Real antisymmetric operators

This section is a bridge between the last two major topics of the course. First, we use the Jordan canonical form theorem to derive a canonical form for antisymmetric matrices. Then we discuss the isomorphism between antisymmetric matrices and vectors in three dimensions, as a foretaste of the subject of general antisymmetric tensors — which we will reach precisely as we run out of time in the course. (See Chapter 8 of Bowen & Wang.)

Canonical form

The Jordan theorem applies to *complex* vector spaces. When $\mathcal{F} = \mathbf{R}$, an operator $\underline{A}$ may not have a JCF. For example, $\det(\underline{A} - \lambda)$ may have complex roots, which can't possibly correspond to real eigenvectors of a real matrix.

Consider a real, *antisymmetric* $\underline{A}$; that is, $\underline{A}^* = -\underline{A}$, where the adjoint equals the transpose since $\mathcal{F} = \mathbf{R}$. With respect to an ON basis, $A^j{}_k = -A^k{}_j$. The antisymmetric matrix A of course defines an antisymmetric operator $\underline{A}\colon \mathbf{R}^N \to \mathbf{R}^N$. Note, however, that this same matrix also defines an anti-Hermitian operator $\underline{A}_{\mathbf{C}}\colon \mathbf{C}^N \to \mathbf{C}^N$. For $\underline{A}_{\mathbf{C}}$ we could just define $\underline{B} \equiv -i\underline{A}_{\mathbf{C}}$ to get an *Hermitian* operator: $\underline{B}^* = \underline{B}$ [where the adjoint is now in a complex space, hence *not* the same thing as a transpose]. Then $\underline{B}$ is (unitarily) diagonalizable, with real eigenvalues $\{\lambda_j\}$. Thus $\underline{A}_{\mathbf{C}} = i\underline{B}$ is diagonalizable, with pure imaginary eigenvalues $\{i\lambda_j\}$.

Since the eigenvectors belong to $\mathbf{C}^N$, it makes sense to complex-conjugate them. [This would not be true of an abstract Hilbert space.] Let $\vec{v}_j$ be one of the eigenvectors:

$$\underline{A}_{\mathbf{C}}\vec{v}_j = i\lambda_j\vec{v}_j\,.$$

Complex-conjugate this equation (identifying $\underline{A}_{\mathbf{C}}$ with its matrix, which is real):

$$\underline{A}_{\mathbf{C}}\vec{v}_j^{\,*} = -i\lambda_j\vec{v}_j^{\,*}.$$

Suppose for the moment that all eigenvalues of $\underline{A}_{\mathbf{C}}$ are nonzero. Then we see that they come in complex-conjugate pairs $\{i\lambda_j, -i\lambda_j\}$ with equal multiplicities within a pair. (Henceforth $\{\lambda_j\}$ stands for that half of the original set $\{\lambda_j\}$ consisting of positive numbers.)

Introduce new basis vectors

$$\begin{aligned} \vec{v}_{jR} &\equiv \frac{1}{\sqrt{2}}(\vec{v}_j + \vec{v}_j^{\,*}),\\ \vec{v}_{jI} &\equiv \frac{-i}{\sqrt{2}}(\vec{v}_j - \vec{v}_j^{\,*}). \end{aligned} \tag{$\natural$}$$

These vectors are *real* (members of $\mathbf{R}^N \subset \mathbf{C}^N$; note that $\mathbf{C}^N = \mathbf{R}^N \oplus i\mathbf{R}^N$ as a real vector space). The action of $\underline{A}_{\mathbf{C}}$ on them is

$$\underline{A}_{\mathbf{C}}\vec{v}_{jR} = -\lambda_j \vec{v}_{jI}\,, \qquad \underline{A}_{\mathbf{C}}\vec{v}_{jI} = +\lambda_j \vec{v}_{jR}\,.$$

So, with respect to such a basis, $\underline{A}_{\mathbf{C}}$ takes the form

$$A = \begin{pmatrix} \begin{matrix} 0 & \lambda_1 \\ -\lambda_1 & 0 \end{matrix} & 0 & 0 \\ 0 & \begin{matrix} 0 & \lambda_2 \\ -\lambda_2 & 0 \end{matrix} & 0 \\ 0 & 0 & \ddots \end{pmatrix}.$$

Since the basis vectors and the matrix are real, we can discard the "$i\mathbf{R}^N$" part of $\mathbf{C}^N$ and think of this as the matrix of a linear mapping $\mathbf{R}^N \to \mathbf{R}^N$. It is the original operator $\underline{A}$ in a new basis! Indeed, the composition of the unitary change-of-basis matrix implied by ($\natural$) with the unitary matrix which diagonalized $\underline{B}$ is a *real* unitary matrix, since it maps a real basis to another real basis. Thus we've shown that $\underline{A}$ can be put into 2×2-block-diagonal form by a similarity transformation with an orthogonal matrix.

