def.: Let $z \in \mathbb{C}$. Define real powers of z by

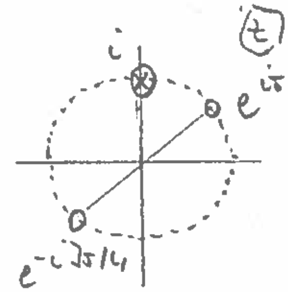
$$\boxed{z^x := r^x e^{i\varphi x}} \qquad x \in \mathbb{R}$$

remark: (6) This is consistent with §2.2 corollary (b).

(7) z^x is not unique for $x \notin \mathbb{N}$. In particular, $x = 1/n$ with $n \in \mathbb{N}$ yields n different values called roots.

example: $z = i = 1 \times e^{i\pi/2} \equiv e^{i(\frac{\pi}{2}+2\pi)}$

$i^{1/2} = e^{i\pi/4}$ or $e^{i(\frac{\pi}{4}+\pi)} = e^{i5\pi/4} \equiv e^{-i3\pi/4}$



§4 Vector spaces and linear spaces

4.1 Vector spaces

def. 1: Let $(V,+)$ be an addition group with neutral element $\mathcal{O}$, and let k be a field. Let there be an exterior multiplication $\varphi: k \times V \to V$ that is

(i) bilinear

(ii) associative in the sense $(\lambda\mu)x = \lambda(\mu x)$ $\forall x \in V$, $\lambda,\mu \in k$

(iii) obeys $1_k x = x$ $\forall x \in V$, where 1_k is the multiplicative neutral element of k.

Then we call V a vector space or linear space over k, or a k-vector space.

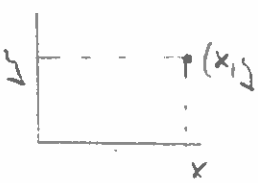
remark: (0) For simplicity we will assume that k is commutative, so $\lambda\mu = \mu\lambda$ $\forall \lambda,\mu \in k$

(1) The elements of V are called vectors; the elements of k are often referred to as scalars.

<u>exmple</u>: (1) $V = k = \mathbb{R}$ with the usual addition of numbers as the interior operation on V, and the usual multiplication as the exterior one. Then $\mathbb{R}$ is in particular a $\mathbb{R}$-vector space (in addition to being a group and a field).

(2) $V = \mathbb{C}$, $k = \mathbb{R}$, with addition on V as defined in §3.3 and $\lambda z = \lambda z_1 + i\lambda z_2$ the multiplication. Then $\mathbb{C}$ is an $\mathbb{R}$-vector space.

(3) Ditto for $V = \mathbb{R}\times\mathbb{R}$, which is isomorphic to $\mathbb{C}$. This is the intuitive "blackboard space". The "vectors" are points in the plane.



Problem 13
Euclidian space
10/10/16

(4) $\mathbb{R}^n$ is an $\mathbb{R}$-vector space. (See §4.8.3 for a generalization known as <u>n-dimensional Euclidian space</u>).

(5) k^n with k an arbitrary field is a k-vector space if we define
$(\lambda_1,\dots,\lambda_n)+(l_1,\dots,l_n) =: (\lambda_1+l_1,\dots,\lambda_n+l_n)$ $(\lambda_i, l_i \in k)$
and $\lambda(l_1,\dots,l_n) := (\lambda l_1,\dots,\lambda l_n)$ $(\lambda, l_i \in k)$

4.2 Basis sets

Let V be a k-vector space.

<u>def 1</u>: If there are finitely many vectors $p_1,\dots,p_n \in V$ such that any $x \in V$ can be expressed as a linear combination of the p_i; i.e., if for any $x \in V$ $\exists\ \lambda_1,\dots,\lambda_n \in k$: $x = \sum_{i=1}^{n} \lambda_i p_i$

then we say that the vectors $p_1,\dots,p_n$ <u>span the space</u>, and we call V <u>finite-dimensional</u>.

<u>exmple</u>: (1) Consider $\mathbb{R}_2$ as an $\mathbb{R}$-vector space. Then

$$p_1 = (1,0),\quad p_2 = (0,2) \quad \text{span the space.}$$

So do $\widetilde{p_1} = (1,0)$, $\widetilde{p_2} = (0,5)$, $\widetilde{p_3} = (1,1)$.

<u>def 2</u>: If any of the p_i $(i=1,\dots,n)$ can be expressed as linear combinations of the remaining $n-1$ vectors p_i, then we call the set of p_i linearly

example: (2) The second set in ex. (1) is linearly dependent, since

$$\tilde p_3 = (1,1) = \tilde p_1 + \tilde p_2$$

def. 3: A minimal set of vectors p_i that span V, none of which can be expressed in terms of the others, is called linearly independent, and are said to form a basis of V. They are also called a set of basis vectors. We denote basis sets by

$$e_1, \dots, e_n$$

and say that V is n-dimensional

example: (3) In example (1), $e_1 = p_1$, $e_2 = p_2$ is a basis. So is $e_1 = \tilde p_1$, $e_2 = \tilde p_2$, or $e_1 = \tilde p_1$, $e_2 = \tilde p_3$.

proposition 1: If n vectors $p_1, \dots, p_n$ are linearly independent, then

$$\sum_{i=1}^{n} \lambda_i p_i = 0 \quad \text{implies} \quad \lambda_i = 0 \ \forall i$$

proof: Suppose $\lambda_{i_0} \neq 0$. $\rightarrow$ $p_{i_0} = \frac{-1}{\lambda_{i_0}} \sum_{i \neq i_0} \lambda_i p_i$

$\rightarrow$ The p_i are not linearly independent $\square$

proposition 2. If $e_1, \dots, e_n$ are a basis of V, then any vector $x \in V$ can be written

$$x = \sum_{i=1}^{n} \lambda_i e_i \qquad (*)$$

with $\lambda_1, \dots, \lambda_n \in$ a unique set of scalars.

remark. (1) We refer to $(*)$ as "expanding x in the basis $\{e_i\}$", and we say that the set of scalars $\{\lambda_i, i = 1, \dots, n\}$ is a representation of the vector x.

proof. The $\{e_i\}$ span V $\Rightarrow$ x can be written in the form $(*)$.

