

Chapter 3

Vectors and Linear Algebra

3.1 Introduction

Linear algebra is the language describing systems in finite (or countably infinite) dimensions, where *dimension* represents the number of variables at hand. This appears naturally in systems with more than one degrees of freedom or in approximate descriptions of complex systems in a discrete set of variables. Linear algebra also gives the basic framework of quantum mechanics, describing observables in terms of eigenvalues. In two and more dimensions, much of linear algebra is illustrated by vectors and their linear transformations.

Angular momentum \mathbf{J} is a vector that appears in numerous problems. Like energy and linear momentum, total angular momentum is a conserved quantity. For freely rotating rigid bodies, angular momentum is in proportion to angular velocity

$$\mathbf{J} = \Omega \mathbf{n}. \tag{3.1}$$

The length Ω denotes the rate of rotation and the direction denotes its orientation. According to Mach's principle, angular velocity is commonly defined relative to the distant stars, where most of the mass is. For periodic motion with period P , the angular velocity satisfies

$$\Omega = \frac{2\pi}{P}. \tag{3.2}$$

For motion at a separation \mathbf{r} about to a given axis, the instantaneous velocity is a tangent

$$\mathbf{v} = \frac{d\mathbf{r}}{dt} = \mathbf{\Omega} \times \mathbf{r}. \tag{3.3}$$

The associated linear momentum is the vector

$$\mathbf{p} = \frac{d}{dt}m\mathbf{r} = m\mathbf{v}, \quad (3.4)$$

when the mass m is time-independent. The associated angular momentum of a particle with linear momentum \mathbf{p} is

$$\mathbf{J} = \mathbf{r} \times \mathbf{p}. \quad (3.5)$$

Thus, \mathbf{J} is a vector formed out of \mathbf{r} and \mathbf{p} . Transformations of \mathbf{J} follow the rules for vectors, e.g., when considering translation or rotation of a coordinate system.

Maps in linear algebra are represented by $n \times m$ matrices, from a linear vector space of dimension m to a linear vector space of dimension n . These vector spaces are often over the real or complex numbers, e.g., \mathbb{R}^n or \mathbb{C}^n . As such, matrices are comprised of row and column vectors of length m and, respectively, n . An $n \times m$ matrix is said to be of dimension $n \times m$.

Consider, for instance, a 2×2 matrix

$$C = \begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix}. \quad (3.6)$$

Its rows $\mathbf{r}_{(i)}$ and columns $\mathbf{c}_{(i)}$ ($i = 1, 2$) can be schematically indicated as

$$C = \left(\begin{array}{c} \text{---} \\ \text{---} \end{array} \right) = \begin{pmatrix} \mathbf{r}_1^T \\ \mathbf{r}_2^T \end{pmatrix}, \quad C = \left(\begin{array}{c|c} | & | \\ | & | \end{array} \right) = (\mathbf{c}_1 \ \mathbf{c}_2) \quad (3.7)$$

with

$$\mathbf{r}_1^T = (1 \ 2), \quad \mathbf{r}_2^T = (2 \ 0), \quad \mathbf{c}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad \mathbf{c}_2 = \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \quad (3.8)$$

where we explicitly include the *transpose* T to denote row vectors according to

$$\begin{pmatrix} x \\ y \end{pmatrix}^T = (x \ y). \quad (3.9)$$

These 2×2 matrices describe various transformations in the two-dimensional plane, such as reflections, rotations and coordinate permutations. Key properties are eigenvalues and the associated eigenvectors, much of which depends on their determinants and symmetry properties.

3.2 Inner and Outer Products

Two linearly independent vectors \mathbf{a} and \mathbf{b} span a parallelogram. The *projection* of \mathbf{a} onto \mathbf{b} defines the inner product (further Sect. 3.5) $\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos \theta$ with

$$\cos \theta = \angle(\mathbf{a}, \mathbf{b}) = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}||\mathbf{b}|} \quad (3.10)$$

denoting the cosine of the angle between the two, where $|\mathbf{a}|$ refers to the length of \mathbf{a} satisfying $|\mathbf{a}| = \sqrt{\mathbf{a} \cdot \mathbf{a}}$. Referenced to a Cartesian coordinate system with basis vectors $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$, expressions obtain in component form, In three dimensions, we have

$$\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}, \quad \mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k} \quad (3.11)$$

and so

$$\mathbf{a} \cdot \mathbf{b} = a_1b_1 + a_2b_2 + a_3b_3. \quad (3.12)$$

The outer product represents the area element, in area and orientation, of the parallelogram, represented by the normal vector

$$(a_2b_3 - a_3b_2)\mathbf{i} + (a_3b_1 - a_1b_3)\mathbf{j} + (a_1b_2 - a_2b_1)\mathbf{k} = \begin{pmatrix} a_2b_3 - a_3b_2 \\ a_3b_1 - a_1b_3 \\ a_1b_2 - a_2b_1 \end{pmatrix}, \quad (3.13)$$

where we used the right handed rule in the direction of movement of a corkscrew turned from \mathbf{a} to \mathbf{b} . Its length equals the area of the parallelogram

$$|\mathbf{a} \times \mathbf{b}| = |\mathbf{a}||\mathbf{b}| \sin \theta. \quad (3.14)$$

3.3 Angular Momentum Vector

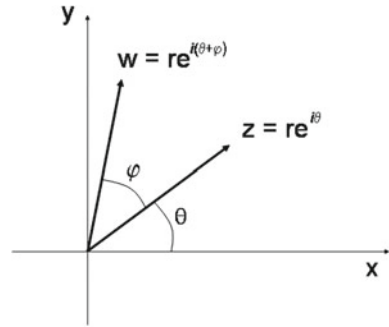
In circular motion, angular momentum \mathbf{J} is a vector with the same orientation as the angular velocity $\boldsymbol{\Omega}$. By the vector identity

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \mathbf{b}(\mathbf{a} \cdot \mathbf{c}) - \mathbf{c}(\mathbf{a} \cdot \mathbf{b}) \quad (3.15)$$

between vectors \mathbf{a} , \mathbf{b} , \mathbf{c} , circular motion gives the *specific angular momentum* (angular momentum per unit mass)

$$\mathbf{j} = \mathbf{r} \times \mathbf{v} = \mathbf{r} \times (\boldsymbol{\Omega} \times \mathbf{r}) = r^2\boldsymbol{\Omega} = r^2\boldsymbol{\Omega}\mathbf{n}, \quad (3.16)$$

Fig. 3.1 Rotation in the (x, y) -plane over an angle φ obtains by multiplication by a 2×2 matrix $R(\varphi)$ of vectors $\mathbf{z} = x\mathbf{i}_x + y\mathbf{i}_y$. In the complex plane, it corresponds to multiplication by $e^{i\varphi}$



since $\mathbf{r} \cdot \mathbf{r} = r^2$ and $\mathbf{r} \cdot \boldsymbol{\Omega} = 0$. With (3.4, 3.5), our model problem of circular motion, therefore, implies

$$\mathbf{j} = r^2 \frac{2\pi}{P} = 2 \frac{\pi r^2}{P} = 2 \frac{dA}{dt} \mathbf{n}, \tag{3.17}$$

that is, \mathbf{j} represents twice the rate-of-change of surface area traced out by the radius \mathbf{r} in the orbital motion. Based on (3.4–3.17), this is a geometrical identify, not restricted to circular motion, familiar as Kepler’s third law in planetary motion.

3.3.1 Rotations

If $z = r e^{i\theta}$, then

$$w = z e^{i\varphi} = r e^{i(\theta+\varphi)} = r (\cos(\theta + \varphi) + i \sin(\theta + \varphi)), \tag{3.18}$$

as illustrated in Fig. 3.1. That is

$$w = r e^{i\theta} e^{i\varphi} = r [\cos \theta \cos \phi - \sin \theta \sin \phi + i (\sin \theta \cos \phi + \cos \theta \sin \theta)]. \tag{3.19}$$

Applied to the basic vectors $\{\mathbf{i}_x, \mathbf{i}_y\}$, we have

$$\mathbf{i}'_x = \mathbf{i}_x \cos \theta + \mathbf{i}_y \sin \theta, \quad \mathbf{i}'_y = -\mathbf{i}_x \sin \theta + \mathbf{i}_y \cos \theta. \tag{3.20}$$

Example 3.1. Consider a basis $\{\mathbf{i}, \mathbf{j}\}$ rotating along with a point (x, y) over the unit circle S^1 . That is, \mathbf{i} points to (x, y) with local tangent \mathbf{j} to S^1 with counter-clockwise orientation. Moving along S^1 at constant angular velocity ω , $\theta = \omega t$ as a function time t and (3.20) implies

$$\frac{d\mathbf{i}}{dt} = \omega [-\mathbf{i} \sin \theta + \mathbf{j} \cos \theta], \quad \frac{d\mathbf{j}}{dt} = \omega [-\mathbf{i} \cos \theta - \mathbf{j} \sin \theta], \quad (3.21)$$

and so

$$\frac{d\mathbf{i}}{dt} = \omega \mathbf{j}, \quad \frac{d\mathbf{j}}{dt} = -\omega \mathbf{i}. \quad (3.22)$$

3.3.2 Angular Momentum and Mach's Principle

Following (3.5) and (3.23), circular particle motion satisfies

$$\mathbf{J} = I \mathbf{n}, \quad I = m\Omega r^2, \quad (3.23)$$

where I denotes the moment of inertia about \mathbf{n} .¹ Evidently, (3.23) implies that $J = 0$ whenever $\Omega = 0$ and visa-versa. In an astronomical context, we may follow Mach² and define the angular velocity as the rate of change of angles *measured relative to the distant stars*. Does (3.23) hold in general?

It turns out that angular momentum is sensitive to matter in the universe anywhere. While (3.23) holds true to great precision under ordinary circumstances when Ω is defined relative to the distant stars, deviations appear in the proximity of massive rotating bodies. This can be detected in tracking the orientation \mathbf{n} of a freely suspended gyroscope relative to a distant star. Recently, the NASA satellite Gravity Probe B³ did just that, and measured an angular velocity in \mathbf{n} at a minute rate of

$$\omega = -39 \text{ mas yr}^{-1} = -6 \times 10^{-16} \text{ rad s}^{-1}. \quad (3.24)$$

It agrees within a 20% window of uncertainty with the *frame-dragging* angular velocity of space-time around the earth, induced by Earth's angular momentum according to the theory of general relativity. According to the exact solution of rotating black holes in general relativity [3], (3.24) is the frame-dragging angular velocity at about 5 million Schwarzschild radii around a maximally spinning black hole with the same angular as the Earth (and 27 times its mass).

Though small, (3.24) defines a key result in our views on the relation between rotation and angular momentum, that comes out non-trivially in curved space-time predicted by the theory of general relativity. In particular, it changes our perception

¹Formally I_{nn} , since I is generally a two-index tensor.

²Ernst Mach (1838–1916).

³Everitt et al. [1] a local measurement on the Riemann tensor. LAGEOS II detected frame dragging in the orientation of polar orbits [2].

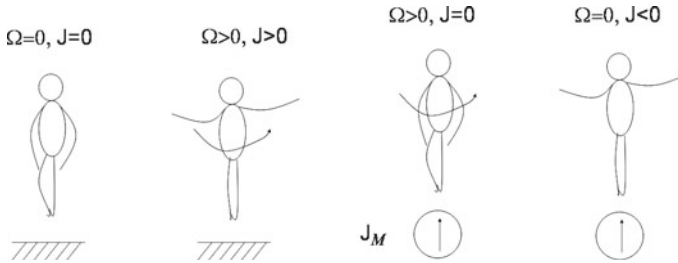


Fig. 3.2 (Left) In flat space-time, the ballerina effect a correspondence between zero angular velocity Ω relative to the distant stars and zero angular momentum J (Mach’s principle). (Right.) In curved space-time, the ballerina effect is different. Here, $J = (\Omega - \omega)I$, where I denotes the moment of inertia and ω is the frame-dragging angular velocity along the angular momentum J_M of a massive object nearby. As a result, $\Omega = \omega$ for $J = 0$ and $J < 0$ when $\Omega = 0$. Mach’s principle is to be generalized include all matter, including massive objects in a local neighborhood

of the ballerina effect (Fig. 3.2). In reality, a ballerina standing *still* with respect to the distant stars experiences a slight lifting of her arms *up*, due to her non-zero angular momentum imparted by frame-dragging around Earth.⁴ In a twist to the original formulation of Mach’s principle, she would experience co-rotation with an angular velocity (3.24) for her arms to be *down* in a fully relaxed state.

Frame dragging (3.24) induced by the angular momentum of the Earth is manifest also in energetic spin-spin interactions.⁵ In response, particles with angular momentum J_p about the spin axis of the Earth experience a potential energy [4]

$$E = \omega J_p, \tag{3.25}$$

that represents a line-integral of *Papapetrou forces* [5] mediated by ω . The energy (3.25) is notoriously small for J_p of classical objects. However, for charged particles like electrons or protons in magnetic fields around black holes, E can be huge and reach energies on the scale of Ultra High Energy Cosmic Rays (UHECRs). Measurement of (3.25) around the Earth awaits future satellite experiments.

3.3.3 Energy and Torque

Angular momentum $\mathbf{J} = J \mathbf{n}$ can be changed by application of a torque, defined as

$$\mathbf{T} = \frac{d}{dt} \mathbf{J} = \mathbf{n} \frac{d}{dt} J + J \frac{d}{dt} \mathbf{n}. \tag{3.26}$$

⁴A related result affects her weight by Papapetrou forces [5], here expressed in Eq. (3.25).

⁵The complete set of frame dragging induced interactions is described by the Riemann tensor.

