Vector Spaces and Tensors

Chapter 0 Vector-space basics

Vector spaces form the backbone of tensor mathematics. This section provides a short review of needed vector space topics so the development does not form a distraction later during deep discussions of general relativity concepts. Theorems that are well-known but which I do not take the time to prove here are labeled as “Facts”.

Definition For these notes, **scalars** are either the reals, $\mathbb{R}$, or the complex numbers, $\mathbb{C}$. A **vector space** is a set **V**, whose elements are called **vectors**, in which the operations addition and scalar-multiplication have been defined and satisfy the following natural rules.

1. If v, w, and x are vectors then v + w = w + v and v + (w + x) = (v + w) + x
2. V has a unique vector 0 such that v + 0 = v for every v ∈ V
3. To each v ∈ V there is a vector -v such that v + (-v) = 0
4. For a scalar α, there is a vector αv ∈ V for each v ∈ V
5. 1v = v, α (βv) = ($\alpha\beta$)v (where 1 is the unit scalar)
6. α (v +w) = αv + αw (0-1)

Definitions

- A set of vectors $\mathcal{V} = \{\boldsymbol{v}_i\}$ is **linearly independent** if $\sum \alpha_i \boldsymbol{v}_i = 0 \Rightarrow \alpha_i = 0$ for all i. (0-2)
- Otherwise we say that $\mathcal{V}$ is **linearly dependent**.
- A set of vectors $\mathcal{W} = \{\boldsymbol{w}_i\}$ **spans V** if whenever $\boldsymbol{v} \in$ **V**, ∃ scalars α_i such that $\boldsymbol{v} = \sum \alpha_i \boldsymbol{w}_i$. (0-3)
- A **basis for a vector space V** is a set $\mathcal{B} = \{\mathbf{e}_i\}$ of linearly independent vectors that spans **V**. The **dimension of V** is the number of elements in $\mathcal{B}$. (0-4)

What this means is that no basis vector can be expressed as a linear combination of the other basis vectors, and that every vector in **V** can be expressed as a linear combination of the basis vectors.

Definition A **linear transformation** of a vector space **V** *into* a vector space **W** is a mapping T: **V**→**W** : $T(\alpha \boldsymbol{v} + \beta \boldsymbol{w}) = \alpha\, T(\boldsymbol{v}) + \beta\, T(\boldsymbol{w})$ for all vectors $\boldsymbol{v}, \boldsymbol{w} \in$ **V** and all scalars α and β. If T is a 1-1 map from **V** *onto* **W** we call T an **isomorphism**. (0-5)

A linear transformation preserves the vector space structure, meaning that addition and scalar multiplication match up:

$$\alpha \boldsymbol{v} + \beta \boldsymbol{w} \;\leftrightarrow\; \alpha\, T(\boldsymbol{v}) + \beta\, T(\boldsymbol{w})$$

However, if T is not *onto*, there will be vectors in **W** with no counterpart in **V**. In that case, $T(\mathbf{V})$ is a proper subspace of **W.**

If T is not *1-1,* there will be non-zero vectors $\boldsymbol{v} \in \mathbf{V}$ such that $T(\boldsymbol{v}) = 0$. These vectors have no counterpart in **W**.

In an isomorphism, all vectors have counterparts, and from a vector space perspective **V** and **W** are identical. In this case, T^{-1}: $\mathbf{W} \to \mathbf{V}$ is also an isomorphism.

Fact 1 A linear transformation defined on the basis vectors of **V** has a unique extension to a linear transformation from all of **V** to **W**. That is, if T is defined on a basis of **V**, we can consider it to be defined on all of **V**.

The next definition uses Fact 1 to show how a matrix can be considered to be a linear transformation, and vice-versa.

Definition Let $\boldsymbol{V}$ be any N-dimensional vector space with a basis $\{\mathbf{e}_i\}$, $M = (m_{ij})$ an $N \times N$ matrix whose entries are scalars, and T the linear transformation from $\mathbf{V} \to \mathbf{V}$ defined by

$$T(\mathbf{e}_i) = \sum_j m_{ij}\, \mathbf{e}_j \,.$$

We say that **the matrix *M* is associated with *T***. (0-6)

When working with matrices, we represent vectors $\boldsymbol{v} = \begin{pmatrix} \vdots \\ v_j \\ \vdots \end{pmatrix}$ as column vectors. We then consider the matrix expression $M\boldsymbol{v} = \boldsymbol{w}$ to be interchangeable with the *associated* linear transformation expression $T(\boldsymbol{v}) = \boldsymbol{w}$.