Now consider the case that $0 \in \sigma(\underline{A}_{\mathbf{C}})$. Then

$$\underline{A}_{\mathbf{C}}\vec{v}_j = \vec{0}, \qquad \underline{A}_{\mathbf{C}}\vec{v}_j^{\,*} = \vec{0}.$$

There are two possibilities:

(1) $\vec{v}_j$ and $\vec{v}_j^{\,*}$ are linearly independent. Then the foregoing argument still applies, and it yields a diagonal block $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ in the canonical form.

(2) $\vec{v}_j$ and $\vec{v}_j^{\,*}$ are dependent. (If N is odd, this case must occur at least once.) Then $\vec{v}_{jR}$ and $\vec{v}_{jI}$ are dependent; choose *one* of them, normalized, to get a *real* basis vector $\vec{v}$ such that $\underline{A}_{(\mathbf{C})}\vec{v} = \vec{0}$.

Thus, for odd N we have the canonical form

$$\begin{pmatrix} 0 & 0 & 0 & \cdots & & \\ 0 & 0 & \lambda_1 & & & \\ 0 & -\lambda_1 & 0 & & & 0 \\ \vdots & & & 0 & \lambda_2 & \\ & & & -\lambda_2 & 0 & \\ & 0 & & & & \ddots \end{pmatrix},$$

where some of the λ's may be 0; for even N we have the same canonical form as before, with some of the λ's possibly 0. *We take this to be the "canonical" form of a real antisymmetric operator.* [This is the **theorem**.]

A direct proof (involving no complex numbers) is possible.

A similar theorem holds giving a canonical form for real *orthogonal* operators (Bowen & Wang, Exercise 30.4, p. 165). In fact, there is a direct relationship between orthogonal and antisymmetric operators, which we'll clarify in the next subsection.)

Rotations, cross product, and curl

Relationship between rotations and antisymmetric operators

Consider the matrix

$$A(t) = \begin{pmatrix} \cos\omega t & -\sin\omega t & 0 \\ \sin\omega t & \cos\omega t & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

which represents a rotation about the z axis. If t is intepreted as time, this family of matrices describes a rotational *motion* at a uniform rate (with angular velocity ω): $\vec{x}(t) = A(t)\vec{x}(0)$.

We can calculate

$$\frac{dA}{dt} = \begin{pmatrix} -\omega\sin\omega t & -\omega\cos\omega t & 0 \\ \omega\cos\omega t & -\omega\sin\omega t & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\left.\frac{dA}{dt}\right|_{t=0} = \begin{pmatrix} 0 & -\omega & 0 \\ \omega & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \equiv \Omega,$$

and we note that Ω is antisymmetric. More generally,

$$\frac{dA(t)}{dt} = \Omega\, A(t) \qquad \text{for all } t$$

(Ω independent of t).

Conversely,

$$A(t) = e^{\Omega t}.$$

This is clear from the differential equation we have just derived, and it can be verified by calculating the exponential matrix by a power series.

We observe that $A(t)$ satisfies the *group properties*

$$A(t)A(s) = A(t+s), \tag{1}$$

$$A(0) = \underline{1}, \tag{2}$$

as well as the orthogonality condition

$$A(t)^* A(t) = \underline{1}. \tag{0}$$

Furthermore, the uniform rotational motion about any other axis will be represented by the matrices $A' \equiv OAO^{-1}$ for some orthogonal O. They will also satisfy the conditions (0)–(2). Conversely, as will become clear presently, any 3×3 family of real matrices (or the operators they represent) satisfying (0)–(2) is a uniform rotational motion about some axis; it is called a *one-parameter group* of rotations or orthogonal transformations.

Lemma.

$$\frac{dA^{-1}}{dt} = -A^{-1}\frac{dA}{dt}A^{-1}$$

for **any** *family of invertible matrices (or operators) depending on a parameter t.*

Proof: Differentiate $A^{-1}A = \underline{1}$:

$$\underline{0} = \frac{d}{dt}(A^{-1}A) = \frac{dA^{-1}}{dt}\,A + A^{-1}\frac{dA}{dt}.$$

Solve for $\dfrac{dA^{-1}}{dt}$.