The dimension of torque is energy, as follows from $[J]=\text{g cm}^2 \text{ s}^{-1}$ (mass times rate of change of area). Because angular momentum is a vector, (3.26) shows the appearance of a torque already when changing its orientation, even when keeping its magnitude constant. In this case, (3.26) may be due to a rotation, i.e.,

$$\Delta \mathbf{T} = J (\mathcal{R} - I) \mathbf{n} \tag{3.27}$$

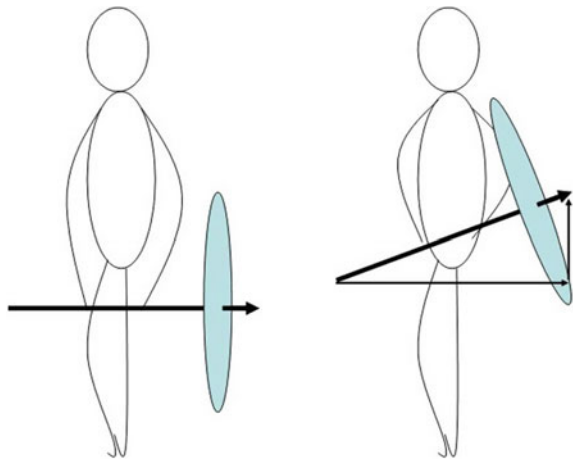
where \mathcal{R} is a rotation matrix. (More on matrices in Sect. 3.5) For a rotation over an angle φ about the x -axis, for example, we have (Sect. 3.4.2)

$$\mathcal{R} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi & -\sin \varphi \\ 0 & \sin \varphi & \cos \varphi \end{pmatrix} \tag{3.28}$$

Feynman [6] gives an illustrative set-up that can be performed using a bicycle wheel attached freely to a rod. In this event, \mathbf{n} is along the y -axis when the rod is initially held horizontally. Attempting to rotate the rod along the x -axis in an effort to move the wheel overhead is described by (3.27), see Fig. 3.3. By (3.28), it introduces a component of $\Delta \mathbf{T}$ along the z -axis. The person performing the rotation will experience a tendency to start rotating in the opposite direction to the angular momentum of the wheel, by conservation of total angular momentum in all three dimensions (in each of the three components x , y and z), i.e.,

$$\mathbf{J}_{\text{wheel}} + \mathbf{J}_{\text{person}} = \mathbf{0}. \tag{3.29}$$

Fig. 3.3 Changing the orientation \mathbf{n} of the angular momentum of a spinning wheel by a rotation introduces a component in an orthogonal direction, here along the vertical direction. Since angular momentum is conserved, a corresponding negative amount of angular momentum along the vertical direction is imparted by the person holding the wheel. The person will experience a counter-torque along the vertical axis



Since *power* is a scalar of dimension energy s^{-1} , the power delivered to or extracted from a rotating object is given by the inner product of torque and angular velocity, i.e.,

$$P = \boldsymbol{\Omega} \cdot \mathbf{T}. \quad (3.30)$$

For our circular motion, we have $\mathbf{T} = \frac{d}{dt}\mathbf{J} = I \frac{d}{dt}\boldsymbol{\Omega}$, and hence

$$P = \frac{d}{dt} \left(\frac{1}{2} \Omega^2 \right) \quad (3.31)$$

It follows that the rotational energy in case of $\mathbf{J} = I\boldsymbol{\Omega}$ satisfies

$$E_{rot} = \frac{1}{2} \Omega^2 I = \frac{1}{2} \boldsymbol{\Omega} \cdot \mathbf{J}. \quad (3.32)$$

Although (3.32) applies to non-relativistic mechanics such as spinning tops, somewhat remarkably it gives a fairly good approximation also to the rotational energy $E_{rot} = k \boldsymbol{\Omega} \cdot \mathbf{J}$, $k^{-1} = 2 \cos^2(\lambda/4)$, of a rotating black hole with non-dimensional angular momentum $\sin \lambda$, since

$$\frac{1}{2} \leq k \leq 0.5858. \quad (3.33)$$

To exemplify angular momentum conservation, consider the problem of the Moon's migration, in absorbing angular momentum in the Earth's spin due to a gravitational tidal torque.

Example 3.2. Some 4.52 Gyr ago, the Earth's spin period at birth was $P = 5.4$ h before the Moon was born. The Earth's normalized angular velocity

$$A_0 = \left(\frac{\Omega}{\Omega_b} \right)_{\oplus} \quad (3.34)$$

then (4.52 Gyr ago) was very similar to the same for Jupiter today, where

$$\Omega = \frac{2\pi}{P}, \quad \Omega_b = \sqrt{\frac{GM}{R^3}} \quad (3.35)$$

denote the actual and, respectively, break-up angular velocity for a planet of mass M and radius R , and G is Newton's constant. Some data:

$$\text{Earth: } M_{\oplus} = 5.97 \times 10^{27} \text{ g}, \quad R_{\oplus} = 6000 \text{ km}, \quad P_{\oplus} = 24 \text{ h}, \quad (3.36)$$

$$\text{Jupiter: } M \simeq 320M_{\oplus}, \quad R \simeq 11R_{\oplus}, \quad P \simeq 0.5P_{\oplus}.$$

The above follows from the following.

- For the Earth's $\Omega_{\oplus,b}$ and today's value $\Omega_{\oplus} = 2\pi/P_{\oplus}$, we have

$$A_1 = \left(\frac{\Omega}{\Omega_b} \right)_{\oplus}. \quad (3.37)$$

- The change P_{\oplus} to 5.4 h from 24 h today satisfies the scaling

$$\left(\frac{\Omega}{\Omega_b} \right)_{\oplus} \propto P_{\oplus}^{-1}. \quad (3.38)$$

- Consequently, the spin angular velocity relative to break up at birth satisfies

$$A_0 = \left(\frac{5.4 \text{ h}}{24 \text{ h}} \right)^{-1} A_1. \quad (3.39)$$

that may be compared to the same ratio of Jupiter today.

With Newton's constant $G = 6.67 \times 10^{-8} \text{ g}^{-1} \text{ cm}^3 \text{ s}^{-2}$ (recall that $G\rho$ has dimension angular velocity squared, i.e., s^{-2}), we have by explicit calculation

$$\Omega_{\oplus} = 7.27 \times 10^{-5} \text{ rad s}^{-1}, \quad \Omega_{\oplus,b} = 1.36 \times 10^{-3} \text{ rad s}^{-1} \quad (3.40)$$

and hence the ratio

$$A_1 = 0.0536. \quad (3.41)$$

By aforementioned scaling with P_{\oplus} , we have

$$A_0 = \left(\frac{24 \text{ h}}{5.4 \text{ h}} \right) A_1 = 4.44 A_1 \simeq 0.2380 \quad (3.42)$$

Repeating the above for Jupiter,

$$B_1 = \left(\frac{\Omega}{\Omega_b} \right)_J = 0.2185, \quad (3.43)$$

that is, our A_0 4.52 Gyr ago and Jupiter's B_1 today are very similar. As a consequence, we expect the weather of the Earth at birth to be very similar to that of Jupiter today, essentially a permanent storm by exceedingly large Coriolis forces. Recall that Coriolis forces scale with $\Omega_{\oplus}^2 \propto P_{\oplus}^{-2}$. They were initially some 20 times stronger than they are now. Thanks, in part, to spin down by the Moon, we can enjoy today's clement climate [7].

3.3.4 Coriolis Forces

Conservation of angular momentum gives rise to apparent forces when moving things around by external forces that leave the angular momentum invariant, as in the absence of any frictional forces. The specific angular momentum in the presence of an angular velocity ω is

$$j = \omega\sigma^2 : \omega = \frac{j}{\sigma^2}, \quad (3.44)$$

where σ denotes the distance to the axis of rotation. Moving a fluid element along the radial direction changes ω , as when the ballerina moves stretched arms inwards, according to $\delta\omega = -2j\sigma^{-3}\delta\sigma$. It comes with a change in azimuthal velocity $\delta v_\varphi = \sigma\delta\omega$ seen in a corotating frame, satisfying

$$\delta v_\varphi = -2\omega\delta\sigma. \quad (3.45)$$

In vector form, (3.45) is

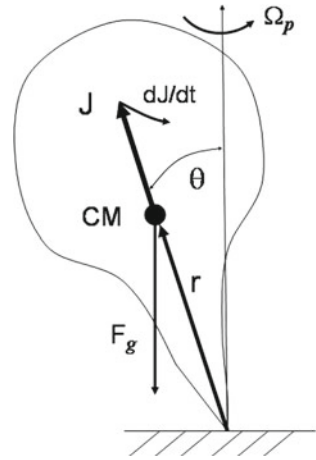
$$\frac{d}{dt}\mathbf{v}_\varphi = -2\boldsymbol{\omega} \times \mathbf{v}_\sigma. \quad (3.46)$$

This result is commonly expressed in terms of the Coriolis force

$$F_c = m \frac{d}{dt}\mathbf{v}_\varphi = 2m\mathbf{v} \times \boldsymbol{\omega} \quad (3.47)$$

Coriolis forces are particularly relevant when working in a rotating frame of reference. In particular, all of us terrestrial inhabitants living with the rotating frame fixed to Earth's surface. Air moving to a different latitude is subject to (3.47), since it changes the distance σ to the Earth's axis of rotation, which is approximately polar. Let Ω denote the absolute angular velocity of the Earth (relative to the distant stars), and express the angular velocity of the air $\omega' = \omega - \Omega$ relative to it, as measured in this rotating frame. Since $\delta\omega' = \delta\omega$, moving air from, say, in the direction of the equator produces a retrograde azimuthal velocity (rotation at an angular velocity $\omega < \Omega$). Moving it a constant angular velocity towards the equator produces a curved trajectory in response to the (retrograde) constant Coriolis force (3.47). This may give rise to large scale circulation patterns in combination with pressure gradients.

Fig. 3.4 Precession of the spinning top causes a velocity $\dot{\mathbf{n}}$ in the orientation \mathbf{n} of the angular momentum, such that $d\mathbf{J}/dt = J\dot{\mathbf{n}}$ absorbs the torque due to the gravitational force \mathbf{F}_g applied at its center of mass CM. In the idealized friction-free set-up, this process involves no exchange of energy or dissipation



3.3.5 Spinning Top

The motion of a spinning top tilted at an angle θ exemplifies the interaction of angular momentum as a vector with a torque, \mathbf{T} , applied continuously by the Earth's gravitational force \mathbf{F}_g as illustrated in Fig. 3.4. In general, we have the relations

$$\mathbf{T} = \frac{d}{dt}\mathbf{J} = \mathbf{r} \times \frac{d}{dt}\mathbf{p} = \mathbf{r} \times \mathbf{F}_g. \tag{3.48}$$

For a top that spins with no friction, the magnitude of its angular momentum vector is conserved. By (3.26, 3.27), the top precesses at an angular velocity Ω_p about the z -axis, $\Omega_p = d\phi/dt$, satisfying

$$\frac{d}{dt}\mathbf{J} = J \frac{d}{dt}\mathbf{n} = J\Omega_p \times \mathbf{n} = \Omega_p \times \mathbf{J}. \tag{3.49}$$

By (3.48), $T = \Omega_p J \sin \theta = rW \sin \theta$, and hence the angular velocity of precession about the vertical axis satisfies

$$\Omega_p J = rW, \tag{3.50}$$

where W denotes the weight of the top and r the distance of its center of mass away from its pivot on the table.

Example 3.3. Illustrative for some vector calculations is a more explicit calculation of the precession frequency (3.50). To this end, Fig. 3.5 shows a massive ring of radius R spinning at an angular velocity ω , whereby it attains an angular momentum per unit mass $J = \omega R^2$. Suppose it is mounted to one end of a rod,

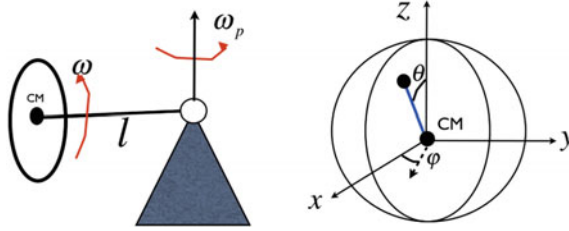


Fig. 3.5 Shown is a ring of radius R rotating at an angular velocity ω about a *horizontal axis* of length l , supported by a pivot that allow rotation at a precession angular velocity ω_p about the *vertical axis*. Upon translation of the center of mass (CM) to the origin of a spherical coordinate system, mass elements on the ring move over the surface of a sphere of radius R , parameterized by a poloidal and azimuthal angle θ and, respectively, φ , wherein $d\theta/dt = \omega$ and $d\varphi/dt = \omega_p$

that is suspended at a pivot at the other end. An approximately horizontal rod hereby precesses with an angular velocity ω_p about the vertical axis without dropping to a vertical position, satisfying (3.49). This result is invariant under linear translation of the CM. Precession is entirely due to the motion of mass-elements about the ring's CM, allowing us to place the CM at the origin of a spherical coordinate system (r, θ, φ) , as if the CM where placed at the pivot. With $\boldsymbol{\omega} = \omega \mathbf{i}_z$, the outer product $\boldsymbol{\omega} \times \mathbf{r}$ is the rotational velocity \mathbf{v}_ϕ of the end point of a vector \mathbf{r} and that $\mathbf{v}_\phi = \omega \sigma$, where $\sigma = b \sin \theta$ is the distance to the axis of rotation. A mass element $\delta m = (M/2\pi)\delta\theta$ in the ring herein assumes an angular momentum $\delta \mathbf{J} = \mathbf{r} \times \delta \mathbf{p} = \delta m \mathbf{r} \times \mathbf{v}$ with position vector

$$\mathbf{r} = b \begin{pmatrix} \sin \theta \cos \varphi \\ \sin \theta \sin \varphi \\ \cos \theta \end{pmatrix}, \quad \omega = \frac{d\theta}{dt}, \quad \omega_p = \frac{d\varphi}{dt}, \quad (3.51)$$

and associated velocity $\mathbf{v} = d\mathbf{r}/dt$,

$$\mathbf{v} = b \begin{pmatrix} \cos \theta \cos \varphi \\ \cos \theta \sin \varphi \\ -\sin \theta \end{pmatrix} \omega + b \begin{pmatrix} -\sin \theta \sin \varphi \\ \sin \theta \cos \varphi \\ 0 \end{pmatrix} \omega_p, \quad (3.52)$$

and acceleration $\mathbf{a} = d\mathbf{v}/dt$,

$$\mathbf{a} = -b \begin{pmatrix} \sin \theta \cos \varphi \\ \sin \theta \sin \varphi \\ \cos \theta \end{pmatrix} \omega^2 - b \begin{pmatrix} \sin \theta \cos \varphi \\ \sin \theta \sin \varphi \\ 0 \end{pmatrix} \omega_p^2 \quad (3.53)$$

$$+ 2b \begin{pmatrix} -\cos \theta \sin \varphi \\ \cos \theta \cos \varphi \\ 0 \end{pmatrix} \omega \omega_p. \quad (3.54)$$

