1 Vectors and Vector Spaces

§ 1.3 Rotations in 3D and SO(3) Groups

Example 1.19. Consider a vector $v \in \mathbb{R}^3$. Suppose that v is rotated into $v' \in \mathbb{R}^3$ via some rotation R.

This can be represented by *matrix-vector multiplication*. In other words, R can be represented as 3×3 matrix.

$$\begin{bmatrix} (v')^1 \\ (v')^2 \\ (v')^3 \end{bmatrix} = \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{bmatrix} \begin{bmatrix} v^1 \\ v^2 \\ v^3 \end{bmatrix}$$

Separating by each *row* of the vector v', we can express the equation above as

$$(v')^\mu = R^\mu{}_\nu v^\nu$$

Note that we changed the location of index consistent to the vector components.

Such matrix R is an element of SO(3) group.

- O(3) group contains *isometries*[1] in $\mathbb{R}^3$.

- The SO(3) group contains O(3) rotations with $\det R = +1$.

<div align="center">൦ൠൟൡ൫</div>

Example 1.20. In **Example 1.6**, we mentioned that the basis is not unique. For example, the following two basis

$$\{e_\mu\} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}, \quad \{e'_\nu\} = \left\{ \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}, \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

which span $\mathbb{R}^3$. Note that these two basis set are **orthonormal**: $e_\mu \cdot e_\nu = \delta_{\mu\nu}$.

The point is - even if we change the basis that describes $\mathbb{R}^3$, the vector itself **does not change**. In other words, changing our point of view does not alter identity of the object.

$$v = v^j e_j = (v')^i e'_i$$

From **Example 1.15** (rotation of components),

$$\begin{aligned} (v')^i e'_i &= (R^i{}_j v^j) e'_i \\ &= v^j (R^i{}_j e'_i) := v^j e_j \end{aligned}$$

we get transformation rule of basis, $e_j = R^i{}_j e'_i$.

Remark. This **passive transformation** point of view enables us to change the coordinate system without altering the nature of object itself.

---

[1] Actually, $M \in$ O(3) means $M^\mathsf{T} M = M M^\mathsf{T} = \mathbb{I}_3$. We will prove this soon.

Example 1.21. In case of standard Euclidean basis(which is orthonormal), the following holds.

$$e_j = R^i{}_j e'_i, \ R^\mathsf{T} R = \mathbb{I}_3, \ \det R = +1, \ e_i \cdot e_j = \delta_{ij}$$

Here, R is a 3×3 representation of SO(3) group. In case of **general linear coordinates**, some of the conditions are lifted.

$$e_j = S^i{}_j e'_i, \ S^\mathsf{T} S \neq \mathbb{I}_3, \ \det S > 0, \ e_i \cdot e_j = g_{ij}$$

The basis is transformed by some 3×3 matrix S. But S is not necessarily element of SO(3). In general linear coordinates, orthonormality of basis is lifted up; off-diagonal terms may not vanish. We can gather all $e_i \cdot e_j$ to form a 3×3 matrix g_{ij}, the **metric**.

❦

The metric defines how we should *inner product* two vectors.

Example 1.22. Assuming the standard basis in $\mathbb{R}^3$, **interval** between two points (x, y, z) and $(x + \mathrm{d}x, y + \mathrm{d}y, z + \mathrm{d}z)$ is defined by

$$\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 = \begin{bmatrix} \mathrm{d}x & \mathrm{d}y & \mathrm{d}z \end{bmatrix} \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix} \begin{bmatrix} \mathrm{d}x \\ \mathrm{d}y \\ \mathrm{d}z \end{bmatrix}$$

The sandwiched matrix is actually: Kroenecker delta - metric of Euclidean space.

$$\mathrm{d}s^2 = \mathrm{d}x^\mu \delta_{\mu\nu} \mathrm{d}x^\nu = \delta_{\mu\nu} \mathrm{d}x^\mu \mathrm{d}x^\nu$$

In general linear coordinates, Kroenecker delta changes into metric tensor of that coordinate system.