Definition The **transpose** of a column vector $\boldsymbol{v}$ is the row vector $\boldsymbol{v}^T = (\cdots\ v_j\ \cdots)$. (0-7)

Even though general relativity is based upon vector spaces whose scalars are real, the Fundamental Theorem of algebra shows that complex numbers can arise as solutions to polynomial equations that have only real coefficients. For this reason, we must temporarily work with vector spaces over the complex numbers to solve certain problems.

The facts and definitions below are for vector spaces over ℂ but also apply to vector spaces over ℝ.

Fact 2 Let **V** be a vector space over ℂ, M be an NxN complex matrix, and T: **V** → **V** the associated linear transformation. Then the following are equivalent.

1. M is singular (i.e., dim[T(**V**)] < N)
2. M is non-invertible (i.e., there is no inverse matrix)
3. det $M = 0$
4. There is a non-zero vector $\boldsymbol{v}$ in **V** such that $M\boldsymbol{v} = 0$

Definition Let $M = (m_{ij})$ be a complex-valued $N \times N$ matrix.

The **transpose** of M is the matrix $M^T = (m_{ji})$. (0-8)

Matrix M is **symmetric** if $M = M^T$; i.e., if $m_{ij} = m_{ji}$ for all i and j.

If z = $x + y\,i$, let $\overline{z} = x - y\,i$ represent its complex conjugate.

The **conjugate** of M is the matrix $M^* = (\overline{m}_{ij})$. (0-9)

We will write M^{T^*} as a shortcut for $(M^T)^*$.

Definition Let $M = (m_{ij})$ be a complex-valued $N \times N$ matrix. We say that $\boldsymbol{v}$ is an **eigenvector of *M*** if there is a scalar $\lambda \in \mathbb{C}$ such that $M\boldsymbol{v} = \lambda\boldsymbol{v}$, and λ is called an **eigenvalue**.

Theorem 0.1 All eigenvectors and eigenvalues of a real symmetric matrix M are real.

Proof. Suppose $M\boldsymbol{v} = \lambda\boldsymbol{v}$ where M has real entries, $\lambda \in \mathbb{C}$, and $\boldsymbol{v}$ has complex components. Since M is real, $M^* = M$. Since M is symmetric, $M^T = M$. Thus, $M = M^{T^*}$. So,

$$\boldsymbol{v}^{T^*} M = \boldsymbol{v}^{T^*} M^{T^*} = (M\boldsymbol{v})^{T^*} = (\lambda\boldsymbol{v})^{T^*} = \lambda^* (\boldsymbol{v}^{T^*}) = \lambda^* \boldsymbol{v}^{T^*} \tag{0-10}$$

$$\boldsymbol{v}^{T^*} M\boldsymbol{v} = \boldsymbol{v}^{T^*}(M\boldsymbol{v}) = \boldsymbol{v}^{T^*}(\lambda\boldsymbol{v}) = \lambda\, \boldsymbol{v}^{T^*}\boldsymbol{v}$$

$$\boldsymbol{v}^{T^*} M\boldsymbol{v} = (\boldsymbol{v}^{T^*} M)\, \boldsymbol{v} \overset{(0-10)}{=} (\lambda^*\, \boldsymbol{v}^{T^*})\, \boldsymbol{v} = \lambda^*\, \boldsymbol{v}^{T^*}\boldsymbol{v}$$

$$\Rightarrow \quad \lambda = \lambda^* \quad \Rightarrow \quad \lambda \text{ is real}$$

Because M is a real matrix and its eigenvalues are also real, the eigenvectors must also be real. That is because $M\boldsymbol{v} = \lambda\boldsymbol{v}$ constitutes a system of real linear equations, and solving them only involves addition, subtraction, multiplication, and division. It does not involve taking complex conjugates. ■

Definition The **dot product** of two (column or row) vectors $\boldsymbol{v} = (v_i)$ and $\boldsymbol{w} = (w_i)$ is

$\boldsymbol{v} \cdot \boldsymbol{w} = \Sigma\, v_i\, w_i$.

Vectors are **orthogonal** if $\boldsymbol{v} \cdot \boldsymbol{w} = 0$. They are **orthonormal** if in addition they are unit vectors; i.e., $\sum v_i{}^2 = 1$. The **norm** of a vector is

$$\|\boldsymbol{v}\| = \sqrt{\boldsymbol{v} \cdot \boldsymbol{v}}\ .$$

The remainder of this section is devoted to proving the following theorem (needed in proof of Theorem 2.4.2 that every point has a coordinate system where the metric tensor matrix is a diagonal matrix having only +1's and -1's on the diagonal).

Theorem 0.2 A real symmetric matrix $M = (m_{ij})$ has a collection of eigenvectors that constitute an orthonormal basis for **V**.