Its inertia introduces a torque

$$\delta \mathbf{T} = \frac{d\delta \mathbf{J}}{dt} = \delta m \left(\frac{d}{dt} \mathbf{r} \times \mathbf{v} + \mathbf{r} \times \frac{d}{dt} \mathbf{v} \right) = \delta m \mathbf{r} \times \mathbf{a} \quad (3.55)$$

that evaluates to

$$\delta \mathbf{T} = \delta m b \left(\omega_p^2 \mathbf{r} \times \mathbf{i}_z + 2\omega\omega_p \mathbf{r} \times \begin{pmatrix} -\cos\theta \sin\varphi \\ \cos\theta \cos\varphi \\ 0 \end{pmatrix} \right) \quad (3.56)$$

To finalize, we integrate (3.56) over all mass elements δm . Making use of the following averages over the *fast angle* θ ,

$$\langle \mathbf{r} \times \mathbf{i}_z \rangle = \frac{1}{2\pi} \int_0^{2\pi} \mathbf{r} \times \mathbf{i}_z d\theta = \frac{b}{2\pi} \int_0^{2\pi} \begin{pmatrix} \sin\theta \sin\varphi \\ -\sin\theta \cos\varphi \\ 0 \end{pmatrix} d\theta = \mathbf{0} \quad (3.57)$$

and

$$\left\langle \mathbf{r} \times \begin{pmatrix} -\cos\theta \sin\varphi \\ \cos\theta \cos\varphi \\ 0 \end{pmatrix} \right\rangle = \frac{1}{2\pi} \int_0^{2\pi} \mathbf{r} \times \begin{pmatrix} -\cos\theta \sin\varphi \\ \cos\theta \cos\varphi \\ 0 \end{pmatrix} d\theta = \quad (3.58)$$

$$\frac{b}{2\pi} \int_0^{2\pi} \begin{pmatrix} -\cos^2\theta \cos\varphi \\ -\cos^2\theta \sin\varphi \\ 2\sin\theta \cos\theta \end{pmatrix} d\theta = -\frac{b}{2} \begin{pmatrix} \cos\varphi \\ \sin\varphi \\ 0 \end{pmatrix}, \quad (3.59)$$

we arrive at a total inertial torque $\mathbf{T} = \int_0^{2\pi} \delta \mathbf{T}$,

$$\mathbf{T} = \frac{M}{2\pi} \int_0^{2\pi} \mathbf{r} \times \mathbf{a} d\theta = Mb \left(\omega_p^2 \langle \mathbf{r} \times \mathbf{i}_z \rangle + 2\omega\omega_p \left\langle \mathbf{r} \times \begin{pmatrix} -\cos\theta \sin\varphi \\ \cos\theta \cos\varphi \\ 0 \end{pmatrix} \right\rangle \right). \quad (3.60)$$

With $J = I\omega$ expressed in the moment of inertia $I = Mb^2$, the latter reduces to $T = \omega\omega_p Mb^2 = \omega_p J$, i.e., our vector identity (3.49).

In Fig. 3.5, if the bar holding the rotating wheel is initially suspended horizontally at the pivot with zero angular momentum about the z -axis, then the onset of precession ω_p —balancing inertial to gravitational torque $gM\sigma$ —produces a finite angular momentum $J_z = M\sigma^2\omega_p$ about the z -axis (upwards, say), where $\sigma = l \cos\alpha$ is the arm length to the z -axis, now at a dip angle α . Since the total angular momentum about the z -axis remains zero, $J_z = J \sin\theta$ (pointing downwards). Given $\omega_p J = Mg\sigma$, it follows that (cf. Exercise 3.3)

$$\tan \alpha = (\omega_p / \Omega)^2, \quad (3.61)$$

where $\Omega = \sqrt{g/\sigma}$.

3.4 Elementary Transformations in the Plane

In the two-dimensional plane with Cartesian coordinates (x, y) , transformations describe a map

$$\mathbf{z} = x\mathbf{i}_x + y\mathbf{i}_y \rightarrow \mathbf{w} = x'\mathbf{i}_x + y'\mathbf{i}_y. \quad (3.62)$$

When linear, such map is a matrix multiplication $\mathbf{w} = C\mathbf{z}$,

$$\mathbf{z} = \begin{pmatrix} x \\ y \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (3.63)$$

with

$$C\mathbf{z} = \begin{pmatrix} \text{---} \\ \text{---} \end{pmatrix} \mathbf{z} = \begin{pmatrix} \mathbf{r}_1^T \mathbf{z} \\ \mathbf{r}_2^T \mathbf{z} \end{pmatrix} \quad (3.64)$$

and

$$\mathbf{r}^T = (a \ b) : \mathbf{r}^T \mathbf{z} = ax + by. \quad (3.65)$$

Equivalently, we have

$$C\mathbf{z} = \left(\begin{array}{c} | \\ | \\ | \end{array} \right) \left\{ x \begin{pmatrix} 1 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\} = x\mathbf{c}_1 + y\mathbf{c}_2. \quad (3.66)$$

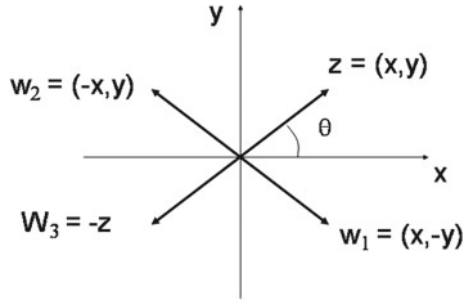
These two views (3.64–3.66) explicitly bring about linearity in the row and column vectors of C .

When working in the two-dimensional plane, we note that (3.62) is equivalent to a map of complex numbers $z = x + iy \rightarrow w = x' + iy'$, that is occasionally useful when working with conformal transformations $w = w(z)$ ($w'(z) \neq 0$).

3.4.1 Reflection Matrix

Figure 3.6 illustrates reflections in the two-dimensional plane about the x -axis, the y -axis and through the origin, $\mathcal{O} = (0, 0)$. Reflection about the x -axis is described by

Fig. 3.6 Reflections in the (x, y) -plane about the x -axis, the y -axis and through the origin, take $z = (x, y)$ to, respectively, $w_1 = (x, -y)$, $w_2 = (-x, y)$ and $w_3 = -z$. Each transformation is described by a 2×2 matrix acting on the vector $\mathbf{z} = x\mathbf{i}_x + y\mathbf{i}_y$.



$$\mathbf{z} = x\mathbf{i}_x + y\mathbf{i}_y \rightarrow \mathbf{w} = x\mathbf{i}_x - y\mathbf{i}_y. \tag{3.67}$$

The same transformation can be written as a matrix equation for the equations $x' = x$ and $y' = -y$ as follows

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \tag{3.68}$$

Reflection about the y -axis is described by

$$\mathbf{z} = x\mathbf{i}_x + y\mathbf{i}_y \rightarrow \mathbf{w} = -x\mathbf{i}_x + y\mathbf{i}_y. \tag{3.69}$$

The same transformation can be written as a matrix equation for $x' = -x$ and $y' = y$ as follows

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \tag{3.70}$$

As mentioned above, (3.67–3.69) are equivalent to taking $z \in \mathbb{C}$ into, respectively,

$$w_1 = \bar{z} = x - iy, \quad w_2 = -\bar{z} = -x + iy. \tag{3.71}$$

Reflection about the origin is described by

$$\mathbf{z} = x\mathbf{i}_x + y\mathbf{i}_y \rightarrow \mathbf{w} = -x\mathbf{i}_x - y\mathbf{i}_y, \quad w_3 = -z, \tag{3.72}$$

The same transformation can be written as a matrix equation for the equations $x' = -x$ and $y' = -y$ as follows

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \tag{3.73}$$

The *identity matrix* is defined by the transformation which leaves \mathbf{z} the same, i.e.,

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (3.74)$$

3.4.2 Rotation Matrix

The above can be extended to continuous transformations such as rotations. The rotation matrix can be derived from the multiplication of complex numbers following (3.18) and (3.62). With $\mathbf{z} = r \cos \theta \mathbf{i}_x + r \sin \theta \mathbf{i}_y$, we have

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = R(\varphi) \begin{pmatrix} x \\ y \end{pmatrix}, \quad (3.75)$$

in terms of the rotation matrix

$$R(\varphi) = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}. \quad (3.76)$$

Evidently, it satisfies

$$\det R = 1, \quad R(-\varphi) = R^{-1}(\varphi) = R^T(\varphi), \quad (3.77)$$

where R^{-1} refers to the inverse of R , R^T refers to the *transpose* and

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21} \quad (3.78)$$

defines the determinant of a 2×2 matrix.

For what follows, we shall generalize (3.9) to matrices. For a square $n \times n$ matrix A , the transpose obtains by interchanging the off-diagonal components a_{ij} ($i \neq j$) about the *principle diagonal* containing the a_{ii} . Schematically, if L refers to the *upper diagonal elements* and U refers to the *lower diagonal elements*, then

$$A = \begin{pmatrix} a_{11} & & & U \\ & a_{22} & & \\ & & \dots & \\ L & & & a_{nn} \end{pmatrix} \rightarrow A^T = \begin{pmatrix} a_{11} & & & L \\ & a_{22} & & \\ & & \dots & \\ U & & & a_{nn} \end{pmatrix}. \quad (3.79)$$

The rotation matrix $R(\varphi)$ in (3.76) is anti-symmetric in its off-diagonal elements, i.e., $U = -L$. A square matrix is said to be anti-symmetric, if $U = -L$ and the

elements on the principle diagonal are zero. Since the diagonal elements in (3.76) are non-zero, $R(\varphi)$ is not an anti-symmetric matrix.

Example 3.4. A symmetric matrix, satisfying $U = L$ as defined in (3.79), is the Lorentz boost

$$\Lambda(\mu) = \begin{pmatrix} \cosh \mu & \sinh \mu \\ \sinh \mu & \cosh \mu \end{pmatrix}, \quad (3.80)$$

that appears in the transformation of four-momenta in Malinowski space. Both $R(\varphi)$ and $\Lambda(\mu)$ have determinant one,

$$\det R(\varphi) = \cos^2 \varphi + \sin^2 \varphi = 1, \quad \det \Lambda(\mu) = \cosh^2 \mu - \sinh^2 \mu = 1. \quad (3.81)$$

3.5 Matrix Algebra

Multiplication of two matrices A of dimension $p \times m$ and B of dimension $m \times q$ produces a new matrix $C = AB$ of dimension $p \times q$. Each entry of C is the inner product of a row from A and a column from B . Schematically, the product C of two 2×2 matrices is

$$C = \begin{pmatrix} \text{---} \\ \text{---} \end{pmatrix} \begin{pmatrix} | \\ | \\ | \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}, \quad (3.82)$$

upon considering A in terms of its rows and B in terms of its columns. The entries of C satisfy

$$c_{ij} = (a_{i1} \ a_{i2}) \begin{pmatrix} b_{1j} \\ b_{2j} \end{pmatrix} = a_{i1}b_{1j} + a_{i2}b_{2j}. \quad (3.83)$$

The product $D = BA$ of the same 2×2 matrices satisfies

$$D = \begin{pmatrix} \text{---} \\ \text{---} \end{pmatrix} \begin{pmatrix} | \\ | \\ | \end{pmatrix} = \begin{pmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{pmatrix} \quad (3.84)$$

upon considering B in terms of its rows and A in terms of its columns, so that

$$d_{ij} = (b_{i1} \ b_{i2}) \begin{pmatrix} a_{j1} \\ a_{j2} \end{pmatrix} = b_{i1}a_{j1} + b_{i2}a_{j2}. \quad (3.85)$$

It is easy to see that in general $D \neq C$, i.e., matrix multiplication does not commute,

$$[A, B] = AB - BA \neq 0, \quad (3.86)$$

where the notation $[\cdot, \cdot]$ refers to the commutator.

