

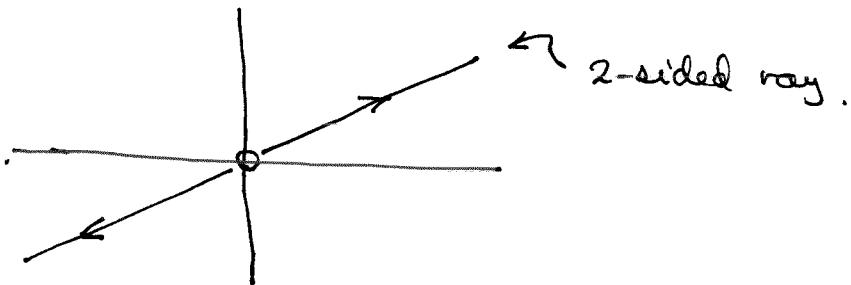
(1)

$\mathbb{R}P^n$. It is defined this way. Define \sim on $\mathbb{R}^{n+1} - \{0\}$ by

$$x, y \in \mathbb{R}^{n+1} - \{0\}.$$

$$x \sim y \text{ if } x = ay, \quad a = \text{real}, \quad a \neq 0.$$

This is not quite the same as what we did above when modding out by the amplitude of \vec{E} , because there a was required to be positive. Thus now the equivalence classes are 2 sided rays in \mathbb{R}^{n+1} , e.g. for $n=1$,

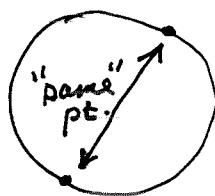


so if we draw a sphere S^n in \mathbb{R}^{n+1} , it intersects each ray in 2 points, antipodal points. Thus,

$$\frac{\mathbb{R}^{n+1} - \{0\}}{\sim} = \mathbb{R}P^n = \frac{S^n}{\sim (\text{antipodal})}.$$

↑
this is defn
of $\mathbb{R}P^n$

For example, with $n=1$, $\mathbb{R}P^1$ is S^1 with antipodal points identified,



We can eliminate most of the double point pairs by cutting the circle in half.



Then along the upper arc, each point corresponds to a single equiv. class (2-sided ray) in \mathbb{R}^2 . At the endpoints, however, there is still one equivalence class represented by 2 points. But we can bring these points together to make them into one point, whereupon we obtain a circle again, but only $\frac{1}{2}$ as big as the first one. So,

$$\mathbb{R}P^1 = S^1$$

Similarly, $\mathbb{R}P^2$ is S^2 with antipodal points identified. By cutting S^2 at the equator and throwing away the southern hemisphere, we get a hemisphere in which



antipodal points on the equator are identified. Sim. for $\mathbb{R}P^3$ etc.

Most of the examples so far involve equivalence relations that are produced by group actions. (More on group actions later.) Here are some equivalence relations that are not produced by group actions.

(7) $X = \text{square with edges identified.}$ (Equivalence classes imply glueing rules.)

$$\begin{matrix} & B \\ & \downarrow \\ A & \xrightarrow{\quad} & B \\ & \downarrow \\ & A \end{matrix} = \frac{X}{\sim} = \text{segment of cylinder.}$$

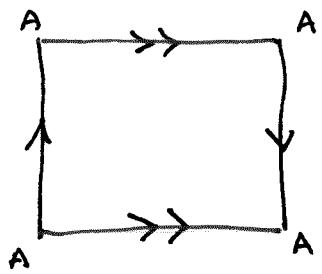
(Each point $x \in X$ is in an equivalence class by itself, except for points on the right and left sides, which are identified in pairs.) Variations on this:

$$\begin{matrix} & B \\ & \downarrow \\ A & \xrightarrow{\quad} & B \\ & \downarrow \\ & A \end{matrix} = \text{Möbius strip}$$

think

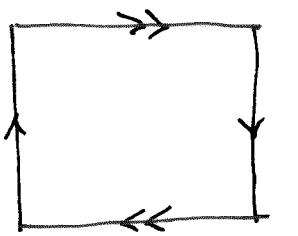
$$\begin{matrix} & A \\ & \rightarrow \\ A & \xrightarrow{\quad} & A \\ & \uparrow \\ & A \end{matrix} = \text{2-torus, } T^2$$

← points on corners are in a 4-point equiv. class. Other points on sides are in 2-point equiv. classes. Other (interior) points are in 1-point equiv. classes.

