

Elements of Geometry for Robotics



Tomas Pajdla
pajdla@cmp.felk.cvut.cz

Tuesday 21st October, 2014

Contents

1	Notation	1
2	Linear algebra	2
2.1	Change of coordinates induced by the change of basis	2
2.2	Determinant	5
2.2.1	Permutation	5
2.2.2	Determinant	6
2.3	Vector product	7
2.4	Dual space and dual basis	9
2.5	Operations with matrices	12
3	Affine space	15
3.1	Vectors	15
3.1.1	Geometric scalars	16
3.1.2	Geometric vectors	17
3.1.3	Bound vectors	17
3.2	Linear space	18
3.3	Free vectors	19
3.4	Affine space	20
3.5	Coordinate system in affine space	23
3.6	An example of affine space	24
3.6.1	Affine space of solutions of a system of linear equations	24
4	Motion	26
4.1	Change of position vector coordinates induced by motion	26
4.2	Rotation matrix	27
4.3	Coordinate vectors	31
5	Rotation	32
5.1	Properties of rotation matrix	32
5.1.1	Inverse of \mathbf{R}	32
5.1.2	Eigenvalues of \mathbf{R}	32
5.1.3	Eigenvectors of \mathbf{R}	33
5.1.4	Rotation axis	37
5.1.5	Rotation angle	38
5.1.6	Matrix $(\mathbf{R} - \mathbf{I})$	39
6	Axis of Motion	40

7	Rotation representation and parameterization	43
7.1	Angle-axis representation of rotation	43
7.2	Euler vector representation and the exponential map	47
7.3	Quaternion representation of rotation	48
7.4	Topology of the space of rotations	55
	Index	59

1 Notation

\emptyset	...	the empty set [1]
$\exp U$...	the set of all subsets of set U [1]
$U \times V$...	Cartesian product of sets U and V [1]
\mathbb{Z}	...	whole numbers [1]
\mathbb{Q}	...	rational numbers [2]
\mathbb{R}	...	real numbers [2]
i	...	imaginary unit [2]
$(S, +, \cdot)$...	space of geometric scalars
A	...	affine space (space of geometric vectors)
(A_o, \oplus, \odot)	...	space of geometric vectors bound to point o
(V, \boxplus, \boxminus)	...	space of free vectors
\mathbb{A}^2	...	real affine plane
\mathbb{A}^3	...	three-dimensional real affine space
\mathbb{P}^2	...	real projective plane
\mathbb{P}^3	...	three-dimensional real projective space
\vec{x}	...	vector
A	...	matrix
A_{ij}	...	ij element of A
A^\top	...	transpose of A
$ A $...	determinant of A
I	...	identity matrix
R	...	rotation matrix
\otimes	...	Kronecker product of matrices
$\beta = [\vec{b}_1, \vec{b}_2, \vec{b}_3]$...	basis (an ordered triple of independent generator vectors)
$\beta^*, \tilde{\beta}$...	the dual basis to basis β
\vec{x}_β	...	column matrix of coordinates of \vec{x} w.r.t. the basis β
$\vec{x} \cdot \vec{y}$...	Euclidean scalar product of \vec{x} and \vec{y} ($\vec{x} \cdot \vec{y} = \vec{x}_\beta^\top \vec{y}_\beta$ in an orthonormal basis β)
$\vec{x} \times \vec{y}$...	cross (vector) product of \vec{x} and \vec{y}
$[\vec{x}]_\times$...	the matrix such that $[\vec{x}]_\times \vec{y} = \vec{x} \times \vec{y}$
$\ \vec{x}\ $...	Euclidean norm of \vec{x} ($\ \vec{x}\ = \sqrt{\vec{x} \cdot \vec{x}}$)
orthogonal vectors	...	mutually perpendicular and all of equal length
orthonormal vectors	...	unit orthogonal vectors
$P \circ l$...	point P is incident to line l
$P \vee Q$...	line(s) incident to points P and Q
$k \wedge l$...	point(s) incident to lines k and l

2 Linear algebra

We rely on linear algebra [3, 4, 5, 6, 7, 8]. We recommend excellent text books [6, 3] for acquiring basic as well as more advanced elements of the topic. Monograph [4] provides a number of examples and applications and provides a link to numerical and computational aspects of linear algebra. We will next review the most crucial topics needed in this text.

2.1 Change of coordinates induced by the change of basis

Let us discuss the relationship between the coordinates of a vector in a linear space, which is induced by passing from one basis to another. We shall derive the relationship between the coordinates in a three-dimensional linear space over real numbers, which is the most important when modeling the geometry around us. The formulas for all other n -dimensional spaces are obtained by passing from 3 to n .