Example 3.3. To illustrate, consider the two matrices

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \quad (3.87)$$

The commutation $[A, B]$ then evaluates to

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = 2 \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (3.88)$$

3.6 Eigenvalue Problems

Eigenvalue problems are defined by the equation

$$A\mathbf{a} = \lambda\mathbf{a}, \quad (3.89)$$

where \mathbf{a} refers to an eigenvector associated with the eigenvalue λ . Equivalently, \mathbf{a} is in the *null-space* (is a right null-vector) of $A - \lambda I$:

$$(A - \lambda I)\mathbf{a} = \mathbf{0}. \quad (3.90)$$

For (3.90) to have a non-trivial solution \mathbf{a} , we must have

$$\det(A - \lambda I) = 0. \quad (3.91)$$

3.6.1 Eigenvalues of $R(\varphi)$

Let us explore (3.90, 3.91) for the rotation matrix $R(\varphi)$,

$$0 = |R - \lambda I| = (\cos \varphi - \lambda)^2 + \sin^2 \varphi : \lambda_{\pm} = \cos \varphi + i \sin \varphi = e^{\pm i\varphi}. \quad (3.92)$$

The eigenvalues are S^1 . It is a consequence of the fact that rotation is unitary (see Sect. 3.7). Also, the eigenvalues satisfy⁶

⁶The product of the eigenvalues equals the determinant of the matrix, as follows from, e.g., the Jordan decomposition theorem. The same theorem shows that the *trace* of a matrix given by the sum of the elements on the principle diagonal equals the sum of the eigenvalues.

$$\lambda_1 \lambda_2 = |R| = 1. \quad (3.93)$$

The associated eigenvectors

$$\mathbf{a} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} \quad (3.94)$$

satisfy (3.90). To be definite, (3.90) defines two homogeneous equations in the two unknown coefficients (α_1, α_2) ,

$$\begin{cases} \alpha_1 \cos \varphi - \alpha_2 \sin \varphi - \lambda \alpha_1 = 0, \\ \alpha_1 \sin \varphi + \alpha_2 \cos \varphi - \lambda \alpha_2 = 0. \end{cases} \quad (3.95)$$

For the eigenvalues satisfying (3.93), these two equations are *linearly dependent*. It suffices to take one of them, to solve for α_1 and α_2 ,

$$\alpha_1 (\cos \varphi - \lambda) - \alpha_2 \sin \varphi = 0 : \alpha_1 = -i \alpha_2, \quad \alpha_1 = i \alpha_2 \quad (3.96)$$

for $\lambda = e^{i\varphi}$ and, respectively, $\lambda = e^{-i\varphi}$. We thus arrive at the eigenvector-eigenvalue pairs

$$\left\{ e^{i\varphi}, \begin{pmatrix} 1 \\ -i \end{pmatrix} \right\}, \left\{ e^{-i\varphi}, \begin{pmatrix} 1 \\ i \end{pmatrix} \right\}. \quad (3.97)$$

These two pairs are complex conjugates. This is no surprise since $R(\varphi)$ is a real matrix, whose determinant $|R - \lambda I|$ defines a quadratic polynomial in λ . With real coefficients, its roots are either both real or a pair of complex conjugates.

3.6.2 Eigenvalues of a Real-Symmetric Matrix

The matrix

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} \quad (3.98)$$

is real-symmetric with eigenvalues-eigenvectors $(\lambda_{\pm}, \mathbf{x}_{\pm})$

$$\left\{ 1 + \sqrt{2}, \begin{pmatrix} 1 + \sqrt{2} \\ 1 \end{pmatrix} \right\}, \left\{ 1 - \sqrt{2}, \begin{pmatrix} 1 - \sqrt{2} \\ 1 \end{pmatrix} \right\}. \quad (3.99)$$

It is readily seen that \mathbf{x}_{\pm} are *orthogonal*:

$$\mathbf{x}_{+}^T \mathbf{x}_{-} = 0. \quad (3.100)$$

We can normalize the eigenvectors to

$$\mathbf{e}_+ = \frac{1}{\sqrt{4 + 2\sqrt{2}}} \begin{pmatrix} 1 + \sqrt{2} \\ 1 \end{pmatrix}, \quad \mathbf{e}_- = \frac{1}{\sqrt{4 - 2\sqrt{2}}} \begin{pmatrix} 1 - \sqrt{2} \\ 1 \end{pmatrix}, \quad (3.101)$$

so that $(\mathbf{e}_+, \mathbf{e}_-)$ forms a new *orthonormal* basis set complementary to (\mathbf{i}, \mathbf{j}) along the x - and y -axis. Hence, we have the general decompositions

$$\mathbf{x} = x\mathbf{i} + y\mathbf{j} = a\mathbf{e}_+ + b\mathbf{e}_-, \quad (3.102)$$

where $x = \mathbf{i} \cdot \mathbf{x}$ and $y = \mathbf{j} \cdot \mathbf{y}$. The coefficients a and b can be read off using multiplication by \mathbf{e}_\pm :

$$a = x\mathbf{i} \cdot \mathbf{e}_+ + y\mathbf{j} \cdot \mathbf{e}_+, \quad b = x\mathbf{i} \cdot \mathbf{e}_- + y\mathbf{j} \cdot \mathbf{e}_-. \quad (3.103)$$

Note that (3.89) defines the eigenvectors as invariant subspaces. We now arrive at a new look at A as an operator on \mathbf{x} in terms of multiplications by eigenvectors along the directions given by the associated eigenvectors,

$$A\mathbf{x} = a\lambda_+\mathbf{e}_+ + b\lambda_-\mathbf{e}_-. \quad (3.104)$$

3.6.3 Hermitian Matrices

Let \dagger denote the *Hermitian conjugate*,⁷ defined as the complex conjugate of the transpose of a matrix element, a column or row vector or a matrix. We define the scalar product of two vectors \mathbf{a} and \mathbf{b} in an n -dimensional vector space by

$$\mathbf{a}^\dagger \mathbf{b} = \bar{a}_1 b_1 + \bar{a}_2 b_2 + \cdots + \bar{a}_n b_n. \quad (3.105)$$

Real-symmetric matrices generalize to complex valued matrices with the same properties of having real eigenvalues and mutually orthogonal eigenvectors associated with different eigenvalues according to (3.116) and, respectively, (3.119). Following the steps of the previous section, these are the self-adjoint or Hermitian matrices satisfying

$$H^\dagger = H, \quad (3.106)$$

defined by transformation of the entries $H_{ij}^\dagger = \bar{H}_{ji}$. Note that applying \dagger twice is an identity operation, i.e., $(A^\dagger)^\dagger = A$ for any $n \times m$ matrix A . Hence, if H is an $n \times n$ matrix, we have

⁷Also referred to as the Hermitian transpose or the conjugate transpose.

$$H = \begin{pmatrix} a_{11} & & L^\dagger \\ & a_{22} & \dots \\ & & \dots \\ L & & a_{nn} \end{pmatrix} \quad (3.107)$$

with real diagonal elements a_{ii} ($i = 1, 2, \dots, n$).

Example 3.5. For instance, the rotation matrix $R(i\mu)$ with imaginary angle $\varphi = i\mu$,

$$H = \begin{pmatrix} \cosh \mu & -i \sinh \mu \\ i \sinh \mu & \cosh \mu \end{pmatrix}, \quad (3.108)$$

is Hermitian. Since $|R(\varphi)| = 1$ for all φ , we have $|H| = 1$ by analytic continuation, which also follows by inspection,

$$|H| = \cosh^2 \mu - \sinh^2 \mu = 1. \quad (3.109)$$

The eigenvalue-eigenvectors obtain by analytic continuation of (3.97), i.e.,

$$\left\{ e^{-\mu}, \begin{pmatrix} 1 \\ -i \end{pmatrix} \right\}, \left\{ e^{\mu}, \begin{pmatrix} 1 \\ i \end{pmatrix} \right\}. \quad (3.110)$$

According to (3.105), the scalar product between the two eigenvectors satisfies

$$\begin{pmatrix} 1 \\ -i \end{pmatrix}^\dagger \begin{pmatrix} 1 \\ i \end{pmatrix} = (1 \ i) \begin{pmatrix} 1 \\ i \end{pmatrix} = 1 + i^2 = 0. \quad (3.111)$$

This result of Example 3.5 is expected, since (3.117–3.119) continues to hold upon replacing T by \dagger , i.e.,

$$\lambda_1 = \lambda_2 \text{ or } \mathbf{a}_1^\dagger \mathbf{a}_2 = 0. \quad (3.112)$$

For a Hermitian matrix, the *eigenvectors of distinct eigenvalues are mutually orthogonal*, where orthogonality is defined according to the inner product (3.105).

Very similar properties of the eigenvalue problem (3.89) appear in the real-symmetric matrix $\Lambda(\mu)$ of (3.80). Again, we will find that the eigenvalues are real and distinct, whose accompanying eigenvectors are mutually orthogonal. These properties hold true for all real-symmetric matrices, as shown by the following.

Consider an eigenvalue-eigenvector pair (λ, \mathbf{a}) to a Hermitian matrix A . Then

$$\mathbf{a}^\dagger A \mathbf{a} = \lambda \mathbf{a}^\dagger \mathbf{a}. \quad (3.113)$$

Here, $\mathbf{a}^\dagger \mathbf{a}$ is real, obtained from the summation of the squared norms of the entries of \mathbf{a} . For (3.94), for example, we have

$$\mathbf{a}^\dagger \mathbf{a} = \bar{\alpha}_1 \alpha_1 + \bar{\alpha}_2 \alpha_2 \geq 0. \quad (3.114)$$

The transpose of the left hand side of (3.113) satisfies

$$\lambda \mathbf{a}^\dagger \mathbf{a} = \mathbf{a}^\dagger A \mathbf{a} = (\mathbf{a}^\dagger A \mathbf{a})^T = \mathbf{a}^T A^T \bar{\mathbf{a}} = \mathbf{a}^T \overline{A^\dagger \mathbf{a}} = \mathbf{a}^T \overline{A \mathbf{a}} = \bar{\lambda} \mathbf{a}^T \bar{\mathbf{a}}. \quad (3.115)$$

and hence $\lambda \mathbf{a}^\dagger \mathbf{a} = \bar{\lambda} \mathbf{a}^T \bar{\mathbf{a}} = \bar{\lambda} \mathbf{a}^\dagger \mathbf{a}$. It follows that the *eigenvalues of a Hermitian matrix are real*:

$$\bar{\lambda} = \lambda, \quad (3.116)$$

since $\mathbf{a}^T \bar{\mathbf{a}} \equiv \mathbf{a}^\dagger \mathbf{a}$.

Following similar arguments, consider

$$A \mathbf{a}_2 = \lambda_2 \mathbf{a}_2 : \mathbf{a}_1^\dagger A \mathbf{a}_2 = \lambda_2 \mathbf{a}_1^\dagger \mathbf{a}_2. \quad (3.117)$$

For a Hermitian A , we have

$$\left(\mathbf{a}_1^\dagger A \mathbf{a}_2 \right)^\dagger = \mathbf{a}_2^\dagger A^\dagger \mathbf{a}_1 = \mathbf{a}_2^\dagger A \mathbf{a}_1 = \lambda_1 \mathbf{a}_2^\dagger \mathbf{a}_1. \quad (3.118)$$

By (3.117, 3.118), we have $\lambda_2 \mathbf{a}_1^\dagger \mathbf{a}_2 = \lambda_1 \mathbf{a}_2^\dagger \mathbf{a}_1$. Since $\mathbf{a}_1^\dagger \mathbf{a}_2 = \mathbf{a}_2^\dagger \mathbf{a}_1$, it follows that

$$\lambda_1 = \lambda_2 \text{ or } \mathbf{a}_1^\dagger \mathbf{a}_2 = 0. \quad (3.119)$$

For a Hermitian matrix, the *eigenvectors of distinct eigenvalues are mutually orthogonal*.

Let us now turn to the example matrix $\Lambda(\mu)$ in (3.89). Its eigenvalues are defined by (3.91) with $A = \Lambda$, that is,

$$0 = |\Lambda - \lambda I| = (\cosh \mu - \lambda)^2 - \sinh^2 \mu, \quad (3.120)$$

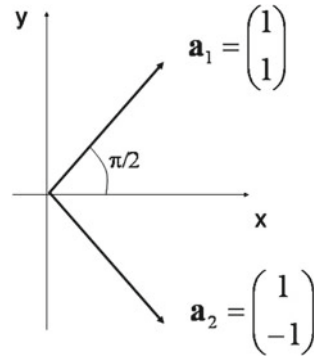
whereby

$$\lambda_{\pm} = \cosh \mu \pm \sinh \mu = \begin{cases} e^{\mu} \\ e^{-\mu} \end{cases}. \quad (3.121)$$

Similar to (3.92), we note

$$\lambda_1 \lambda_2 = |\Lambda| = 1. \quad (3.122)$$

Fig. 3.7 The matrix Λ in (3.80) is real-symmetric. With two distinct eigenvalues, its eigenvectors are orthogonal. As shown, a second eigenvector $\mathbf{a}_2 = \mathbf{i}_x - \mathbf{i}_y$ hereby follows immediately from orthogonality to the first $\mathbf{a}_1 = \mathbf{i}_x + \mathbf{i}_y$



The equation for the eigenvectors (3.95) in terms of (α_1, α_2) are again a linearly dependent system of equations when λ assumes one of the eigenvalues (3.121). Considering the first of (3.95) with $\lambda = e^\mu$,

$$\alpha_1(\cosh \mu - e^\mu) + \alpha_2 \sinh \mu = 0 : \alpha_1 = \alpha_2, \tag{3.123}$$

we obtain the eigenvalue-eigenvector pair

$$\left\{ e^\mu, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}. \tag{3.124}$$

According to (3.119), the eigenvector associated with $\lambda = e^{-\mu}$ is orthogonal to that of (3.124). Since we are working in two dimensions, the second eigenvalue-eigenvector pair is therefore

$$\left\{ e^{-\mu}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\} \tag{3.125}$$

as illustrated in Fig. 3.7. The same obtains by solving (3.123) with e^μ replaced by $e^{-\mu}$.

3.7 Unitary Matrices and Invariants

The reflections and rotations shown in Fig. 3.6 *preserve norm and angles*. If \mathbf{a} and \mathbf{b} are two real vectors and \mathbf{a}' and \mathbf{b}' are their images, e.g.,

$$\mathbf{a}' = R(\varphi)\mathbf{a}, \quad \mathbf{b}' = R(\varphi)\mathbf{b} \tag{3.126}$$

then the inner product

$$\rho = \mathbf{a}^T \mathbf{b} \quad (3.127)$$

is preserved, since

$$(\mathbf{a}')^T \mathbf{b}' = (R(\varphi)\mathbf{a})^T R(\varphi)\mathbf{b} = \mathbf{a}^T R(\varphi)^T R(\varphi)\mathbf{b} = \mathbf{a}^T \mathbf{b} \quad (3.128)$$

by the property of unitarity

$$R(\varphi)^T R(\varphi) = R(-\varphi)R(\varphi) = I. \quad (3.129)$$

In particular, $|\mathbf{a}'|^2 = (\mathbf{a}')^T \mathbf{a}' = |\mathbf{a}|^2$ and, likewise, $|\mathbf{b}'|^2 = |\mathbf{b}|^2$, showing that their norms are preserved. If θ and θ' refer to the angle between (\mathbf{a}, \mathbf{b}) and, respectively, $(\mathbf{a}', \mathbf{b}')$, then

$$|\mathbf{a}'||\mathbf{b}'| \cos \theta' = |\mathbf{a}'||\mathbf{b}'| \cos \theta' = (\mathbf{a}')^T \mathbf{b}' = \mathbf{a}^T \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos \theta, \quad (3.130)$$

which shows that $\cos \theta' = \cos \theta$. Since the norms and angles (between two vectors) are *invariant* under rotations, we say that $R(\phi)$ is *unitary*, defined by the property (3.129).

