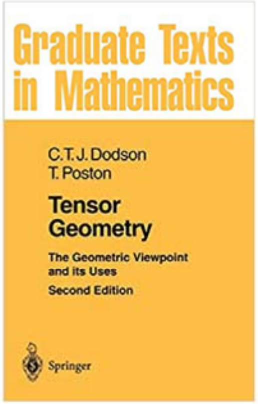
Differential Geometry

Chapter IV. Metric Vector Spaces

IV.4. Diagonalizing Symmetric Operators —Proofs of Theorems



Lemma IV.4.02

Lemma IV.4.02. If $\mathbf{x}$ is a maximal vector of a symmetric operator $\mathbf{A}$ on an inner product space $(X, \mathbf{G})$ then $\mathbf{x}$ is an eigenvector of the operator $\mathbf{A}^2$, belonging to the eigenvalue $\|\mathbf{A}\|^2$.

Proof. We have

$$\begin{aligned}
\|\mathbf{A}\|^2 &= \|\mathbf{A}\mathbf{x}\|^2 \text{ since } \mathbf{x} \text{ is a maximal vector} \\
&= \mathbf{A}\mathbf{x}\cdot\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{A}\mathbf{x}\cdot\mathbf{x} \text{ by definition of transpose} \\
&= \mathbf{A}^2\mathbf{x}\cdot\mathbf{x} \text{ since } \mathbf{A} \text{ is symmetric, } \mathbf{A}^T = \mathbf{A} \\
&\le \|\mathbf{A}^x\mathbf{x}\|\|\mathbf{x}\| \text{ by Schwarz's Inequality (Lemma IV.1.07)} \\
&= \|\mathbf{A}^2\mathbf{x}\| \text{ since } \|\mathbf{x}\| = 1 \\
&= \|\mathbf{A}(\mathbf{A}\mathbf{x})\| \le \|\mathbf{A}\|\|\mathbf{A}\mathbf{x}\| \text{ by Exercise IV.4.1} \\
&\le \|\mathbf{A}\|(\|\mathbf{A}\|\|\mathbf{x}\|) \text{ by Exercise IV.4.1} \\
&= \|\mathbf{A}\|^2 \text{ since } \|\mathbf{x}\| = 1.
\end{aligned}$$

Lemma IV.4.02 (continued)

Lemma IV.4.02. If $\mathbf{x}$ is a maximal vector of a symmetric operator $\mathbf{A}$ on an inner product space $(X, \mathbf{G})$ then $\mathbf{x}$ is an eigenvector of the operator $\mathbf{A}^2$, belonging to the eigenvalue $\|\mathbf{A}\|^2$.

Proof (continued). But the all inequalities must in fact be equalities. This means $\mathbf{A}^2\mathbf{x}\cdot\mathbf{x} = \|\mathbf{A}^2\mathbf{x}\|\|\mathbf{x}\|$ so that we have equality in Schwarz's Inequality and hence (by Lemma IV.1.07) we have $\mathbf{A}^2\mathbf{x} = \mathbf{x}1$ for some $a \in \mathbb{R}$. So $\mathbf{x}$ is an eigenvector of $\mathbf{A}^2$ with eigenvalue a where, by the equalities above, $a = a(\mathbf{x}\cdot\mathbf{x}) = (\mathbf{x}a)\cdot\mathbf{x} = (\mathbf{A}^2\mathbf{x})\cdot\mathbf{x} = \|\mathbf{A}\|^2$, as claimed. □

Lemma IV.4.03

Lemma IV.4.03. A symmetric operator $\mathbf{A}$ on a finite dimensional inner product space has an eigenvector belonging to an eigenvalue $+\|\mathbf{A}\|$ on $-\|\mathbf{A}\|$.