Now suppose x can also be written as

$$x = \sum_{i=1}^{n} \alpha_i e_i \,, \quad \alpha_i \in k$$

$$\Rightarrow \quad 0 = x - x = \sum_{i=1}^{n} (\alpha_i - \lambda_i) e_i$$

But the set $\{e_i\}$ is linearly independent $\Rightarrow$ $\alpha_i - \lambda_i = 0$

by prop. 1 $\Rightarrow$ $\underline{\alpha_i = \lambda_i \ \forall i}$ $\qquad \square$

remark: (2) One often denotes the $\lambda_i \in k$ by x^i and calls them the components or coordinates of the vector x in the basis $\{e_i\}$:

$$(*) \qquad x = \sum_{i=1}^{n} x^i e_i \equiv x^i e_i \qquad (\text{"summation convention"})$$

example: (4) The space $\mathbb{R}^n$ from §4.1 example (4) has a basis

$$\{ e_1 = (1,0,\dots,0),\ e_2 = (0,1,0,\dots,0),\ \dots,\ e_n = (0,\dots,0,1) \}$$

called the cartesian basis.

remark: (3) Prop. 2 provides a one-to-one correspondence between x and its components $\Rightarrow$ All n-dimensional $\mathbb{R}$-vector spaces are isomorphic to $\mathbb{R}^n$.

4.3 Tensor spaces

def. 1: (a) A mapping $f: V \to k$ is called a linear form if

(i) $f(x+y) = f(x) + f(y) \ \forall x,y \in V$ $\quad$ (ii) $f(\lambda x) = \lambda f(x) \ \forall x \in V, \lambda \in k$

(b) A mapping $f: V \times V \to k$ is called a bilinear form if

(i) $f(x+y,z) = f(x,z) + f(y,z)$ $\quad$ (ii) $f(x,y+z) = f(x,y) + f(x,z)$

(iii) $f(\lambda x, y) = \lambda f(x,y) = f(x, \lambda y)$ $\quad \forall x,y,z \in V,\ \lambda \in k$

def. 2: The scalars $t_{ij} := f(e_i, e_j)$ are called the coordinates (or components) of the bilinear form f in the basis $\{e_i\}$.

proposition: The coordinates t_{ij} completely determine the form f.

proof: Let $x, y \in V$ any pair of vectors. Then

$$f(x,y) = f(x^i e_i, y^j e_j) = x^i y^j f(e_i, e_j) = t_{ij} x^i y^j$$

$\rightarrow$ If we know the t_{ij}, then we know $f(x,y)$ for any $x, y \in V$.

def. 3: The n^2 scalars $t_{ij} \in K$ are called the coordinates of the rank-2 tensor t (which is equivalent to the bilinear form f).

theorem: The set of rank-2 tensors forms a vector space of dimension n^2 over K.

Problem 14

Prove this

proof: Assignment #4 Problem 13.

remark: (2) Analogously, one can consider multilinear forms $f(x_1, x_2, x_3)$, $f(x_1, x_2, x_3, x_4)$ etc. to construct tensors of rank 3, rank 4, etc., with coordinates t_{ijk}, t_{ijkl}, etc.

week 3

3(9,10,11,12)

example: (1) Consider R_3 with its cartesian basis $\{e_1, e_2, e_3\}$. The rank-3 tensor defined by

Problem 15

Cross product

$$\varepsilon: R_3 \times R_3 \times R_3 \to R, \quad \varepsilon(e_i, e_j, e_k) = \varepsilon_{ijk} =: \begin{cases} +1 & \\ -1 & \\ 0 & \end{cases}$$

with ...

$$\varepsilon_{ijk} := \begin{cases} +1 & \text{if } \begin{pmatrix} i & j & k \\ 1 & 2 & 3 \end{pmatrix} \text{ is even} \\ -1 & \text{if } \begin{pmatrix} i & j & k \\ 1 & 2 & 3 \end{pmatrix} \text{ is odd} \\ 0 & \text{if } (ijk) \text{ is not a permutation of } (123) \end{cases}$$

is called Levi-Civita tensor or completely antisymmetric tensor of rank 3.

10/11/17

remark: (3) The such defined tensor ε has the components $\varepsilon(e_i, e_j, e_k) = \varepsilon_{ijk}$ only with respect to the right-handed cartesian basis $e_1 = (1,0,0)$, $e_2 = (0,1,0)$, $e_3 = (0,0,1)$. If we evaluate ε on a different basis $\{\hat{e}_i\}$, the result will be different, $\varepsilon(\hat{e}_i, \hat{e}_j, \hat{e}_k) \neq \varepsilon(e_i, e_j, e_k)$. See §5 for an elaboration on this.

(4) By contrast, one often uses the Levi-Civita symbol $\varepsilon_{ijk} = \operatorname{sgn} \sigma \begin{pmatrix} i & j & k \\ 1 & 2 & 3 \end{pmatrix}$ independent of the basis chosen. The ε_{ijk} then are not the coordinates of the tensor ε: $\varepsilon(e_i, e_j, e_k) = \varepsilon_{ijk}$ only for a special basis, see remark (3) above.

example: (2) Consider R_n with its cartesian basis $\{e_i;\ i = 1, \dots, n\}$. The rank-2 tensor defined by

Problem 16

symmetric tensors

$$\delta : R_n \times R_n \to R, \quad \delta(e_i, e_j) \equiv \delta_{ij} := \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

is called (Euclidean) Kronecker delta.

