 1 (the derivative of the translation) times the derivative of f at the point $p + q = T_q(p)$:

$$\frac{\mathrm{d}}{\mathrm{d}x}\bigg|_{x=p} f(x+q) = f'(x+q)|_{x=p} \cdot (x+q)'|_{x=p} = f'(p+q) \cdot 1 = f'(p+q).$$

Corollary 9. *All rational functions of one variable are differentiable at each point their domain, i.e., wherever the denominator does not vanish.*

One can easily multiply the number of useful examples by considering pre- and postcompositions with affine maps whose differentials are obvious without any computations.

Proof of the Chain Rule, Theorem 6. Consider the particular case when $p = 0$, $q = 0$ (note that these may be different zeros in spaces of different dimensions)! If $F(x) = L \cdot x + \|x\| \cdot E(x)$ and $G(y) = My + \|y\| \cdot E'(y)$, then the composition $H = G \circ F$ can be computed using the linearity:

$$\begin{aligned} H(x) &= M\big(Lx + \|x\|E(x)\big) \\ &\quad + \big\|Lx + \|x\|E(x)\big\|E'(y) \\ &= MLx + \|x\|\,E''(x). \end{aligned}$$

Since $\|E(x)\| \to 0$ and $E(y(x)) \to 0$ as $\|x\| \to 0$, we see that the terms denoted by $E''(x)$ tend to zero as $x \to 0$. □

Problem 19. Write up the complete proof in the general case where $p \neq 0$ and $q = F(p) \neq 0$.

3.1. **Inversion.** Consider the transformation of *inversion*: this is the map $J\colon \mathbb{R}^2 \smallsetminus \{0\} \to \mathbb{R}^2$ defined by the formulas

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} \dfrac{x}{x^2+y^2} \\ \\ \dfrac{y}{x^2+y^2}. \end{pmatrix} \tag{17}$$

This map possesses very interesting properties:

(1) The image of any ray $\{(ta, tb) : t \in \mathbb{R}^1+\}$ belongs to this ray;
(2) The image of any circle centered at the origin $\{x^2 + y^2 = r^2\}$ is another circle, also centered around the origin;
(3) This map is "symmetry-like": $J \circ J$ is the identical map;
(4) J leaves all points on the unit circle $\{x^2 + y^2 = 1\}$ on their places;
(5) Preimage by J of any line not passing through the origin, is a circle passing through the origin;
(6) Image by J of any line not passing through the origin, is a circle;
(7) The (pre)image by J of any circle not passing through the origin, is a circle;
(8) The (pre)image by J of any circle passing through the origin, is a line not passing through the origin.

In addition, this map keeps angles between vectors: if $(p, q) \neq (0, 0)$ is a point on the plane, then the corresponding differential $\mathrm{d}J(p, q)$ is the linear map which has the form

$$\begin{pmatrix} v \\ w \end{pmatrix} \mapsto \begin{pmatrix} av + bw \\ -bv + aw \end{pmatrix}$$

with the numbers a, b depending on (p, q). Such linear map is the composition of the rotation by some angle and a homothety (stretching) by $R = \sqrt{a^2 + b^2}$.

Problem 20. Prove all the statements above and compute the numbers a, b, R at the point (p, q).