5

Change of Basis

In many applications, we may need to switch between two or more different bases for a vector space. So it would be helpful to have formulas for converting the components of a vector with respect to one basis into the corresponding components of the vector (or matrix of the operator) with respect to the other basis. The theory and tools for quickly determining these "change of basis formulas" will be developed in these notes.

5.1 Unitary and Orthogonal Matrices

Definitions

Unitary and orthogonal matrices will naturally arise in change of basis formulas. They are defined as follows:

$$\mathbf{A} \text{ is a } \textit{unitary matrix} \quad \Longleftrightarrow \quad \mathbf{A} \text{ is an invertible matrix with } \mathbf{A}^{-1} = \mathbf{A}^{\dagger} \quad ,$$

and

$$\begin{aligned} \mathbf{A} \text{ is an } \textit{orthogonal matrix} \quad &\Longleftrightarrow \quad \mathbf{A} \text{ is an invertible real matrix with } \mathbf{A}^{-1} = \mathbf{A}^{\mathsf{T}} \\ &\Longleftrightarrow \quad \mathbf{A} \text{ is a real unitary matrix} \quad . \end{aligned}$$

Because an orthogonal matrix is simply a unitary matrix with real-valued entries, we will mainly consider unitary matrices (keeping in mind that anything derived for unitary matrices will also hold for orthogonal matrices after replacing $\mathbf{A}^{\dagger}$ with $\mathbf{A}^{\mathsf{T}}$).

The basic test for determining if a square matrix is unitary is to simply compute $\mathbf{A}^{\dagger}$ and see if it is the inverse of $\mathbf{A}$; that is, see if

$$\mathbf{A}\mathbf{A}^{\dagger} = \mathbf{I} \quad .$$

!►*Example 5.1:* *Let*

$$\mathbf{A} = \begin{bmatrix} \frac{3}{5} & \frac{4}{5}i \\ \frac{4}{5}i & \frac{3}{5} \end{bmatrix}$$

Then

$$\mathbf{A}^{\dagger} = \begin{bmatrix} \frac{3}{5} & -\frac{4}{5}i \\ -\frac{4}{5}i & \frac{3}{5} \end{bmatrix}$$

and

$$\mathbf{A}\mathbf{A}^{\dagger} = \begin{bmatrix} \frac{3}{5} & \frac{4}{5}i \\ \frac{4}{5}i & \frac{3}{5} \end{bmatrix} \begin{bmatrix} \frac{3}{5} & -\frac{4}{5}i \\ -\frac{4}{5}i & \frac{3}{5} \end{bmatrix} = \cdots = \mathbf{I} \quad .$$

So $\mathbf{A}$ *is unitary.*

Obviously, for a matrix to be unitary, it must be square. It should also be fairly clear that, if $\mathbf{A}$ is a unitary (or orthogonal) matrix, then so are $\mathbf{A}^*$, $\mathbf{A}^{\mathsf{T}}$, $\mathbf{A}^{\dagger}$ and $\mathbf{A}^{-1}$.

?►*Exercise 5.1:* *Prove that, if* $\mathbf{A}$ *is a unitary (or orthogonal) matrix, then so are* $\mathbf{A}^*$, $\mathbf{A}^{\mathsf{T}}$, $\mathbf{A}^{\dagger}$ *and* $\mathbf{A}^{-1}$.

The term "unitary" comes from the value of the determinant. To see this, first observe that, if $\mathbf{A}$ is unitary then

$$\mathbf{I} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{A}\mathbf{A}^{\dagger} \quad .$$

Using this and already discussed properties of determinants, we have

$$1 = \det(\mathbf{I}) = \det(\mathbf{A}\mathbf{A}^{\dagger}) = \det(\mathbf{A})\,\det(\mathbf{A}^{\dagger}) = \det(\mathbf{A})\,\det(\mathbf{A})^* = |\det(\mathbf{A})|^2 \quad .$$

Thus,

$$\mathbf{A} \text{ is unitary} \quad\Longrightarrow\quad |\det \mathbf{A}| = 1 \quad .$$

And since there are only two real numbers which have magnitude 1, it immediately follows that

$$\mathbf{A} \text{ is orthogonal} \quad\Longrightarrow\quad \det \mathbf{A} = \pm 1 \quad .$$

An immediate consequence of this is that if the absolute value of the determinant of a matrix is not 1, then that matrix cannot be unitary.

Why the term "orthogonal" is appropriate will become obvious later.

Rows and Columns of Unitary Matrices

Let

$$\mathbf{U} = \begin{bmatrix} u_{11} & u_{12} & u_{13} & \cdots & u_{1N} \\ u_{21} & u_{22} & u_{23} & \cdots & u_{2N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ u_{N1} & u_{N2} & u_{N3} & \cdots & u_{NN} \end{bmatrix}$$

be a square matrix. By definition, then,

$$\mathbf{U}^{\dagger} = \begin{bmatrix} {u_{11}}^* & {u_{21}}^* & \cdots & {u_{N1}}^* \\ {u_{12}}^* & {u_{22}}^* & \cdots & {u_{N2}}^* \\ {u_{13}}^* & {u_{23}}^* & \cdots & {u_{N3}}^* \\ \vdots & \vdots & \ddots & \vdots \\ {u_{1N}}^* & {u_{2N}}^* & \cdots & {u_{NN}}^* \end{bmatrix} \quad .$$

More concisely,

$$[\mathbf{U}]_{jk} = u_{jk} \qquad \text{and} \qquad \left[\mathbf{U}^\dagger\right]_{jk} = \left([\mathbf{U}]_{kj}\right)^* = u_{kj}{}^* \quad .$$

Now observe:

$$\begin{aligned}
& & \mathbf{U} &\text{ is unitary} \\
&\iff & \mathbf{U}^\dagger &= \mathbf{U}^{-1} \\
&\iff & \mathbf{U}^\dagger\mathbf{U} &= \mathbf{I} \\
&\iff & \left[\mathbf{U}^\dagger\mathbf{U}\right]_{jk} &= [\mathbf{I}]_{jk} \\
&\iff & \sum_m \left[\mathbf{U}^\dagger\right]_{jm} [\mathbf{U}]_{mk} &= \delta_{jk} \qquad \text{“for all } (j,k)\text{”} \\
&\iff & \sum_m u_{mj}{}^* u_{mk} &= \delta_{jk} \qquad \text{“for all } (j,k)\text{”}
\end{aligned}$$

The righthand side of the last equation is simply the formula for computing the standard matrix inner product of the column matrices

$$\begin{bmatrix} u_{1j} \\ u_{2j} \\ \vdots \\ u_{Nj} \end{bmatrix} \qquad \text{and} \qquad \begin{bmatrix} u_{1k} \\ u_{2k} \\ \vdots \\ u_{Nk} \end{bmatrix} \quad ,$$

and the last line tells us that

$$\text{this inner product} = \begin{cases} 1 & \text{if} \quad j = k \\ 0 & \text{if} \quad j \neq k \end{cases} \quad .$$

In other words, that line states that

$$\left\{ \begin{bmatrix} u_{11} \\ u_{21} \\ \vdots \\ u_{N1} \end{bmatrix} , \begin{bmatrix} u_{12} \\ u_{22} \\ \vdots \\ u_{N2} \end{bmatrix} , \begin{bmatrix} u_{13} \\ u_{23} \\ \vdots \\ u_{N3} \end{bmatrix} , \ldots , \begin{bmatrix} u_{1N} \\ u_{2N} \\ \vdots \\ u_{NN} \end{bmatrix} \right\}$$

is an orthonormal set of column matrices. But these column matrices are simply the columns of our original matrix $\mathbf{U}$. Consequently, the above set of observations (starting with “$\mathbf{U}$ is unitary”) reduces to

> *A square matrix $\mathbf{U}$ is unitary if and only if its columns form an orthonormal set of column matrices (using the standard column matrix inner product).*

You can verify that a similar statement holds using the rows of $\mathbf{U}$ instead of its columns.