10/12/16 remark: (5) We will come back to δ_{ij} in §§4.4 and 4.8.

4.4 Dual spaces

Let V be an n-dimensional k-vector space with basis $\{e_1, \dots, e_n\}$ and let f be a linear form on V. Expand $x \in V$ in the basis: $x = x^i e_i$ and consider

$$f(x) = f(x^1 e_1 + \dots + x^n e_n) = f(e_1)x^1 + \dots + f(e_n)x^n =: u_1 x^1 + \dots + u_n x^n \equiv u_i x^i$$

where $u_i := f(e_i) \in K$ $\qquad (*)$

remark: (1) Every linear form on V can be written in the form $(*)$, i.e., the scalars u_i uniquely determine the form.

(2) The $(u_1, \dots, u_n)$, and hence the linear forms on V, form a vector space V^* that is isomorphic to K^n and hence to V.

def. 1: (a) The space V^* of linear forms on V is called *dual to V*.

(b) The elements of V^* are called *covectors*. They are one-to-one correspondent to the vectors that are the elements of V.

remark: (3) Covectors are defined via linear forms, and rank-n tensors are defined via n-linear forms (§4.3).
$\to$ Covectors can be considered tensors of rank 1.

def. 2: The scalar $f(x) \in K$ is called the *scalar product* between the covector u that corresponds to f and the vector x.
We write $\qquad u \cdot x := u_i x^i$

remark: (3') Covectors are also called *covariant vectors*, in which case vectors are referred to as *contravariant vectors*.

Remark. (4) Since V^* is isomorphic to V, we do not have to distinguish between the two spaces and can consider

$$y_i := u_i$$

the covariant components of the vector y that corresponds to the covector u under the isomorphism between V and V^*. The contravariant components of the same vector are the y^i. Thus

$$u \cdot x \equiv y \cdot x = y_i x^i \quad \text{(see also §4.8)}$$

(5) In Physics, one often denotes vectors by $|x\rangle$ and covectors by $\langle y|$ and writes the scalar product

$$\langle y|x\rangle := y_i x^i \quad \text{(see also §4.7)}$$

(6) The covectors $e^1 = (1,0,\dots,0)$, $e^2 = (0,1,0,\dots,0)$, ..., $e^n = (0,\dots,0,1)$ form a cartesian basis of V^* that corresponds to the cartesian basis of V and is called cobasis. (Here 1 and 0 are the multiplicative and additive neutral elements, respectively, in K).

(7) The scalar product $e^i \cdot e_j =: \delta^i_j = \boxed{\begin{cases} 1 & \text{if } i=j \\ 0 & \text{otherwise} \end{cases}}$ (8) is identical to the Euclidian Kronecker delta δ_{ij} defined in §4.3 example 2.

(8) δ^i_j is always defined by (8). However, $\delta^i_j = \delta_{ij}$ only if $\mathbb{R}_n$ is considered a Euclidian space. More generally, $\delta^i_j \neq \delta_{ij}$, see §4.8.

def. 3: (c) Bilinear forms $f: V^* \times V^* \to K$ acting on the co-basis define <u>contravariant tensors</u> of rank 2:

$$f(e^i, e^j) = t^{ij}$$

and analogously for higher-rank tensors.

(b) Multilinear forms acting on mixed sets of basis and co-basis vectors define <u>mixed tensors</u>. E.g., $f: V^* \times V \times V^* \to K$ defines

$$f(e^i, e_j, e^l) = t^i{}_j{}^l$$

remark: (9) Vectors can be considered contravariant tensors of rank 1.

example: (1) The object $\delta^i{}_j$ from remark (7) is a mixed tensor of rank 2.

remark: (10) $\delta^i{}_j$ takes the co-basis vector e^i and computes its j^{th} coordinate with respect to the basis: $(e^i)_j = \delta^i{}_j$. Similarly, $(e_i)^j = \delta_i{}^j$. $\delta^i{}_j = \delta_j{}^i$, which why some authors (e.g. LL) write δ^i_j.

def. 4: A contravariant tensor whose components are given by the product of the components of two contravariant vectors x and y is called the <u>tensor product</u> of x and y and denoted by

$$t = x \otimes y, \quad t^{ij} = x^i y^j$$

Analogously, $t_{ij} = x_i y_j$, $t_i{}^j = x_i y^j$, $t^i{}_j = x^i y_j$

remark: (11) We do <u>not</u> yet know how the basis vectors are related to the co-basis vectors, or covariant vectors to contravariant ones. In <u>Euclidean space</u>, $e^i = e_i$ and $x^i = x_i$, and for higher-rank tensors we don't have to distinguish between co- and contravariant indices (e.g., $\delta^i{}_j = \delta_{ij}$, see remark (7)). However, this property represents an additional postulate that [illegible]

4.5 Metric spaces

def. 1. Let M be a set, and $\rho : M \times M \to \mathbb{R}$ a mapping with properties

(i) $\rho(x,y) \geq 0 \ \forall x,y \in M$ and $\rho(x,y) = 0$ iff $x = y$ (positive definiteness)

(ii) $\rho(x,y) = \rho(y,x) \ \forall x,y \in M$ (symmetry)

(iii) $\rho(x,z) \leq \rho(x,y) + \rho(y,z)$ (triangle inequality)

Then we call M a metric space with metric ρ.

remark: (1) M can have additional properties (e.g., it can be a group, or a field, or a vector space), but this is no necessary.

example: (1) $M = \mathbb{R}$ with $\rho(x,y) = |x-y| := \begin{cases} x-y & \text{if } x-y \geq 0 \\ -(x-y) & \text{if } x-y < 0 \end{cases}$

is a metric space.