Proof. Let I be the identity matrix. $p(\lambda) \equiv \det(M - \lambda I)$ is a polynomial in λ known as the characteristic polynomial of M. By the Fundamental Theorem of Algebra, the polynomial $p(\lambda)$ can be factored,

$$p(\lambda) = \pm (\lambda - \lambda_1) \cdots (\lambda - \lambda_N),$$

where λ_i are complex roots (even though the elements of M are real). It is possible that some or all of the λ_i are the same, but we have shown that $p(\lambda)$ has at least one characteristic root, λ_1.

We next show that M has at least one eigenvector; namely, an eigenvector $\boldsymbol{v}$ whose eigenvalue is the characteristic root λ_1. We seek a non-zero vector $\boldsymbol{v}$ that satisfies $M\boldsymbol{v} = \lambda_1 \boldsymbol{v}$. Denote $\boldsymbol{v}$ as the column vector (v_j).

Let $\{\mathbf{e}_i\}$ be a basis for **V**. By the definition above, the matrix $M = (m_{ij})$ is associated with the linear transformation T defined on the basis vectors of $\boldsymbol{V}$ by

$$T : \mathbf{V} \to \mathbf{V} : T(\mathbf{e}_i) = m_{ij}\, \mathbf{e}_j\ .$$

Define a matrix $H \equiv M - \lambda_1 I$, where I is the identity matrix (ones on the main diagonal, zeroes elsewhere). So, $\det(H) = \det(M - \lambda_1 I) = p(\lambda_1) = 0$. By Fact 2, there is a non-zero, possibly complex vector $\boldsymbol{w}$ in $\boldsymbol{V}$ such that $H\boldsymbol{w} = 0$. That is,

$$M\boldsymbol{w} - \lambda_1 \boldsymbol{w} = (M - \lambda_1)\, \boldsymbol{w} = H\boldsymbol{w} = 0 \qquad \Rightarrow \qquad M\boldsymbol{w} = \lambda_1 \boldsymbol{w}.$$

Thus, $\boldsymbol{w}$ is an eigenvector of M having the characteristic root λ_1 as its eigenvalue. By Theorem 0.1, $\boldsymbol{v}$ and λ_1 are real.

Having found one eigenvector, we now extend it into an orthonormal basis for $\boldsymbol{V}$. We set $\boldsymbol{v}_1 = \frac{\boldsymbol{w}}{\|\boldsymbol{w}\|}$, a unit vector having λ_1 as its eigenvalue. Define the **null space** of $\boldsymbol{v}_1$:

$$N_1 = \{\, \boldsymbol{v} : \boldsymbol{v} \bullet \boldsymbol{v}_1 = 0\}.$$

It is easy to confirm that N_1 satisfies the definition, above, of a vector space. N_1 is the subspace of vectors that are orthogonal to $\boldsymbol{v}_1$. Claim dim $(N_1) = N - 1$:

Using the Gram-Schmidt orthogonalization process, $\boldsymbol{v}_1$ can be extended to an orthonormal basis $\{\boldsymbol{v}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_N\}$ of $\boldsymbol{V}$. Thus, $\boldsymbol{e}_i \bullet \boldsymbol{v}_1 = 0$ for all $i > 1$ since $\boldsymbol{e}_i$ and $\boldsymbol{v}_1$ are orthonormal basis vectors. Hence, by the definition of N_1, $\{\boldsymbol{e}_2, \cdots, \boldsymbol{e}_N\}$ is contained in N_1, and, thus, forms an $(N - 1)$ dimensional basis for it. ✓

Claim $MN_1 \subseteq N_1$ where $MN_1 = \{M\boldsymbol{v} : \boldsymbol{v} \in N_1\}$:

Let $\boldsymbol{v} \in N_1$. We need to show that $M\boldsymbol{v} \in N_1$. Since $\boldsymbol{v} \bullet \boldsymbol{v}_1 = \boldsymbol{v}^T \boldsymbol{v}_1$ and $M = M^{\mathrm{T}^*}$,

$$M\boldsymbol{v} \bullet \boldsymbol{v}_1 = (M\boldsymbol{v})^{\mathrm{T}^*} \boldsymbol{v}_1 = \boldsymbol{v}^{\mathrm{T}^*} M \boldsymbol{v}_1 = \boldsymbol{v}^{\mathrm{T}^*} \lambda_1\, \boldsymbol{v}_1 = \lambda_1\, \boldsymbol{v}^T \boldsymbol{v}_1 = \lambda_1\, \boldsymbol{v} \bullet \boldsymbol{v}_1 = 0 \qquad ✓$$