?►*Exercise 5.2:* *Show that a square matrix $\mathbf{U}$ is unitary if and only if its rows form an orthonormal set of row matrices (using the row matrix inner product). (Hints: Either consider how the above derivation would have changed if we had used $\mathbf{U}\mathbf{U}^\dagger$ instead of $\mathbf{U}^\dagger\mathbf{U}$, or use the fact just derived along with the fact that $\mathbf{U}$ is unitary if and only if $\mathbf{U}^\mathsf{T}$ is unitary.)*

In summary, we have just proven:

Theorem 5.1 (The Big Theorem on Unitary Matrices)
Let

$$\mathbf{U} = \begin{bmatrix} u_{11} & u_{12} & u_{13} & \cdots & u_{1N} \\ u_{21} & u_{22} & u_{23} & \cdots & u_{2N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ u_{N1} & u_{N2} & u_{N3} & \cdots & u_{NN} \end{bmatrix}$$

be a square matrix. Then the following statements are equivalent (that is, if any one statement is true, then all the statements are true).

1. $\mathbf{U}$ *is unitary.*
2. *The columns of* $\mathbf{U}$ *form an orthonormal set of column matrices (with respect to the usual matrix inner product). That is,*

$$\sum_m u_{mj}{}^* u_{mk} = \begin{cases} 1 & \text{if } j = k \\ 0 & \text{if } j \neq k \end{cases} \quad .$$

3. *The rows of* $\mathbf{U}$ *form an orthonormal set of row matrices (with respect to the usual matrix inner product). That is,*

$$\sum_m u_{jm}{}^* u_{km} = \begin{cases} 1 & \text{if } j = k \\ 0 & \text{if } j \neq k \end{cases} \quad .$$

?►*Exercise 5.3:* *What is the corresponding "Big Theorem on Orthogonal Matrices"?*

An Important Consequence and Exercise

You can now verify a result that will be important in our change of basis formulas involving orthonormal bases.

?►*Exercise 5.4:* *Let*

$$\mathbf{U} = \begin{bmatrix} u_{11} & u_{12} & u_{13} & \cdots & u_{1N} \\ u_{21} & u_{22} & u_{23} & \cdots & u_{2N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ u_{N1} & u_{N2} & u_{N3} & \cdots & u_{NN} \end{bmatrix}$$

be a square matrix, and let

$$\mathcal{S} = \{ \mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_N \} \qquad \text{and} \qquad \mathcal{B} = \{ \mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_N \}$$

be two sets of vectors (in some vector space) related by

$$\begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \ldots & \mathbf{b}_N \end{bmatrix} = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \ldots & \mathbf{e}_N \end{bmatrix} \mathbf{U} \quad .$$

(I.e.,

$$\mathbf{b}_j = \sum_k \mathbf{e}_k u_{kj} \qquad \text{for} \quad j = 1, 2, \ldots, N \quad .)$$

a: *Show that*

$$\mathcal{S} \text{ is orthonormal and } \mathbf{U} \text{ is a unitary matrix} \quad \Longrightarrow \quad \mathcal{B} \text{ is also orthonormal} \quad .$$

b: *Show that*

$$\mathcal{S} \text{ and } \mathcal{B} \text{ are both orthonormal sets} \quad \Longrightarrow \quad \mathbf{U} \text{ is a unitary matrix} \quad .$$

5.2 Change of Basis for Vector Components: The General Case

Given the tools and theory we've developed, finding and describing the "most general formulas for changing the basis of a vector space" is disgustingly easy (assuming the space is finite dimensional).

So let's assume we have a vector space $\mathcal{V}$ of finite dimension N and with an inner product $\langle \cdot | \cdot \rangle$. Let

$$\mathcal{A} = \{ \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_N \} \qquad \text{and} \qquad \mathcal{B} = \{ \mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_N \} \quad .$$

be two bases for $\mathcal{V}$, and define the corresponding four $N \times N$ matrices

$$\mathbf{M}_{\mathcal{A}\mathcal{A}} \quad , \quad \mathbf{M}_{\mathcal{A}\mathcal{B}} \quad , \quad \mathbf{M}_{\mathcal{B}\mathcal{B}} \quad \text{and} \quad \mathbf{M}_{\mathcal{B}\mathcal{A}}$$

by

$$[\mathbf{M}_{\mathcal{A}\mathcal{A}}]_{jk} = \left\langle \mathbf{a}_j \,\middle|\, \mathbf{a}_k \right\rangle \quad ,$$

$$[\mathbf{M}_{\mathcal{A}\mathcal{B}}]_{jk} = \left\langle \mathbf{a}_j \,\middle|\, \mathbf{b}_k \right\rangle \quad ,$$

$$[\mathbf{M}_{\mathcal{B}\mathcal{B}}]_{jk} = \left\langle \mathbf{b}_j \,\middle|\, \mathbf{b}_k \right\rangle \quad ,$$

and

$$[\mathbf{M}_{\mathcal{B}\mathcal{A}}]_{jk} = \left\langle \mathbf{b}_j \,\middle|\, \mathbf{a}_k \right\rangle \quad .$$

(See the pattern?) These matrices describe how the vectors in $\mathcal{A}$ and $\mathcal{B}$ are related to each other.

Two quick observations should be made about these matrices:

1. The first concerns the relation between $\mathbf{M}_{\mathcal{A}\mathcal{B}}$ and $\mathbf{M}_{\mathcal{B}\mathcal{A}}$. Observe that

$$[\mathbf{M}_{\mathcal{A}\mathcal{B}}]_{jk} = \left\langle \mathbf{a}_j \,\middle|\, \mathbf{b}_k \right\rangle = \left\langle \mathbf{b}_k \,\middle|\, \mathbf{a}_j \right\rangle^* = \left([\mathbf{M}_{\mathcal{B}\mathcal{A}}]_{kj}\right)^* = \left[{\mathbf{M}_{\mathcal{B}\mathcal{A}}}^\dagger\right]_{jk} \quad .$$

So

$$\mathbf{M}_{\mathcal{A}\mathcal{B}} = {\mathbf{M}_{\mathcal{B}\mathcal{A}}}^\dagger \quad . \tag{5.1a}$$

Likewise, of course,

$$\mathbf{M}_{\mathcal{B}\mathcal{A}} = {\mathbf{M}_{\mathcal{A}\mathcal{B}}}^\dagger \quad . \tag{5.1b}$$

2. The second observation is that $\mathbf{M}_{\mathcal{AA}}$ and $\mathbf{M}_{\mathcal{BB}}$ greatly simplify if the bases are orthonormal. If $\mathcal{A}$ is orthonormal, then

$$[\mathbf{M}_{\mathcal{AA}}]_{jk} = \langle \mathbf{a}_j \mid \mathbf{a}_k \rangle = \delta_{jk} = [\mathbf{I}]_{jk} \quad .$$

So

$$\mathbf{M}_{\mathcal{AA}} = \mathbf{I} \qquad \text{if} \quad \mathcal{A} \text{ is orthonormal} \quad . \tag{5.2a}$$

By exactly the same reasoning, it should be clear that

$$\mathbf{M}_{\mathcal{BB}} = \mathbf{I} \qquad \text{if} \quad \mathcal{B} \text{ is orthonormal} \quad . \tag{5.2b}$$