Problem 17
R as a metric space

proof: Problem 17.

def. 2. Consider an infinite sequence $x_n \in M$ $(n = 1, 2, \ldots)$ of elements. We say $x^* \in M$ is the limit of the sequence, and write $\lim_{n\to\infty} x_n = x^*$, or $x_n \Rightarrow x^*$, or $\lim_{n\to\infty} \rho(x_n, x^*) = 0$, if

For any $\varepsilon > 0 \ \exists N \in \mathbb{N} : \ \rho(x_n, x^*) < \varepsilon \ \forall n > N$.

We also say 'the sequence x_n converges to x^*'.

proposition 1: A sequence can have at most one limit.

Problem 18
limits 10/17/16

proof: Problem 18

def. 3: Let x_n be a sequence. If for any $\varepsilon > 0 \ \exists N \in \mathbb{N}$: $\rho(x_n, x_m) < \varepsilon$ $\forall n,m > N$, then we call the sequence a Cauchy sequence.

remark: (2) less formally, one writes $\rho(x_n, x_m) \to 0$ for $n, m \to \infty$

proposition 2: Every sequence with a limit is a Cauchy sequence.

proof: Problem 18

remark: (3) The converse is not true!

example: (2) Let $M = \mathbb{Q}$ and $x_n = (1+1/n)^n$. Then $\lim_{n\to\infty} x_n = e \notin \mathbb{Q}$

(3) Let $M = \mathbb{R}$ (with ρ from example (1)). Then every Cauchy sequence has a limit.

sketch of proof: (i) Every Cauchy sequence is bounded.

(ii) A Cauchy sequence converges iff it has a convergent subsequence.

(iii) In $\mathbb{R}$, every bounded sequence has a convergent subsequence (Bolzano-Weierstrass).

def 4: A metric space in which every Cauchy sequence has a limit is called complete.

remark: (4) The mechanism demonstrated in example (2) is the only one by which a Cauchy sequence can avoid having a limit.

proposition 3: A metric space that is not complete can always be made complete by adding a suitable set of elements. The completion is unique up to isomorphism.

proof: difficult.

4.6 Banach spaces

def. 1: Let $\mathcal{B}$ be a k-vector space (with null vector 0 and $k = \mathbb{R}$ or $\mathbb{C}$), and let $\|\ldots\| : \mathcal{B} \to \mathbb{R}$ be a mapping with the properties (NB: $\mathbb{R}$, not k!)

(i) $\|x\| \geq 0 \;\; \forall\, x \in \mathcal{B}$, and $\|x\| = 0$ iff $x = 0$.
(positive semi-definiteness)

(ii) $\|x+y\| \leq \|x\| + \|y\| \;\; \forall\, x,y \in \mathcal{B}$ (triangle inequality)

(iii) $\|a x\| = |a| \cdot \|x\| \;\; \forall\, x \in \mathcal{B},\, a \in k$ (linearity)

Then we say that $\|\ldots\|$ is a norm on $\mathcal{B}$, and $\|x\|$ is the norm of $x \in \mathcal{B}$.

def. 2: Define $d : \mathcal{B} \times \mathcal{B} \to \mathbb{R}$ by $d(x,y) := \|x-y\| \;\; \forall\, x,y \in \mathcal{B}$.
Then we call $d(x,y)$ the distance between x and y.

remark: (1) d is a metric in the sense of §4.5.

(2) $\|x\| = d(x,0)$.

(3) Every linear space with a norm is in particular a metric space.

def. 3: A linear space with a norm that is complete is called a Banach space or $\mathcal{B}$-space.

example: (1) $\mathbb{R}$ as a vector space with $\|x\| := |x|$ is a $\mathcal{B}$-space.

(2) $\mathbb{C}$ as a vector space with $\|z\| = |z| := \sqrt{z_1^2 + z_2^2}$ $(z = z_1 + i z_2)$ is a $\mathcal{B}$-space.

def. 4 : Let B be a B-space over $\mathbb{C}$, and let $\ell : B \to \mathbb{C}$ be a linear form in the sense of §4.3. Then the norm of ℓ is defined as

$$\|\ell\| := \sup_{\|x\|=1} \{|\ell(x)|\} .$$

remark: (4) The space B^* of linear forms ℓ is the space dual to B in the sense of §4.4.

proposition 1 : On B^*, the norm of linear forms (def. 4) defines a norm in the sense of def. 1.

proof : Problem 19

Problem 19

Proof of prop

theorem : B^* is complete and therefore a B-space.

proof : difficult (see books)

4.7 Hilbert spaces

def. 1 : Let H be a linear space over $\mathbb{C}$. Let $(,) : H \times H \to \mathbb{C}$ be a mapping such that

(i) $(x,y) = (y,x)^* \quad \forall\, x,y \in H$

(ii) $(x+y, z) = (x,z) + (y,z) \quad \forall\, x,y,z \in H$

(iii) $(x,x) \geq 0$, and $(x,x) = 0$ iff $x = \emptyset$ with $\emptyset$ the null vector

(iv) $(\lambda x, y) = \lambda^* (x,y) \quad \forall\, x,y \in H,\ \lambda \in \mathbb{C}$

Then we define the norm of $x \in H$ by $\|x\| := (x,x)^{1/2}$

Week 4

4/13.14.15.16.

remark: (1) Note the subtle difference between the definition of $(,)$ and the definition of a bilinear form in §4.3 def. 1 (iii) lemma: Cauchy-Schwartz inequality

$$\boxed{|(x,y)|^2 \leq (x,x)\cdot(y,y)} \qquad (*)$$

remark: (2) In fairness to Victor Bunyakovsky (1804-1889) this should be called the BCS inequality, but in the Western literature it's almost always just CS, and in the Russian literature it's just B.

proof: $(*)$ obviously holds as an equality if $y=0$.