Proof. Let $\mathbf{x}$ be a maximal vector of $\mathbf{A}$ (which exists since the inner product space if finite dimensional). Then by Lemma IV.4.02, $\mathbf{x}$ is an eigenvector of $\mathbf{A}^2$ with eigenvalue $\|\mathbf{A}\|^2$, so $\mathbf{A}^2\mathbf{x} = \mathbf{x}\|\mathbf{A}\|^2$ and $(\mathbf{A} - \|\mathbf{A}\|^2\mathbf{I})\mathbf{x} = \mathbf{0}$. Hence $(\mathbf{A} + \|\mathbf{A}\|\mathbf{I})(\mathbf{A} - \|\mathbf{A}\|\mathbf{I})\mathbf{x} = \mathbf{0}$. So either $(\mathbf{A} - \|\mathbf{A}|\mathbf{I}) = \mathbf{0}$, in which case $\mathbf{x}$ is an eigenvector of $\mathbf{A}$ with eigenvalue $\|\mathbf{A}\|$, or $(\mathbf{A} - \|\mathbf{A}\|\mathbf{I})\mathbf{x}$ is an eigenvector of $\mathbf{A}$ with eigenvalue $-\|\mathbf{A}\|$, as claimed. □

Theorem IV.4.04

Lemma IV.4.04. If $\mathbf{X}$ is an eigenvector of a self-adjoint linear operator $\mathbf{A}$ on a metric vector space then $\mathbf{x} \cdot \mathbf{y} = 0$ implies $\mathbf{x} \cdot \mathbf{A}\mathbf{y} = 0$. That is, $\mathbf{A}(\mathbf{x}^\perp) \subseteq \mathbf{x}^\perp$ and so the map $\mathbf{y} \mapsto \mathbf{A}\mathbf{y}$ is an operator on $\mathbf{x}^\perp$, called the operator on $\mathbf{x}^\perp$ *induced* by $\mathbf{A}$.

Proof. Let λ be the eigenvalue corresponding to eigenvector $\mathbf{x}$. Then $\mathbf{x} \cdot \mathbf{y} = 0$ implies $\lambda(\mathbf{x} \cdot \mathbf{y}) = 0$, or $(\mathbf{x}\lambda) \cdot \mathbf{y} = 0$ or $(\mathbf{A}\mathbf{x}) \cdot \mathbf{y} = 0$. Therefore $\mathbf{x} \cdot \mathbf{A}^T\mathbf{y} = 0$ by the definition of $\mathbf{A}^T$ and, since $\mathbf{A}$ is hypothesized to be symmetric, $\mathbf{x} \cdot (\mathbf{A}\mathbf{y}) = 0$, as claimed. □

Theorem IV.4.05

Theorem IV.4.05. If $\mathbf{A}$ is a symmetric linear operator on a finite dimensional inner product space X, then X has an orthonormal basis of eigenvectors of $\mathbf{A}$.

Proof. Let $\dim(X) = n$. By Lemma IV.4.03 there is some eigenvector $\mathbf{x}_1$ of $\mathbf{A}$ corresponding to real eigenvalue $\pm\|\mathbf{A}\|$. Set $\mathbf{b} = \mathbf{x}_1/\|\mathbf{x}_1\|$ so that $\mathbf{b}$ is a unit eigenvector. Since $\mathbf{A}$ is symmetric by hypothesis, then it is self-adjoint by the definition of symmetric. By Lemma IV.4.04 there is a linear operator $\mathbf{A}' : \mathbf{x}^\perp \to \mathbf{b}^\perp$ defined as $\mathbf{A}'(\mathbf{x}) = \mathbf{A}\mathbf{x}$. Since $\mathbf{A}$ is symmetric on X then $\mathbf{A}'$ (which is just $\mathbf{A}$ restricted to $\mathbf{B}^\perp$) is symmetric on $\mathbf{b}^\perp$ (with respect to the inner product restricted to $\mathbf{b}^\perp$). By Lemma IV.4.03 applied to inner product space $\mathbf{b}^\perp$ and linear operator $\mathbf{A}'$ there is an eigenvector $\mathbf{b}_2$ (a unit vector, without loss of generality, since eigenvectors are by definition nonzero) of $\mathbf{A}'$ corresponding to real eigenvalue $\pm\|\mathbf{A}'\|$. Since $\mathbf{b}_2 \in \mathbf{b}^\perp$ then $\mathbf{A}'\mathbf{b}_2 = \mathbf{A}\mathbf{b}_2$ and so $\mathbf{b}_2$ is also an eigenvector fo $\mathbf{A}$ with the same eigenvalue.