§ 1 Coordinates Let us consider an ordered basis $\beta = [\vec{b}_1 \ \vec{b}_2 \ \vec{b}_3]$ of a three-dimensional vector space V^3 over scalars \mathbb{R} . A vector $\vec{v} \in V^3$ is uniquely expressed as a linear combination of basic vectors of V^3 by its *coordinates* $x, y, z \in \mathbb{R}$, i.e. $\vec{v} = x \vec{b}_1 + y \vec{b}_2 + z \vec{b}_3$, and can be represented as an ordered triple of coordinates, i.e. as $\vec{v}_\beta = [x \ y \ z]^\top$.

We see that an ordered triple of scalars can be understood as a triple of coordinates of a vector in V^3 w.r.t. a basis of V^3 . However, at the same time, the set of ordered triples $[x \ y \ z]^\top$ is also a three-dimensional *coordinate linear space* \mathbb{R}^3 over \mathbb{R} with $[x_1 \ y_1 \ z_1]^\top + [x_2 \ y_2 \ z_2]^\top = [x_1 + x_2 \ y_1 + y_2 \ z_1 + z_2]^\top$ and $s [x \ y \ z]^\top = [sx \ sy \ sz]^\top$ for $s \in \mathbb{R}$. Moreover, the ordered triple of the following three particular coordinate vectors

$$\sigma = \left[\begin{array}{c|c|c} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \end{array} \right] \quad (2.1)$$

forms an ordered basis of \mathbb{R}^3 , the *standard basis*, and therefore a vector $\vec{v} = [x \ y \ z]^\top$ is represented by $\vec{v}_\sigma = [x \ y \ z]^\top$ w.r.t. the standard basis in \mathbb{R}^3 . It is noticeable that the vector \vec{v} and the coordinate vector \vec{v}_σ of its coordinates w.r.t. the standard basis of \mathbb{R}^3 , are identical.

§ 2 Two bases Having two ordered bases $\beta = [\vec{b}_1 \ \vec{b}_2 \ \vec{b}_3]$ and $\beta' = [\vec{b}'_1 \ \vec{b}'_2 \ \vec{b}'_3]$ leads to expressing one vector \vec{x} in two ways as $\vec{x} = x \vec{b}_1 + y \vec{b}_2 + z \vec{b}_3$ and $\vec{x} = x' \vec{b}'_1 + y' \vec{b}'_2 + z' \vec{b}'_3$. The vectors of the basis β can also be expressed in the basis β'

using their coordinates. Let us introduce

$$\begin{aligned}\vec{b}_1 &= a_{11} \vec{b}'_1 + a_{21} \vec{b}'_2 + a_{31} \vec{b}'_3 \\ \vec{b}_2 &= a_{12} \vec{b}'_1 + a_{22} \vec{b}'_2 + a_{32} \vec{b}'_3 \\ \vec{b}_3 &= a_{13} \vec{b}'_1 + a_{23} \vec{b}'_2 + a_{33} \vec{b}'_3\end{aligned}\tag{2.2}$$

§3 Change of coordinates We will next use the above equations to relate the coordinates of \vec{x} w.r.t. the basis β to the coordinates of \vec{x} w.r.t. the basis β'

$$\begin{aligned}\vec{x} &= x \vec{b}_1 + y \vec{b}_2 + z \vec{b}_3 \\ &= x (a_{11} \vec{b}'_1 + a_{21} \vec{b}'_2 + a_{31} \vec{b}'_3) + y (a_{12} \vec{b}'_1 + a_{22} \vec{b}'_2 + a_{32} \vec{b}'_3) + z (a_{13} \vec{b}'_1 + a_{23} \vec{b}'_2 + a_{33} \vec{b}'_3) \\ &= (a_{11} x + a_{12} y + a_{13} z) \vec{b}'_1 + (a_{21} x + a_{22} y + a_{23} z) \vec{b}'_2 + (a_{31} x + a_{32} y + a_{33} z) \vec{b}'_3 \\ &= x' \vec{b}'_1 + y' \vec{b}'_2 + z' \vec{b}'_3\end{aligned}\tag{2.3}$$

Since coordinates are unique, we get

$$x' = a_{11} x + a_{12} y + a_{13} z\tag{2.4}$$

$$y' = a_{21} x + a_{22} y + a_{23} z\tag{2.5}$$

$$z' = a_{31} x + a_{32} y + a_{33} z\tag{2.6}$$

Coordinate vectors \vec{x}_β and $\vec{x}_{\beta'}$ are thus related by the following matrix multiplication