Now let $y \neq 0 \rightsquigarrow (y,y) > 0$

define $t := x - \frac{(x,y)}{(y,y)}y \rightsquigarrow (t,y) = (x,y) - \frac{(x,y)}{(y,y)}(y,y) = 0$

$\rightsquigarrow x = t + \frac{(x,y)}{(y,y)}y$

$$\rightsquigarrow (x,x) = (t,t) + \frac{(x,y)}{(y,y)}\underbrace{(t,y)}_{=0} + \frac{(y,x)}{(y,y)}\underbrace{(y,t)}_{=0} + \frac{(x,y)(y,x)}{(y,y)^2}(y,y)$$

$$= \underbrace{(t,t)}_{\geq 0} + \frac{(x,y)(y,x)}{(y,y)} \geq \frac{(x,y)(y,x)}{(y,y)}$$

$$\rightsquigarrow \underline{\underline{(x,x)(y,y)}} \geq (x,y)(y,x) = \underline{\underline{|(x,y)|^2}} \qquad \square$$

remark: (3) An analogous inequality holds for any scalar product in any linear space; $(\ ,\ )$ in $\mathcal{H}$ is just a particular example.

proposition 1 : The so defined norm is a norm in the sense of §4.6 def. 1.

proof : Problem 20

Problem 20
Proof of prop. 1

def. 2 : On H, a metric is defined by $\rho(x,y) := \|x-y\|$, and Cauchy sequences are defined by §4.5 def. 3.

proposition 2 : The so defined metric is a metric in the sense of §4.5

def. 3 : If H is complete, then it is called a Hilbert space or H-space.

remark : (4) Every H-space is in particular a B-space.

10/19/16

def. 4 : For every fixed $y \in H$, we define a linear form l by

$$l(x) := (y,x) \quad \forall x \in H$$

proposition 3 : The so defined l are linear forms in the sense of §4.3 def. 1 (a).

proof : Problem 20

proposition 4 . Every linear form on H can be uniquely written in this form ; i.e., for every l $\exists!\, y \in H : l(x) = (y,x)$

proof : difficult (books) .

corollary . The space of linear forms is a dual space H^* in the sense of §4.4 that is isomorphic to H. In particular, H^* is a H-space.

def. 5: A mapping $\langle\ |\ \rangle : H^* \times H \to \mathbb{C}$ is defined by $\langle l|x\rangle := l(x)$

remark: (5) For every $l \in H^*$ $\exists\, y \in H : l(x) = (y,x) \rightsquigarrow \langle l|x\rangle = (y,x)$

(6) Since H^* is isomorphic to H, one often does not distinguish between l and y and writes, sloppily, $\langle y|x\rangle := \langle l|x\rangle = (y,x)$.

(7) In quantum mechanics, the state of a system is represented by a vector in a Hilbert space.

4.8 Generalized metrics; Minkowski space

4.8.1 Scalar product

def. 1: Let V be an n-dim. vector space over $\mathbb{R}$, and let $g : V\times V \to \mathbb{R}$ be a bilinear form that is symmetric, $g(x,y) = g(y,x)$, and then define a symmetric rank-2 tensor $g_{ij} = g(e_i, e_j) = g_{ji}$ (see Problem 16). Let g have an inverse g^{-1}, corresponding to a tensor g^{ij}, in the sense

$$g_{ij}\, g^{jl} = \delta_i^{\ l}$$

Then we call the real number

$$\boxed{g(x,y) \equiv x\cdot y \equiv xy = x^i g_{ij} y^j}$$

the (generalized) scalar product of x and y, and g the (generalized) metric or the metric tensor.

remark: (1) This is not a metric in the sense of §4.5. For instance, there is no guarantee that $g(x,y) \geq 0$, or even that $g(x,x) \geq 0$

(2) §4.2 remark (3) $\rightsquigarrow$ V is isomorphic to $\mathbb{R}_n$ $\rightsquigarrow$ we will discuss $\mathbb{R}_n$ endowed with the metric g. Let $\{e_1, \dots, e_n\}$ be a basis

def. 2: Define an adjoint basis or co-basis $\{e^1, \dots, e^n\}$ by

$$\boxed{e^i = g^{ij} e_j}$$

remark: (3) Note that $e^i \in V$, whereas the co-basis vectors (also called e^i) in §4.4 were elements of V^*. However, since $V^* \cong V$ we might as well develop the co-structure on V.

(4) The relation between e^i and e_j can be inverted:

$$\underline{e_i} = \delta_i^{\ j} e_j = g_{ik} g^{kj} e_j = \underline{g_{ik} e^k}$$

def. 3: The coordinates x^i of x in the basis, $x = x^i e_i$, are called contravariant; the coordinates x_i of x in the co-basis, $x = x_i e^i$, are called covariant.

remark: (5) All of this is consistent with §4.4. However, now we have specified the relation between the basis and the co-basis, which we had left unspecified in §4.4, see §4.4 remark (11).

proposition 1: The contravariant and covariant coordinates are related:

$$\boxed{x_i = g_{ij} x^j, \quad x^i = g^{ij} x_j}$$

proof: $x = x^i e_i = x^i g_{ij} e^j = x_j e^j \Rightarrow \underline{x_j} = x^i g_{ij} = \underline{g_{ji} x^i}$

$\underline{x^i} = \delta^i_{\ j} x^j = g^{ik} g_{kj} x^j = \underline{g^{ik} x_k}$ □

<u>remark</u>: (0) The δ^i_j are <u>not</u> coordinates of a tensor, and whether an index is up or down has no significance except in the context of the summation convention. Hence, $\delta^i_j = \delta_{ij} = \delta_i{}^j$. It is crucial, however, whether an index is left (row) or right (column).

<u>corollary</u>: The scalar product can be written

$$\boxed{g(x,y) \equiv x \cdot y = x^i y_i = x_j y^j}$$

<u>remark</u>: (6) This is consistent with §4.4 remark (4).