Now let $\mathbf{v}$ be any vector in $\mathcal{V}$ and, for convenience, let us denote the components of $\mathbf{v}$ with respect to $\mathcal{A}$ and $\mathcal{B}$ using α_j's and β_j's respectively,

$$|\mathbf{v}\rangle_{\mathcal{A}} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_N \end{bmatrix} \qquad \text{and} \qquad |\mathbf{v}\rangle_{\mathcal{B}} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_N \end{bmatrix} \quad .$$

Remember, this means

$$\mathbf{v} = \sum_k \alpha_k \mathbf{a}_k = \sum_k \beta_k \mathbf{b}_k \quad . \tag{5.3}$$

Our goal is to find the relations between the α_j's and the β_j's so that we can find one set given the other. One set of relations can be found by taking the inner product of $\mathbf{v}$ with each $\mathbf{a}_j$ and using equations (5.3). Doing so:

$$\langle \mathbf{a}_j \mid \mathbf{v} \rangle = \left\langle \mathbf{a}_j \,\middle|\, \sum_k \alpha_k \mathbf{a}_k \right\rangle = \left\langle \mathbf{a}_j \,\middle|\, \sum_k \beta_k \mathbf{b}_k \right\rangle$$

$$\Longrightarrow \qquad \langle \mathbf{a}_j \mid \mathbf{v} \rangle = \sum_k \alpha_k \langle \mathbf{a}_j \mid \mathbf{a}_k \rangle = \sum_k \beta_k \langle \mathbf{a}_j \mid \mathbf{b}_k \rangle$$

$$\Longrightarrow \qquad \langle \mathbf{a}_j \mid \mathbf{v} \rangle = \sum_k \langle \mathbf{a}_j \mid \mathbf{a}_k \rangle \alpha_k = \sum_k \langle \mathbf{a}_j \mid \mathbf{b}_k \rangle \beta_k$$

$$\Longrightarrow \qquad \langle \mathbf{a}_j \mid \mathbf{v} \rangle = \sum_k [\mathbf{M}_{\mathcal{AA}}]_{jk}\, \alpha_k = \sum_k [\mathbf{M}_{\mathcal{AB}}]_{jk}\, \beta_k$$

The formulas in the second equation in the last line are simply formulas for the j^{th} entry in the products of $\mathbf{M}_{\mathcal{AA}}$ and $\mathbf{M}_{\mathcal{AB}}$ with the column matrices of α_k's and β's. So that equation tells us that

$$\begin{bmatrix} \langle \mathbf{a}_1 \mid \mathbf{v} \rangle \\ \langle \mathbf{a}_2 \mid \mathbf{v} \rangle \\ \vdots \\ \langle \mathbf{a}_N \mid \mathbf{v} \rangle \end{bmatrix} = \mathbf{M}_{\mathcal{AA}} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_N \end{bmatrix} = \mathbf{M}_{\mathcal{AB}} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_N \end{bmatrix}$$

Recalling what the column matrices of α_k's and β's are, we see that this reduces to

$$\begin{bmatrix} \langle\, \mathbf{a}_1 \mid \mathbf{v}\,\rangle \\ \langle\, \mathbf{a}_2 \mid \mathbf{v}\,\rangle \\ \vdots \\ \langle\, \mathbf{a}_N \mid \mathbf{v}\,\rangle \end{bmatrix} = \mathbf{M}_{\mathcal{A}\mathcal{A}} \, |\mathbf{v}\rangle_{\mathcal{A}} = \mathbf{M}_{\mathcal{A}\mathcal{B}} \, |\mathbf{v}\rangle_{\mathcal{B}} \quad . \tag{5.4}$$

?►*Exercise 5.5 (semi-optional):* *Using the same assumptions as were used to derive (5.4), derive that*

$$\begin{bmatrix} \langle\, \mathbf{b}_1 \mid \mathbf{v}\,\rangle \\ \langle\, \mathbf{b}_2 \mid \mathbf{v}\,\rangle \\ \vdots \\ \langle\, \mathbf{b}_N \mid \mathbf{v}\,\rangle \end{bmatrix} = \mathbf{M}_{\mathcal{B}\mathcal{A}} \, |\mathbf{v}\rangle_{\mathcal{A}} = \mathbf{M}_{\mathcal{B}\mathcal{B}} \, |\mathbf{v}\rangle_{\mathcal{B}} \quad . \tag{5.5}$$

Equations (5.4) and (5.5) give the relations between the components of a vector with respect to the two different bases, as well as the relations between these components and the inner products of the vector with each basis vector. Our current main interest is in the relations between the different components (i.e., the rightmost two-thirds of (5.4) and (5.5)). For future reference, let us summarize what the above tells us in that regard.

Lemma 5.2 (The Big Lemma on General Change of Bases)
Let

$$\mathcal{A} = \{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_N\} \qquad \text{and} \qquad \mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_N\}$$

be any two bases for an N-dimensional vector space $\mathcal{V}$. Then, for any $\mathbf{v}$ in $\mathcal{V}$,

$$\mathbf{M}_{\mathcal{A}\mathcal{A}} \, |\mathbf{v}\rangle_{\mathcal{A}} = \mathbf{M}_{\mathcal{A}\mathcal{B}} \, |\mathbf{v}\rangle_{\mathcal{B}} \tag{5.6a}$$

and

$$\mathbf{M}_{\mathcal{B}\mathcal{A}} \, |\mathbf{v}\rangle_{\mathcal{A}} = \mathbf{M}_{\mathcal{B}\mathcal{B}} \, |\mathbf{v}\rangle_{\mathcal{B}} \tag{5.6b}$$

where $\mathbf{M}_{\mathcal{A}\mathcal{A}}$, $\mathbf{M}_{\mathcal{A}\mathcal{B}}$, $\mathbf{M}_{\mathcal{B}\mathcal{B}}$ and $\mathbf{M}_{\mathcal{B}\mathcal{A}}$ are the four $N \times N$ matrices given by

$$[\mathbf{M}_{\mathcal{A}\mathcal{A}}]_{jk} = \left\langle\, \mathbf{a}_j \;\middle|\; \mathbf{a}_k \,\right\rangle \quad , \qquad [\mathbf{M}_{\mathcal{A}\mathcal{B}}]_{jk} = \left\langle\, \mathbf{a}_j \;\middle|\; \mathbf{b}_k \,\right\rangle \quad ,$$

$$[\mathbf{M}_{\mathcal{B}\mathcal{B}}]_{jk} = \left\langle\, \mathbf{b}_j \;\middle|\; \mathbf{b}_k \,\right\rangle \quad \text{and} \quad [\mathbf{M}_{\mathcal{B}\mathcal{A}}]_{jk} = \left\langle\, \mathbf{b}_j \;\middle|\; \mathbf{a}_k \,\right\rangle \quad .$$

Naturally, the above formulas simplify considerably when the two bases are orthonormal. That will be of particular interest to us. Before going there, however, let us observe an number of "little results" that immediately follow from the above lemma and its derivation, and which apply whether or not either basis is orthonormal:

1. Solving equations (5.6a) and (5.6b) for $|\mathbf{v}\rangle_{\mathcal{B}}$ yields the change of basis formulas

$$|\mathbf{v}\rangle_{\mathcal{B}} = \mathbf{M}_{\mathcal{A}\mathcal{B}}{}^{-1} \, \mathbf{M}_{\mathcal{A}\mathcal{A}} \, |\mathbf{v}\rangle_{\mathcal{A}}$$

and

$$|\mathbf{v}\rangle_{\mathcal{B}} = \mathbf{M}_{\mathcal{B}\mathcal{B}}{}^{-1} \, \mathbf{M}_{\mathcal{B}\mathcal{A}} \, |\mathbf{v}\rangle_{\mathcal{A}} \quad .$$