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}\tag{2.7}$$

which we concisely write as

$$\vec{x}_{\beta'} = \mathbf{A} \vec{x}_\beta\tag{2.8}$$

The columns of matrix \mathbf{A} can be viewed as vectors of coordinates of basic vectors, $\vec{b}_1, \vec{b}_2, \vec{b}_3$ of β in the basis β'

$$\mathbf{A} = \begin{bmatrix} \left| \begin{array}{c} \vec{b}_1 \\ \vec{b}_{1\beta'} \end{array} \right| & \left| \begin{array}{c} \vec{b}_2 \\ \vec{b}_{2\beta'} \end{array} \right| & \left| \begin{array}{c} \vec{b}_3 \\ \vec{b}_{3\beta'} \end{array} \right| \end{bmatrix}\tag{2.9}$$

and the matrix multiplication can be interpreted as a linear combination of the columns of \mathbf{A} by coordinates of \vec{x} w.r.t. β

$$\vec{x}_{\beta'} = x \vec{b}_{1\beta'} + y \vec{b}_{2\beta'} + z \vec{b}_{3\beta'}\tag{2.10}$$

Matrix \mathbf{A} plays such an important role here that it deserves its own name. Matrix \mathbf{A} is very often called the *change of basis matrix from basis β to β'* or the *transition matrix from basis β to basis β'* [4, 9] since it can be used to pass from coordinates w.r.t. β to coordinates w.r.t. β' by Equation 2.8.

However, literature [5, 10] calls \mathbf{A} the *change of basis matrix from basis β' to β* , i.e. it (seemingly illogically) swaps the bases. This choice is motivated by the fact

that \mathbf{A} relates vectors of β and vectors of β' by Equation 2.2 as

$$\begin{bmatrix} \vec{b}_1 & \vec{b}_2 & \vec{b}_3 \end{bmatrix} = \begin{bmatrix} a_{11}\vec{b}'_1 + a_{21}\vec{b}'_2 + a_{31}\vec{b}'_3 & a_{12}\vec{b}'_1 + a_{22}\vec{b}'_2 + a_{32}\vec{b}'_3 & a_{13}\vec{b}'_1 + a_{23}\vec{b}'_2 + a_{33}\vec{b}'_3 \end{bmatrix} \quad (2.11)$$

$$\begin{bmatrix} \vec{b}_1 & \vec{b}_2 & \vec{b}_3 \end{bmatrix} = \begin{bmatrix} \vec{b}'_1 & \vec{b}'_2 & \vec{b}'_3 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \quad (2.12)$$

$$(2.13)$$

and therefore giving

$$\begin{bmatrix} \vec{b}_1 & \vec{b}_2 & \vec{b}_3 \end{bmatrix} = \begin{bmatrix} \vec{b}'_1 & \vec{b}'_2 & \vec{b}'_3 \end{bmatrix} \mathbf{A} \quad (2.14)$$

or equivalently

$$\begin{bmatrix} \vec{b}'_1 & \vec{b}'_2 & \vec{b}'_3 \end{bmatrix} = \begin{bmatrix} \vec{b}_1 & \vec{b}_2 & \vec{b}_3 \end{bmatrix} \mathbf{A}^{-1} \quad (2.15)$$

where the multiplication of a row of column vectors by a matrix from the right in Equation 2.14 has the meaning given by Equation 2.11 above. Yet another variation of the naming appeared in [7, 8] where \mathbf{A}^{-1} was named the *change of basis matrix from basis β to β'* .

We have to conclude that the meaning associated with the *change of basis matrix* varies in the literature and hence we will avoid this confusing name and talk about \mathbf{A} as about the *matrix transforming coordinates of a vector from basis β to basis β'* .

There is the following interesting variation of Equation 2.14

$$\begin{bmatrix} \vec{b}'_1 \\ \vec{b}'_2 \\ \vec{b}'_3 \end{bmatrix} = \mathbf{A}^{-\top} \begin{bmatrix} \vec{b}_1 \\ \vec{b}_2 \\ \vec{b}_3 \end{bmatrix} \quad (2.16)$$

where the basic vectors of β and β' are understood as elements of column vectors. For instance, vector \vec{b}'_1 is obtained as

$$\vec{b}'_1 = a_{11}^* \vec{b}_1 + a_{12}^* \vec{b}_2 + a_{13}^* \vec{b}_3 \quad (2.17)$$

where $[a_{11}^*, a_{12}^*, a_{13}^*]$ is the first row of $\mathbf{A}^{-\top}$.