Let T_2 be the linear transformation generated by restricting T to N_1, the $(N - 1)$ dimensional null space of $\boldsymbol{v}_1$, and let M_2 be the matrix associated with T_2. Repeating our logic above, $p(\lambda) \equiv \det(M_2 - \lambda\, I)$ has a real root λ_2 that is an eigenvalue of M_2, and λ_2 has a corresponding unit eigenvector $\boldsymbol{v}_2$. Because $\boldsymbol{v}_2 \in N_1$, we have $\boldsymbol{v}_2 \bullet \boldsymbol{v}_1 = 0$.
$\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ forms an orthonormal set. (So, even though λ_2 might equal λ_1, we have that $\boldsymbol{v}_2 \neq \boldsymbol{v}_1$.)

We have to do this one more time. Let T_3 be the linear transformation generated by restricting T_2 to N_2, the $(N - 2)$ dimensional null space of $\boldsymbol{v}_2$. Let M_3 be the matrix associated with T_3. As above, we generate a real eigenvalue λ_3 having a corresponding unit eigenvector $\boldsymbol{v}_3$ that is orthogonal to $\boldsymbol{v}_2$, and since $\boldsymbol{v}_3 \in N_2 \subseteq N_1$, we also have that it is orthogonal to $\boldsymbol{v}_1$. Thus, $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ form an orthonormal set.

Continuing this process, we eventually obtain the orthonormal basis $\{\boldsymbol{v}_i\}$ consisting of eigenvectors of M that have corresponding real eigenvalues. ■

Appendix C2 Dual spaces

Dual space theory provides the foundation of tensors, so we present it next.

Definition Let M be a manifold. The vector space at a point P∈M generated by the tangent basis is called the **tangent space**, $\mathbf{T}_P = \{\lambda^a \boldsymbol{e}_a\}$, and the vector space generated by the gradient basis is called the **cotangent space**, designated $\mathbf{T}_P{}^* = \{\mu_b \mathbf{e}^b\}$.

The definitions along with graphical illustrations of **tangent bases** and **cotangents bases**, also called natural and dual bases, are found in Section 1.1 Spherical and cylindrical natural and dual bases.

The objective in this subsection is to define abstract dual spaces and show that the cotangent space is the dual space of the tangent space, and vice-versa.

Definition Let **V** be an abstract *n*-dimensional vector space and

$\mathbf{V}^* = \{f : \mathbf{V} \to \mathbb{R} : f \text{ is linear}\}$

$\mathbf{V}^{**} = \{\omega : \mathbf{V}^* \to \mathbb{R} : \omega \text{ is linear}\}$

V* is called the **dual space of V**. **V**** is the dual space of **V*** and is called the **2nd dual space of V**. While the members of **V** are referred to as **vectors**, the members of **V*** are called **dual vectors** or **covectors**.

Members $\boldsymbol{v}$ of **V** are written in boldface, but members f of **V*** and ω of **V**** are functions and are not bolded.

Theorem C2.1 Function spaces are vector spaces

Proof. It is straight-forward to show that function spaces like **V*** and **V**** satisfy the vector space conditions when the following natural definitions are made:

Zero function: $0(\boldsymbol{v}) \equiv \mathbf{0}$ for all $\boldsymbol{v}$

Additive inverse of T: $-T(\boldsymbol{v}) \equiv -\boldsymbol{v}$

Addition: $(S + T)(\boldsymbol{v}) \equiv S(\boldsymbol{v}) + T(\boldsymbol{v})$

Scalar Multiplication: $(\alpha T)(\boldsymbol{v}) \equiv \alpha\, T(\boldsymbol{v})$ ∎

Theorem C2.2 Let

$\mathcal{B} = \{\mathbf{e}_a : a = 1, \cdots, N\}$ be a basis for **V** and

$$\mathcal{B}^* \equiv \{e^b : \mathbf{V} \to \mathbb{R} : e^b(\mathbf{e}_a) = \delta^b_a\}. \quad \text{(C2-1)}$$

Then $\mathcal{B}^*$ is a basis for **V***

Proof. We need to show that $\{e^b\}$ is linearly independent and spans V*.

<u>Linearly independent</u> $\lambda_b\, e^b = 0 \;\Rightarrow\; \lambda_a = \lambda_b \delta^b_a = \lambda_b\, e^b(\mathbf{e}_a) = 0$ for $a = 1, \cdots, n$ ✓

<u>Spans V*</u> Let $f \in \mathbf{V}^*$. $f(\mathbf{e}_a) \in \mathbb{R}$. Denote it as $f(\mathbf{e}_a) = \epsilon_a$ for $a = 1, \cdots, n$.