Generalized to complex valued matrices, we say that A is unitary if

$$A^\dagger A = I, \quad (3.131)$$

by which A is norm and angle preserving following (3.126–3.130) with \dagger replacing T . In a unitary matrix, therefore, the columns and rows form orthonormal sets. This is evident by inspection in the rotation matrix $R(\varphi)$: its row

$$\mathbf{r}_1 = (\cos \varphi \ - \sin \varphi), \ \mathbf{r}_2 = (\sin \varphi \ \cos \varphi) \quad (3.132)$$

and column vectors

$$\mathbf{c}_1 = \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix}, \ \mathbf{c}_2 = \begin{pmatrix} -\sin \varphi \\ \cos \varphi \end{pmatrix} \quad (3.133)$$

satisfy

$$\mathbf{r}_i \mathbf{r}_j^T = \delta_{ij}, \ \mathbf{c}_i^T \mathbf{c}_j = \delta_{ij}, \quad (3.134)$$

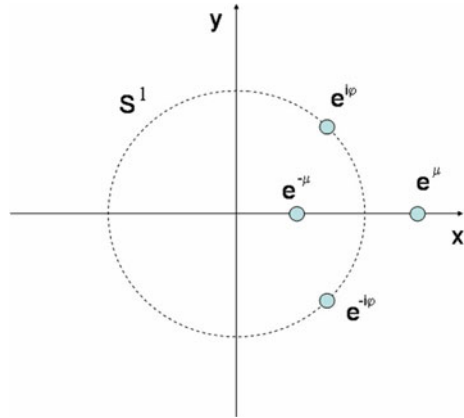
where δ_{ij} denotes the Kronecker delta symbol ($\delta_{ij} = 1$ ($i = j$), $\delta_{ij} = 0$ ($i \neq j$)).

The eigenvalues of a unitary matrix (3.131) are on the unit circle, as follows from

$$\mathbf{a}^\dagger \mathbf{a} = \mathbf{a}^\dagger A^\dagger A \mathbf{a} = (\lambda \mathbf{a})^\dagger (\lambda \mathbf{a}) = |\lambda|^2 \mathbf{a}^\dagger \mathbf{a}, \quad (3.135)$$

where \mathbf{a} denotes an eigenvector of the eigenvalue λ .

Fig. 3.8 Show are the eigenvalues $\lambda_{\pm} = e^{\pm i\varphi}$ of the rotation matrix $R(\varphi)$ on the unit circle S^1 . Since R has a complete orthonormal set of eigenvectors, it is unitary. Shown are also $\lambda_{\pm} = e^{\pm i\mu}$ of Λ . Away from S^1 , Λ is not unitary



The $n \times n$ unitary matrices are $U(n)$. $U(n)$ is a group in that (i) $C = AB$ is in $U(n)$ for any $A, B \in U(n)$, (ii) every $A \in U(n)$ has an inverse $A^{-1} \in U(n)$ and (hence) (iii) $U(n)$ contains the identity matrix I . (Specifically, $AI = IA = A$.) For any two matrices A and B , we have

$$\det AB = \det A \det B = \det BA, \tag{3.136}$$

that we state here without proof. (It may be seen from the fact that the determinant equals the product of eigenvalues.) According to (3.131), unitary matrices hereby satisfy

$$|\det A| = 1. \tag{3.137}$$

Elements of $U(n)$ have a complete set of orthonormal eigenvectors with eigenvalues on the unit circle (Fig. 3.8). The special unitary group $SU(n) \subset U(n)$ have unit determinant,

$$\det A = 1, \tag{3.138}$$

exemplified by the rotation matrices $R(\varphi)$ in (3.129).

In contrast, Hermitian matrices have a complete set of orthonormal eigenvectors with eigenvalues on the real axis. A matrix can be both unitary and Hermitian only if its eigenvalues are ± 1 . Examples are $n \times n$ Householder matrices representing reflections across the plane normal to \mathbf{u} in \mathbb{R}^n ,

$$H = I - 2\mathbf{u}\mathbf{u}^\dagger, \quad \mathbf{u}^\dagger \mathbf{u} = 1. \tag{3.139}$$

3.8 Hermitian Structure of Minkowski Spacetime

A Hermitian $n \times n$ matrix A on \mathbb{C}^n introduces a metric structure through an inner product defined by their real eigenvalues λ_i ,

$$(\mathbf{a}, \mathbf{b}) = \mathbf{a}^\dagger A \mathbf{b} = \sum_{i=1}^n \lambda_i \mathbf{a}^\dagger \mathbf{b}. \quad (3.140)$$

If all eigenvalues are positive, this metric structure introduces a norm equivalent to the Euclidean norm on \mathbf{R}^{2n} ,

$$|a|_* = \sqrt{(\mathbf{a}, \mathbf{a})} = \sqrt{\sum_{i=1}^n \lambda_i |a|^2}. \quad (3.141)$$

The Lorentz metric of Sect. 1.5 is an example of a real-symmetric matrix on \mathbb{R}^4 with signature $(1, -1, -1, -1)$, referring to one positive and three negative eigenvalues. The metric structure it introduces follows (3.140) with A given by η_{ab} , that is referred to as hyperbolic rather than Euclidean. We next strengthen this association to Hermitian matrices with some interesting consequences.

By dimension, we are at liberty to introduce complex combinations of the real 3+1 space-time components of a vector in terms of two complex-valued component vectors. Embedding the latter into a 2×2 Hermitian matrix, (a) the Lorentz metric obtains by the determinant of the matrix and (b) Lorentz transformations correspond to unitary transformations by unimodular matrices from $SL(2, \mathbb{C})$. Remarkably, the unimodular matrices giving a unitary transformation are effectively square roots of Lorentz transformations of four-vectors. For rotations on the unit sphere,⁸ it gives a double cover of the rotations on the unit sphere S^2 .⁹

The causal structure of Minkowski spacetime, defined by the Lorentz metric, is given geometrically by Lorentz invariant light cones. The generators of light cones are light rays. Light rays are integral curves of null-vectors with length zero. This refers to the fact that the change in total phase along a light ray is zero by definition—light rays define the propagation of wave fronts carrying constant total phase of electromagnetic radiation. They carry information on direction, but not distance. Projection of light rays onto the celestial sphere defines a one-to-one map of directions onto S^2 . *Light rays hereby have two degrees of freedom*,¹⁰ and are either future- or past-oriented with opposite signs of their angular velocity in the propagation of an electromagnetic wave. Since the dimension of Minkowski space is four, this suggests a formulation in two null-vectors.

⁸The celestial sphere in the language of cosmology.

⁹Commonly referred to as $SO(3)$, described by rotation matrices with determinant +1.

¹⁰Photons carry an additional degree of freedom in polarization.

Expressed in terms of complex variables, a 2×2 formulation is realized by *spinors* of $(\epsilon_{AB}, \mathbb{C}^2)$, where ϵ_{AB} refers to the metric spinor as follows.

Given a four-vector $k^b = (k^t, k^x, k^y, k^z)$, consider the Hermitian matrix

$$K = \begin{pmatrix} k^t + k^z & k^x - ik^y \\ k^x + ik^y & k^t - k^z \end{pmatrix} = k^t \sigma_t + k^x \sigma_x + k^y \sigma_y + k^z \sigma_z, \quad (3.142)$$

expanded in terms of the Hermitian Pauli spin matrices¹¹

$$\sigma_t = I, \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (3.143)$$

with respective eigenvalues $\lambda = \pm 1$, $\lambda = \pm i$, $\lambda = \pm i$ and $\lambda = \pm i$. The Pauli spin matrices embed the basis vectors of Minkowski space,

$$\begin{aligned} \sigma_t &= K \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\}, & \sigma_x &= K \left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\}, \\ \sigma_y &= K \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right\}, & \sigma_z &= K \left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}. \end{aligned} \quad (3.144)$$

Notice that the σ_i ($i = x, y, z$) are trace-free. From the determinant of Z ,

$$\det K = (k^t)^2 - (k^x)^2 - (k^y)^2 - (k^z)^2, \quad (3.145)$$

the length of k^b in Minkowski space satisfies

$$s^2 = \det K, \quad (3.146)$$

incorporating the line-element $s^2 = \eta_{ab} k^a k^b$ with Minkowski metric

$$\eta_{ab} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \quad (3.147)$$

Here, we use the Einstein summation convention of summing over all index values $a = t, x, y, z$ in combinations of covariant and contravariant indices. In Exercise 1.11, we noticed that η_{ab} reduced to 1+1 is invariant under Lorentz boosts. It is

¹¹Wolfgang Pauli 1900–1958.

not difficult to ascertain that η_{ab} is invariant under general Lorentz transformations including rotations. As such, η_{ab} is a Lorentz invariant tensor.

Consider an element $L \in \text{SL}(2, \mathbb{C})$, mentioned above. These unimodular elements have $8 - 2 = 6$ degrees of freedom. Then

$$K \rightarrow L^\dagger K L \quad (3.148)$$

preserves (3.146), since

$$\det L^\dagger K L = \det K. \quad (3.149)$$

Notice that (3.149) holds true also for $L = -I$, showing that the sign of L is not determined by a given Lorentz transformation of k^b . Even so, a given L from $\text{SL}(2, \mathbb{C})$ in (3.148) defines a Lorentz transformation of k^b .

Example 3.6. Consider a Lorentz boost with rapidity μ of $k^b = (1, 0, 0, 0)^T$ to $(\cosh \mu, 0, 0, \sinh \mu)^T$ along the z -axis,

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} \cosh \mu - \sinh \mu & 0 \\ 0 & \cosh \mu + \sinh \mu \end{pmatrix} = \begin{pmatrix} e^{-\mu} & 0 \\ 0 & e^\mu \end{pmatrix}. \quad (3.150)$$

It obtains by a boost

$$L(\mu) = \begin{pmatrix} e^{-\frac{1}{2}\mu} & 0 \\ 0 & -e^{\frac{1}{2}\mu} \end{pmatrix}, \quad (3.151)$$

whereas a rotation of $k^b = (0, 1, 0, 0)$ to $(0, \cos \theta, \sin \theta, 0)$ about the z -axis,

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & \cos \theta + i \sin \theta \\ \cos \theta - i \sin \theta & 0 \end{pmatrix} = \begin{pmatrix} 0 & e^{i\theta} \\ e^{-i\theta} & 0 \end{pmatrix}, \quad (3.152)$$

obtains by a rotation

$$L(\theta) = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{i\frac{1}{2}\theta} & i e^{-i\frac{1}{2}\theta} \\ i e^{i\frac{1}{2}\theta} & e^{-i\frac{1}{2}\theta} \end{pmatrix}. \quad (3.153)$$

Viewed by continuation starting from the identity matrix, (3.153) goes at the heart of spinors to be introduced below: a rotation over 2π in physical space gives rise to a change in sign in $L(\theta)$. A continuing rotation over 4π restores the original sign. The $L(\theta)$ in (3.153) are elements of $\text{SU}(2)$, since $L^\dagger L = I$. Accordingly, $\text{SU}(2)$ is a two-fold cover of $\text{SO}(3)$.

Light cones are described by null-rays k^b , satisfying

$$s^2 = \det K = 0. \quad (3.154)$$

Their embedding (3.142) is therefore in rank-one matrices¹² (3.142) of the form

$$Z = \begin{pmatrix} \xi \bar{\xi} & \eta \bar{\xi} \\ \bar{\eta} \xi & \eta \bar{\eta} \end{pmatrix} = \begin{pmatrix} \bar{\xi} \\ \bar{\eta} \end{pmatrix} (\xi \eta), \quad (3.155)$$

whose determinant is identically equal to zero. Here, right hand side expresses the *spinor* and its transpose indicated by a primed index

$$\kappa^A = (\xi \eta), \quad \kappa^{A'} = \begin{pmatrix} \xi \\ \eta \end{pmatrix} \quad (\xi, \eta \in \mathbb{C}) \quad (3.156)$$

following the convention, to using unprimed and primed indices for a row and, respectively, column vector notation. Accordingly, we write

$$Z^{AA'} = \kappa^A \bar{\kappa}^{A'}, \quad (3.157)$$

where $\bar{\kappa}^{A'}$ is the Hermitian transpose of κ^A .

Let κ denote the row vector κ^A in (3.156). Then (3.148) implies a corresponding Lorentz transformation of a spinor κ ,

$$Z \rightarrow (\kappa L)^\dagger (\kappa L) : \kappa \rightarrow \kappa L. \quad (3.158)$$

Rotation over 2π in real space now has a corresponding sign change in the spinor.

Now write the determinant of $K = K^{AA'}$ in (3.142) as

$$\det K = K^{11} K^{22} - K^{21} K^{12} \equiv \epsilon_{AB} \epsilon_{A'B'} K^{AA'} K^{BB'} \quad (3.159)$$

in terms of the anti-symmetric metric spinor $\epsilon_{AB} = -\epsilon_{BA}$, $\epsilon_{A'B'} = -\epsilon_{B'A'}$ with $\epsilon_{01} = \epsilon_{0'1'} = 1$, i.e.,

$$\epsilon_{AB} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \epsilon_{A'B'} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \quad (3.160)$$

Then

$$\epsilon_{AB} \epsilon_{A'B'} K^{AA'} K^{BB'} = g_{ab} k^a k^b : \epsilon_{AB} \epsilon_{A'B'} = g_{ab} \quad (3.161)$$

with implicit reference to the basis elements (3.170) to be discussed below. In (3.159), note that the incomplete contraction $\epsilon_{A'B'} K^{AA'} K^{BB'}$ is an antisymmetric tensor in

¹²The image space is comprised of multiples of one vector $(\xi \eta)^\dagger$.

our two-dimensional spinor space \mathbb{C}^2 . Since the antisymmetric elements of $L(2, \mathbb{C})$ are spanned by ϵ_{AB} ,

$$\epsilon_{A'B'} K^{AA'} K^{BB'} = \frac{1}{2} \epsilon^{AB} \det K, \quad (3.162)$$

taking into account (3.159) and $\epsilon_{AB} \epsilon^{AB} = 2$.