§ 4 Example We demonstrate the relationship between vectors and bases on a concrete example. Consider two bases α and β represented by coordinate vectors, which we write into matrices

$$\alpha = \begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \vec{a}_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \quad (2.18)$$

$$\beta = \begin{bmatrix} \vec{b}_1 & \vec{b}_2 & \vec{b}_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \quad (2.19)$$

and a vector \vec{x} with coordinates w.r.t. the basis α

$$\vec{x}_\alpha = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad (2.20)$$

We see that basic vectors of α can be obtained as the following linear combinations of basic vectors of β

$$\vec{a}_1 = +1\vec{b}_1 + 0\vec{b}_2 + 0\vec{b}_3 \quad (2.21)$$

$$\vec{a}_2 = +1\vec{b}_1 - 1\vec{b}_2 + 1\vec{b}_3 \quad (2.22)$$

$$\vec{a}_3 = -1\vec{b}_1 + 0\vec{b}_2 + 1\vec{b}_3 \quad (2.23)$$

$$(2.24)$$

or equivalently

$$[\vec{a}_1 \ \vec{a}_2 \ \vec{a}_3] = [\vec{b}_1 \ \vec{b}_2 \ \vec{b}_3] \begin{bmatrix} 1 & 1 & -1 \\ 0 & -1 & 0 \\ 0 & 1 & 1 \end{bmatrix} = [\vec{b}_1 \ \vec{b}_2 \ \vec{b}_3] \mathbf{A} \quad (2.25)$$

Coordinates of \vec{x} w.r.t. β are hence obtained as

$$\vec{x}_\beta = \mathbf{A} \vec{x}_\alpha, \quad \mathbf{A} = \begin{bmatrix} 1 & 1 & -1 \\ 0 & -1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \quad (2.26)$$

$$\begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & -1 \\ 0 & -1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad (2.27)$$

We see that

$$\alpha = \beta \mathbf{A} \quad (2.28)$$

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & -1 \\ 0 & -1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \quad (2.29)$$

The following questions arises: When are the coordinates of a vector \vec{x} (Equation 2.8) and the basic vectors themselves (Equation 2.16) transformed in the same way? In other words, when $\mathbf{A} = \mathbf{A}^{-\top}$. We shall give the answer to this question later in paragraph 2.4.

2.2 Determinant

Determinat [3] of a matrix \mathbf{A} , denoted by $|\mathbf{A}|$, is a very interesting and useful concept. It can be, for instance, used to check the linear independence of a set of vectors or to define an orientation of the space.

2.2.1 Permutation

A *permutation* [3] π on the set $[n] = \{1, \dots, n\}$ of integers is a one-to-one function from $[n]$ onto $[n]$. The identity permutation will be denoted by ϵ , i.e. $\epsilon(i) = i$ for all $i \in [n]$.

§ 5 Composition of permutations Let σ and π be two permutations on $[n]$. Then, their composition, i.e. $\pi(\sigma)$, is also a permutation on $[n]$ since a composition of two one-to-one onto functions is a one-to-one onto function.

§ 6 Sign of a permutation We will now introduce another important concept related to permutations. *Sign*, $\text{sgn}(\pi)$, of a permutation π is defined as

$$\text{sgn}(\pi) = (-1)^{N(\pi)} \quad (2.30)$$

where $N(\pi)$ is equal to the number of *inversions* in π , i.e. the number of pairs $[i, j]$ such that $i, j \in [n]$, $i < j$ and $\pi(i) > \pi(j)$.

2.2.2 Determinant

Let S_n be the set of all permutations on $[n]$ and \mathbf{A} be an $n \times n$ matrix. Then, *determinant* $|\mathbf{A}|$ of \mathbf{A} is defined by the formula

$$|\mathbf{A}| = \sum_{\pi \in S_n} \text{sgn}(\pi) \mathbf{A}_{1,\pi(1)} \mathbf{A}_{2,\pi(2)} \cdots \mathbf{A}_{n,\pi(n)} \quad (2.31)$$

Notice that for every $\pi \in S_n$ and for $j \in [n]$ there is exactly one $i \in [n]$ such that $j = \pi(i)$. Hence

$$\{[1, \pi(1)], [2, \pi(2)], \dots, [n, \pi(n)]\} = \{[\pi^{-1}(1), 1], [\pi^{-1}(2), 2], \dots, [\pi^{-1}(n), n]\} \quad (2.32)$$

and since the multiplication of elements of \mathbf{A} is commutative we get

$$|\mathbf{A}| = \sum_{\pi \in S_n} \text{sgn}(\pi) \mathbf{A}_{\pi^{-1}(1),1} \mathbf{A}_{\pi^{-1}(2),2} \cdots \mathbf{A}_{\pi^{-1}(n),n} \quad (2.33)$$