Theorem IV.4.05 (continued)

Theorem IV.4.05. If $\mathbf{A}$ is a symmetric linear operator on a finite dimensional inner product space X, then X has an orthonormal basis of eigenvectors of $\mathbf{A}$.

Proof (continued). So we have an orthonormal set $\{\mathbf{b}_1, \mathbf{b}_2\}$ of eigenvectors of $\mathbf{A}$. Let $S_1 = (\text{span}(\mathbf{b}_1))^\perp$ and $S_2 = (\text{span}(\mathbf{b}_1, \mathbf{b}_2))^\perp$, so that, by Corollary IV.2.05, $\dim(S_1) = n - 1$ and $\dim(S_2) = n - 2$. We can now inductively find eigenvectors $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n$ of $\mathbf{A}$ with real eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, respectively (and subspaces $S_1, S_2, \ldots, S_n$ of X such that $\dim(S_k) = n - k$), just as we did for $\mathbf{b}_1$ and $\mathbf{b}_2$. Since $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ is an orthonormal set of eigenvectors in X where $\dim(X) = n$, then $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ is an orthonormal basis of eigenvectors of $\mathbf{A}$ for X. □

Corollary IV.4.07

Corollary IV.4.07. If $\mathbf{A}$ is a symmetric linear operator on a finite dimensional inner product space with orthonormal basis β and if μ is a root of multiplicity m of the characteristic equation $\det([\mathbf{A} - \lambda\mathbf{I}]_\beta^\beta) = 0$ then the eigenspace belonging to μ has dimension m.

Proof. By Corollary IV.4.06, $[\mathbf{A}]_\beta^\beta$ is a diagonal matrix with diagonal entries $\lambda_1, \lambda_2, \ldots, \lambda_n$ (the not-necessarily-distinct eigenvalues of $\mathbf{A}$). Then by the Fundamental Theorem of Algebra

$$\det([\mathbf{A} - \lambda\mathbf{I}]_\beta^\beta) = \det([\mathbf{A}]_\beta^\beta - \lambda[\mathbf{I}]_\beta^\beta) = (\lambda_1 - \lambda)(\lambda_2 - \lambda)\cdots(\lambda_n - \lambda).$$

Since μ is a root of the characteristic equation of multiplicity m, then say $\lambda_{j_1}\lambda_{j_2} = \cdots = \lambda_{j_m} = \mu$. Then the eigenspace of μ is spanned by $\mathbf{b}_{j_1}, \mathbf{b}_{j_2}, \ldots, \mathbf{b}_{j_m}$ and since these vectors are orthonormal then they are linearly independent and hence a basis for the eigenspace of μ. So the dimensional of this eigenspace is m, as claimed. □

Corollary IV.4.09

Corollary IV.4.09. In an inner product space $(X, \mathbf{G})$ for any symmetric bilinear form $\mathbf{h}$ on X we can find an orthonormal basis $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n$ for X such that $\mathbf{h}(\mathbf{b}_i, \mathbf{b}_j) = 0$ for $i \neq j$.

Proof. Let $\mathbf{x} \in X$ and define $\mathbf{h_x} : X \to \mathbb{R}$ as $\mathbf{h_x}(\mathbf{y}) = \mathbf{h}(\mathbf{x}, \mathbf{y})$ for all $\mathbf{y} \in X$ (so $\mathbf{h_x} \in X^*$). Next define $\mathbf{A_h} : X \to X$ as $\mathbf{A_h}(\mathbf{x}) = \mathbf{G}_\uparrow(\mathbf{h_x})$ for all $\mathbf{x} \in X$. Then we have for all $\mathbf{y} \in X$,