$$\boxed{\mathrm{d}s^2 = \mathrm{d}x^\mu g_{\mu\nu} \mathrm{d}x^\nu}$$

Observation 1.23. Consider a Euclidean space with Euclidean metric $\delta_{\mu\nu}$. Rotating basis should yield the same interval $\mathrm{d}s^2$ indeed. Hence, $(\mathrm{d}s^2)' = \mathrm{d}s^2$ gives

$$\delta_{\mu\nu} \mathrm{d}x'^\mu \mathrm{d}x'^\nu = \delta_{\mu\nu} R^\mu{}_\lambda \mathrm{d}x^\lambda R^\nu{}_\sigma \mathrm{d}x^\sigma$$
$$= (R^\mu{}_\lambda \delta_{\mu\nu} R^\nu{}_\sigma) \mathrm{d}x^\lambda \mathrm{d}x^\sigma := \delta_{\lambda\sigma} \mathrm{d}x^\lambda \mathrm{d}x^\sigma$$

In terms of matrices,

$$R^\mathsf{T} \cdot \mathbb{I}_3 \cdot R = \mathbb{I}_3 \quad \Longrightarrow \quad R^\mathsf{T} R = \mathbb{I}_3$$

Hence

$$\mathrm{O}(3) := \{ R \in \mathrm{GL}(3, \mathbb{R}) \mid R^\mathsf{T} R = \mathbb{I}_3 = R R^\mathsf{T} \}$$

Observation 1.24. Consider the **Minkowski space-time**. Now space and time should be considered at once:

$$\mathrm{d}\vec{x} = \begin{bmatrix} \mathrm{d}t & \mathrm{d}x & \mathrm{d}y & \mathrm{d}z \end{bmatrix}^{\mathsf{T}}$$

The spacetime is equipped with **Minkowski metric** η.

$$\eta_{\mu\nu} = \mathrm{diag}(-1, 1, 1, 1)$$

Rotations in $\mathbb{R}^3$ corresponds to **Lorentz transformations** in space-time: the transformation $L \in \mathrm{GL}(4, \mathbb{R})$ which connects different inertial frames!

$$(x')^{\mu} = L^{\mu}{}_{\nu} x^{\nu}$$

Lorentz transformation conserves the **space-time interval**[2] $\mathrm{d}s^2 = (\mathrm{d}s')^2$. This gives

$$\eta_{\mu\nu}\, \mathrm{d}x'^{\mu}\, \mathrm{d}x'^{\nu} = \eta_{\mu\nu} L^{\mu}{}_{\lambda}\, \mathrm{d}x^{\lambda} L^{\nu}{}_{\sigma}\, \mathrm{d}x^{\sigma}$$
$$= L^{\mu}{}_{\lambda} \eta_{\mu\nu} L^{\nu}{}_{\sigma} \cdot \mathrm{d}x^{\lambda}\, \mathrm{d}x^{\sigma} := \eta_{\lambda\sigma}\, \mathrm{d}x^{\lambda}\, \mathrm{d}x^{\sigma}$$

which reduces into $L^{\mathsf{T}} \cdot \eta \cdot L = \eta$. Similar to SO(3) group

$$\mathrm{SO}(3) = \{ R \in \mathrm{GL}(3, \mathbb{R}) \mid R^{\mathsf{T}} R = \mathbb{I}_3,\ \det R = +1 \}$$

we can define the group of all Lorentz transformations.

$$\mathrm{SO}(1, 3) = \{ R \in \mathrm{GL}(4, \mathbb{R}) \mid L^{\mathsf{T}} \eta L = \eta,\ \det L = +1,\ L^{0}{}_{0} > 0 \}$$

The last two conditions are required to define proper, orthochronous Lorentz transformations, but we will not go deeper here.

---

[2]This is derived from the assumption of special relativity.

2 Tensors and Tensor Products

§ 2.1 Tensors: Definition

Definition 2.1 ((1,0)-tensors)**.** We (re)identify vectors as **(1,0)-tensors**.

$$v = v^\mu e_\mu \in V$$

Definition 2.2 ((0,1)-tensors)**.** We (re)identify dual vectors as **(0,1)-tensors**.

$$w = w_\mu e^\mu \in V^*$$

where dual basis e^μ is defined by $e^\mu(e_\nu) = \delta^\mu{}_\nu$.

In other words, for $w \in V^*$ and $v \in V$, $w(v) \in \mathbb{R}$.

Definition 2.3 (Natural pairing)**.** The **natural pairing** $\langle \bullet, \bullet \rangle \,:\, V^* \times V \to \mathbb{R}$ is defined by

$$\langle w, v \rangle = w(v) \in \mathbb{R}$$

Note that this is not an inner product.