Let us next define a submatrix of \mathbf{A} and find its determinant. Consider $k \leq n$ and two one-to-one *monotonic* functions $\rho, \nu: [k] \rightarrow [n]$, $i < j \Rightarrow \rho(i) < \rho(j)$, $\nu(i) < \nu(j)$. We define $k \times k$ submatrix $\mathbf{A}^{\rho,\nu}$ of an $n \times n$ matrix \mathbf{A} by

$$\mathbf{A}_{i,j}^{\rho,\nu} = \mathbf{A}_{\rho(i),\nu(j)} \quad \text{for } i, j \in [k] \quad (2.34)$$

We get the determinant of $\mathbf{A}^{\rho,\nu}$ as follows

$$|\mathbf{A}^{\rho,\nu}| = \sum_{\pi \in S_k} \text{sgn}(\pi) \mathbf{A}_{1,\pi(1)}^{\rho,\nu} \mathbf{A}_{2,\pi(2)}^{\rho,\nu} \cdots \mathbf{A}_{k,\pi(k)}^{\rho,\nu} \quad (2.35)$$

$$= \sum_{\pi \in S_k} \text{sgn}(\pi) \mathbf{A}_{\rho(1),\nu(\pi(1))} \mathbf{A}_{\rho(2),\nu(\pi(2))} \cdots \mathbf{A}_{\rho(k),\nu(\pi(k))} \quad (2.36)$$

Let us next split the rows of the matrix \mathbf{A} into two groups of k and m rows and find the relationship between $|\mathbf{A}|$ and the determinants of certain $k \times k$ and $m \times m$ submatrices of \mathbf{A} . Take $1 \leq k, m \leq n$ such that $k + m = n$ and define a one-to-one function $\rho: [m] \rightarrow [k + 1, n] = \{k + 1, \dots, n\}$, by $\rho(i) = k + i$. Next, let $\Omega \subseteq \exp [n]$ be the set of all subsets of $[n]$ of size k . Let $\omega \in \Omega$. Then, there is exactly one one-to-one monotonic function φ_ω from $[k]$ onto ω since $[k]$ and ω are finite sets of integers of the same size. Let $\bar{\omega} = [n] \setminus \omega$. Then, there is exactly one one-to-one monotonic function $\varphi_{\bar{\omega}}$ from $[k + 1, n]$ onto $\bar{\omega}$. Let further there be $\pi_k \in S_k$ and

$\pi_m \in S_m$. With the notation introduced above, we are getting a version of the generalized Laplace expansion of the determinant [11, 12]

$$|\mathbf{A}| = \sum_{\omega \in \Omega} \left(\prod_{i \in [k], j \in [k+1, n]} \text{sgn}(\varphi_{\bar{\omega}}(j) - \varphi_{\omega}(i)) \right) |\mathbf{A}^{\varepsilon, \varphi_{\omega}}| |\mathbf{A}^{\rho, \varphi_{\bar{\omega}}(\rho)}| \quad (2.37)$$

2.3 Vector product

Let us look at an interesting mapping from $\mathbb{R}^3 \times \mathbb{R}^3$ to \mathbb{R}^3 , the *vector product* in \mathbb{R}^3 [6] (which it also often called the cross product [4]). Vector product has interesting geometrical properties but we shall motivate it by its connection to systems of linear equations.

§7 Vector product Assume two linearly independent coordinate vectors $\vec{x} = [x_1 \ x_2 \ x_3]^\top$ and $\vec{y} = [y_1 \ y_2 \ y_3]^\top$ in \mathbb{R}^3 . The following system of linear equations

$$\begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{bmatrix} \vec{z} = 0 \quad (2.38)$$

has a one-dimensional subspace V of solutions in \mathbb{R}^3 . The solutions can be written as multiples of one non-zero vector \vec{w} , the basis of V , i.e.

$$\vec{z} = \lambda \vec{w}, \quad \lambda \in \mathbb{R} \quad (2.39)$$

Let us see how we can construct \vec{w} in a convenient way from vectors \vec{x}, \vec{y} .