(7) In particular, $g(e^i, e_j) = g^i{}_j = e^i \cdot e_j \equiv \delta^i{}_j$ from §4.4 remark (7)

<u>This</u> is <u>always</u> true, irrespective of what g is.

However, $\delta_{ij} = g_{ik}\delta^k{}_j = g_{ij}$, which in general is <u>not</u> equal to $\delta^i{}_j$.

People often write (sloppily) $\delta_{ij} := \delta^i{}_j$ irrespective of the metric g_{ij}. δ_{ij} then is <u>not</u> a tensor in the space characterized by g, but strictly a symbol that means $\delta_{ij} = \begin{cases} 1 & \text{if } i=j \\ 0 & \text{otherwise} \end{cases}$. See also the difference between the Levi-Civita tensor and the Levi-Civita symbol in §4.3. In Euclidean space these confusing issues do not exist, see §4.8.3 example (1).

4.8.2 Basis transformations

<u>def. 1</u>: (a) An $n \times n$ array of real numbers $D^{\overset{\text{row}}{i}}{}_{\overset{\text{column}}{j}}$ that are arranged in a square pattern of rows and columns we call a (real) <u>$n\times n$ matrix D</u>.

$$\begin{pmatrix} D^1{}_1 & D^1{}_2 & \dots & D^1{}_n \\ D^2{}_1 & D^2{}_2 & \dots & D^2{}_n \\ \vdots & & & \vdots \\ D^n{}_1 & D^n{}_2 & \dots & D^n{}_n \end{pmatrix}$$

The $D^i{}_j$ are called <u>matrix elements</u>.

→

p 38 f.p.

(e) The *determinant* of a $n\times n$ matrix D is defined as

$$\underline{\det D} \equiv \begin{vmatrix} D_{11} & D_{12} & \cdots & D_{1n} \\ D_{21} & D_{22} & \cdots & D_{2n} \\ \vdots & & & \vdots \\ D_{n1} & D_{n2} & \cdots & D_{nn} \end{vmatrix} := \underline{\sum_{\sigma} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} D_{i,\sigma(i)}} \longrightarrow$$

example 1: $\underline{n=2}$ $\qquad D = \begin{pmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{pmatrix}$

$$\underline{\det D} = \begin{vmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{vmatrix} = \operatorname{sgn}\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} D_{11} D_{22} + \operatorname{sgn}\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} D_{12} D_{21} = \underline{D_{11}D_{22} - D_{12}D_{21}}$$

(b) We call D <u>invertible</u> if a matrix D^{-1} exists such that

$$D^i{}_j\,(D^{-1})^j{}_k = (D^{-1})^i{}_j\,D^j{}_k = \delta^i{}_k$$

or $\quad D D^{-1} = \mathbb{1}_n \quad$ (see part (d) below)

with $\mathbb{1}_n$ the $n\times n$ <u>unit matrix</u> whose elements are $(\mathbb{1}_n)^i{}_j = \delta^i_j$

row ← column

(c) The matrix D^T with elements $D_j{}^i$ is called <u>transposed of D</u> and denoted by D^T, i.e.,

$$(D^T)^i{}_j = D_j{}^i$$

$$D^T = \begin{pmatrix} D^1{}_1 & D^2{}_1 & \dots & D^n{}_1 \\ D^1{}_2 & D^2{}_2 & \dots & D^n{}_2 \\ \vdots & \vdots & & \vdots \\ D^1{}_n & D^2{}_n & \dots & D^n{}_n \end{pmatrix}$$

(d) The <u>product</u> of two matrices A, B is defined by

$$(AB)^i{}_j := A^i{}_k\,B^k{}_j$$

10/24/16

<u>proposition 1</u>: (i) $(AB)^T = B^T A^T$

<u>proof</u>: $((AB)^T)^i{}_j = (AB)_j{}^i = A_j{}^k B_k{}^i = (B^T)^i{}_k (A^T)^k{}_j = (B^T A^T)^i{}_j \quad \square$

(ii) $(D^{-1})^T = (D^T)^{-1}$ <u>proof</u>: $D^T (D^{-1})^T \overset{(i)}{=} D^{-1} D = \mathbb{1}_n \quad \square$

(iii) $\det(AB) = \det A \cdot \det B$, $\det D^{-1} = 1/\det D$, $\det D^T = \det D$

<u>proof</u>: books (e.g., Smirnov vol. III/1)

<u>def. 2</u>: Consider $\mathbb{R}_n$ endowed with a metric g, let $\{e_i;\ i=1,\dots,n\}$ be a basis, and let D be an invertible matrix. Then we define a new basis $\{\hat{e}_i\}$ through the <u>basis transformation</u>

$$\boxed{\hat{e}_i := e_j\,(D^{-1})^j{}_i}$$

<u>remark</u>: (1) The inverse basis transformation is given by D, as expected

$$\underline{\hat{e}_i D^i{}_j} = e_l (D^{-1})^l{}_i D^i{}_j = e_l \delta^l{}_j = \underline{e_j}$$

<u>proposition 2</u>: $\{\hat{e}_i\}$ is indeed a basis in the sense of §4.2.

<u>proof</u>: let $x = x^i e_i$ be an arbitrary vector.

remark (1) $\rightsquigarrow$ $x = x^i \hat{e}_j D^j{}_i = D^j{}_i x^i \hat{e}_j$

$\rightsquigarrow$ <u>$\{\hat{e}_i\}$ spans the space</u>

Now let $\hat{\lambda}^i \in \mathbb{R}$ $(i = 1, \dots, n)$ and consider

$$\hat{\lambda}^i \hat{e}_i = \hat{\lambda}^i e_j (D^{-1})^j{}_i =: \lambda^j e_j \quad (*)$$

with $\lambda^j = \hat{\lambda}^i (D^{-1})^j{}_i$ $\rightsquigarrow$ $\hat{\lambda}^i = D^i{}_j \lambda^j$ $(**)$

$\{e_i\}$ is linearly independent $\rightsquigarrow$ $\lambda^i e_i = 0$ implies $\lambda^i = 0$ $\forall i$.