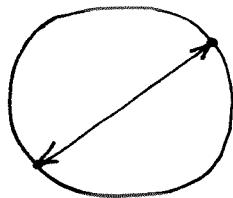


= Klein bottle.

a 2-dim. surface that
cannot be imbedded in \mathbb{R}^3
without ~~of~~ self-intersections.



=



= \mathbb{RP}^2

disk with opposite
points on boundary
identified

x-y

Note: the 2-disk D^2 is the set $x^2+y^2 \leq 1$ in the plane, i.e. it is the interior of a circle plus the boundary.

(B) A different identification for the disk D^2 . Let interior points be in equivalence classes, let all boundary points be in a single equiv. class.

$$X = D^2 = \text{shaded circle}. \quad \text{Let } x \sim y \text{ if } |x| = |y| = 1.$$

Then

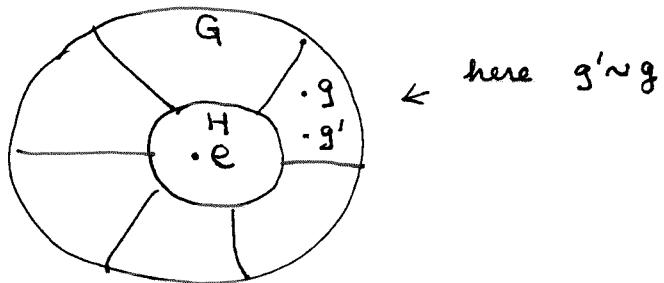
$$\frac{X}{\sim} = \frac{D^2}{\sim}, \quad \text{shaded circle} \rightarrow \text{solid torus} \rightarrow \text{circle} = S^2.$$

$$\frac{D^2}{\sim} = S^2.$$

Another example of equiv. classes, quotient spaces, this time from abstract group theory. Let G = a group, H a subgroup, $H \subset G$. Let $g, g' \in G$, and define $g' \sim g$ if $\exists h \in H$ such that $g' = gh$. Easy to show from def'n that this is an equivalence relation; can also show this by using group actions, as in HW problem. Now identify equivalence classes by representative elements. Start with $[e]$, which is

$$[e] = \{eh \mid h \in H\} = H.$$

The equiv. class of the identity is the subgp. H itself. But since equiv. classes are disjoint, G gets divided up like this:



* Consider $[g]$ for any element $g \in G$. This is the set

$$[g] = \{gh \mid h \in H\} = gH$$

where the notation gH means, multiply each element of H by g to get a new list of elements of G , which constitute the equiv. class $[g]$. This equiv. class is called a left coset (of H).

One can also define right cosets. These are the sets

$$\{hg \mid h \in H\} = Hg,$$

and they correspond to a different equivalence relation, $g' \sim g$

when $g' = hg$ for some $h \in H$. In general, the right cosets are different from the left cosets (the two equivalence relations divide G into equiv. classes in two different ways).

Return to left cosets. We can define the quotient space

$$\frac{G}{\sim} = \frac{G}{H} = \text{space of left cosets} = \underline{\text{coset space}}$$

↑ this is one notation for it.

In using notation like this, the equiv. relation intended must be supplied by context (left cosets, right etc.). Here we are thinking of left cosets. The quotient space is also called a "homogeneous space." It is a space in which each point represents a coset in G .

Does G/H inherit a group structure from G ? That is, is it meaningful to multiply cosets? If $a, b \in G$ and $[a], [b]$ are cosets, then the logical definition of the product of two cosets is

$$[a][b] = [ab].$$

But this is only meaningful if the equivalence class on the RHS is independent of the representative elements a, b on the LHS. That is, let $a' = ah_1, b' = bh_2$ where $h_1, h_2 \in H$, so that a', b' are new representative elements on the LHS. Is it true that $a'b' \sim ab$, i.e. that $a'b' = abh_3$, some $h_3 \in H$; i.e. that $(ab)^{-1}a'b' \in H$? But

$$(ab)^{-1}a'b' = b^{-1}a^{-1}ah_1bh_2 = b^{-1}h_1bh_2,$$

and this $\in H$ iff $b^{-1}h_1b \in H$. Thus we can say, the proposed multiplication law of cosets is meaningful if for all

(6)

$b \in G, h \in H, b^{-1}hb \in H$. But this implies
 $b^{-1}Hb = H$ (with the set theoretic interpretation of multiplying
a group element times H). This leads to a definition:

Def. A subgroup $H \subset G$ is said to be normal if $b^{-1}Hb = H$,
 $\forall b \in G$.