Consider determinants of two matrices constructed from the matrix of the system (2.38) by adjoining its first, resp. second, row to the matrix of the system (2.38)

$$\left| \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ x_1 & x_2 & x_3 \end{bmatrix} \right| = 0 \quad \left| \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ y_1 & y_2 & y_3 \end{bmatrix} \right| = 0 \quad (2.40)$$

which gives

$$x_1(x_2 y_3 - x_3 y_2) + x_2(x_3 y_1 - x_1 y_3) + x_3(x_1 y_2 - x_2 y_1) = 0 \quad (2.41)$$

$$y_1(x_2 y_3 - x_3 y_2) + y_2(x_3 y_1 - x_1 y_3) + y_3(x_1 y_2 - x_2 y_1) = 0 \quad (2.42)$$

and can be rewritten as

$$\begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{bmatrix} \begin{bmatrix} x_2 y_3 - x_3 y_2 \\ -x_1 y_3 + x_3 y_1 \\ x_1 y_2 - x_2 y_1 \end{bmatrix} = 0 \quad (2.43)$$

We see that vector

$$\vec{w} = \begin{bmatrix} x_2 y_3 - x_3 y_2 \\ -x_1 y_3 + x_3 y_1 \\ x_1 y_2 - x_2 y_1 \end{bmatrix} \quad (2.44)$$

solves Equation 2.38.

Notice that elements of \vec{w} are the three two by two minors of the matrix of the system (2.38). The rank of the matrix is two, which means that at least one of the minors is non-zero, and hence \vec{w} is also non-zero. We see that \vec{w} is a basic vector of V . Formula 2.44 is known as the *vector product* in \mathbb{R}^3 and \vec{w} is also often denoted by $\vec{x} \times \vec{y}$.

§ 8 Vector product under the change of basis Let us next study the behavior of the vector product under the change of basis in \mathbb{R}^3 . Let us have two bases β, β' in \mathbb{R}^3 and two vectors \vec{x}, \vec{y} with coordinates $\vec{x}_\beta = [x_1 \ x_2 \ x_3]^\top, \vec{y}_\beta = [y_1 \ y_2 \ y_3]^\top$ and $\vec{x}_{\beta'} = [x'_1 \ x'_2 \ x'_3]^\top, \vec{y}_{\beta'} = [y'_1 \ y'_2 \ y'_3]^\top$. We introduce

$$\vec{x}_\beta \times \vec{y}_\beta = \begin{bmatrix} x_2 y_3 - x_3 y_2 \\ -x_1 y_3 + x_3 y_1 \\ x_1 y_2 - x_2 y_1 \end{bmatrix} \quad \vec{x}_{\beta'} \times \vec{y}_{\beta'} = \begin{bmatrix} x'_2 y'_3 - x'_3 y'_2 \\ -x'_1 y'_3 + x'_3 y'_1 \\ x'_1 y'_2 - x'_2 y'_1 \end{bmatrix} \quad (2.45)$$

To find the relationship between $\vec{x}_\beta \times \vec{y}_\beta$ and $\vec{x}_{\beta'} \times \vec{y}_{\beta'}$, we will use the following fact. For every three vectors $\vec{x} = [x_1 \ x_2 \ x_3]^\top, \vec{y} = [y_1 \ y_2 \ y_3]^\top, \vec{z} = [z_1 \ z_2 \ z_3]^\top$ in \mathbb{R}^3 there holds

$$\vec{z}^\top (\vec{x} \times \vec{y}) = [z_1 \ z_2 \ z_3] \begin{bmatrix} x_2 y_3 - x_3 y_2 \\ -x_1 y_3 + x_3 y_1 \\ x_1 y_2 - x_2 y_1 \end{bmatrix} = \begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix} = \begin{vmatrix} \vec{x}^\top \\ \vec{y}^\top \\ \vec{z}^\top \end{vmatrix} \quad (2.46)$$