But $(*)$ $\rightsquigarrow$ $\hat{\lambda}^i \hat{e}_i = 0$ implies $\lambda^i e_i = 0$, which implies $\lambda^i = 0$,

which by $(**)$ implies $\hat{\lambda}^i = 0$ $\forall i$

$\rightsquigarrow$ <u>$\{\hat{e}_i\}$ is linearly independent</u> $\square$

<u>proposition 3</u>: let $x \in R_n$ be a vector whose contravariant coordinates with respect to the basis $\{e_i\}$ are x^i. Then its coordinates with respect to $\{\hat{e}_i\}$ are

$$\boxed{\hat{x}^i = D^i{}_j x^j} \quad \text{or} \quad \boxed{\hat{x} = Dx}$$

<u>proof</u>: $x = x^i e_i = x^i \hat{e}_j D^j{}_i = x^i D^j{}_i \hat{e}_j = \hat{x}^j \hat{e}_j$ with $\underline{\hat{x}^j = D^j{}_i x^i}$

<u>remark</u>: (2) The inverse relation is $x^i = (D^{-1})^i{}_j \hat{x}^j$

(3) D applied to vectors is called a <u>coordinate transformation</u>.

proposition 4: Let $g_{ij} = e_i \cdot e_j$ be the metric in the basis $\{e_i\}$, and let Λ^{-1} be a basis transformation: $\tilde{e}_i = e_j (\Lambda^{-1})^j{}_i$.
Then the metric $\tilde{g}$ in the basis $\{\tilde{e}_i\}$ is given by

$$\tilde{g}_{ij} = \left((\Lambda^{-1})^T\right)_i{}^k g_{kl} (\Lambda^{-1})^l{}_j \quad \text{or} \quad \boxed{\tilde{g} = (\Lambda^{-1})^T g \Lambda^{-1}}$$

and the inverse relation is $\boxed{g = \Lambda^T \tilde{g} \Lambda}$

proof: $\underline{\tilde{g}_{ij}} = \tilde{e}_i \cdot \tilde{e}_j = e_k (\Lambda^{-1})^k{}_i \cdot e_l (\Lambda^{-1})^l{}_j = (\Lambda^{-1})^k{}_i (\Lambda^{-1})^l{}_j \underbrace{e_k \cdot e_l}_{= g_{kl}}$

$$= (\Lambda^{-1})^k{}_i \, g_{kl} (\Lambda^{-1})^l{}_j = \underline{\left((\Lambda^{-1})^T\right)_i{}^k g_{kl} (\Lambda^{-1})^l{}_j}$$

$$\rightsquigarrow \underline{(\Lambda^{-1})^T g \Lambda^{-1} = \tilde{g}} \quad \rightsquigarrow \underline{\Lambda^T \tilde{g} \Lambda} = \Lambda^T (\Lambda^{-1})^T g \Lambda^{-1} \Lambda \overset{\text{prop 1}}{=} \underline{g}$$

4.8.3 Normal coordinate systems

lemma: For any symmetric matrix $M^i{}_j = M_j{}^i$ that has an inverse there exists a transformation Λ such that

$$\boxed{\tilde{M}^i{}_j = (\Lambda^T M \Lambda)^i{}_j = m^i \delta^i_j} \quad (\text{no summation})$$

proof: books

remark: (1) This result of linear Algebra is sometimes called the finite-dimensional spectral theorem

corollary: Let g_{ij} be a metric on $\mathbb{R}^n$. Then there exists a coordinate transformation Λ such that

$$\boxed{\tilde{g}_{ij} = \lambda_i \delta_{ij}}$$ (no summation) with $\lambda_i \neq 0 \ \forall i$ and δ_{ij} the Kronecker

Week 5
25 (13.12.2022)

Theorem: There exists a coordinate transformation Δ such that

$$\boxed{\hat{g} = \begin{pmatrix} 1 & & & & & 0 \\ & \ddots & & & & \\ & & 1 & & & \\ & & & -1 & & \\ & & & & \ddots & \\ 0 & & & & & -1 \end{pmatrix} \begin{matrix} \} \ m \text{ times } +1 \\ \\ \} \ n-m \text{ times } -1 \end{matrix}} \quad (0 \leq m \leq n) \qquad (*)$$

proof: corollary $\Rightarrow$ we can achieve $\hat{g}_{ij} = \lambda_i \delta_{ij}$, $\lambda_i \neq 0$

$\Rightarrow$ we can relabel the basis vectors such that

$$\lambda_1, \ldots, \lambda_m > 0, \quad \lambda_{m+1}, \ldots, \lambda_n < 0$$

Define $(\Delta^{-1})^i{}_j = \delta^i{}_j / |\lambda_i|^{1/2}$ (no summation; NB: $\lambda_i \neq 0$)

$$\Rightarrow \underline{\hat{\hat{g}}_{ij}} = (\Delta^{-1})^T{}_i{}^\lambda \hat{g}_{\lambda\ell} (\Delta^{-1})^\ell{}_j = \frac{\lambda_\ell}{\sqrt{|\lambda_i| \cdot |\lambda_\ell|}} \delta_i{}^\lambda \delta_{\lambda\ell} \delta^\ell{}_j$$

$$= \frac{\lambda_\ell}{\sqrt{|\lambda_i| \cdot |\lambda_\ell|}} \delta_{i\ell} \delta^\ell{}_j = \frac{\lambda_i}{|\lambda_i|} \delta_{ij} = \underline{\delta_{ij} \times \begin{cases} +1 & \text{if } \lambda_i > 0 \\ -1 & \text{if } \lambda_i < 0 \end{cases}}$$

def. 1: Basis sets in which the metric has the form $(*)$ are called <u>normal coordinate systems</u>.

remark: (2) The number m is characteristic of the metric vector space and independent of the basis ('<u>Sylvester's rigidity theorem</u>').

example: (1) $\underline{m = n}$ $\quad g = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$, $g_{ij} = \delta_{ij}$

remark: (3) This is a metric in the sense of § 4.5.
$\mathbb{R}_n$ endowed with this metric is called n-dimensional <u>Euclidean space</u> E

The normal coordinate systems are called <u>cartesian</u>.