❧

Definition 2.4 (Bilinear form)**.** Consider a **bilinear form** $T \,:\, V \times V \to \mathbb{R}$. For $u, v \in V$, $T(u, v) \in \mathbb{R}$ and

$$T(a_1 u_1 + a_2 u_2, v) = a_1 T(u_1, v) + a_2 T(u_2, v)$$
$$T(u, b_1 v_1 + b_2 v_2) = b_1 T(u, v_1) + b_2 T(u, v_2)$$

for $a_1, a_2, b_1, b_2 \in F$ and $u_1, u_2, u, v_1, v_2, v \in V$.

Definition 2.5 ((0,2)-tensor)**.** We (re)identify bilinear forms as **(0,2)-tensors**. Components of (0,2)-tensor T is defined by

$$T(u, v) = T(u^i e_i, v^j e_j) = u^i v^j T(e_i, e_j) := T_{ij} u^i v^j$$

Remark. Where does (0,2)-tensor belong? Is it $(V^*)^2$? No, it is not.

Definition 2.6 (Symmetric and antisymmetric tensors)**.** A (0,2)-tensor is **symmetric** if it does not change sign upon changing two inputs.

$$T(u, v) = T(v, u) \quad (u, v \in V) \implies T_{ij} = T_{ji} \quad \text{(in components)}$$

An **antisymmetric** tensor changes sign.

$$T(u, v) = -T(v, u) \quad (u, v \in V) \implies T_{ij} = -T_{ji} \quad \text{(in components)}$$

Remark. This definition can be extended into higher-order tensors with multiple indices. In that case, the tensor is symmetric/antisymmetric with respect to certain indices.

Inner product is an operation such that takes two vectors to yield a scalar - it is a (0,2) tensor!

Definition 2.7 (Inner product). Inner product is a (0,2)-tensor $g : V \times V \to \mathbb{R}$ such that

(1) $g_{\mu\nu} = g_{\nu\mu}$ (g is a symmetric tensor)

(2) $\det g \neq 0$

(3) $g(v,v) \geq 0$ for all $v \in V$. Especially, $g(v,v) = 0 \iff v = 0$.

<center>❧</center>

Definition 2.8 ((q, r)-tensors). A (q,r)-tensor $T^{(q,r)} : (V^*)^q \times V^r \to \mathbb{R}$ has q vector-like parts and r covector-like parts. Note that covectors get vectors as inputs[3]. Since vector components have superscript indices and dual vector components have subscript indices, the component of (q,r)-tensor is

$$T^{i_1 \cdots i_q}{}_{j_1 \cdots j_r} = T(e^{i_1}, \cdots, e^{i_q}, e_{j_1}, \cdots, e_{j_r})$$

Here, we are hiding how can we construct such (q,r)-tensors. We relyed on several linear algebra analogies until now. Before we proceed, let's remark one thing:

Remark. We can view the bilinear form (or a (0,2)-tensor) $T : V \times V \to \mathbb{R}$ in another perspective. If we feed one input in advance, $T(\bullet, v)$ $(v \in V)$ acts like a covector(it takes one vector to yield a scalar).

$$T(\bullet, v) : V \to \mathbb{R} \quad \implies \quad T : V \to V^* \text{ ???}$$

It seems that one vector space V has moved right hand side of the arrow, by adding the asterisk! This kind of **partial insertion view** is quite useful understanding higher order tensors. In components,

$$T_{ij} v^j := w_i$$

Pre-substituting one vector(note the superscript index) gives w_i: the subscript on the component indicates that w is a covector.

---

[3]Even if vectors do not take covectors as *inputs*, we may rely on duality for our better understanding.

§ 2.2 Tensor Products

Example 2.9. Suppose that we have two (0,1)-tensors $v = v_\mu e^\mu$ and $w = w_\nu e^\nu$. How can we construct a (0,2)-tensor from v and w? The answer is **tensor product**(denoted by $\otimes$). The new (0,2)-tensor $v \otimes w$ is defined by

$$(v \otimes w)(e_i, e_j) := v_i w_j$$

Although we do not proof this claim, *given the basis $\{e^\mu\}$, $\{e^\mu \otimes e^\nu\}$ forms basis of any (0,2)-tensor.*

$$T = T_{\mu\nu} e^\mu \otimes e^\nu$$

Now, we denote the space of all (0,2)-tensors as $\boxed{V^* \otimes V^*}$.