Apply the lemma to a family with properties (0) and (2):

$$\left.\frac{dA^*}{dt}\right|_{t=0} = -\left.\frac{dA}{dt}\right|_{t=0} \equiv -\Omega,$$

but also

$$\left.\frac{dA^*}{dt}\right|_{t=0} = \left.\left(\frac{dA}{dt}\right)^*\right|_{t=0} = \Omega^*.$$

Thus Ω is antisymmetric. Using (1), one can now show that $dA/dt = \Omega A$ for all t. Thus $A(t) = e^{\Omega t}$, which is a rotational motion about some axis (the direction of the 0-eigenvector of A).

This discussion can be summarized by the statement that *one-parameter groups of orthogonal operators (or matrices) are in one-to-one correspondence with antisymmetric operators.* The antisymmetric object Ω is called the *(infinitesimal) generator* of the one-parameter group $A(t)$.

This treatment has been for $\mathbf{R}^3$. Most of it carries over to arbitrary dimension, but the concept of "axis" does not. Our normal-form theorem for antisymmetric operators, together with a straightforward generalization of the foregoing discussion, shows that a general rotation is built out of rotations in a number of two-dimensional subspaces (orthogonal to each other). If the dimension is *odd*, there is one direction in space left over, which may be called the axis.

Relationship between antisymmetric operators and the vector cross product

From now on we are restricted purely to dimension 3. The most general antisymmetric matrix is

$$\Omega = \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix},$$

where the rationale for the numbering of the elements will soon become clear. The action of $\underline{\Omega}$ on a vector is given by

$$\Omega\vec{y} = \begin{pmatrix} -\omega_3 y_2 + \omega_2 y_3 \\ \omega_3 y_1 - \omega_1 y_3 \\ -\omega_2 y_1 + \omega_1 y_2 \end{pmatrix} \equiv \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ \omega_1 & \omega_2 & \omega_3 \\ y_1 & y_2 & y_3 \end{vmatrix} \equiv \vec{\omega} \times \vec{y}$$

in the notations of classical vector algebra. Thus *every antisymmetric matrix is associated with a vector in* $\mathbf{R}^3$*, and vice versa.* In fact, *the association is an isomorphism of vector spaces.* The action of the antisymmetric operator on other vectors corresponds to the cross product of the vector with the other vectors.

Any time you see a cross product in an application, there is at least one antisymmetric operator hiding in the problem somewhere, disguised as a vector!

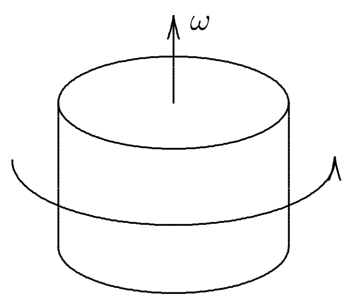
The prototype example is *angular velocity.* Rotation of a body is associated with a vector $\vec{\omega}$ along the axis of rotation. The (linear) velocity of a point $\vec{y}$ on the body is $\vec{v} = \vec{\omega} \times \vec{y}$. ($\vec{v}$ and $\vec{y}$ are functions of time.) This just means

$$\frac{d\vec{y}}{dt} = \Omega\vec{y},$$

where

$$\vec{y}(t) = A(t)\vec{y}(0)$$

— in accordance with the discussions above.



Other physical quantities which are antisymmetric operators disguised as vectors include angular momentum, torque, and magnetic field. Except for the last, all of these are clearly associated with rotations. The appearance of magnetic field in the list will be elucidated later.

Let us further recall the elementary geometrical definition of the vector cross product: $\vec{u} \times \vec{v}$ is the vector (1) perpendicular to the plane of $\vec{u}$ and $\vec{v}$, (2) with length $\|\vec{u}\|\|\vec{v}\| \sin\theta$ ($\theta \equiv$ angle between $\vec{u}$ and $\vec{v}$), and (3) (if $\sin\theta \neq 0$) with direction (sign) such that $\vec{u}$, $\vec{v}$, $\vec{u} \times \vec{v}$ form a basis of the same handedness as $\hat{\imath}$, $\hat{\jmath}$, $\hat{k}$. Note that only (3) refers to a coordinate system. Thus $\vec{u} \times \vec{v}$ is independent of coordinate system *except for handedness.*

If $\vec{u}$ and $\vec{v}$ are ordinary ("true") vectors, then $\vec{u} \times \vec{v}$ is a *pseudovector* whose sign depends on the handedness of the coordinate system.

On the other hand, angular velocity, magnetic field, and so on are *already* pseudovectors, since their physical definitions involve "right-hand rules". Therefore, when such a vector appears as a factor in a cross product, the relevant statement is: If $\vec{\omega}$ is a pseudovector and $\vec{y}$ is a true vector, then $\vec{\omega} \times \vec{y}$ is a true vector. Thus in applications we often see two cross products together, one to create a pseudovector and another to undo it.