2. Solving equations (5.6a) and (5.6b) for $|\mathbf{v}\rangle_{\mathcal{A}}$ yields the change of basis formulas

$$|\mathbf{v}\rangle_{\mathcal{A}} = {\mathbf{M}_{\mathcal{A}\mathcal{A}}}^{-1}\mathbf{M}_{\mathcal{A}\mathcal{B}}|\mathbf{v}\rangle_{\mathcal{B}}$$

and

$$|\mathbf{v}\rangle_{\mathcal{A}} = {\mathbf{M}_{\mathcal{B}\mathcal{A}}}^{-1}\mathbf{M}_{\mathcal{B}\mathcal{B}}|\mathbf{v}\rangle_{\mathcal{B}} \quad .$$

3. From equation (5.4) we have that

$$|\mathbf{v}\rangle_{\mathcal{A}} = {\mathbf{M}_{\mathcal{A}\mathcal{A}}}^{-1}\begin{bmatrix}\langle\,\mathbf{a}_1\,|\,\mathbf{v}\,\rangle\\ \langle\,\mathbf{a}_2\,|\,\mathbf{v}\,\rangle\\ \vdots\\ \langle\,\mathbf{a}_N\,|\,\mathbf{v}\,\rangle\end{bmatrix} \quad .$$

Don't memorize these formulas. They are nice formulas and can save a little work when dealing with arbitrary matrices. However, since we will attempt to restrict ourselves to orthonormal (or at least orthogonal) bases, we won't have a great need for these formulas as written. If you must memorize formulas, memorize those in the Big Lemma — they are much more easily memorized. And besides the above change of basis formulas are easily derived from the formulas in that lemma.

5.3 Change of Basis for Vector Components: When the Bases Are Orthonormal

The Main Result

Now consider how the formulas in the Big Lemma (lemma 5.2 on page 5–7) simplify when the bases

$$\mathcal{A} = \{\,\mathbf{a}_1,\,\mathbf{a}_2,\,\ldots,\,\mathbf{a}_N\,\} \qquad \text{and} \qquad \mathcal{B} = \{\,\mathbf{b}_1,\,\mathbf{b}_2,\,\ldots,\,\mathbf{b}_N\,\}$$

are orthonormal. Then, as noted on page 5–5 (equation set (5.1)),

$$\mathbf{M}_{\mathcal{A}\mathcal{A}} = \mathbf{I} \qquad \text{and} \qquad \mathbf{M}_{\mathcal{B}\mathcal{B}} = \mathbf{I} \quad .$$

Thus, equation (5.6a) in the Big Lemma reduces to

$$\mathbf{M}_{\mathcal{A}\mathcal{B}}|\mathbf{v}\rangle_{\mathcal{B}} = |\mathbf{v}\rangle_{\mathcal{A}} \quad , \tag{5.7a}$$

and equation (5.6b) in the Big Lemma reduces to

$$|\mathbf{v}\rangle_{\mathcal{B}} = \mathbf{M}_{\mathcal{B}\mathcal{A}}|\mathbf{v}\rangle_{\mathcal{A}} \quad . \tag{5.7b}$$

Equations (5.7a) and (5.7b) are certainly nicer than the original equations in the Big Lemma, but look at what we get when they are combined:

$$|\mathbf{v}\rangle_{\mathcal{B}} = \mathbf{M}_{\mathcal{B}\mathcal{A}}|\mathbf{v}\rangle_{\mathcal{A}} = \mathbf{M}_{\mathcal{B}\mathcal{A}}\underbrace{\mathbf{M}_{\mathcal{A}\mathcal{B}}|\mathbf{v}\rangle_{\mathcal{B}}}_{|\mathbf{v}\rangle_{\mathcal{A}}}$$

That is,

$$|\mathbf{v}\rangle_{\mathcal{B}} = [\mathbf{M}_{\mathcal{B}\mathcal{A}}\mathbf{M}_{\mathcal{A}\mathcal{B}}]\,|\mathbf{v}\rangle_{\mathcal{B}} \qquad \text{for each} \quad \mathbf{v}\in\mathcal{V} \quad ,$$

which is only possible if

$$\mathbf{M}_{\mathcal{BA}}\mathbf{M}_{\mathcal{AB}} = \mathbf{I} \quad ,$$

and which, in turn, means that

$$\mathbf{M}_{\mathcal{BA}}{}^{-1} = \mathbf{M}_{\mathcal{AB}} \quad \text{and} \quad \mathbf{M}_{\mathcal{AB}}{}^{-1} = \mathbf{M}_{\mathcal{BA}} \quad .$$

But remember (equation set (5.1) on page 5–5) that

$$\mathbf{M}_{\mathcal{AB}} = \mathbf{M}_{\mathcal{BA}}{}^{\dagger} \quad \text{and} \quad \mathbf{M}_{\mathcal{BA}} = \mathbf{M}_{\mathcal{AB}}{}^{\dagger} \quad .$$

Combining this with the previous line gives us

$$\mathbf{M}_{\mathcal{BA}}{}^{\dagger} = \mathbf{M}_{\mathcal{BA}}{}^{-1} \quad \text{and} \quad \mathbf{M}_{\mathcal{AB}}{}^{\dagger} = \mathbf{M}_{\mathcal{AB}}{}^{-1} \quad .$$

So $\mathbf{M}_{\mathcal{BA}}$ and its adjoint $\mathbf{M}_{\mathcal{AB}}$ are unitary matrices.

What the above tells us it that all the matrices involved are easily computed via the adjoint once we know one of the matrices.

We will make one more small set of observation regarding matrices $\mathbf{M}_{\mathcal{AB}}$ and $\mathbf{M}_{\mathcal{BA}}$ and the components of the vectors in basis $\mathcal{B}$ with respect to basis $\mathcal{A}$. Remember, since $\mathcal{A}$ is an orthonormal basis, the k^{th} component of $\mathbf{b}_j$ with respect to $\mathcal{A}$ is given by $\langle \mathbf{a}_k \,|\, \mathbf{b}_j \rangle$. That is,

$$\mathbf{b}_j = \sum_k \langle \mathbf{a}_k \mid \mathbf{b}_j \rangle \mathbf{a}_k = \sum_k \mathbf{a}_k \langle \mathbf{a}_k \mid \mathbf{b}_j \rangle = \sum_k \mathbf{a}_k \, [\mathbf{M}_{\mathcal{AB}}]_{kj} \quad .$$

This is just the formula for the matrix product

$$\begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_N \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_N \end{bmatrix} \mathbf{M}_{\mathcal{AB}} \quad .$$

In summary, we have the following:

Theorem 5.3 (The Big Theorem on Change of Orthonormal Bases)
Let $\mathcal{V}$ be an N-dimensional vector space with orthonormal bases

$$\mathcal{A} = \{ \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_N \} \quad \textit{and} \quad \mathcal{B} = \{ \mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_N \} \quad .$$

Let $\mathbf{M}_{\mathcal{AB}}$ and $\mathbf{M}_{\mathcal{BA}}$ be a pair of $N \times N$ matrices which are adjoints of each other and which satisfy any one of the following sets of conditions:

1. $[\mathbf{M}_{\mathcal{AB}}]_{jk} = \langle \mathbf{a}_j \mid \mathbf{b}_k \rangle$.
2. $[\mathbf{M}_{\mathcal{BA}}]_{jk} = \langle \mathbf{b}_j \mid \mathbf{a}_k \rangle$.
3. $\begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_N \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_N \end{bmatrix} \mathbf{M}_{\mathcal{AB}}$.
4. $\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_N \end{bmatrix} = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_N \end{bmatrix} \mathbf{M}_{\mathcal{BA}}$.
5. $|\mathbf{v}\rangle_{\mathcal{A}} = \mathbf{M}_{\mathcal{AB}} \, |\mathbf{v}\rangle_{\mathcal{B}}$ *for each* $\mathbf{v} \in \mathcal{V}$.
6. $|\mathbf{v}\rangle_{\mathcal{B}} = \mathbf{M}_{\mathcal{BA}} \, |\mathbf{v}\rangle_{\mathcal{A}}$ *for each* $\mathbf{v} \in \mathcal{V}$.