If $\boldsymbol{v} \in \mathbf{V}$, then $\boldsymbol{v} = \lambda^a \mathbf{e}_a$ since $\boldsymbol{\mathcal{B}} = \{\mathbf{e}_a\}$ is a basis for **V**. So, for all $\boldsymbol{v}$ we have

$$f(\boldsymbol{v}) = f(\lambda^a \mathbf{e}_a) \overset{(*)}{=} \lambda^a f(\mathbf{e}_a) = \lambda^a\, \epsilon_a = \lambda^a\, \epsilon_b\, \delta^b_a \overset{(C2-1)}{=} \lambda^a\, \epsilon_b\, e^b(\mathbf{e}_a) = \epsilon_b\, e^b(\lambda^a \mathbf{e}_a) = \epsilon_b\, e^b(\boldsymbol{v})$$

$\Rightarrow\; f = \epsilon_b\, e^b$; i.e., f is a linear combination of the e^b. ✓ ■

The justification for step (*), above, is that, by definition of **V***, f is a linear function.

Example C2-1 Let **V** be an *N*-dimensional vector space of column vectors and **W** an *N*-dimensional vector space of row vectors. Basis vectors of **V** and **W**, respectively, are

$$\mathbf{e}_a = \begin{pmatrix} 0 \\ \vdots \\ 1_a \\ \vdots \\ 0 \end{pmatrix} \qquad \text{and} \qquad \mathbf{e}^b = (0 \ldots 1^b \ldots 0)\,.$$

Define a real-valued function e^b on the basis vectors of **V** by

$$e^b(\mathbf{e}_a) = \delta^b_a$$

By Theorem C2.2, $\mathcal{B}^* = \{e^b\}$ is a basis for **V*** and, so, the 1-1 and onto mapping $e^b \mapsto \mathbf{e}^b$ generates an isomorphism from **V*** to **W**. This constitutes the proof of the next theorem.

Theorem C2.3 Row vectors can be regarded as the dual space of column vectors, and column vectors can be regarded as the dual space of row vectors

Proof. Let $\boldsymbol{V}^\top$ is a space of row vectors that are transposes of column vectors:

$$\boldsymbol{v}^\top = (v_1 \ldots v_N) \in \boldsymbol{V}^\top \quad \Leftrightarrow \quad \boldsymbol{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_N \end{pmatrix} \in \mathbf{V}.$$

By Theorem C2.2, $\mathcal{B}^* = \{e^a\}$ is a basis for **V***, the dual space for the column vectors **V**. By Example C2-1, $\boldsymbol{\mathcal{B}}^\top = \{\mathbf{e}^a\}$ is a basis for $\boldsymbol{V}^\top$, and so the mapping $e^a \mapsto \mathbf{e}^a$ defines an isomorphism between the dual space of **V** and the space $\boldsymbol{V}^\top$ of row vectors. The converse statement can be proven similarly. ■

In General Relativity, $\mathbf{T_P}^*$ is defined as the cotangent space, generated by the gradient vectors. In vector space theory, the same symbol $\mathbf{T_P}^*$ is defined as the dual of the tangent vector space. Fortunately, by Theorem C2.4 below, they are isomorphic, and we are free to change between them on-the-fly as needed.

Theorem C2.4 The cotangent space $\mathbf{T}_P^*$ can be considered to be the dual, $\mathbf{T_P}^*$, of the tangent space, and the tangent space $\mathbf{T}_P$ can be considered to be the dual, $(\mathbf{T}_P^*)^*$, of cotangent space.

Proof.

Tangent space: $\mathbf{T}_P = \{\lambda^a \boldsymbol{e}_a\}$ (members are vectors, bolded)

Cotangent space: $\mathbf{T_P}^* = \{\lambda_a \boldsymbol{e}^a\}$ (members are vectors, bolded)

Dual to tangent space: $\mathbf{T_P}^* = \{\lambda_a e^a\}$ (members are functions, not bolded)

The mapping $\boldsymbol{e}^a \mapsto e^a$ of the basis vectors generates the isomorphism $\mathbf{T}_P^* \cong \mathbf{T_P}^*$ ✓

Conversely, the tangent space can be regarded as a vector space of column vectors:

$$\mathbf{T}_P = \left\{ \begin{pmatrix} \lambda^1 \\ \vdots \\ \lambda^N \end{pmatrix} \right\}, \text{ having basis } \{\boldsymbol{\epsilon}^a\} \text{ where } \boldsymbol{\epsilon}^a = \begin{pmatrix} 0 \\ \vdots \\ 1^a \\ \vdots \\ 0 \end{pmatrix}.$$

The cotangent space can be regarded as a vector space of row vectors:

$\mathbf{T_P}^* = \{ (\lambda_1 \ldots \lambda_N) \}$ having basis $\{\boldsymbol{\epsilon}_a\}$ where $\boldsymbol{\epsilon}_a = (0 \ldots 1_a \ldots 0)$.