$$\begin{aligned}(\mathbf{A_h x}) \cdot \mathbf{y} &= \mathbf{G}_\uparrow(\mathbf{h_x}) \cdot \mathbf{y} = \mathbf{G}(\mathbf{G}_\uparrow(\mathbf{h_x}), \mathbf{y}) \\ &= \mathbf{h_x}(\mathbf{y}) \text{ since } \mathbf{G}_\uparrow(\mathbf{x}^* \text{ where } \mathbf{x}^*(\mathbf{y}) = \mathbf{G}(\mathbf{x}, \mathbf{y}) = \mathbf{x} \cdot \mathbf{y} \text{ for all} \\ &\quad \mathbf{y} \in X\text{; see Theorem IV.1.09 and Note IV.1.A (here } \mathbf{x}^* = \mathbf{h_x}) \\ &= \mathbf{h}(\mathbf{x}, \mathbf{y}) = \mathbf{h}(\mathbf{y}, \mathbf{x}) \text{ since } \mathbf{h} \text{ is symmetric by hypothesis} \\ &= \mathbf{y}(\mathbf{x}) = (\mathbf{A_h y}) \cdot \mathbf{x} \text{ as just established} \\ &\quad \text{(with } \mathbf{x} \text{ and } \mathbf{y} \text{ interchanged)} \\ &= \mathbf{x} \cdot (\mathbf{A_h y}) \text{ since } \mathbf{G} \text{ (and so dot product) is symmetric because} \\ &\quad (X, \mathbf{G}) \text{ is an inner product space.}\end{aligned}$$

Corollary IV.4.09 (continued)

Corollary IV.4.09. In an inner product space $(X, \mathbf{G})$ for any symmetric bilinear form $\mathbf{h}$ on X we can find an orthonormal basis $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n$ for X such that $\mathbf{h}(\mathbf{b}_i, \mathbf{b}_j) = 0$ for $i \neq j$.

Proof (continued). Since $(\mathbf{A_h x}) \cdot \mathbf{y} = \mathbf{x} \cdot (\mathbf{A_h y})$ for all $\mathbf{x}, \mathbf{y} \in X$ then $\mathbf{A_h}$ is self-adjoint and since $(X, \mathbf{G})$ is an inner product space then, by definition, $\mathbf{A_h}$ is a symmetric linear operator. By Theorem IV.4.05, there is an orthonormal basis $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ for X of eigenvectors of $\mathbf{A_h}$; say $\mathbf{A_h b}_i = \lambda_i \mathbf{b}_i$ for each i. Then for this orthonormal basis,

$$\begin{aligned}\mathbf{h}(\mathbf{b}_i, \mathbf{b}_j) &= (\mathbf{A_h b}_i) \cdot \mathbf{b}_j \text{ since } (\mathbf{A_h x}) \cdot \mathbf{y} = \mathbf{h}(\mathbf{x}, \mathbf{y}), \text{ as established above} \\ &= (\lambda_i \mathbf{b}_i) \cdot \mathbf{b}_j = \lambda_i (\mathbf{b}_i \cdot \mathbf{b}_j) = 0 \text{ if } i \neq j,\end{aligned}$$

as claimed. □

Lemma IV.4.11

Lemma IV.4.11. If $\mathbf{h}$ is isotropic, then $\mathbf{h} = \lambda \mathbf{G}$ for some $\lambda \in \mathbb{R}$ and $\mathbf{A_h} = \lambda \mathbf{I}$.

Proof. If $\mathbf{h}$ is isotropic then there is $\lambda \in \mathbb{R}$ (λ is real by Corollary IV.2.08, since $\mathbf{A_h}$ is symmetric as established in the proof of Theorem IV.4.09) such that $\mathbf{A_h x} = \lambda \mathbf{x}$ for all $\mathbf{x} \in X$. Then $\mathbf{A_h x} - \lambda \mathbf{x} = (\mathbf{A_h} - \lambda \mathbf{I})\mathbf{x} = \mathbf{0}$ for all $\mathbf{x} \in X$; that is, $\mathbf{A_h} = \lambda \mathbf{I} = \mathbf{0}$ (the $\mathbf{0}$ operator) and $\mathbf{A_h} = \lambda \mathbf{I}$, as claimed.

As shown in the proof of Theorem IV.4.09, $\mathbf{h}(\mathbf{x}, \mathbf{y}) = (\mathbf{A_h x}) \cdot \mathbf{y}$ for all $\mathbf{x}, \mathbf{y} \in X$, so $\mathbf{h}(\mathbf{x}, \mathbf{y}) = \lambda \mathbf{x}) \cdot \mathbf{y} = \lambda(\mathbf{x}, \mathbf{y}) = \lambda \mathbf{G}(\mathbf{x}, \mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in X$ and hence $\mathbf{h} = \lambda \mathbf{G}$, as claimed. □

Lemma IV.4.13

Lemma IV.4.13. If a self-adjoint linear operator $\mathbf{A}$ on Lorentz space $\mathbb{L}^4$ has a timelike eigenvector $\mathbf{v}$ (i.e., $\mathbf{v} \cdot \mathbf{v} > 0$), then $\mathbb{L}^4$ has an orthonormal basis of eigenvectors of $\mathbf{A}$.