Then $\mathbf{M}_{\mathcal{AB}}$ and $\mathbf{M}_{\mathcal{BA}}$ are unitary matrices satisfying all of the above conditions, as well as

$$\mathbf{M}_{\mathcal{AB}} = \mathbf{M}_{\mathcal{BA}}{}^{\dagger} = \mathbf{M}_{\mathcal{BA}}{}^{-1} \quad \textit{and} \quad \mathbf{M}_{\mathcal{BA}} = \mathbf{M}_{\mathcal{AB}}{}^{\dagger} = \mathbf{M}_{\mathcal{AB}}{}^{-1} \quad .$$

Finding the Matrices

If you keep in mind the equation

$$\begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_N \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_N \end{bmatrix} \mathbf{M}_{\mathcal{AB}}$$

from the Big Theorem (theorem 5.3, above), and have the formulas for the $\mathbf{b}_k$'s in terms of the $\mathbf{a}_j$'s,

$$\mathbf{b}_k = \beta_{k1}\mathbf{a}_1 + \beta_{k2}\mathbf{a}_2 + \cdots + \beta_{kN}\mathbf{a}_N \qquad \text{for} \quad j = 1, 2, \ldots, N \quad ,$$

then you can obtain the matrix $\mathbf{M}_{\mathcal{AB}}$ easily by simply noting that each of these equations can be written as

$$\mathbf{b}_k = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_N \end{bmatrix} \begin{bmatrix} \beta_{k1} \\ \beta_{k2} \\ \vdots \\ \beta_{kN} \end{bmatrix} \quad .$$

Comparing the last equation with the one a few lines earlier for $\begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_N \end{bmatrix}$, it should be clear that the above column matrix on the right must be the k^{th} column in $\mathbf{M}_{\mathcal{AB}}$.

See, no computations are needed (provided the bases are orthonormal).

!►*Example 5.2:* *Let $\mathcal{V}$ be a three-dimensional space of traditional vectors with a standard basis*

$$\mathcal{A} = \{\mathbf{i}, \mathbf{j}, \mathbf{k}\} \quad .$$

Let

$$\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$$

where

$$\mathbf{b}_1 = \frac{1}{\sqrt{2}}[\mathbf{i} + \mathbf{k}] \quad , \quad \mathbf{b}_2 = \frac{1}{\sqrt{2}}[\mathbf{i} - \mathbf{k}] \quad \text{and} \quad \mathbf{b}_3 = \mathbf{j} \quad .$$

"By inspection", it should be clear that $\mathcal{B}$ is also an orthonormal basis for $\mathcal{V}$.

Now observe that the formulas for the $\mathbf{b}_k$'s can be rewritten as

$$\mathbf{b}_1 = \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix} \quad , \quad \mathbf{b}_2 = \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ -\frac{1}{\sqrt{2}} \end{bmatrix}$$

and

$$\mathbf{b}_3 = \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad ,$$

which can be written even more consisely as

$$\begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \mathbf{b}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \end{bmatrix} \quad .$$

As noted above, we then must have

$$\mathbf{M}_{\mathcal{AB}} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \end{bmatrix} \quad .$$

Multiple Change of Basis

Suppose, now, we have three orthonormal bases

$$\mathcal{A} = \{ \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_N \} \quad , \quad \mathcal{B} = \{ \mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_N \}$$

and

$$\mathcal{C} = \{ \mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_N \}$$

Then, in addition to $\mathbf{M}_{\mathcal{BA}}$ and $\mathbf{M}_{\mathcal{AB}}$ we also have $\mathbf{M}_{\mathcal{AC}}$, $\mathbf{M}_{\mathcal{BA}}$, $\mathbf{M}_{\mathcal{CA}}$ and $\mathbf{M}_{\mathcal{CB}}$, all defined analogously to the way we defined $\mathbf{M}_{\mathcal{BA}}$ and $\mathbf{M}_{\mathcal{AB}}$. Applying the above theorem we see that, for every $\mathbf{v}$ in $\mathcal{V}$,

$$\mathbf{M}_{\mathcal{CA}} |\mathbf{v}\rangle_{\mathcal{A}} = |\mathbf{v}\rangle_{\mathcal{C}} = \mathbf{M}_{\mathcal{CB}} |\mathbf{v}\rangle_{\mathcal{B}} = \mathbf{M}_{\mathcal{CB}} \left[\mathbf{M}_{\mathcal{BA}} |\mathbf{v}\rangle_{\mathcal{A}} \right] \quad .$$

Thus,

$$\mathbf{M}_{\mathcal{CA}} |\mathbf{v}\rangle_{\mathcal{A}} = \mathbf{M}_{\mathcal{CB}} \mathbf{M}_{\mathcal{BA}} |\mathbf{v}\rangle_{\mathcal{A}} \qquad \text{for every} \quad \mathbf{v} \in \mathcal{V} \quad .$$

From this, it immediately follows that

$$\mathbf{M}_{\mathcal{CA}} = \mathbf{M}_{\mathcal{CB}} \mathbf{M}_{\mathcal{BA}} \quad . \tag{5.8}$$

Remember, this is assuming the bases are all orthonormal.

5.4 Traditional Rotated and Flipped Bases

Let us briefly restrict ourselves to a two- or three-dimensional traditional vector space $\mathcal{V}$, with orthonormal bases

$$\mathcal{A} = \{ \mathbf{a}_1, \mathbf{a}_2 \} \quad \text{and} \quad \mathcal{B} = \{ \mathbf{b}_1, \mathbf{b}_2 \}$$

or

$$\mathcal{A} = \{ \mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3 \} \quad \text{and} \quad \mathcal{B} = \{ \mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3 \} \quad .$$

Direction Cosines

Since we are assuming a space of traditional vectors, the scalars are real numbers, the inner product is the traditional dot product, and the change of bases matrices $\mathbf{M}_{\mathcal{AB}}$ and $\mathbf{M}_{\mathcal{BA}}$ are orthogonal and are transposes of each other. In this case,

$$[\mathbf{M}_{\mathcal{AB}}]_{jk} = \langle \mathbf{a}_j \mid \mathbf{b}_k \rangle = \mathbf{a}_j \cdot \mathbf{b}_k = \|\mathbf{a}_j\| \, \|\mathbf{b}_k\| \cos\bigl(\theta(\mathbf{a}_j, \mathbf{b}_k)\bigr) \quad .$$

And because the $\mathbf{a}_j$'s and $\mathbf{b}_k$'s are unit vectors, this reduces to

$$[\mathbf{M}_{\mathcal{AB}}]_{jk} = \cos\bigl(\theta_{jk}\bigr) \qquad \text{where} \quad \theta_{jk} = \text{angle between } \mathbf{a}_j \text{ and } \mathbf{b}_k \quad .$$

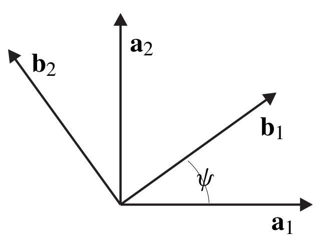
These are called the *direction cosines* relating the two bases.