By Theorem C2.3, the dual to a space of row vectors can be regarded a space of column vectors:

$$(\mathbf{T}_P^*)^* = \left\{ \begin{pmatrix} \omega^1 \\ \vdots \\ \omega^N \end{pmatrix} \right\}, \text{ having basis } \{\epsilon^a\} \text{ where } \epsilon^a = \begin{pmatrix} 0 \\ \vdots \\ 1^a \\ \vdots \\ 0 \end{pmatrix}$$

The mapping $\boldsymbol{\epsilon}^a \mapsto \epsilon^a$ of the basis vectors generates the required isomorphism. ■

Theorem C2.4 underscores that cotangent vectors can be regarded as functions of tangent vectors, and vice versa. This is key to the definition of tensor products.

Theorem C2.5 Let $\boldsymbol{\lambda} = \lambda^a \mathbf{e}_a$ be a tangent vector, $\boldsymbol{\mu} = \mu_b \mathbf{e}^b$ a cotangent vector, and $\lambda = \lambda^a e_a$ and $\mu = \mu_b e^b$ their functional equivalents. Then

$$\mu(\boldsymbol{\lambda}) = \lambda(\boldsymbol{\mu}) = \lambda^a \mu_a \tag{C2-3}$$

Proof. e^b and e_a, respectively, were defined on basis vectors as $e^b(\mathbf{e}_a) = \delta^b_a$ and also $e_a(\mathbf{e}^b) = \delta^b_a$. So,

$$\lambda^a \mu_a = \lambda^a \mu_b \delta^b_a = \lambda^a \mu_b \, e^b(\mathbf{e}_a) = \mu_b e^b(\lambda^a \mathbf{e}_a) = \mu(\boldsymbol{\lambda}) \quad ✓$$

$$\lambda^a \mu_b \delta^b_a = \lambda^a \mu_b \, e_a(\mathbf{e}^b) = \lambda^a e_a(\mu_b \mathbf{e}^b) = \lambda(\boldsymbol{\mu}) \quad ✓ \qquad ■$$

Appendix C3 Tensors

Definition The **product of two abstract vectors spaces** is

$$\mathbf{V} \times \mathbf{W} \equiv \{(\boldsymbol{v},\boldsymbol{w}): \boldsymbol{v} \in \mathbf{V},\ \boldsymbol{w} \in \mathbf{W}\}\ . \tag{C3-1}$$

The **tensor product of V and W** is

$$\mathbf{V} \otimes \mathbf{W} \equiv \{T: \mathbf{V}^* \times \mathbf{W}^* \to \mathbb{R}: T \text{ is bilinear}\}\ . \tag{C3-2}$$

Bilinear means that the function T is linear in each component, separately:

$$T(\alpha\, \boldsymbol{v} + \beta\, \boldsymbol{u},\, \boldsymbol{w}) = \alpha\, T(\boldsymbol{v},\, \boldsymbol{w}) + \beta\, T(\boldsymbol{u},\, \boldsymbol{w}) \tag{C3-3}$$

$$T(\boldsymbol{v},\, \alpha\, \boldsymbol{w} + \beta\, \boldsymbol{x}) = \alpha\, T(\boldsymbol{v},\, \boldsymbol{w}) + \beta\, T(\boldsymbol{v},\, \boldsymbol{x}) \tag{C3-4}$$

Observe that the tensor product $\mathbf{V} \otimes \mathbf{W}$ is a function space.

Theorem C3.1 The tensor product space $\mathbf{V} \otimes \mathbf{W}$ is a vector space.

Proof. From Theorem C2.1, we know that function spaces are vector spaces when using the natural definitions for the zero function, the additive inverse, addition, and scalar multiplication. It only remains to confirm that 0, $-T$, $S+T$, and αT are also bilinear functions, and this is straight-forward. ■

Definition The functions that are members of a tensor space are known as **tensors**.

In this definition of tensor product, **V** and **W** can be any vector spaces, and that includes the duals of vector spaces. Thus, the definition encompasses $\mathbf{V}\otimes\mathbf{W}^*$, $\mathbf{V}^*\otimes\mathbf{W}$, and $\mathbf{V}^*\otimes\mathbf{W}^*$. For example, the tensor product definition for $\mathbf{V}^*$ and $\mathbf{W}^*$ is

$$\mathbf{V}^* \otimes \mathbf{W}^* \equiv \{T: \mathbf{V} \times \mathbf{W} \to \mathbb{R}: T \text{ is bilinear}\}$$

because we identify **V** with $\mathbf{V}^{**}$ and **W** with $\mathbf{W}^{**}$.