This does not mean that each individual element of H is invariant under conjugation by $b \in G$, just that if you conjugate all the elements of H by b , you get the same list all over again, perhaps in a different order. As we say, H is an invariant subgroup under conjugation by any element of G .

So the logic above says that multiplication of cosets is defined if H is normal. The converse is also true.

The quotient (or coset) space $\frac{G}{H}$ is defined as a set regardless of the properties of H , but it is a group (the quotient group) iff H is normal.

If H is normal, then the left cosets and right cosets are identical, so it doesn't matter which equiv. relation is used to construct the quotient group. If H is not normal, then the 2 spaces of cosets (left and right) are not the same.

If G is Abelian, then $b^{-1}Hb = Hb^{-1}b = H$, and all quotient spaces G/H are groups (all subgroups are normal).

Now for some selected topics in linear algebra, with a slight geometrical flavor.

Let V be a vector space over a field K . In practice, K usually = \mathbb{R} or \mathbb{C} (real or complex vector spaces); this means that the coefficients used in forming linear combinations of vectors (scalars) are either real or complex numbers. (In general it is not meaningful to say whether the vectors themselves are "real" or "complex".) Assume you know definitions of vector space, basis, linear (in)dependence, span, dimension, etc. We will (for now) deal only with finite dimensional vector spaces.

Psychological problem: Physicists have a tendency to assume that all vector spaces possess a metric, i.e., a definition of a scalar product, but for many problems the vector spaces you encounter do not possess any metric that is ~~is~~ natural to the problem at hand. (Example: the (x, p) phase space of a mechanical problem in 1D.) Of course you can always introduce an arbitrary metric, but this is almost always a bad idea unless the metric is a natural outcome of the structure presented by your problem, or ^(occasionally) unless you can show that results don't depend on the choice of the metric.

Instead it is better to develop those structures of linear algebra that can be developed without reference to any metric, and to understand those. Then we introduce a metric and see what new structures become available. This is what we shall do.

- First, let V be a vector space and $\{e_1, e_2, \dots, e_n\}$ be a basis in V (thus, $\dim V = n$). If $v \in V$, then we can write in a unique way,

$$v = \sum_{i=1}^n v^i e_i,$$

where $v^i \in K$ are the components of v w.r.t. the basis $\{e_i\}$.

The upper (contravariant) index on v is deliberate. Note, $v^i \in K$ but $e_i \in V$.

Now consider linear maps, that is, linear homomorphisms between vector spaces, $f: V \rightarrow W$. These satisfy

$$\begin{aligned} f(v_1 + v_2) &= f(v_1) + f(v_2), & \forall v_1, v_2 \in V. \\ f(kv) &= kf(v) & \forall v \in V, k \in K \end{aligned}$$