Definition 2.10 (Tensor Product Space)**.** The space of all (q, r)-tensors is denoted by

$$\bigotimes_q V \otimes \bigotimes_r V^*$$

It has q components that transform like vectors, and r components that transform like covectors. Note that

$$T^{(q,r)} : \bigotimes_q V^* \otimes \bigotimes_r V \to \mathbb{R}$$

Vectors take covectors as inputs[4] and covectors take vectors as inputs.

Now we know how higher-order tensors are constructed. Let's downgrade them now.

Definition 2.11 (Contraction)**.** Given (q, r)-tensor $T^{(q,r)}$, **contraction** decreases vector and covector indices by 1 (yielding $(q-1, r-1)$-tensor). Practically,

$$T = (\cdots, e^\mu, \cdots, e_\mu, \cdots)$$

putting same index for some vector and covector part and summing - is a contraction.

Example 2.12. Contraction utilizes the natural pairing between dual basis and original vector space basis.

(1) Consider a (1,1)-tensor T. Since it has one vector-like component and one covector-like component, it should be represented by

$$T^\mu{}_\nu (e_\mu \otimes e^\nu) \in V \otimes V^*$$

From the definition of dual basis $e^\nu(e_\mu) = \langle e^\nu, e_\mu \rangle = \delta^\nu{}_\mu$, contraction gives

$$T^\mu{}_\nu \delta^\nu{}_\mu = T^\mu{}_\mu \in \mathbb{R}$$

Note that (1) contracting a (1,1)-tensor yielded (0,0)-tensor(a scalar) and (2) contraction is similar to **taking a trace** through certain selected axis.

(2) Consider a (1,2)-tensor $T^a{}_{bc}(e_a \otimes e^b \otimes e^c) \in V \otimes V^* \otimes V^*$. Then we have two possible options for contraction:

$$\langle e^b, e_a \rangle = \delta^b{}_a \implies T^a{}_{bc} \delta^b{}_a = T^a{}_{ac} \in V^*$$

$$\langle e^c, e_a \rangle = \delta^c{}_a \implies T^a{}_{bc} \delta^c{}_a = T^a{}_{ba} \in V^*$$

---

[4]Do not take this message too seriously!

What does the location of indices mean?

- If the component index is **superscript**, it means that it transforms like vector components.

- If the component index is **subscript**, it means that it transforms like covector components.

Definition 2.13 (Raising & lowering indices). Consider $u = u^\mu e_\mu \in V$. The **metric tensor**

$$g = g_{\mu\nu} e^\mu \otimes e^\nu \in V^* \otimes V^*$$

is used for raising and lowering of indices. Tensor product with metric tensor followed by contraction yields

$$(u^\mu e_\mu) \otimes (g_{\nu\lambda} e^\nu \otimes e^\lambda) \implies (u^\mu g_{\nu\lambda} \delta^\nu{}_\mu) e^\lambda = \underbrace{(u^\mu g_{\mu\lambda})}_{:= u_\lambda} e^\lambda$$

since $\langle e^\nu, e_\mu \rangle = \delta^\nu{}_\mu$. The same procedure gives raising of indices.

$$u_\mu = g_{\mu\nu} u^\nu, \ u^\mu = g^{\mu\nu} u_\nu$$

Definition 2.14 (Inverse of metric tensors). We know metric tensor as (0,2)-tensor, which has two covector-like components $g_{\mu\nu}$. Then what is $g^{\mu\nu}$? $g^{\mu\nu}$ is a **name** given to inverse of $g_{\mu\nu}$.

$$g_{\mu\nu} g^{\nu\lambda} = g_\mu^\lambda \quad \text{(in index raising perspective, shorthand notation)}$$
$$= \delta_\mu{}^\lambda \quad \text{(since } g g^{-1} = \mathbb{I})$$

The *awkward* index positions(considering the original shape of a tensor) are result of *shorthand* notations. For example, a (2,0)-tensor $T^{\mu\nu}$,

$$T^\mu{}_\lambda = T^{\mu\nu} g_{\nu\lambda}$$

the LHS is a shorthand for the RHS.