$x_i = g_{ij} x^j = \delta_{ij} x^j = x^i$ covariant coordinates = contravariant coordinates

$x \cdot x = (x_1)^2 + \ldots + (x_n)^2$ 'Pythagorean theorem'

(2) $\underline{m=1}$ $(n \geq 2)$ $\quad g = \begin{pmatrix} 1 & & 0 \\ & -1 \ddots & \\ 0 & & -1 \end{pmatrix}$ remark: (4) This is a generalized metric

$\mathbb{R}_n$ endowed with this metric is called Minkowski space M_n.
The normal coordinate systems are called inertial frames.

$x_1 = x^1, \quad x_i = -x^i \quad (i = 2, \dots, n)$

$x \cdot x = x_i x^i = (x_1)^2 - \sum_{i=2}^{n} (x_i)^2$

remark: (5) Special Relativity postulates that classical mechanical systems can be described as point masses moving in a Minkowski space with $n=4$.

(6) In Physics, one often labels $x = (x_0, x_1, x_2, x_3)$ with $x_0 = ct$ ("time"), $(x_1, x_2, x_3) = \vec{x}$ ("space") c is a characteristic velocity ("speed of light in vacuum").

Problem 21
contractions in M_2

(7) Galilean Mechanics postulates that classical mechanical systems can be described as point masses moving in a Euclidean space with $n=4$.
This turned out to be incompatible with both experiment (Michelson-Morley), and with Maxwell's theory of classical E&M.

(8) Why did Nature choose $m=1$? No answer within Physics.

10/26/16

4.8.4 Normal coordinate transformations

def. 1: A coordinate transformation is called normal if it leaves invariant a normal coordinate system. That is, Λ is normal iff

$$\boxed{g = (\Lambda^{-1})^T g \Lambda^{-1}} \quad \text{or} \quad \boxed{g_{ij} = (\Lambda^{-1})^k{}_i\, g_{kl} (\Lambda^{-1})^l{}_j}$$

examples: (1) Euclidean space: $\mathbb{1}_n = (\Lambda^{-1})^T \Lambda^{-1} \Leftrightarrow \Lambda^T = \Lambda^{-1}$
The normal coordinate transformations are called orthogonal.

(2) Minkowski space: The normal coordinate transformations are called Lorentz transformations.

lemma: (i) If Λ is a normal coordinate trafo, then so is Λ^{-1}.

proof. $g = (\Lambda^{-1})^T g \Lambda^{-1} \Rightarrow \Lambda^T g \Lambda = \Lambda^T (\Lambda^{-1})^T g \Lambda^{-1} \Lambda \overset{§4.8.2\text{ prop. 1}}{=} g$ □

(ii) If Λ_1 and Λ_2 are normal, then so is $\Lambda_1 \Lambda_2$.

proof: $g = \Lambda_2^T g \Lambda_2 = \Lambda_2^T \Lambda_1^T g \Lambda_1 \Lambda_2 = (\Lambda_1 \Lambda_2)^T g (\Lambda_1 \Lambda_2)$ □

theorem: The set of normal coordinate transformations forms a (e.g. non-abelian) group under matrix multiplication.

proof: Closure ✓ by lemma (ii)
Neutral element $\mathbb{1}_n$ is a normal trafo
Matrix multiplication is associative
Inverse exists by lemma (i) □

Problem 22
time-like + space-like vectors

Problem 23
special Lorentz trafos

remarks: (1) The group of normal trafos in the Euclidean space is called the orthogonal group $O(n)$.

Problem 24

(2) The group of normal trafos in Minkowski space is called

proposition: Let Δ be a normal coordinate transformation. Then $\boxed{\det \Delta = \pm 1}$

proof: def. 1 $\Rightarrow$ $\det g = \det(\Delta^T g \Delta) \overset{\S 4.8.2\,\text{prop}\,1}{=} \det g \, (\det \Delta)^2$
$\Rightarrow (\det \Delta)^2 = 1 \Rightarrow \det \Delta = \pm 1$ □

§5 Tensor fields

5.1 Tensor fields

Let V be $\mathbb{R}_n$ endowed with a (generalized) metric g as defined in §4.8.1. Let Δ be a normal coordinate transformation from a normal coordinate system cs to a normal coordinate system $\widehat{cs}$:

$$\hat{x}^i = \Delta^i{}_j \, x^j$$

def. 1: For every $x \in V$, consider a rank-N tensor $t^{i_1 \dots i_N}(x)$.
We call $t^{i_1 \dots i_N}(x)$ a <u>tensor field</u> if, under a normal coordinate transformation,

$$\boxed{\hat{t}^{i_1 \dots i_N}(\hat{x}) = \Delta^{i_1}{}_{j_1} \dots \Delta^{i_N}{}_{j_N} \, t^{j_1 \dots j_N}(x)} \qquad (*)$$

remark: (1) This generalizes the tensor concept of §4.3 by assigning a different tensor to every vector in V.
(2) Tensor fields are tensor-valued functions on V.

proposition: For the special case of homogeneous tensor fields, i.e., if $t^{i_1 \dots i_N}(x) \equiv t^{i_1 \dots i_N}$ independent of x, we recover the usual tensor concept of §4.3