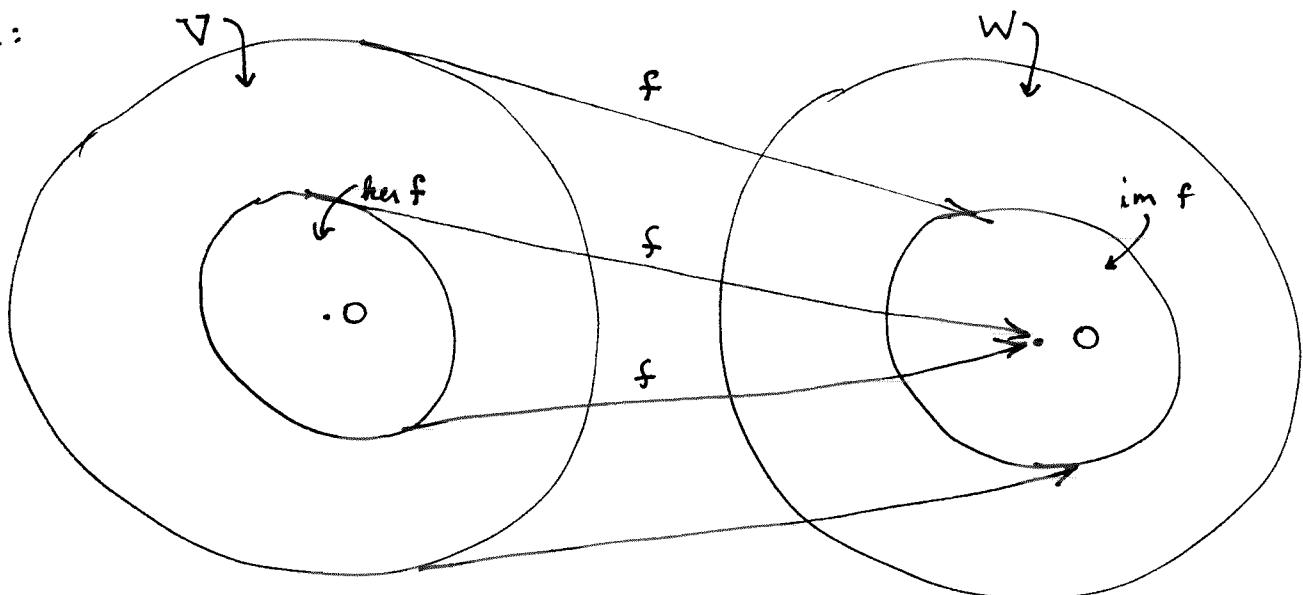
the field

There are two spaces that can be defined using only the linear structure:

- $\ker f = \{v \in V \mid f(v) = 0\}$ = set of all vectors in V annihilated by f
- $\text{im } f = \{w \in W \mid w = f(v) \text{ for some } v \in V\}$ = usual def'n. of image.

Note, $\begin{cases} \ker f \subset V \\ \text{im } f \subset W \end{cases}$, note also, $\ker f$ is a vector subspace of V
 $\text{im } f$ is a vector subspace of W

Picture:



Basic theorem (important):

$$\dim \ker f + \dim \text{im } f = \dim V$$

Context: $f: V \rightarrow W$, linear map.
field K ($= \mathbb{R}$ or \mathbb{C})

Proof: Let $\{g_1, \dots, g_r\}$ be a basis for $\ker f$ ($g_i \in V$)

Let $\{h'_1, \dots, h'_s\}$ be a basis for $\text{im } f$ ($h'_i \in W$).

Let v = any vector in V . Note $f(v) \in \text{im } f$, so it can be expanded in the basis $\{h'_i\}$:

$$f(v) = \sum_{i=1}^s c^i h'_i, \quad c^i = \text{coefficients}, \quad c^i \in K.$$

But for each i , \exists some vector $h_i \in V$ such that $h'_i = f(h_i)$ ($f: h_i \mapsto h'_i$) (h_i is not unique, in general). So

$$f(v) = \sum_{i=1}^s c^i f(h_i) = f\left(\sum_{i=1}^s c^i h_i\right),$$

$$\text{or } f\left(v - \sum_{i=1}^s c^i h_i\right) = 0. \quad \text{Thus } v - \sum_{i=1}^s c^i h_i \in \ker f,$$

and this vector can be expanded in the basis $\{g_i\}$, say $\sum_{i=1}^r d^i g_i$, $d^i \in K$

$$\text{thus, } v = \sum_{i=1}^r d^i g_i + \sum_{i=1}^s c^i h_i,$$

and we see that any $v \in V$ can be written as a lin. comb. of $\{g_i\}, \{h_i\}$.

Now, are these vectors lin. indep? To show that they are, let

$$\sum_{i=1}^r a^i g_i + \sum_{i=1}^s b^i h_i = 0.$$

Apply f to both sides, use fact that $f(g_i) = 0$, $f(h_i) = h'_i$. Gives

$$\sum_{i=1}^s b^i h'_i = 0 \Rightarrow b^i = 0 \text{ since } \{h'_i\} \text{ are lin. indep.}$$

But this $\Rightarrow \sum_{i=1}^r a^i g_i = 0 \Rightarrow a^i = 0$ since the $\{g_i\}$ are lin. indep. So all

coeffs a^i, b^i vanish, and the set $\{g_i, h_i\}$ are lin. indep. and span V (they form a basis for V).

Thus, $\dim V = r+s = \dim \ker f + \dim \text{im } f$. QED

- Some intuition about this theorem: f acts on V and annihilates some vectors (those in $\ker f$). The ones it doesn't annihilate go into $\text{im } f$. Therefore $\dim V = \dim \ker f + \dim \text{im } f$. (at least this is a way of remembering the theorem.)