We can write

$$\begin{aligned} \vec{x}_{\beta'} \times \vec{y}_{\beta'} &= \begin{bmatrix} [1 \ 0 \ 0] (\vec{x}_{\beta'} \times \vec{y}_{\beta'}) \\ [0 \ 1 \ 0] (\vec{x}_{\beta'} \times \vec{y}_{\beta'}) \\ [0 \ 0 \ 1] (\vec{x}_{\beta'} \times \vec{y}_{\beta'}) \end{bmatrix} = \begin{bmatrix} \left\| \begin{bmatrix} \vec{x}_{\beta'}^\top \\ \vec{y}_{\beta'}^\top \\ 1 \ 0 \ 0 \end{bmatrix} \right\| \\ \left\| \begin{bmatrix} \vec{x}_{\beta'}^\top \\ \vec{y}_{\beta'}^\top \\ 0 \ 1 \ 0 \end{bmatrix} \right\| \\ \left\| \begin{bmatrix} \vec{x}_{\beta'}^\top \\ \vec{y}_{\beta'}^\top \\ 0 \ 0 \ 1 \end{bmatrix} \right\| \end{bmatrix}^\top \\ &= \begin{bmatrix} \left\| \begin{bmatrix} \vec{x}_\beta^\top \mathbf{A}^\top \\ \vec{y}_\beta^\top \mathbf{A}^\top \\ 1 \ 0 \ 0 \end{bmatrix} \right\| \\ \left\| \begin{bmatrix} \vec{x}_\beta^\top \mathbf{A}^\top \\ \vec{y}_\beta^\top \mathbf{A}^\top \\ 0 \ 1 \ 0 \end{bmatrix} \right\| \\ \left\| \begin{bmatrix} \vec{x}_\beta^\top \mathbf{A}^\top \\ \vec{y}_\beta^\top \mathbf{A}^\top \\ 0 \ 0 \ 1 \end{bmatrix} \right\| \end{bmatrix}^\top \\ &= \begin{bmatrix} \left\| \begin{bmatrix} \vec{x}_\beta^\top \\ \vec{y}_\beta^\top \\ [1 \ 0 \ 0] \mathbf{A}^{-\top} \end{bmatrix} \mathbf{A}^\top \right\| \\ \left\| \begin{bmatrix} \vec{x}_\beta^\top \\ \vec{y}_\beta^\top \\ [0 \ 1 \ 0] \mathbf{A}^{-\top} \end{bmatrix} \mathbf{A}^\top \right\| \\ \left\| \begin{bmatrix} \vec{x}_\beta^\top \\ \vec{y}_\beta^\top \\ [0 \ 0 \ 1] \mathbf{A}^{-\top} \end{bmatrix} \mathbf{A}^\top \right\| \end{bmatrix}^\top \\ &= \begin{bmatrix} [1 \ 0 \ 0] \mathbf{A}^{-\top} (\vec{x}_\beta \times \vec{y}_\beta) \\ [0 \ 1 \ 0] \mathbf{A}^{-\top} (\vec{x}_\beta \times \vec{y}_\beta) \\ [0 \ 0 \ 1] \mathbf{A}^{-\top} (\vec{x}_\beta \times \vec{y}_\beta) \end{bmatrix} |\mathbf{A}^\top| \\ &= \frac{\mathbf{A}^{-\top}}{|\mathbf{A}^{-\top}|} (\vec{x}_\beta \times \vec{y}_\beta) \end{aligned} \quad (2.47)$$

§ 9 Vector product as a linear mapping It is interesting to see that for all $\vec{x}, \vec{y} \in \mathbb{R}^3$ there holds

$$\vec{x} \times \vec{y} = \begin{bmatrix} x_2 y_3 - x_3 y_2 \\ -x_1 y_3 + x_3 y_1 \\ x_1 y_2 - x_2 y_1 \end{bmatrix} = \begin{bmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \quad (2.48)$$

and thus we can introduce matrix

$$[\vec{x}]_\times = \begin{bmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{bmatrix} \quad (2.49)$$

and write

$$\vec{x} \times \vec{y} = [\vec{x}]_\times \vec{y} \quad (2.50)$$

Notice also that $[\vec{x}]_x^\top = -[\vec{x}]_x$ and therefore

$$(\vec{x} \times \vec{y})^\top = ([\vec{x}]_x \vec{y})^\top = -\vec{y}^\top [\vec{x}]_x \quad (2.51)$$

The result of §8 can also be written in the formalism of this paragraph. We can write for every $\vec{x}, \vec{y} \in \mathbb{R}^3$

$$[\mathbf{A} \vec{x}_\beta]_x \mathbf{A} \vec{y}_\beta = (\mathbf{A} \vec{x}_\beta) \times (\mathbf{A} \vec{y}_\beta) = \frac{\mathbf{A}^{-\top}}{|\mathbf{A}^{-\top}|} (\vec{x}_\beta \times \vec{y}_\beta) = \frac{\mathbf{A}^{-\top}}{|\mathbf{A}^{-\top}|} [\vec{x}_\beta]_x \vec{y}_\beta \quad (2.52)$$

and hence we get for every $\vec{x} \in \mathbb{R}^3$

$$[\mathbf{A} \vec{x}_\beta]_x \mathbf{A} = \frac{\mathbf{A}^{-\top}}{|\mathbf{A}^{-\top}|} [\vec{x}_\beta]_x \quad (2.53)$$

2.4 Dual space and dual basis

Let us start with a three-dimensional linear space L over scalars S and consider the set L^* of all linear functions $f: L \rightarrow S$, i.e. the functions on L for which the following holds true

$$f(a\vec{x} + b\vec{y}) = a f(\vec{x}) + b f(\vec{y}) \quad (2.54)$$

for all $a, b \in S$ and all $\vec{x}, \vec{y} \in L$.