The metric spinor allows lowering and raising indices

$$\kappa_A = \kappa^B \epsilon_{BA}, \quad \kappa^B = \epsilon^{BA} \kappa_A. \quad (3.163)$$

Lowering and raising is by multiplication from the left and, respectively, right. The same rules apply to A' . Since the metric spinor is skew symmetric, we automatically have that spinors are null,

$$\kappa^A \kappa_A = \kappa^A \kappa^B \epsilon_{BA} = \kappa^2 \kappa^1 - \kappa^1 \kappa^2 = 0. \quad (3.164)$$

In practical terms, the spinor κ^A is a square root of a null-vector k^b in $Z^{AA'}$.

Consider two null-vectors k^b and l^b represented by spinors o^A and ι^A . Then

$$k^c l_c = (o^A \bar{o}^{A'}) (\iota_A \bar{\iota}^{A'}) = (o^A \iota_A) (o^A \iota_A)^\dagger = |o^A \iota_A|^2 \geq 0, \quad (3.165)$$

whereby $k^c l_c \geq 0$, i.e., k^b and l^b share the same direction in time, e.g., are future-oriented. Choosing two distinct null-vectors, we may insist

$$k^c l_c = 1 : o^A \iota_A = 1. \quad (3.166)$$

As members of \mathbb{C}^2 , choosing such pair as a basis gives

$$\epsilon_{AB} = o_A \iota_B - \iota_A o_B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \quad (3.167)$$

To be explicit, consider

$$o^A = (1 \ 0), \quad \iota^A = (0 \ 1). \quad (3.168)$$

Then $\iota_A = \epsilon_{AB} \iota^B = (1 \ 0)$, whereby (3.166) is satisfied, and

$$\begin{aligned} o^A \bar{o}^{A'} &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \iota^A \bar{\iota}^{A'} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \\ o^A \bar{\iota}^{A'} &= \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \iota^A \bar{o}^{A'} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \end{aligned} \quad (3.169)$$

The result identifies the Pauli spin-matrices with metric spin-tensors

$$\begin{aligned}\sigma_t^{AA'} &= o^A \bar{o}^{A'} + \iota^A \bar{\iota}^{A'}, & \sigma_x^{AA'} &= o^A \bar{\iota}^{A'} + \iota^A \bar{o}^{A'}, \\ \sigma_y^{AA'} &= i (o^A \bar{\iota}^{A'} - \iota^A \bar{o}^{A'}), & \sigma_z^{AA'} &= o^A \bar{o}^{A'} - \iota^A \bar{\iota}^{A'}.\end{aligned}\tag{3.170}$$

In the notation of linear algebra, note that $\bar{o}^{A'} = (o^A)^\dagger$, etc. It recovers Pauli matrices in (3.142) as a basis of four-vectors in 3+1 Minkowski space. Note that (3.170) also introduces an algebraic map of a complex second-rank spinor, e.g., $\phi_{AA'}$, to four-vectors with possibly complex valued components.

3.9 Eigenvectors of Hermitian Matrices

For $n \times n$ Hermitian matrix A , $A^\dagger = A$, let (λ, \mathbf{a}) denote one of its eigenvalue-eigenvector pairs. The latter always exists by virtue of eigenvalue solutions to (3.91). Let

$$\hat{\mathbf{a}} = \frac{\mathbf{a}}{\sqrt{\mathbf{a}^\dagger \mathbf{a}}}\tag{3.171}$$

denote the normalized eigenvector, satisfying $\hat{\mathbf{a}}^\dagger \mathbf{a} = 1$. For instance, we have

$$\mathbf{a} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \rightarrow \hat{\mathbf{a}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.\tag{3.172}$$

Let \mathbf{u} be any vector. We may decompose it orthogonally as

$$\mathbf{u} = \mathbf{u}_\parallel + \mathbf{u}_\perp.\tag{3.173}$$

Here, \mathbf{u}_\parallel and \mathbf{u}_\perp are parallel and orthogonal to \mathbf{a} , obtained from the projection operator

$$P = I - \hat{\mathbf{a}}\hat{\mathbf{a}}^\dagger : \mathbf{u}_\perp = P\mathbf{u}, \mathbf{u}_\parallel = (I - P)\mathbf{u}.\tag{3.174}$$

Geometrically, the *image space* of PA consists of all vectors orthogonal to u_\parallel ,

$$\text{Im}PA = (u_\parallel)^\perp.\tag{3.175}$$

We also note that \mathbf{u}_\perp is in the plane with normal \mathbf{a} . The expansion (3.173) hereby satisfies

$$\mathbf{u} = (I - P)\mathbf{u} + P\mathbf{u},\tag{3.176}$$

and hence

$$A\mathbf{u} = \lambda(I - P)\mathbf{u} + AP\mathbf{u}. \quad (3.177)$$

Since $u_{\parallel} = (I - P)\mathbf{u}$, being parallel to \mathbf{a} , it is an eigenvector of A with eigenvalue λ . Since \mathbf{u} in (3.177) is arbitrary, it follows that

$$A = \lambda(I - P) + AP. \quad (3.178)$$

Since A is Hermitian with λ real, and $P^{\dagger} = P$, we have

$$\lambda(I - P) + AP = A = A^{\dagger} = \lambda(I - P) + PA^{\dagger} = \lambda(I - P) + PA. \quad (3.179)$$

We thus find that A and P commute, i.e.,

$$AP = PA. \quad (3.180)$$

It follows that in particular that

$$\text{Im}PA = PA = (u_{\parallel})^{\perp}. \quad (3.181)$$

Since I and A commute trivially, also $(I - P)$ and A commute: $[(I - P), A] = 0$ and the Hermitian matrix A operates completely independently on the one-dimensional subspace of vectors along an eigenvector \mathbf{a} and on the subspace of vectors in the $n - 1$ dimensional hypersurface normal to \mathbf{a} . Equation (3.180) also shows that PA is Hermitian:

$$(PA)^{\dagger} = A^{\dagger}P^{\dagger} = AP = PA. \quad (3.182)$$

Therefore, we can repeat all the steps (3.171–3.180) for an eigenvector \mathbf{a}' of $A_1 = AP$. Since $AP\mathbf{a}' = PA\mathbf{a}'$, this eigenvector is in the image space of P , and hence it is orthogonal to \mathbf{a} . By this orthogonality, P' associated with \mathbf{a}' commutes with P . It follows that $A_1P' = APP'$ is Hermitian. Continuing in this fashion, we ultimately arrive at n mutually orthogonal eigenvectors $\mathbf{a}, \mathbf{a}', \dots, \mathbf{a}'''\dots'$.

Example 3.7. The 2×2 Hermitian matrix

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} \quad (3.183)$$

has eigenvalue-eigenvector pairs $\{(\lambda_1, \mathbf{a}_1), (\lambda_2, \mathbf{a}_2)\}$ with

$$\left\{ 1 \pm \sqrt{2}, \begin{pmatrix} 1 \pm \sqrt{2} \\ 1 \end{pmatrix} \right\}. \quad (3.184)$$

We wish to view the operation of A on a vector \mathbf{u} as the sum of linearly independent operations associated with the directions \mathbf{a}_1 and \mathbf{a}_2 . We first rewrite (3.185) in terms of the equivalent orthonormal pair

$$\hat{\mathbf{a}}_1 = \frac{\mathbf{a}_1}{\sqrt{\mathbf{a}_1^\dagger \mathbf{a}_1}}, \quad \hat{\mathbf{a}}_2 = \frac{\mathbf{a}_2}{\sqrt{\mathbf{a}_2^\dagger \mathbf{a}_2}} \quad (3.185)$$

following (3.171). The $\{\hat{\mathbf{a}}_1, \hat{\mathbf{a}}_2\}$ form an *orthonormal basis* (a *complete* set of orthonormal vectors) for vectors \mathbf{u} in our two-dimensional space. They satisfy the property

$$\hat{\mathbf{a}}_i^\dagger \hat{\mathbf{a}}_j = \delta_{ij} = \begin{cases} 1 & (i = j) \\ 0 & (i \neq j) \end{cases} \quad (3.186)$$

Here, δ_{ij} is the commonly used Kronecker delta symbol. For (3.183), we have

$$\hat{\mathbf{a}}_{1,2} = \frac{1}{2(2 \pm \sqrt{2})} \begin{pmatrix} 1 \pm \sqrt{2} \\ 1 \end{pmatrix}. \quad (3.187)$$

For an arbitrary vector, we can write

$$\mathbf{u} = \alpha \hat{\mathbf{a}}_1 + \beta \hat{\mathbf{a}}_2. \quad (3.188)$$

Multiplication by $\hat{\mathbf{a}}_{1,2}$ from the left obtains

$$\mathbf{a}_1^\dagger \mathbf{u} = \alpha \mathbf{a}_1^\dagger \hat{\mathbf{a}}_1 + \beta \mathbf{a}_1^\dagger \hat{\mathbf{a}}_2 = \alpha \quad (3.189)$$

$$\mathbf{a}_2^\dagger \mathbf{u} = \alpha \mathbf{a}_2^\dagger \hat{\mathbf{a}}_1 + \beta \mathbf{a}_2^\dagger \hat{\mathbf{a}}_2 = \beta. \quad (3.190)$$

Substitution of (3.190) into (3.188) gives the explicit expression

$$\mathbf{u} = \hat{\mathbf{a}}_1 \hat{\mathbf{a}}_1^\dagger \mathbf{u} + \hat{\mathbf{a}}_2 \hat{\mathbf{a}}_2^\dagger \mathbf{u}. \quad (3.191)$$

This represents the *Gram-Schmidt* orthogonal decomposition of \mathbf{u} with respect to the eigenvectors of A . Accordingly, $A\mathbf{u}$ satisfies

$$A\mathbf{u} = \alpha A\hat{\mathbf{a}}_1 + \beta A\hat{\mathbf{a}}_2 = \alpha \lambda_1 \hat{\mathbf{a}}_1 + \beta \lambda_2 \hat{\mathbf{a}}_2 = \lambda_1 \hat{\mathbf{a}}_1 \hat{\mathbf{a}}_1^\dagger \mathbf{u} + \lambda_2 \hat{\mathbf{a}}_2 \hat{\mathbf{a}}_2^\dagger \mathbf{u}. \quad (3.192)$$

Since \mathbf{u} is arbitrary, we conclude

$$A = \lambda_1 \hat{\mathbf{a}}_1 \hat{\mathbf{a}}_1^\dagger + \lambda_2 \hat{\mathbf{a}}_2 \hat{\mathbf{a}}_2^\dagger. \tag{3.193}$$

The same follows from $A = AI$, $I = \hat{\mathbf{a}}_1 \hat{\mathbf{a}}_1^\dagger + \hat{\mathbf{a}}_2 \hat{\mathbf{a}}_2^\dagger$.

Example 3.8. For Λ in (3.80) we have, according to (3.124, 3.125), a normalized pair of eigenvalues-eigenvectors given by

$$\left\{ e^{\pm\mu}, \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \pm 1 \end{pmatrix} \right\}. \tag{3.194}$$

Following (3.194), we consider

$$\lambda_1 \hat{\mathbf{a}}_1 \hat{\mathbf{a}}_1^\dagger = \frac{e^\mu}{\sqrt{2}} (1 \ 1) \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{e^\mu}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \tag{3.195}$$

$$\lambda_2 \hat{\mathbf{a}}_2 \hat{\mathbf{a}}_2^\dagger = \frac{e^{-\mu}}{\sqrt{2}} (1 \ -1) \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{e^{-\mu}}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

Adding these expressions gives

$$\frac{e^\mu}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \frac{e^{-\mu}}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} e^\mu + e^{-\mu} & e^\mu - e^{-\mu} \\ e^\mu - e^{-\mu} & e^\mu + e^{-\mu} \end{pmatrix}, \tag{3.196}$$

i.e., we recover our definition of a Lorentz boost,

$$A = \begin{pmatrix} \cosh \mu & \sinh \mu \\ \sinh \mu & \cosh \mu \end{pmatrix}. \tag{3.197}$$

3.10 QR Factorization

In viewing an $n \times m$ matrix A as a *linear map* from the vector space \mathbb{C}^m to \mathbb{C}^n , we frequently encounter the question if A is imaging \mathbb{C}^m onto all of \mathbb{C}^n or just a linear subspace of it. Similarly, A may map all nonzero vectors from \mathbb{C}^m to nonzero vectors or map some linear subspace of \mathbb{C}^m to the origin $\mathbf{0}$ in \mathbb{C}^n .