EXAMPLE: Magnetic field.

1. Biot–Savart law:

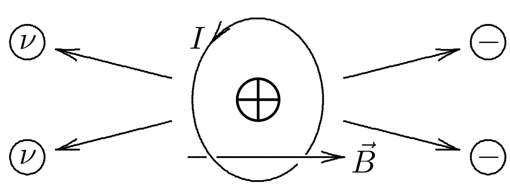
$$d\vec{B} = \frac{I\, d\vec{l} \times \vec{y}}{\|\vec{y}\|^3}\,;$$

$$\vec{B} = I \oint_{\text{circuit}} \frac{d\vec{l} \times \vec{y}}{\|\vec{y}\|^3}\,.$$

(For a single charge, $\vec{B} = e\,\dfrac{\vec{v} \times \vec{y}}{\|\vec{y}\|^3} = \vec{v} \times \vec{E}$.)

2. Lorentz force law: $\vec{F} = e(\vec{E} + \vec{v} \times \vec{B})$ $\quad$ ($\vec{v} =$ velocity of *another* charge).

Thus the sign of $\vec{B}$ is purely a convention and cancels out of the final answer for the magnetic force between two charges. [However, in some nuclear decays, particles are emitted preferentially along the direction of an applied magnetic field. This shows that some laws of nature *do* make a distinction between left and right ("overthrow of parity" — 1956).]



From this discussion we expect that the vectors $\vec{\omega}$ associated to antisymmetric matrices via the cross product are pseudovectors, not true vectors. In other words, the isomorphism between antisymmetric operators and vectors involves an arbitrary sign convention; returning to the determinantal definition of the cross product, we see that this sign is hidden in the ordering of $\{\hat{\imath}, \hat{\jmath}, \hat{k}\}$ there.

Here is another way of looking at this: If $\Omega \equiv L_{\vec{\omega}}$ is the antisymmetric matrix corresponding to a coordinate vector $\vec{\omega}$, then under a change of coordinates $\vec{x} \mapsto O\vec{x} + \vec{x}_0$ (where $O^{-1} = O^*$), we should have

$$OL_{\vec{\omega}}O^{-1} = (\det\, O)L_{O\vec{\omega}}\,,$$

where the determinant is ± 1 depending on whether the handedness changes. This relation can be verified algebraically by a tedious calculation which involves equating O^* to O^{-1}, calculated as

$$\frac{\operatorname{adj} O}{\det\, O}\,.$$

Relationship of these matters to the curl

The curl operation on vector fields is usually defined by applying the cross product formally to the gradient "vector",

$$\nabla \equiv \frac{\partial}{\partial x_1}\,\hat{\imath} + \cdots.$$

That is, if $\vec{\omega}(\vec{x}) = \nabla \times \vec{A}$, then

$$\omega_2 = \frac{\partial A_1}{\partial x_3} - \frac{\partial A_3}{\partial x_1}\,, \qquad \text{etc.}$$

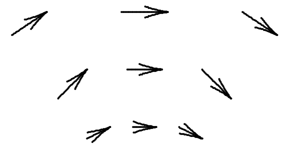
We see that $\omega_2 = (L_{\vec{\omega}})_{13}$ (and similarly for the other components), where

$$L_{(\nabla\times\vec{A})} = J_{\vec{A}} - J^*_{\vec{A}}$$

= twice the antisymmetric part of the Jacobian matrix of $\vec{A}$ (the table of all first-order partial derivatives of all component functions of $\vec{A}$). Clearly, the antisymmetric matrix is more fundamental here than the vector.

Under a change of coordinates, J transforms to OJO^{-1}, evaluated at a suitably transformed argument vector. (The O is the rotation of the components of $\vec{A}$; the O^{-1} comes from the chain rule for the transformation of the partial derivatives.) It follows that $\nabla \times \vec{A}$ transforms as a pseudovector. In other words, the formula for $\nabla \times \vec{A}$ is, by definition, the same in all coordinate systems, but it does not define quite the same vector: There is a sign change when you switch from right-handed to left-handed coordinates. The formula *does* always define the same antisymmetric operator.

Since a vector field with nonzero curl has a rotational character, we have come full circle in our discussion of four types of mathematical objects whose relationship is left tantalizingly obscure in "classical" textbooks.



Chapter 8 of Bowen & Wang, especially Sec. 41 on “Duality”, introduces one to the 20th-century generalization of these matters to dimensions other than 3.