Two-Dimensional Case

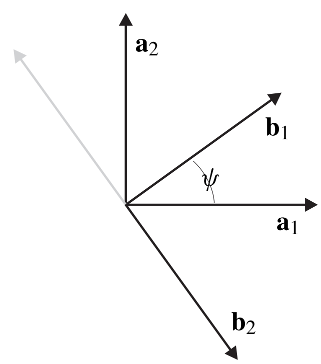
If $\mathcal{V}$ is two dimensional, then the possible geometric relationships between

$$\mathcal{A} = \{\mathbf{a}_1, \mathbf{a}_2\} \qquad \text{and} \qquad \mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$$

are easily sketched:



or



Clearly, we can only have one of the following two situations:

1. $\{\mathbf{b}_1, \mathbf{b}_2\}$ are the vectors that can be obtained by "rotating" $\{\mathbf{a}_1, \mathbf{a}_2\}$ through some angle ψ, in which case the matrix relating $\begin{bmatrix}\mathbf{b}_1 & \mathbf{b}_2\end{bmatrix}$ to $\begin{bmatrix}\mathbf{a}_1 & \mathbf{a}_2\end{bmatrix}$ is a "rotation matrix" $\mathbf{R}_\psi$ (see problem ***K*** in *Homework Handout IV*).

or

2. $\{\mathbf{b}_1, \mathbf{b}_2\}$ are the vectors that can be obtained by "rotating" $\{\mathbf{a}_1, \mathbf{a}_2\}$ through some angle ψ, and then "flipping" the direction of the second vector. In this case, the matrix relating $\begin{bmatrix}\mathbf{b}_1 & \mathbf{b}_2\end{bmatrix}$ to $\begin{bmatrix}\mathbf{a}_1 & \mathbf{a}_2\end{bmatrix}$ is given by a "rotation matrix" $\mathbf{R}_\psi$ followed by a "flip the direction of the second vector matrix". (Equivalently, we could first rotate by a slightly different angle and then flip the direction of the first rotated vector — or do the "flipping" first to either $\mathbf{a}_1$ or $\mathbf{a}_2$ and then rotate.)

With a little thought, it should be clear that the matrices for the "flips" in the second case are simply

$$\mathbf{F}_1 = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \text{and} \qquad \mathbf{F}_2 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad .$$

Clearly, then the matrix $\mathbf{M}_{\mathcal{BA}}$ can be written as either a single rotation matrix, or as the product of a rotation matrix with one of these "flip" matrices. It's not at all hard to show that

$$\det(\text{any rotation matrix}) = 1 \qquad \text{and} \qquad \det(\text{either flip matrix}) = -1 \quad .$$

Consequently, at least if we are considering orthonormal bases $\mathcal{A}$ and $\mathcal{B}$ for a two-dimensional, traditional vector space,

$$\mathcal{A} \text{ and } \mathcal{B} \text{ are rotated images of each other} \quad \Longleftrightarrow \quad \det(\mathbf{M}_{\mathcal{BA}}) = +1 \quad .$$

Three-Dimensional Case

Likewise, if

$$\mathcal{A} = \{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\} \qquad \text{and} \qquad \mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$$

are any two orthonormal bases for a traditional three-dimensional space, then $\mathbf{M}_{\mathcal{BA}}$ will either be a matrix for a "rotation by some angle about some vector", or the product of such a rotation matrix with one of the flip matrices

$$\mathbf{F}_1 = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad , \quad \mathbf{F}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{F}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad .$$

And, again, it can be shown that (still assuming $\mathcal{A}$ and $\mathcal{B}$ are orthonormal bases)

$$\mathcal{A} \text{ and } \mathcal{B} \text{ are rotated images of each other} \quad \Longleftrightarrow \quad \det(\mathbf{M}_{\mathcal{BA}}) = +1 \quad .$$

Consequently,

$$\mathcal{A} \text{ and } \mathcal{B} \text{ are both right-handed, or are both left-handed} \quad \Longleftrightarrow \quad \det(\mathbf{M}_{\mathcal{BA}}) = +1 \quad .$$

It can be shown that the rotation matrix, itself, can be written as a product of three "simple rotation matrices"

$$\mathbf{R}_{3,\gamma}\mathbf{R}_{2,\beta}\mathbf{R}_{3,\alpha}$$

where $\mathbf{R}_{j,\phi}$ corresponds to a rotation of angle ϕ about the j^{th} basis vector. The angles α , β , and γ are the infamous "Euler angles". Now turn to pages 139, 140 and 141 in Arfken, Weber and Harris and skim their discussion of rotations and Euler angles.

Warnings:

1. There are subtle differences between a matrix for a rotation operator (to be discussed in the next chapter) and the change of basis matrix for "rotated bases". Both may be called rotation matrices, but may differ in the signs of some of the corresponding entries.
2. I have not carefully checked Arken and Weber's discussion so I will not guarantee whether their product
$$\mathbf{R}_{3,\gamma}\mathbf{R}_{2,\beta}\mathbf{R}_{3,\alpha}$$
is actually what we are calling $\mathbf{M}_{\mathcal{BA}}$ or if the signs of some of the entries in some of the matrices need to be switched. In other words, I have not verified whether each of Arfken, Weber and Harris's "rotation matrices" are describing the operation of rotating a vector in space, or describing a change of basis when one basis is a rotation of the other.
3. Different authors use different conventions for the Euler angles — different axes of rotation and different order of operations. Be careful, the computations based on two different conventions may well lead to two different and incompatible results.

5.5 Sidenote: Vectors Defined by a Transformation Law

Suppose we have a huge collection of bases

$$\left\{ \mathcal{B}^1,\ \mathcal{B}^2,\ \mathcal{B}^3,\ \dots,\ \right\} \qquad \text{with} \quad \mathcal{B}^m \ = \ \left\{ \mathbf{b}_1^m,\ \mathbf{b}_2^m,\ \dots,\ \mathbf{b}_1^m \right\}$$

for our N-dimensional vector space $\mathcal{V}$ (for example, this might be the collection of all "rotations" of some favorite orthonormal basis). Then each $\mathbf{v}$ in $\mathcal{B}$ has components with respect to each of these $\mathcal{B}^m$'s,

$$|\mathbf{v}\rangle_{\mathcal{B}^m} \ = \ \begin{bmatrix} v_1^m \\ v_2^m \\ \vdots \\ v_N^m \end{bmatrix} \qquad \text{where} \quad \mathbf{v} \ = \ \sum_{j=1}^{N} v_j^m \mathbf{b}_j^m \quad .$$

Also, for each pair of bases $\mathcal{B}^m$ and $\mathcal{B}^n$, there will be a corresponding "change of basis" matrix $\mathbf{M}_{\mathcal{B}^m\mathcal{B}^n}$ such that

$$|\mathbf{v}\rangle_{\mathcal{B}^m} \ = \ \mathbf{M}_{\mathcal{B}^m\mathcal{B}^n} \, |\mathbf{v}\rangle_{\mathcal{B}^n} \qquad \text{for each} \quad \mathbf{v} \in \mathcal{V} \quad . \tag{5.9}$$

This last expression is sometimes called a "transformation law", even though we are not really dealing with a true transform of vectors here.