To streamline this discussion, we introduce the image of A , defined by the linear vector space

$$\text{Im } A = \{ \mathbf{v} \mid \mathbf{v} = A\mathbf{u}, \mathbf{u} \in \mathbb{C}^m \} \tag{3.198}$$

and the *kernel* of A , also known as the null space of A , defined by the linear vector space

$$\text{Ker } A = \{\mathbf{u} \mid \mathbf{A}\mathbf{u} = \mathbf{0}, \mathbf{u} \in \mathbb{C}^m\}. \quad (3.199)$$

The image space is supported by the column vectors $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m$ of A , e.g., for a 2×2 matrix

$$A = \left(\begin{array}{|c|} \hline | \\ \hline | \\ \hline | \\ \hline \end{array} \right) = (\mathbf{a}_1 \ \mathbf{a}_2), \quad (3.200)$$

The image space forms out of linear combinations

$$\mathbf{A}\mathbf{u} = u_1\mathbf{a}_1 + u_2\mathbf{a}_2 + \dots + u_m\mathbf{a}_m \quad (3.201)$$

by choice of \mathbf{u} in \mathbb{C}^m ,

$$\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ \cdot \\ u_m \end{pmatrix}. \quad (3.202)$$

The row space is supported by the row vectors $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n$ of A , e.g., for a 2×2 matrix

$$A = \left(\begin{array}{c} \text{---} \\ \text{---} \end{array} \right) = \begin{pmatrix} \mathbf{b}_1^T \\ \mathbf{b}_2^T \end{pmatrix} \quad (3.203)$$

that forms out of the linear combinations

$$\mathbf{r} = v_1\mathbf{b}_1 + v_2\mathbf{b}_2 + \dots + v_n\mathbf{b}_n. \quad (3.204)$$

by choice of \mathbf{v} in \mathbb{C}^m ,

$$\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \cdot \\ v_n \end{pmatrix}. \quad (3.205)$$

Following (3.199), the kernel of A consists of the vectors that are orthogonal to all of its row vectors \mathbf{b}_i ($i = 1, 2, \dots, n$), commonly written as

$$\text{Ker } A = (\text{Im } A^\dagger)^\perp, \quad (3.206)$$

where \perp refers to orthogonality with respect to the inner product $\mathbf{a} \cdot \mathbf{b} = \mathbf{a}^\dagger \mathbf{b}$ for vectors in \mathbb{C}^m .

3.10.1 Examples of Image and Null Space

Let A be a nonsingular 2×2 matrix. We read off the columns vectors following (3.200), e.g.,

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix}: \mathbf{a}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \mathbf{a}_2 = \begin{pmatrix} 2 \\ 0 \end{pmatrix}. \quad (3.207)$$

Hence, $\text{Im } A$ is defined by all vectors obtained from linear combinations of the \mathbf{a}_1 and \mathbf{a}_2 in (3.207). We say

$$\text{Im } A = \text{span}\{\mathbf{a}_1, \mathbf{a}_2\} = \text{span} \left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \end{pmatrix} \right\}. \quad (3.208)$$

Evidently, we have

$$\text{Im } A = \mathbb{R}^2 \quad (3.209)$$

(or \mathbb{C}^2) since the \mathbf{a}_1 and \mathbf{a}_2 in (3.207) point in different directions, whereby they are linearly independent. This may also be inferred from the fact that $\det A \neq 0$.

Alternatively, consider the singular matrix

$$B = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}: \mathbf{a}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \mathbf{a}_2 = \begin{pmatrix} 2 \\ 4 \end{pmatrix}. \quad (3.210)$$

In this event, the second column satisfies

$$\mathbf{a}_2 = 2\mathbf{a}_1 \quad (3.211)$$

and the two columns are linearly dependent, as follows also from the fact that

$$\det B = 0. \quad (3.212)$$

Consequently, the image space is the one dimensional subspace given by

$$\text{Im } B = \text{span}\{\mathbf{a}_1, \mathbf{a}_2\} = \text{span} \left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\}. \quad (3.213)$$

Proceeding with (3.207) above, we have, following (3.203), the row vectors

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix}: \mathbf{b}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \mathbf{b}_2 = \begin{pmatrix} 2 \\ 0 \end{pmatrix}. \quad (3.214)$$

They happen to be the same as the column vectors since A is real-symmetric. $\text{Ker } A$ is defined by vectors that are orthogonal to both \mathbf{b}_1 and \mathbf{b}_2 . Since \mathbf{b}_1 and \mathbf{b}_2 are linearly

independent, we have

$$\text{Ker } A = \mathbf{0}. \quad (3.215)$$

For the alternative (3.210), we have the row vectors

$$B = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}: \quad \mathbf{b}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad \mathbf{b}_2 = \begin{pmatrix} 2 \\ 4 \end{pmatrix}. \quad (3.216)$$

In this event, the second row satisfies $\mathbf{b}_2 = 2\mathbf{b}_1$, whereby they are linearly dependent. Consequently, the null space of A is the one dimensional subspace, given by the vectors orthogonal to \mathbf{b}_1 , i.e.,

$$\text{Ker } B = \begin{pmatrix} 1 \\ 2 \end{pmatrix}^\perp = \text{span} \left\{ \begin{pmatrix} -2 \\ 1 \end{pmatrix} \right\}. \quad (3.217)$$

The matrices (3.207) and (3.210) satisfy, respectively,

$$\dim \text{Im } A + \dim \text{Ker } A = 2 + 0 = 2, \quad (3.218)$$

$$\dim \text{Im } B + \dim \text{Ker } B = 1 + 1 = 2.$$

3.10.2 Dimensions of Image and Null Space

In what follows, we will restrict our discussion to square matrices of size $n \times n$. In this event, the dimension $\dim \text{Im } A$ of (3.198) is n whenever A is of full rank, i.e., when $\det A \neq 0$. Complementary to this, we have that $\dim \text{Ker } A$ of (3.199) is 0 whenever A is of full rank, i.e., when $\det A \neq 0$. However, $\dim \text{Im } A < n$ and $\dim \text{Ker } A > 0$ when $\det A = 0$.

The matrices (3.207) and (3.210) exemplify a general relationship of $n \times n$ matrices, satisfying

$$\dim \text{Im } A + \dim \text{Ker } A = n \quad (3.219)$$

To derive this relationship, we begin by observing the invariance

$$\text{Im } A = \text{Im } A' \quad (3.220)$$

for $A' = (\mathbf{a}'_1 \mathbf{a}'_2 \dots \mathbf{a}'_n)$ obtained from $A = (\mathbf{a}_1 \mathbf{a}_2 \dots \mathbf{a}_n)$ by changing a column vector \mathbf{a}_j by linear superposition with any of the other column vectors. Specifically, this may be by choice of $1 \leq j \leq n$ and a linear superposition

$$\mathbf{a}'_i = \mathbf{a}_i \quad (i \neq j), \quad \mathbf{a}'_j = \mathbf{a}_j - \sum_{i=1}^{j-1} \mu_i \mathbf{a}_i. \quad (3.221)$$

This transformation has a corresponding upper triangular transformation matrix U such that $A' = AU$. For instance, when $n = 3$ and $j = 2$,

$$U_2 = \begin{pmatrix} 1 & -\mu_1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad U_3 = \begin{pmatrix} 1 & 0 & -\mu'_1 \\ 0 & 1 & -\mu'_2 \\ 0 & 0 & 1 \end{pmatrix}. \quad (3.222)$$

Since the AU_2 and AU_2U_3 form as superpositions of the column vectors of A , their image space remains $\text{Im } A$.

The *Gram-Schmidt orthogonalization* of \mathbf{a}_j from A to mutually orthogonalized \mathbf{a}_i ($1 \leq i \leq j - 1$) satisfies

$$\mathbf{a}'_j = \mathbf{a}_j - \sum_{i=1}^{j-1} \mu_i \mathbf{a}_i, \quad \mu_i = \frac{\mathbf{a}'_j \mathbf{a}_i}{\mathbf{a}'_i \mathbf{a}_i}. \quad (3.223)$$

Here, we omit projections $\mu_i \mathbf{a}_i$ whenever $\mathbf{a}_i = \mathbf{0}$. Performing (3.223) for each $j = 2, 3, \dots$ consecutively up to $j = n$ produces $A'^{\dots'}$ with column vectors that are all orthogonal. For 2×2 matrix A , the Gram-Schmidt orthogonalization of its column vectors obtains in one step

$$A' = AU_2 = \left(\begin{array}{c|c} \left| \right| & \left| \right| \\ \left| \right| & \left| \right| \end{array} \right) \begin{pmatrix} 1 & -\mu_1 \\ 0 & 1 \end{pmatrix}, \quad (3.224)$$

If A has full rank, then so has A' . This may also be seen from the product rule

$$\det A' = \det A \det U = \det A, \quad (3.225)$$

since $\det U = 1$. The determinant of A' is nonzero iff the determinant of A is nonzero. For a 3×3 matrix, we apply (3.222). The product $U = U_2U_3$ is upper triangular,

$$U = U_2U_3 = \begin{pmatrix} 1 & -\mu_1 & -\mu'_1 + \mu_1\mu'_2 \\ 0 & 1 & -\mu'_2 \\ 0 & 0 & 1 \end{pmatrix}. \quad (3.226)$$

The above is readily extended to $n \times n$ matrices

$$A' = AU = AU_2U_3 \cdots U_n \quad (3.227)$$

whose columns are mutually orthogonal, where U is upper triangular with unit determinant.

If $\det A = 0$, some of the columns of A' are zero, i.e.,

$$A' = (\mathbf{a}_1 \ \mathbf{a}_2 \ \dots \ \mathbf{0} \ \dots \ \mathbf{a}_j \ \dots \ \mathbf{0} \ \dots \ \mathbf{a}_n). \quad (3.228)$$

The null vectors of A' are of the form $\mathbf{u}' = (0 \ \dots \ 1 \ \dots \ 0)^T$, where 1 appears at a position j where $\mathbf{a}_j = \mathbf{0}$. Since U is invertible, $A = A'U^{-1}$, and hence $\mathbf{u} = U\mathbf{u}'$ is a null vector of A . Our theorem (3.219) now readily follows: the number of non-zero columns in A' define the dimension of the image space of A and the number of zero columns of A' define the dimension of the null space of A .

Example 3.9. Consider the non-singular matrix

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 0 \\ 0 & 2 & 3 \end{pmatrix}. \quad (3.229)$$

The first and second steps in (3.223) produce, respectively,

$$A' = AU_2 = \begin{pmatrix} 1 & 6/5 & 1 \\ 2 & -3/5 & 0 \\ 0 & 2 & 3 \end{pmatrix}, \quad U_2 = \begin{pmatrix} 1 & -4/5 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (3.230)$$

$$A'' = A'U_3 = \begin{pmatrix} 1 & 6/5 & -\frac{20}{29} \\ 2 & -3/5 & \frac{10}{29} \\ 0 & 2 & \frac{15}{29} \end{pmatrix}, \quad U_3 = \begin{pmatrix} 1 & 0 & -1/5 \\ 0 & 1 & -\frac{36}{29} \\ 0 & 0 & 1 \end{pmatrix}. \quad (3.231)$$

It follows that

$$\begin{pmatrix} 1 & 6/5 & -\frac{20}{29} \\ 2 & -3/5 & \frac{10}{29} \\ 0 & 2 & \frac{15}{29} \end{pmatrix} = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 0 \\ 0 & 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & -4/5 & \frac{23}{29} \\ 0 & 1 & -\frac{36}{29} \\ 0 & 0 & 1 \end{pmatrix}, \quad (3.232)$$

where the second matrix on the right hand side is $U = U_2U_3$. Similarly, consider the singular matrix

$$B = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 0 \\ 0 & 0 & 3 \end{pmatrix}. \quad (3.233)$$

The first and second step in (3.223) produce, respectively,

$$B' = BU_2 \begin{pmatrix} 1 & 0 & 1 \\ 2 & 0 & 0 \\ 0 & 0 & 3 \end{pmatrix}, \quad U_2 = \begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (3.234)$$

$$B'' = B'U_3 = \begin{pmatrix} 1 & 0 & 4/5 \\ 2 & 0 & -2/5 \\ 0 & 0 & 3 \end{pmatrix}, \quad U_3 = \begin{pmatrix} 1 & 0 & -1/5 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (3.235)$$

It follows that

$$\begin{pmatrix} 1 & 0 & 4/5 \\ 2 & 0 & -2/5 \\ 0 & 0 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & -2 & -1/5 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (3.236)$$

where the second matrix on the right hand side is $U = U_1U_2$.

Comparing (3.232–3.236), we see that A in (3.229) has full rank with the trivial null space $\text{Ker } A = \mathbf{0}$, whereas B in (3.233) is of rank 2 with the nontrivial null space given by the second column of $U = U_2U_3$, i.e.,

$$\text{Ker } B = \text{span } U \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \text{span } \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}. \quad (3.237)$$

3.10.3 QR Factorization by Gram-Schmidt

The above is more commonly used to derive the QR factorization of a matrix upon including normalization in each step of the Gram-Schmidt procedure,

$$\mathbf{a}'_j \rightarrow \hat{\mathbf{a}}_j = \frac{\mathbf{a}'_j}{\sqrt{(\mathbf{a}')^\dagger \mathbf{a}'}} \quad (3.238)$$

if $\mathbf{a}'_j \neq \mathbf{0}$ (otherwise, we skip this step). The result $A = QR$ has column vectors of Q forming an orthonormal bases for $\text{Im } A$ and R upper triangular. If A is square and invertible, then Q is unitary, $Q^\dagger Q = I$, whereby $R = Q^\dagger A$.

Example 3.10. Consider the QR factorizations of the non-singular 2×2 matrix. Let $A' = AU_2$ be the outcome of the Gram-Schmidt procedure and D_2 denote the diagonal matrix D_2 containing the norms its column vectors. Then $A' = QD_2$ defines

$$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} \sqrt{5} & \frac{4}{\sqrt{5}} \\ 0 & \frac{3}{\sqrt{5}} \end{pmatrix} \equiv QR. \quad (3.239)$$

For a singular matrix, we similarly obtain

$$\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} \sqrt{5} & 2\sqrt{5} \\ 0 & 0 \end{pmatrix}. \quad (3.240)$$

For the A and B in (3.229) and (3.233), the QR factorizations are

$$A = \begin{pmatrix} 1/5 \sqrt{5} & \frac{6}{145} \sqrt{145} & -\frac{4}{29} \sqrt{29} \\ 2/5 \sqrt{5} & -\frac{3}{145} \sqrt{145} & \frac{2}{29} \sqrt{29} \\ 0 & \frac{2}{29} \sqrt{145} & \frac{3}{29} \sqrt{29} \end{pmatrix} \begin{pmatrix} \sqrt{5} & 4/\sqrt{5} & 1/\sqrt{5} \\ 0 & 1/\sqrt{145} & \frac{36}{145} \sqrt{145} \\ 0 & 0 & \frac{5}{29} \sqrt{29} \end{pmatrix}, \quad (3.241)$$

$$B = \begin{pmatrix} 1/\sqrt{5} & 2/\sqrt{5} & 0 \\ 2/\sqrt{5} & -1/\sqrt{5} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{5} & 4/\sqrt{5} & 1/\sqrt{5} \\ 0 & 3/\sqrt{5} & 2/\sqrt{5} \\ 0 & 0 & 3 \end{pmatrix}. \quad (3.242)$$

Table 3.1 summarizes this discussion.