In general relativity this definition is usually applied to tangent and cotangent spaces. For example, the tensor product of $\mathbf{T}_P$ and $\mathbf{T}_P^*$ is

$$\mathbf{T}_P \otimes \mathbf{T}_P^* = \{T: \mathbf{T}_P^* \times \mathbf{T}_P \to \mathbb{R}: T \text{ is bilinear}\}.$$

The tensor definition also encompasses $\mathbf{T}_P \otimes \mathbf{T}_P$, $\mathbf{T}_P^* \otimes \mathbf{T}_P$, $\mathbf{T}_P^* \otimes \mathbf{T}_P^*$, and even $\mathbf{T}_P \otimes \mathbf{T}_Q$ where Q is a different point of M.

In addition to defining tensor products of vector *spaces*, we also define tensor products of individual *vectors*. We keep in mind that when a vector is a function, it is not bolded. However, a vector as an argument of a function is a traditional vector and *is* bolded.

Definition   The **tensor product of a tangent basis vector with a cotangent basis vector** is the function

$$\boxed{e_a \otimes e^b : \mathbf{T}_P{}^* \times \mathbf{T}_P \to \mathbb{R} : e_a \otimes e^b\left(\mu_d\, \mathbf{e}^d,\ \lambda^c\, \mathbf{e}_c\right) \equiv \mu_a\, \lambda^b}\,. \qquad \text{(C3-5)}$$

Notice that a tangent vector operates on a cotangent vector in the first coordinate, and a cotangent vector operates on a tangent vector in the second coordinate. This shows that the tensor symbol ⊗ serves as a bookkeeper, preventing intermingling of the first and second coordinates. It essentially allows us to perform independent operations in two different vector spaces simultaneously. The only slight intermingling that can occur is that scalars can be brought out from one coordinate and then put back into the other coordinate: $T(\alpha \boldsymbol{v},\boldsymbol{w}) = \alpha T(\boldsymbol{v},\boldsymbol{w}) = T(\boldsymbol{v},\alpha \boldsymbol{w})$.

We sometimes denote the basis tensor as

$$\boxed{e_a^b \equiv e_a \otimes e^b}\,.$$

From equation (C2-2), we observe that

$$\boxed{e_a^b\left(\mathbf{e}^d,\ \mathbf{e}_c\right) = \delta_a^d\, \delta_c^b}\,. \qquad \text{(C3-6)}$$

It is a simple matter to show that $e_a \otimes e^b$ is bilinear.

Theorem C3.2 $\boldsymbol{\mathcal{B}} \equiv \{e_a \otimes e^b\}$ is a basis for $\boldsymbol{T}_P \otimes \boldsymbol{T}_P{}^*$.

Proof. Let $T \in \mathbf{T}_P \otimes \mathbf{T}_P{}^*$. We must show that T is a linear combination of terms $e_a \otimes e^b$. Let $\boldsymbol{w}^* = \mu_a \mathbf{e}^a \in \mathbf{T}_P{}^*$, $\boldsymbol{v} = \lambda^b \mathbf{e}_b \in \mathbf{T}_P$, and $\tau_b^a = T(\mathbf{e}^a,\ \mathbf{e}_b)$. Then T: $\mathbf{T}_P{}^* \times \mathbf{T}_P \to \mathbb{R}$ and

$$T(\boldsymbol{w}^*,\boldsymbol{v}) = T(\mu_a \mathbf{e}^a, \lambda^b \mathbf{e}_b) = \mu_a\, \lambda^b\ T(\mathbf{e}^a,\ \mathbf{e}_b) = \mu_a\, \lambda^b\ \tau_b^a = \tau_b^a\ \mathrm{e}_a \otimes \mathrm{e}^b(\mu_d\, \mathbf{e}^d,\ \lambda^c\, \mathbf{e}_c)$$
$$= \tau_b^a\ \mathrm{e}_a \otimes \mathrm{e}^b(\boldsymbol{w}^*,\boldsymbol{v}).$$

Since this holds for all vectors $\boldsymbol{w}^*$ and $\boldsymbol{v}$, then

$$T = \tau_b^a\ \mathrm{e}_a \otimes \mathrm{e}^b \qquad \text{(C3-7)}$$

■

Observe that T is an outer product; it includes terms having $\mathrm{e}_a \otimes \mathrm{e}^b$ for every combination of a and b.