Observation 2.15. Someone may wonder **the reason we use metric for raising and lowering the indices**. The short answer is - metric encodes the geometry of spaces. It tells us how the space is curved. But if it does not satisfy your eager, there exists an alternative explanation. With the partial insertion view, define a new dual vector

$$\tau(x) := (x, \bullet) \in V^*$$

where $(\bullet, \bullet)$ denotes the usual inner product. Here, $g_{\mu\nu} = (e_\mu, e_\nu)$. Then,

$$g^{\mu\lambda} \tau(e_\lambda)(e_\nu) = g^{\mu\lambda}(e_\lambda, e_\nu) = g^{\mu\lambda} g_{\lambda\nu} = \delta^\mu{}_\nu$$

Recall the definition of dual basis: $e^\mu(e_\nu) = \delta^\mu{}_\nu$. Hence, $g^{\mu\lambda} \tau(e_\lambda)$ is the dual basis e^μ! In other words, metric maps e_λ to a dual basis e^μ. For the general vector $X = X^\mu e_\mu \in V$,

$$\tau(X) = (X, \bullet) = X^\mu (e_\mu, \bullet)$$
$$= X^\mu \tau(e_\mu) = X^\mu \delta_\mu{}^\nu \tau(e_\nu)$$
$$= X^\mu g_{\mu\lambda} \underbrace{(g^{\lambda\nu} \tau(e_\nu))}_{= e^\lambda}$$
$$:= X_\lambda e^\lambda$$

We can observe $X^\mu g_{\mu\lambda}$ corresponds to X_λ, considering the previous analogies.

Exercise 2.16. Consider the **energy-momentum tensor of electromagnetic field**

$$T^{ab} = F^{ac}F^b_{\ c} - \frac{1}{4}\eta^{ab}F^{cd}F_{cd}$$

where $\eta = \text{diag}(-1, 1, 1, 1)$ is the Minkowski metric. Note that space-time has dimension 4.

(a) The **electromagnetic tensor** F_{ab} is defined as

$$F_{ab} = \partial_a A_b - \partial_b A_a$$

where $\partial_a := \dfrac{\partial}{\partial x^a}$ and $A_\mu = (-\phi, \vec{A})$. Show that F_{ab} is an antisymmetric tensor[a].

(b) Using the fact that $F_{i0} = E_i$ $(i = 1, 2, 3)$ and $F_{ij} = \epsilon_{ijk}B^k$ $(i, j = 1, 2, 3)$, show that

$$T^{0j} = (\mathbf{E} \times \mathbf{B})^j \ (j = 1, 2, 3)$$

(c) Show that $T^{ab} = T^{ba}$ (T is symmetric).

(d) Show that $T^a_{\ a} = 0$ (T is traceless).

---

[a]Therefore, all diagonal part is zero.

Proof. Note that ∂_a, partial differentiation with respect to x^a, behaves like covector.

(a) Trivial.

(b) Since $\eta^{0j} = 0$ $(j = 1, 2, 3)$,
$$T^{0j} = F^{0i}F^j_{\ i} = -F^{i0}F^j_{\ i}$$

Here,

$$F^{i0} = F_{ab}\eta^{ia}\eta^{0b} = F_{ab}\eta^{ia}(-\delta^{0b}) = -F_{a0}\eta^{ia} = -E_a\eta^{ia} = -E^i$$
$$F^j_{\ i} = F_{bi}\eta^{bj} = \epsilon_{bik}B^k\eta^{bj}$$

Combining these two yields

$$T^{0j} = E^i\epsilon_{bik}B^k\eta^{bj} = (\mathbf{E} \times \mathbf{B})_b\eta^{bj} = (\mathbf{E} \times \mathbf{B})^j$$

(c) Note that $\eta^{ab} = \eta^{ba}$.

$$T^{ba} = F^{bc}F^a_{\ c} - \frac{1}{4}\eta^{ba}F^{cd}F_{cd}$$
$$= F^{bc}F^{ad}\eta_{dc} - \frac{1}{4}\eta^{ab}F^{cd}F_{cd} = F^{ad}F^b_{\ d} - \frac{1}{4}\eta^{ba}F^{cd}F_{cd} = T^{ab}$$

(d) Note that $\eta^{ab}\eta_{ab} = \text{tr}(\text{diag}(-1, 1, 1, 1) \cdot \text{diag}(-1, 1, 1, 1)) = 4$.

$$T^a_{\ a} = T^{ab}\eta_{ab} = F^{ac}F^b_{\ c}\eta_{ab} - \frac{1}{4}\eta^{ab}\eta_{ab}F^{cd}F_{cd} = F^{ac}F_{ac} - \frac{1}{4} \cdot 4 \cdot F^{cd}F_{cd} = 0$$

∎