Now suppose that, in the course of tedious calculations and/or drinking, we obtain, corresponding to each different basis $\mathcal{B}^m$, a corresponding N-tuple of scalars $\left(w_1^m, w_2^m, \dots, w_N^m\right)$. The question naturally arrises as to whether

$$\left(w_1^1, w_2^1, \dots, w_N^1\right) \quad , \quad \left(w_1^2, w_2^2, \dots, w_N^2\right) \quad , \quad \left(w_1^3, w_2^3, \dots, w_N^3\right) \quad , \quad \dots$$

all describe the same vector $\mathbf{w}$ but in the various corresponding bases. Obviously (to us at least) the answer is "yes" if and only if

$$\begin{bmatrix} w_1^m \\ w_2^m \\ \vdots \\ w_N^m \end{bmatrix} \ = \ \mathbf{M}_{\mathcal{B}^m\mathcal{B}^n} \begin{bmatrix} w_1^n \\ w_2^n \\ \vdots \\ w_N^n \end{bmatrix} \qquad \text{for every pair of bases } \mathcal{B}_m \text{ and } \mathcal{B}_n \quad . \tag{5.10}$$

This is because we can set

$$\mathbf{w} \ = \ \sum_{j=1}^{N} w_j^1 \mathbf{b}_j^1$$

and use the fact that (5.10) holds to verify that the change of basis formulas (5.9) will give us, for each $\mathcal{B}^m$, the originally given N-tuple of scalars $\left(w_1^m, w_2^m, \dots, w_N^m\right)$ as the components of $\mathbf{w}$ with respect to that basis.

This $\mathbf{w}$ is called *the vector defined by the transformation law (5.9).*

5.6 "Volumes" of N-Dimensional Hyper-Parallelpipeds (Part II)

Let us now continue our discussion from section 3.6 on "volumes of hyper-parallelepipeds".

Recollections from Section 3.6

Recall the situation and notation: In a N-dimensional Euclidean space we have a "hyper-parallelepiped" $\mathcal{P}_N$ generated by a linearly independent set of N vectors

$$\{ \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_N \} \quad .$$

Our goal is to find a way to compute the N-dimensional volume of this object — denoted by $V_N(\mathcal{P}_N)$ — using the components of the $\mathbf{v}_k$'s with respect to any basis

$$\mathcal{B} = \{ \mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_N \} \quad .$$

For our computations, we are letting $\mathbf{V}_\mathcal{B}$ be the "matrix of components of the $\mathbf{v}_j$'s with respect to basis $\mathcal{B}$" given by

$$\mathbf{V}_\mathcal{B} = \text{“} \left[|\mathbf{v}_1\rangle_\mathcal{B} \quad |\mathbf{v}_2\rangle_\mathcal{B} \quad |\mathbf{v}_3\rangle_\mathcal{B} \quad \cdots \quad |\mathbf{v}_N\rangle_\mathcal{B} \right] \text{”} = \begin{bmatrix} v_{1,1} & v_{2,1} & v_{3,1} & \cdots & v_{N,1} \\ v_{1,2} & v_{2,2} & v_{3,2} & \cdots & v_{N,2} \\ v_{1,3} & v_{2,3} & v_{3,3} & \cdots & v_{N,3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ v_{1,N} & v_{2,N} & v_{3,N} & \cdots & v_{N,N} \end{bmatrix}$$

where, for $j = 1, 2, \dots, N$,

$$\mathbf{v}_j = \sum_{k=1}^{N} v_{j,k} \mathbf{b}_k \quad .$$

Recall, also, that we derived one formula for the volume; namely,

$$V_N(\mathcal{P}_N) = \det(\mathbf{V}_\mathcal{N}) \tag{5.11}$$

where $\mathcal{N}$ is the orthonormal set generated from $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_N\}$ via the Gram-Schmidt procedure (see theorem 3.4 on page 3–27).

General Component Formulas

From lemma 5.2 on page 5–7 on general change of bases formulas, we know

$$\mathbf{M}_{\mathcal{N}\mathcal{N}} \left|\mathbf{v}_j\right\rangle_\mathcal{N} = \mathbf{M}_{\mathcal{N}\mathcal{B}} \left|\mathbf{v}_j\right\rangle_\mathcal{B} \quad .$$

Since $\mathcal{N}$ is orthonormal, however, $\mathbf{M}_{\mathcal{N}\mathcal{N}}$ is just the identity matrix, and the above reduces to

$$\left|\mathbf{v}_j\right\rangle_\mathcal{N} = \mathbf{M}_{\mathcal{N}\mathcal{B}} \left|\mathbf{v}_j\right\rangle_\mathcal{B} \quad .$$

Thus,

$$\begin{aligned}\mathbf{V}_{\mathcal{N}} &= \left[|\mathbf{v}_1\rangle_{\mathcal{N}} \quad |\mathbf{v}_2\rangle_{\mathcal{N}} \quad |\mathbf{v}_3\rangle_{\mathcal{N}} \quad \cdots \quad |\mathbf{v}_N\rangle_{\mathcal{N}}\right] \\ &= \mathbf{M}_{\mathcal{N}\mathcal{B}}\left[|\mathbf{v}_1\rangle_{\mathcal{B}} \quad |\mathbf{v}_2\rangle_{\mathcal{B}} \quad |\mathbf{v}_3\rangle_{\mathcal{B}} \quad \cdots \quad |\mathbf{v}_N\rangle_{\mathcal{B}}\right] = \mathbf{M}_{\mathcal{N}\mathcal{B}}\mathbf{V}_{\mathcal{B}} \quad .\end{aligned}$$

Combining this with formula (5.11) and a property of determinants (equation (4.3) on page 4–11) then yields

$$V_N(\mathcal{P}_N) = \det(\mathbf{V}_{\mathcal{N}}) = \det(\mathbf{M}_{\mathcal{N}\mathcal{B}}\mathbf{V}_{\mathcal{B}}) = (\det \mathbf{M}_{\mathcal{N}\mathcal{B}})(\det \mathbf{V}_{\mathcal{B}}) \quad .$$

Let us simplify matters a little more. Let $\mathcal{U}$ be any orthonormal basis. As noted on page 5–11 (equation (5.8)),

$$\mathbf{M}_{\mathcal{N}\mathcal{B}} = \mathbf{M}_{\mathcal{N}\mathcal{U}}\mathbf{M}_{\mathcal{U}\mathcal{B}} \quad .$$

But, since $\mathcal{U}$ and $\mathcal{N}$ are orthonormal bases of "traditional" vectors, $\mathbf{M}_{\mathcal{N}\mathcal{U}}$ must be an orthogonal matrix (and, thus, have $\det(\mathbf{M}_{\mathcal{N}\mathcal{U}}) = \pm 1$). Consequently,

$$\begin{aligned}\det(\mathbf{M}_{\mathcal{N}\mathcal{B}}) &= \det(\mathbf{M}_{\mathcal{N}\mathcal{U}}\mathbf{M}_{\mathcal{U}\mathcal{B}}) \\ &= \det(\mathbf{M}_{\mathcal{N}\mathcal{U}}) \det(\mathbf{M}_{\mathcal{U}\mathcal{B}}) = \pm\det(\mathbf{M}_{\mathcal{U}\mathcal{B}}) \quad .\end{aligned}$$

Combining this with our last formula for $V_N(\mathcal{P}_N)$, gives us our most general component formula for $V_N(\mathcal{P}_N)$ (and the next theorem).