Also, note that T is sum of terms $\tau^a_b\, e_a \otimes e^b$ and cannot in general be expressed as a singleton tensor product, like $k\, e_c \otimes e^d$. We conclude that even though $\mathcal{B}$ consists only of singleton tensor products ($e_a \otimes e^b$), it generates a vector space that is richer than just a space of singleton products.

Finally, where are the familiar tensors like $\tau^{a\,b}$, $\tau_{a\,b}$, and $\tau^a_{b\,c}$? Ignoring Einstein summation notation for a moment, the answer is that in general relativity, a singleton tensor $T = \lambda^a e_a \otimes \mu_b e^b = \lambda^a \mu_b\, e_a \otimes e^b$ is expressed as $T \overset{(C3-7)}{=} \tau^a_b \equiv \lambda^a \mu_b$ where the bases $e_a \otimes e^b$ are ignored. Thus, the tensor τ^a_b that we are familiar with is actually a function, an element of a function vector space. Equation (C3-7) is its definition. The other familiar tensor expressions are also functions, generalizations of equation (C3-7), developed below.

Definition Members of $\boldsymbol{T}_\mathrm{P} \otimes \boldsymbol{T}_\mathrm{P}^*$ that can be expressed as a single tensor product are called **decomposable.**

Corollary dim ($\boldsymbol{T}_\mathrm{P} \otimes \boldsymbol{T}_\mathrm{P}^*$) = (dim $\boldsymbol{T}_\mathrm{P}$) (dim $\boldsymbol{T}_\mathrm{P}^*$) (C3-8)

The tensor product definition can be extended to include more than just two vector spaces. As before, this definition represents tensor products of vectors, covectors, and a mixture of vectors and covectors.

Definition Let **T** be a vector space and **T*** its dual space. The **tensor product of *k* vector spaces T with ℓ covector spaces T*** is

$$\underbrace{\mathbf{T} \otimes \cdots \otimes \mathbf{T}}_{k\text{-times}} \otimes \underbrace{\mathbf{T}^* \otimes \cdots \otimes \mathbf{T}^*}_{\ell\text{-times}} = \{T\colon \underbrace{\mathbf{T}^* \times \cdots \times \mathbf{T}^*}_{k\text{-times}} \times \underbrace{\mathbf{T} \times \cdots \times \mathbf{T}}_{\ell\text{-times}} \to \mathbb{R}\colon T \text{ is multilinear}\} \quad \text{(C3-9)}$$

Definition

$$\boxed{e^{b_1 \cdots b_\ell}_{a_1 \cdots a_k} \equiv e_{a_1} \otimes \cdots \otimes e_{a_k} \otimes e^{b_1} \otimes \cdots \otimes e^{b_\ell}} \quad \text{(C3-10)}$$

The collection of objects $e^{b_1 \cdots b_\ell}_{a_1 \cdots a_k}$ is a basis, so the tensor (function) T is defined as a bilinear sum of terms:

$$\boxed{T = T^{a_1 \cdots a_k}_{b_1 \cdots b_\ell} \, e^{b_1 \cdots b_\ell}_{a_1 \cdots a_k}} \tag{C3-11}$$

where

$$\boxed{e^{b_1 \cdots b_\ell}_{a_1 \cdots a_k}\left(\mathbf{e}^{d_1}, \cdots, \mathbf{e}^{d_k}, \mathbf{e}_{c_1}, \cdots, \mathbf{e}_{c_\ell}\right) = \delta^{d_1}_{a_1} \cdots \delta^{d_k}_{a_k} \, \delta^{b_1}_{c_1} \cdots \delta^{b_\ell}_{c_\ell}}. \tag{C3-12}$$

Tensors of the form $T^{a_1 \cdots a_k}$ are called **rank (*k*, 0) contravariant tensors**, $T_{b_1 \cdots b_\ell}$ are called **rank (0, ℓ) covariant tensors**, and $T^{a_1 \cdots a_k}_{b_1 \cdots b_\ell}$ are called **rank (*k*, ℓ) mixed tensors**. Rank (1,0) tensors are contravariant vectors and rank (0,1) tensors are covariant vectors. We also extend this definition to include **rank (0,0) tensors**, defined to be scalars. This is consistent because scalars are defined as functions on the coordinates of P. That is,

$\varphi : \mathrm{M} \to \mathbb{R} : \varphi(x^a) \equiv \varphi(x^1, \cdots, x^n) = k$, where k is a vector-space scalar.

We note that the tensor product definition above could have been made for k distinct vector spaces $\mathbf{T}^k$ and ℓ distinct (and unrelated) covector spaces $\mathbf{T}_\ell{}^*$, but the notation would have then gotten rather messy.