Remark: Nakahara calls the subspace of V spanned by the $\{h_i\}$ the "orthogonal complement" to $\ker f$, and he writes $(\ker f)^\perp$ for it. Please ignore this. We don't have a metric on our vector spaces yet, so "orthogonal" is undefined. In any case, the vectors $\{h_i\}$ are not unique, because you can add any element of the kernel to them without changing their definition. That is, if $k \in \ker f$, then $f(h_i+k) = f(h_i) = h_i$, so h_i+k works just as well as h_i .

- So the "space spanned by the $\{h_i\}$ " has no invariant meaning.

Another remark: It turns out, however, that there is a way of defining a space that is "complementary" to $\ker f$, in a certain sense, but it is not a subspace of V , it is a quotient space. See the HW problems.

Now we move to the concept of the dual space to a vector space V . Setup:
Let K be the field over which V is defined ($K = \mathbb{R}$ or \mathbb{C} in practice). Elements of K are called "scalars". We assume V is finite-dimensional.

Let α be a linear map, $\alpha: V \rightarrow K$.

Such maps are called dual vectors or co-vectors.

This is a special case of a linear map, where $W = K$
(a one-dimensional space).

Def. The dual space to V , denoted V^* , is the set of all such linear maps:
(dual vectors).

$$V^* = \{ \alpha \mid \alpha: V \rightarrow K, \text{ linear} \}.$$

First note that V^* (like V) is a vector space, under the obvious definition of multiplication of maps by scalars and the addition of maps:

For $\alpha_1, \alpha_2 \in V^*$ and $c_1, c_2 \in K$, define $c_1\alpha_1 + c_2\alpha_2$ by

$$(c_1\alpha_1 + c_2\alpha_2)(v) = c_1\alpha_1(v) + c_2\alpha_2(v), \quad v \in V.$$

Now it turns out that $\boxed{\dim V = \dim V^*}$ (important fact). Easy way to see this:

Let $\dim V = n$, let $\{e_i, i=1, \dots, n\}$ be a basis in V .

Define $\alpha_i = \alpha(e_i)$, call $\{\alpha_i\}$ the components of $\alpha: V \rightarrow K$.

For given α , the components α_i uniquely specify α . That is, if the n scalars $\alpha_i \in K$ are given, then the action of α on any vector $v \in V$ is determined:

$$\alpha(v) = \alpha\left(\sum_{i=1}^n v^i e_i\right) = \sum_{i=1}^n v^i \alpha(e_i) = \sum_{i=1}^n v^i \alpha_i$$

Conversely, if α is given, the α_i are determined by $\alpha_i = \alpha(e_i)$. So this provides a 1-to-1 map between V^* and K^n ($= \mathbb{R}^n$ or \mathbb{C}^n). Moreover, this map is linear (hence a linear isomorphism). Hence $\dim V^* = n = \dim V$.

Remark 1: The lower index on α_i is deliberate, just as is the upper index on v^i : the lower index is a "covariant" index and the upper is "contravariant". (Standard terminology in tensor analysis.)

Remark 2: The concept of the dual space is important. Often the analysis of a problem becomes clarified when we switch attention from some space to its dual. This will be an important theme in this course.

Return to dual space. We have chosen a basis $\{e_i\}$ in V . What about a basis in V^* ? This would be a set of n linearly independent covectors, call them $\{e^{*i}, i=1, \dots, n\}$, which can be defined by specifying their action on the basis $\{e_i, i=1, \dots, n\}$ in V . The definition

$$e^{*i}(e_j) = \delta_j^i$$

is convenient. It means

$$\alpha = \sum_i \alpha_i e^{*i},$$

where $\alpha_i = \alpha(e_i)$ are the components as defined above. Thus those components are the components in the usual sense of α w.r.t. the basis $\{e^{*i}\}$.

The basis $\{e^{*i}\}$ in V^* is said to be dual to the basis $\{e_i\}$ in V .

"Inner product" notation mentioned by Nakara. N. wants to use the notation,

$$\langle \alpha, v \rangle = \alpha(v)$$

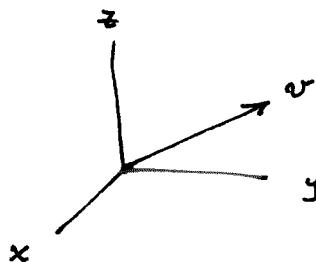
and call it an "inner product." This is ok as long as you don't confuse this with the inner product associated with a metric (which we haven't introduced yet). This \langle , \rangle operation is a map,

$$\xrightarrow{\quad} \langle , \rangle : V^* \times V \rightarrow K : (\alpha, v) \mapsto \alpha(v).$$

But to be safe I'd prefer not to call this an "inner product" because we will define an inner product later that does involve the metric.