Proof. Since $\mathbb{L}^4$ is a metric vector space and $\mathbf{A}$ is self-adjoint, then by Lemma IV.2.04 the restriction of $\mathbf{A}$ to $\mathbf{v}^\perp$ is a linear operator on $\mathbf{v}^\perp$.

We claim that the metric tensor $\mathbf{G}$ on span$(\mathbf{x})$ is non-degenerate. If for $\mathbf{x} \in \text{span}(\mathbf{v})$ we have $\mathbf{G}(\mathbf{x}, \mathbf{v}a) = 0$ for all $\mathbf{v}a \in \text{span}(\mathbf{v})$, or equivalently for all $a \in \mathbb{R}$, implies $\mathbf{G}(\mathbf{v}b, \mathbf{v}a) = 0$ for all $a \in \mathbb{R}$ where $\mathbf{x} = \mathbf{v}b$. So $ab\mathbf{G}(\mathbf{v}, \mathbf{v}) = 0$ for all $a \in \mathbb{R}$. Since $\mathbf{v}$ is timelike then $\mathbf{G}(\mathbf{v}, \mathbf{v}) = \mathbf{v} \cdot \mathbf{v} > 0$, so we must have $b = 0$ and $\mathbf{x} = \mathbf{v}b = \mathbf{0}$. That is, $\mathbf{G}$ is non-degenerate on span$(\mathbf{v})$. So by Corollary IV.2.06, $\mathbf{G}$ is non-degenerate on $\mathbf{v}^\perp$.

With $\beta = \{(1,0,0,0), (0,1,0,0), (0,0,1,0), (0,0,0,1)\}$ as an orthonormal basis of $\mathbb{L}^4$ we have $k = 1$, in the notation of Theorem IV.2.08, since $(1,0,0,0) \cdot (1,0,0,0) = 1$, $(0,1,0,0) \cdot (0,1,0,0) = -1$, $(0,0,1,0) \cdot (0,0,1,0) = -1$, and $(0,0,0,1) \cdot (0,0,0,1) = -1$.

Lemma IV.4.13 (continued)

Lemma IV.4.13. If a self-adjoint linear operator **A** on Lorentz space $\mathbb{L}^4$ has a timelike eigenvector **v** (i.e., $\mathbf{v} \cdot \mathbf{v} > 0$), then $\mathbb{L}^4$ has an orthonormal basis of eigenvectors of **A**.

Proof. So any orthonormal basis of $\mathbb{L}^4$ will have $k = 1$ by Theorem IV.3.08; that is, any orthonormal basis of $\mathbb{L}^4$ will consist of 1 timelike vector and 3 spacelike vectors. As shown in the proof of Theorem IV.3.08 (with $W = \text{span}(\mathbf{v})$ and $N = \mathbf{b}^\perp$ in the notation of the proof) we have that **G** is negative definite on $\mathbf{v}^\perp$; that is, **G** is an inner product on $\mathbf{v}^\perp$ (this is the first time we have used a negative definite inner product). So applying Theorem IV.4.05 to inner product space $(\mathbf{v}^\perp, \mathbf{G}|_{\mathbf{v}^\perp})$ and symmetric linear operator $\mathbf{A}|_{\mathbf{v}^\perp}$ (recall that a symmetric operator is a self-adjoint operator on an inner product space), there are 3 spacelike orthonormal eigenvectors of $\mathbf{A}|_{\mathbf{v}^\perp}$ (and so eigenvectors of **A** since **A** and $\mathbf{A}|_{\mathbf{v}^\perp}$ agree on $\mathbf{v}^\perp$). So these 3 spacelike vectors, along with timelike vector $\mathbf{v}/\|\mathbf{v}\|$, form an orthonormal basis of $\mathbb{L}^4$ consisting of eigenvectors of **A**. □