Theorem 5.4 (general component formula for the volume of a hyper-parallelepiped)
Let $\mathcal{P}_N$ be the hyper-parallelepiped in an generated by a linearly independent set $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_N\}$ of traditional vectors. Also, using any basis $\mathcal{B}$ for the space spanned by the $\mathbf{v}_k$'s, let $\mathbf{V}_{\mathcal{B}}$ be the matrix of components of the $\mathbf{v}_j$'s with respect to basis $\mathcal{B}$,

$$\mathbf{V}_{\mathcal{B}} = \begin{bmatrix} v_{1,1} & v_{2,1} & v_{3,1} & \cdots & v_{N,1} \\ v_{1,2} & v_{2,2} & v_{3,2} & \cdots & v_{N,2} \\ v_{1,3} & v_{2,3} & v_{3,3} & \cdots & v_{N,3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ v_{1,N} & v_{2,N} & v_{3,N} & \cdots & v_{N,N} \end{bmatrix} \quad \textit{with} \quad \begin{bmatrix} v_{k,1} \\ v_{k,2} \\ v_{k,3} \\ \vdots \\ v_{k,N} \end{bmatrix} = |\mathbf{v}_k\rangle_{\mathcal{B}} \quad .$$

Then, letting $V_N(\mathcal{P}_N)$ denote the N-dimensional volume of $\mathcal{P}_N$,

$$V_N(\mathcal{P}_N) = |\det(\mathbf{M}_{\mathcal{U}\mathcal{B}}) \det(\mathbf{V}_{\mathcal{B}})| \tag{5.12}$$

where $\mathcal{U}$ is any orthonormal basis.

Special Cases

$\mathcal{B}$ is orthonormal

In particular, if $\mathcal{B}$ is an orthonormal basis, then we can let $\mathcal{U} = \mathcal{B}$ in formula (5.12). Because of the orthonormality of $\mathcal{B}$ we have

$$\det(\mathbf{M}_{\mathcal{U}\mathcal{B}}) = \det(\mathbf{M}_{\mathcal{B}\mathcal{B}}) = \det(\mathbf{I}) = 1 \quad ,$$

and formula (5.12) reduces to

$$V_N(\mathcal{P}_N) = |\det(\mathbf{V}_\mathcal{B})| = \left|\det\begin{bmatrix} v_{1,1} & v_{2,1} & v_{3,1} & \cdots & v_{N,1} \\ v_{1,2} & v_{2,2} & v_{3,2} & \cdots & v_{N,2} \\ v_{1,3} & v_{2,3} & v_{3,3} & \cdots & v_{N,3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ v_{1,N} & v_{2,N} & v_{3,N} & \cdots & v_{N,N} \end{bmatrix}\right| \quad . \tag{5.13}$$

?►*Exercise 5.6:* *Let $\mathcal{V}$ be a four-dimensional space of traditional vectors with orthonormal basis*

$$\{\mathbf{i}, \mathbf{j}, \mathbf{k}, \mathbf{l}\}$$

and let

$$\mathbf{v}_1 = 3\mathbf{i} \quad ,$$
$$\mathbf{v}_2 = 2\mathbf{i} + 4\mathbf{j} \quad ,$$
$$\mathbf{v}_3 = 8\mathbf{i} - 3\mathbf{j} + 5\mathbf{k}$$

and

$$\mathbf{v}_4 = 4\mathbf{i} + 7\mathbf{j} - 2\mathbf{k} + 2\mathbf{l} \quad .$$

Using formula (5.13), compute the "four-dimensional volume" of the hyper-parallelepiped generated by $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ and $\mathbf{v}_4$. (Compare the work done here to that done for the same problem in exercise set 3.15 on page 3–24.)

$\mathcal{B}$ and $\mathcal{P}_N$ are "Parallel"

Let's now assume each $\mathbf{b}_j$ is parallel to $\mathbf{v}_j$ and "pointing in the same direction". That is

$$v_{jk} = 0 \quad \text{if} \quad j \neq k \quad \text{and} \quad v_{jj} \geq 0 \quad .$$

For "convenience", let

$$\Delta v_k = v_{kk}$$

so that we can write

$$\mathbf{v}_j = \Delta v_j \mathbf{b}_j \quad \text{for} \quad j = 1, 2, \dots, N \quad .$$

Then

$$\det(\mathbf{V}_\mathcal{B}) = \det\begin{bmatrix} \Delta v_1 & 0 & 0 & \cdots & 0 \\ 0 & \Delta v_2 & 0 & \cdots & 0 \\ 0 & 0 & \Delta v_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \Delta v_N \end{bmatrix} = \Delta v_1 \, \Delta v_2 \, \Delta v_3 \cdots \Delta v_N \quad ,$$

and formula (5.12) becomes

$$V_N(\mathcal{P}_N) = |\det(\mathbf{M}_{\mathcal{U}\mathcal{B}})| \, \Delta v_1 \, \Delta v_2 \, \Delta v_3 \cdots \Delta v_N \quad . \tag{5.14}$$

where $\mathcal{U}$ is any orthonormal basis.

Of course, if there is any othogonality, then the above simplifies. In particular, if $\mathcal{B}$ is orthogonal, then you can easily show that

$$V_N(\mathcal{P}_N) = \|\mathbf{b}_1\|\,\|\mathbf{b}_2\|\cdots\|\mathbf{b}_N\|\ \Delta v_1\,\Delta v_2\,\Delta v_3\,\cdots\,\Delta v_N \quad .$$

If you compare formula (5.14) with formulas (3.11) through (3.11), starting on page 3–28, it should be clear that we've already derived geometric formulas for the above $|\det \mathbf{M}_{\mathcal{UB}}|$ when $N = 1$, $N = 2$ and $N = 3$. Since "cut and paste" is so easy, we'll rewrite those geometric formulas:

$$V_1(\mathcal{P}_1) = \cdots = \|\mathbf{b}_1\|\,\Delta v_1 \quad ,$$

$$V_2(\mathcal{P}_2) = \sqrt{\|\mathbf{b}_1\|^2\,\|\mathbf{b}_2\|^2 - (\mathbf{b}_1\cdot\mathbf{b}_2)^2}\ \Delta v_1\,\Delta v_2 \quad ,$$

and

$$V_3(\mathcal{P}_3) = \sqrt{A - B + C}\ \Delta v_1\,\Delta v_2\,\Delta v_3$$

where

$$A = \|\mathbf{b}_1\|^2\,\|\mathbf{b}_2\|^2\,\|\mathbf{b}_3\|^2 \quad ,$$

$$B = \|\mathbf{b}_1\|^2\,(\mathbf{b}_2\cdot\mathbf{b}_3)^2 + \|\mathbf{b}_2\|^2\,(\mathbf{b}_1\cdot\mathbf{b}_2)^2 + \|\mathbf{b}_3\|^2\,(\mathbf{b}_1\cdot\mathbf{b}_2)^2 \quad ,$$

and

$$C = 2\frac{(\mathbf{b}_1\cdot\mathbf{b}_2)^2(\mathbf{b}_1\cdot\mathbf{b}_3)^2}{\|\mathbf{b}_1\|^2} = 2\frac{(\mathbf{b}_2\cdot\mathbf{b}_1)^2(\mathbf{b}_2\cdot\mathbf{b}_3)^2}{\|\mathbf{b}_2\|^2} = \cdots \quad .$$

As we noted in on page 3–29, these formulas reduce even further if the $\mathbf{b}_k$'s are unit vectors, and even further if the set of $\mathbf{b}_k$'s is orthonormal. If you've not already done so, then do the next exercise (which is the same as exercise 3.17):

?►Exercise 5.7: *To what do the above formulas for $V_1(\mathcal{P}_1)$, $V_2(\mathcal{P}_2)$ and $V_3(\mathcal{P}_3)$ reduce*

a: *when the $\mathbf{b}_k$'s are unit vectors.*

b: *when the set of $\mathbf{b}_k$ is orthonormal.*