(different)

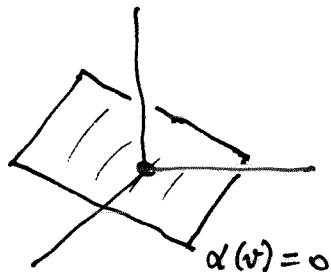
Remark: How to visualize a covector. To visualize a vector $v \in V$ is easy, e.g. if $V = \mathbb{R}^3$,



But what about $\alpha \in V^*$, i.e. $\alpha: V \rightarrow \mathbb{R}$ (we'll take $K = \mathbb{R}$ for this discussion). α of course is a vector in V^* , but how can we visualize it in V ?

Well, α is a real-valued function on V , so we can look at its contour surfaces (level sets), $\alpha(v) = \text{const}$. The value 0 is particularly interesting, the level set $\alpha(v) = 0$ is otherwise just the kernel of α .

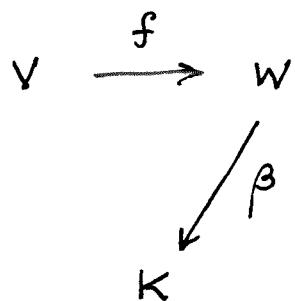
Can easily show that if $\alpha \neq 0$, then $\ker \alpha$ is an $(n-1)$ -dimensional vector subspace of V ($\dim V = n$), i.e. a hyperplane passing thru the origin.



And the surfaces $\alpha(v) = \text{const} \neq 0$ are other hyperplanes parallel to this one. So if you specify $\ker \alpha$ (the hyperplane), and the value of α on any parallel (but different) hyperplane, you have specified α . These hyperplanes, especially $\ker \alpha$, are part of the geometrical interpretation of covectors.

Begin with relation between dual spaces V^* and W^* when we have a map $f: V \rightarrow W$. (There is still no metric.) Does f , which takes a vector $v \in V$ and produces a vector $w = f(v) \in W$, do something similar to V^* and W^* , that is, take a form (=dual vector) in V^* and produce another form in W^* ? The answer is no, in general. But it does allow one to take a form in W^* and produce another form in V^* (in the reverse direction from the action of f itself). This action on forms, going from $W^* \rightarrow V^*$, is called the pull-back.

Suppose we are given $f: V \rightarrow W$ and some $\beta \in W^*$, that is $\beta: W \rightarrow K$ (the scalars). Picture of the maps:



The picture makes it obvious that we can go directly from V to K by composing the maps, that is, let $\alpha \in V^*$ be defined by

$$\alpha = \beta \circ f, \quad \alpha: V \rightarrow K.$$

This specifies a mapping between W^* and V^* which we denote by $f^*: W^* \rightarrow V^*$, called the pull-back of f . That is,

$$f^*: W^* \rightarrow V^*: \beta \mapsto \beta \circ f.$$

An equivalent definition of the pull-back is to specify $f^* \beta$ by its action on vectors in V :

$$(f^* \beta)(v) = \beta(f(v)), \quad \forall v \in V \quad \text{defines } f^* \beta.$$

There is no natural way to define a map: $V^* \rightarrow W^*$ (a "push-forward") unless f is invertible, whereupon you could use f^{-1*} .

We've done almost everything that can be done without a metric, so now let's introduce one. Nakahara's discussion of this is backwards, confused, and wrong in part, so ignore what he says and use the following.

Begin with the case $K = \mathbb{R}$ (real vector spaces), since things are somewhat more complicated when $K = \mathbb{C}$. Idea of metric is measure of distance.

Given real vector space V . A metric or (metric tensor) is a map

$$g: V \times V \rightarrow \mathbb{R},$$

such that:

1) g is linear in both operands,

$$\left. \begin{aligned} g(c_1 u_1 + c_2 u_2, v) &= c_1 g(u_1, v) + c_2 g(u_2, v) \\ g(u, c_1 v_1 + c_2 v_2) &= c_1 g(u, v_1) + c_2 g(u, v_2) \end{aligned} \right\} \begin{array}{l} \forall u_1, u_2, u \in V \\ \forall v_1, v_2, v \in V \\ \forall c_1, c_2 \in \mathbb{R}. \end{array}$$

2) g is positive definite,

$$g(v, v) \geq 0, \quad \forall v \in V$$

$$g(v, v) = 0 \quad \text{iff } v = 0$$

3) g is symmetric,

$$g(u, v) = g(v, u), \quad \forall u, v \in V.$$

Then the quantity $g(u, v)$ is the inner product of 2 vectors, which we may denote by $\langle u, v \rangle$ (another notation for it).