3.11 Exercises

3.1. Let

$$\mathbf{a} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, \quad \mathbf{c} = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}. \quad (3.243)$$

Calculate

Table 3.1 Matrices and some symmetry properties

1. Angular momentum is a vector of dimension mass times area per unit time. In flat spacetime, angular momentum is proportional to angular velocity relative to the distant stars. Angular momentum is a conserved quantity. Torque is the time rate of change of angular momentum.
2. An $n \times m$ matrix A of n row and m column vectors may be real-symmetric ($A^T = A$), skew-symmetric ($A^T = -A$), Hermitian ($A^\dagger = A$) or unitary ($A^\dagger A = I$). A real-symmetric or Hermitian matrix has real eigenvalues and eigenvectors with different eigenvalues are orthogonal.
3. The product of eigenvalues is the determinant of a matrix A . If A is unitary ($A^\dagger A = I$) or Hermitian ($A^\dagger = A$), the eigenvalues are on the unit circle, respectively, the real line. The Householder matrix (3.139) exemplifies a unitary and real-symmetric matrix with eigenvalues ± 1 . An $n \times n$ Hermitian matrix A has a complete set of orthonormal eigenvectors, i.e., an orthonormal basis of \mathbb{R}^n .
4. Every vector \mathbf{u} has a unique Gram-Schmidt decomposition in terms of a sum of vectors along orthonormal eigenvectors $\hat{\mathbf{a}}$ of A . A operates on \mathbf{u} by multiplication along $\hat{\mathbf{a}}$ by its real eigenvalues.
5. The Minkowski metric on four-vectors is invariant under Lorentz transformations, preserving length of and inner products between four-vectors. They are realized by elements of $SL(2, \mathbb{C})$ acting on spinors, given by rows of two complex numbers. Spinors are “square roots” of null-vectors, that change sign following rotation over 2π in Minkowski space.
6. The image space of an $n \times m$ matrix A is spanned by its m column vectors \mathbf{a}_i , $\text{Im } A = \text{span}\{\mathbf{a}_i\}$, and the null space of A is orthogonal to the span of its n row vectors \mathbf{b}_i , $\text{Ker } A = (\text{span}\{\mathbf{b}_i\})^\perp$. Square matrices of $n \times n$ satisfy (3.219), $\dim \text{Im } A + \dim \text{Ker } A = n$. If $\det A = 0$, image space has rank less than n and the null space of A has positive rank: $\dim \text{Im } A \leq n - 1$, $\dim \text{Ker } A \geq 1$.
7. Image and null space obtain by Gram-Schmidt orthogonalization on column vectors of a matrix A , producing $A' = AU$ or $A = QR$ by upper triangular matrices U and R satisfying $\text{Im } A' = \text{Im } A = \text{Im } Q$. The null space of A is given by the orthogonal complement to the span of A 's columns in U , by zero-columns $\mathbf{a}'_i = \mathbf{0}$ in A' .

$$(i) : \mathbf{a} \times \mathbf{b} \cdot \mathbf{c}, \quad \mathbf{a} \cdot \mathbf{b} \times \mathbf{c}; \quad (ii) : \mathbf{a} \times (\mathbf{b} \times \mathbf{c}), \quad (\mathbf{a} \times \mathbf{b}) \times \mathbf{c}. \quad (3.244)$$

Compare your answers to (i) and explain.

3.2. Consider a Cartesian coordinate system (x, y, z) and rotation of a vector \mathbf{r} about the z -axis with angular velocity $\omega = \Omega \mathbf{i}_z$, where \mathbf{i}_z denotes the unit vector along the z -axis. The velocity of \mathbf{r} satisfies $\mathbf{v} = \omega \times \mathbf{r}$, where \times denotes the outer product.

(i) If $\mathbf{r} = 2\mathbf{i}_x + 3\mathbf{i}_z$, calculate the velocity \mathbf{v} .

(ii) Show that $|\mathbf{v}| = \Omega\sigma$, where σ denotes the distance to the axis of rotation.

3.3. Derive the equivalent expression for the dip angle (3.245), given by

$$\sin \theta = \frac{M\sigma^2\omega_p}{J} = \frac{M^2g\sigma^3}{J^2} = \left(\frac{g}{\sigma}\right) \left(\frac{\sigma}{b}\right)^4 \frac{1}{\omega^2}. \quad (3.245)$$

3.4. For the each of the following transformations in the two-dimensional plane, state which are projections, reflections and rotations:

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad A_4 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad (3.246)$$

$$A_5 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

3.5. Show that complex numbers $z = x + iy$ can be written in terms of the matrices

$$A(z) = \begin{pmatrix} x & -y \\ y & x \end{pmatrix} \quad (3.247)$$

satisfying $A(z)A(w) = A(zw)$ by the rules of matrix multiplication. In particular, show that for $z = i$, (3.247) satisfies $I + A^2 = 0$, where I denotes the identity matrix ($z = 1$).

3.6. Consider the matrix

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix}. \quad (3.248)$$

Compute the, determinant, the eigenvalues and eigenvectors.

3.7. Permutation of the x - and y -coordinates is described by

$$\mathbf{z} = x\mathbf{i}_x + y\mathbf{i}_y \rightarrow \mathbf{w} = y\mathbf{i}_x + z\mathbf{i}_y, \quad w = y + ix. \quad (3.249)$$

Show that $w(z)$ is not analytic in z , i.e., the Cauchy-Riemann relations are not satisfied. Derive the equivalent 2×2 matrix equation for $x' = y$ and $y' = x$.

3.8. Consider the matrix

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}. \quad (3.250)$$

Obtain the eigenvalues λ_i and eigenvectors \mathbf{a}_i ($i = 1, 2$) and decompose A in the form

$$A = \lambda_1 A_1 + \lambda_2 A_2, \quad A_i = \hat{\mathbf{a}}_i \hat{\mathbf{a}}_i^\dagger, \quad (3.251)$$

where hat refers to normalization to unit norm.

3.9. Show the orthogonality (3.119).

3.10. If A is both unitary and Hermitian, show that $A = A^{-1}$.

3.11. Consider the Householder matrix (3.139). Show that H is Hermitian ($H^\dagger = H$) and unitary ($H^\dagger H = I$), whence it is *involuntary* (H is its own inverse): $H^2 = I$. In two dimensions, determine its eigenvalue-eigenvector pairs for a general direction \mathbf{u} and interpret the result geometrically. What happens in three dimensions to the multiplicity of the eigenvalues?

3.12. Let A be Hermitian, i.e., $A^\dagger = A$. Show that A is diagonalizable according to

$$A = U \Lambda U^\dagger, \quad (3.252)$$

where U is the unitary matrix satisfying $U^\dagger U = I$. Compute U for A in (3.250). [*Hint*: Compose U from the eigenvectors of A .]

3.13. Consider the matrix

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 4 + a \end{pmatrix}. \quad (3.253)$$

Compute the determinant and determine the *condition number* of A , defined by the ratio of the maximal to the minimal square root of the eigenvalues of $A^\dagger A$. [*Hint*: use (3.252).] What happens when a approaches zero? Compute the solution to the system of equations

$$A\mathbf{u} = \mathbf{v}. \quad (3.254)$$

Show that the solution is regular, respectively, ill-behaved as a approaches zero when

$$\mathbf{v} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \end{pmatrix}. \tag{3.255}$$

What is the condition number of A^2 . How does it generalize to A^n ($n \geq 3$)?

3.14. Show that $U(1)$ can be identified with the tangents of complex numbers on the unit circle S^1 . [Hint: Express elements of S^1 by $e^{i\theta}$ and generalize the Taylor expansion θ ,

$$e^{i\theta} = 1 + i\theta + O(\theta^2), \tag{3.256}$$

about the identity $\theta = 0$ to arbitrary θ .]¹³

3.15. Show that $U(1)$ is *abelian*:

$$AB - BA = 0 \quad (A, B \in U(1)). \tag{3.257}$$

3.16. Show that the elements of $U(2)$ are of the form

$$A = e^{i\theta} \begin{pmatrix} z & -\bar{w} \\ w & \bar{z} \end{pmatrix}, \quad z\bar{z} + w\bar{w} = 1 \tag{3.258}$$

and that they are in general not Hermitian.

3.17. Following (3.258), specialize to $\det A = 1$.¹⁴ Give a general representation of $SU(2)$ in terms of *traceless* 2×2 matrices (the sum of diagonal elements being zero). Determine the number of degrees of freedom in view of the conditions $A^\dagger A = I$ and $\det A = 1$. Show that these traceless matrices are Hermitian and derive their eigenvalues.

3.18. Illustrate a double cover of S^1 by way of a curve on a two-torus.

3.19. From the definition of the inner product of two arbitrary spinors o^A and ι^A and the definition of lowering indices, show that

$$o^A \iota_A = -o_A \iota^A. \tag{3.259}$$

¹³ S^1 is illustrative of a one-dimensional manifold which is compact and simply connected. It has nontrivial topology, since the winding number of a loop in S^1 can take any value in \mathbb{Z} . By homotopy, the topology of S^1 is the same as that of the punctured disk $0 < |z| \leq 1$.

¹⁴ As a result, we say $SU(2) \subset U(2) \cong SU(2) \times U(1)$.

3.20. In (3.165), obtain ι^A from a rotation of o^A and evaluate their inner product as a function of the rotation angle. Next, consider a spinor basis

$$o^A = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \end{pmatrix}, \quad \iota^A = \frac{1}{\sqrt{2}} \begin{pmatrix} i & 1 \end{pmatrix}. \quad (3.260)$$

Show that $o^A \iota_{A'} = 1$ in (3.166). Rank-one matrices of the form $o^A \bar{\iota}^{A'}$ may be expanded as¹⁵

$$o^A \bar{\iota}^{A'} = \frac{1}{2} \left[\begin{pmatrix} -i \\ 1 \end{pmatrix} \begin{pmatrix} 1 & i \end{pmatrix} \right]^T = \begin{pmatrix} -i & 1 \\ 1 & i \end{pmatrix}^T = \begin{pmatrix} -i & 1 \\ 1 & i \end{pmatrix}. \quad (3.261)$$

where T denotes the ordinary matrix transpose. Express the Pauli spin matrices in this new basis similar to (3.170).

3.21. Occasionally, we allow coordinates to become complex. For reference, recall the line-element

$$ds^2 = dx^2 + dy^2 = dr^2 + r^2 d\theta^2 \quad (3.262)$$

of the Euclidean plane, expressed in Cartesian and, respectively, polar coordinates. The Euclidean is flat, like an ordinary sheet of paper. Consider¹⁶

$$ds^2 = -x^2 dt^2 + dx^2. \quad (3.263)$$

Show that (3.263) again is the line-element of a flat two-surface using analytic continuation in t .¹⁷

3.22. Consider the matrix

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 4 + a \end{pmatrix}. \quad (3.264)$$

(i) Obtain the image space and the null space of the matrix for all a . [*Hint*: Distinguish between $a = 0$ and $a \neq 0$.];

(ii) Apply Gram-Schmidt orthogonalization to obtain $A' = AU$, where the column factors of A' are orthogonal and U is upper triangular;

(iii) Obtain the QR factorization of A .

¹⁵The symbols $o^A \bar{\iota}^{A'}$ and $\bar{\iota}^{A'} o^A$ are the same, i.e., there is no ordering between unprimed and primed indices. Only upon expansion into a matrix, a choice of ordering is made.

¹⁶A so-called Rindler space.

¹⁷It can be shown that flatness is preserved under analytic continuation, whereby the Lorentz metric $ds^2 = -dt^2 + dx^2 = (idt)^2 + dx^2$ is trivially flat.

3.23. Following (3.239), obtain the QR factorizations of the 2×2 rotation matrix $R(\varphi)$ and the Lorentz boost $\Lambda(\mu)$ of Example 3.8.

3.24. Obtain the QR factorizations of general 2×2 matrices that are (a) Hermitian or (b) unitary.

3.25. Let $i = 0, 1, 2, 3$ correspond to (t, x, y, z) . For the Pauli spin matrices (3.143), show or calculate

(i) The σ_i are involutory: $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = -i\sigma_1\sigma_2\sigma_3 = I$; $\det \sigma_i = -1$ for $i = 1, 2, 3$ and they are trace-free, $\text{Tr}(\sigma_i) = 0$.

(ii) The σ_i satisfy $\sigma_i\sigma_j + \sigma_j\sigma_i = 2\delta_{ij}I$ ($i, j = 1, 2, 3$).

(iii) The σ_a ($a = 0, 1, 2, 3$) form a basis of the 2×2 Hermitian matrices.

(iv) The eigenvalues and eigenvectors of the σ_i ($i = 1, 2, 3$).

(v) The commutator $[\sigma_i, \sigma_j]$ for all $i, j = 1, 2, 3$.

References

1. Everitt, C.W.F., et al, 2011, Phys. Rev. Lett., 106, 221101, <http://www.einstein.stanford.edu>.
2. Ciufolini, L., & Pavlis, E.C., 2004, Nature, 431, 958.
3. Kerr, R.P., 1963, Phys. Rev. Lett., 11, 237.
4. van Putten, M.H.P.M., 2005, Nuov. Cim. B, 28, 597; van Putten, M.H.P.M., & Gupta, A.C., 2009, Mon. Not. R. Astron. Soc., 394, 2238.
5. Papapetrou A., 1951, Proc. R. Soc., 209, 248.
6. Feynman, R.P., 1963, *Lectures on Physics*, Vol. I (Addison-Wesley Publishing Co.), Ch. 20.
7. van Putten, M.H.P.M., 2017, NewA, 54, 115, [arXiv:1609.07474](https://arxiv.org/abs/1609.07474).