Let $\{e_i\}$ be a basis in V . Then we define

$$g_{ij} = g(e_i, e_j) = \text{component matrix of } g \text{ in the given basis.}$$

Condition 2 implies that g_{ij} is a positive definite matrix, hence that $\det g_{ij} \neq 0$ (g_{ij} is nonsingular, since all of its eigenvalues are positive.)

Note that in relativity theory, we deal with metrics that are not positive definite. In this case, we replace requirement 2) in the definition with the nonsingularity requirement, that $\det g_{ij} \neq 0$. (This condition makes reference to a basis, but is independent of the basis chosen.) A way of writing the nonsingularity condition without reference to a basis is to say, ~~if~~ if $g(u, v) = 0$ for all $u \in V$, then $v = 0$.

A metric, regarded as a distance function on V , induces an association between V and V^* . The latter is an alternative way of looking at a metric. In the expression $g(u, v)$, regard u as fixed and v as variable. To emphasize this, write

$$g_u(v) = g(u, v),$$

thereby defining a function $g_u : V \rightarrow \mathbb{R}$. Such a function is a form, i.e., $g_u \in V^*$. Thus we have a ~~phi~~ mapping, *

$$g : V \rightarrow V^* : u \mapsto g_u. \quad (\text{linear})$$

It's abuse of notation to use the same symbol g for this map as for the distance function, but they're so closely related that everyone does so anyway. Now put this into component language.

~~Let $v \in V$, let $\{e_i\}$ be a basis in V so that $v = \sum_{i=1}^n v^i e_i$, and let $\alpha = g_v$.~~

Let $u \in V$, let $\{e_i\}$ be a basis in V and write $u = \sum_{i=1}^m u^i e_i$, and let $\alpha = g_u \in V^*$. Find components α_i of α w.r.t. the dual basis.

$$\alpha(v) = \sum_j \alpha_j v^j = g_u(v) = \sum_{i,j} u^i g_{ij} v^j, \quad \forall v \in V$$

so $\alpha_i = \sum_j u^i g_{ij}$.

The form $\alpha = g_{ij} u^i$ is so closely associated with u that in applications it is often identified with it, and we just write u_i (with a lower index) instead of α_i or $(g_{ij})_i$. This is called lowering an index.

$$u_i = \sum_j g_{ij} u^j.$$

Now because g_{ij} is invertible, we ~~can~~ have the inverse map: $V^* \rightarrow V$.

Let g^{ij} be the inverse matrix of g_{ij} (std notation), so that

$$\sum_k g_{ik} g^{kj} = \delta_i^j.$$

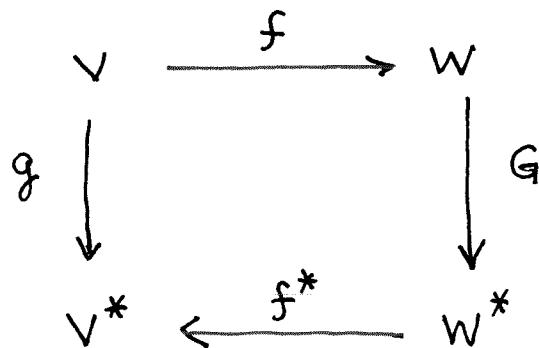
Then the inverse map $g^{-1}: V^* \rightarrow V: \alpha^i \mapsto u^i$ is specified in components by

$$u^i = \sum_j g^{ij} \alpha^j.$$

One often writes α^i instead of u^i (same symbol, but with an upper index) and speaks of raising an index.

summary: A metric g on V induces an ^{real vector space.} ^{linear} isomorphism between V and V^* .

Now examine how metrics interact with maps. Suppose we have a linear map between spaces, each of which possesses a metric. Let g be the metric on V and G the metric on W , and suppose $f: V \rightarrow W$ is linear. Then we have the following picture of maps and spaces,



Notice that there is a route to get from W to V , since g and G are invertible.