Let us next define the addition $+^*: L^* \times L^* \rightarrow L^*$ of linear functions $f, g \in L^*$ and the multiplication $\cdot^*: S \times L^* \rightarrow L^*$ of a linear function $f \in L^*$ by a scalar $a \in S$ such that

$$(f +^* g)(\vec{x}) = f(\vec{x}) + g(\vec{x}) \quad (2.55)$$

$$(a \cdot^* f)(\vec{x}) = a f(\vec{x}) \quad (2.56)$$

holds true for all $a \in S$ and for all $\vec{x} \in L$. One can verify that $(L^*, +^*, \cdot^*)$ over $(S, +, \cdot)$ is itself a linear space [3, 6, 5]. It makes therefore a good sense to use arrows above symbols for linear functions, e.g. \vec{f} instead of f .

The linear space L^* is derived from, and naturally connected to, the linear space L and hence deserves a special name. Linear space L^* is called [3] the *dual (linear) space* to L .

Now, consider a basis $\beta = [\vec{b}_1, \vec{b}_2, \vec{b}_3]$ of L . We will construct a basis β^* of L^* , in a certain natural and useful way. Let us take three linear functions $\vec{b}_1^*, \vec{b}_2^*, \vec{b}_3^* \in L^*$ such that

$$\begin{aligned} \vec{b}_1^*(\vec{b}_1) &= 1 & \vec{b}_1^*(\vec{b}_2) &= 0 & \vec{b}_1^*(\vec{b}_3) &= 0 \\ \vec{b}_2^*(\vec{b}_1) &= 0 & \vec{b}_2^*(\vec{b}_2) &= 1 & \vec{b}_2^*(\vec{b}_3) &= 0 \\ \vec{b}_3^*(\vec{b}_1) &= 0 & \vec{b}_3^*(\vec{b}_2) &= 0 & \vec{b}_3^*(\vec{b}_3) &= 1 \end{aligned} \quad (2.57)$$

where 0 and 1 are the zero and the unit element of S , respectively. First of all, one has to verify [3] that such an assignment is possible with linear functions over L . Secondly one can show [3] that functions $\vec{b}_1^*, \vec{b}_2^*, \vec{b}_3^*$ are determined by this assignment uniquely on all vectors of L . Finally, one can observe [3] that the triple $\beta^* = [\vec{b}_1^*, \vec{b}_2^*, \vec{b}_3^*]$ forms an (ordered) basis of L^* . The basis β^* is called the *dual basis* of L^* , i.e. it is the basis of L^* , which is related in a special (dual) way to the basis β of L .

§ 10 Evaluating linear functions Consider a vector $\vec{x} \in L$ with coordinates $\vec{x}_\beta = [x_1, x_2, x_3]^\top$ w.r.t. a basis $\beta = [\vec{b}_1, \vec{b}_2, \vec{b}_3]$ and a linear function $\vec{h} \in L^*$ with coordinates $\vec{h}_{\beta^*} = [h_1, h_2, h_3]^\top$ w.r.t. the dual basis $\beta^* = [\vec{b}_1^*, \vec{b}_2^*, \vec{b}_3^*]$. The value $\vec{h}(\vec{x}) \in S$ is obtained from the coordinates \vec{x}_β and \vec{h}_{β^*} as

$$\vec{h}(\vec{x}) = \vec{h}(x_1 \vec{b}_1 + x_2 \vec{b}_2 + x_3 \vec{b}_3) \quad (2.58)$$

$$= (h_1 \vec{b}_1^* + h_2 \vec{b}_2^* + h_3 \vec{b}_3^*)(x_1 \vec{b}_1 + x_2 \vec{b}_2 + x_3 \vec{b}_3) \quad (2.59)$$

$$\begin{aligned} &= h_1 \vec{b}_1^*(\vec{b}_1) x_1 + h_1 \vec{b}_1^*(\vec{b}_2) x_2 + h_1 \vec{b}_1^*(\vec{b}_3) x_3 \\ &\quad + h_2 \vec{b}_2^*(\vec{b}_1) x_1 + h_2 \vec{b}_2^*(\vec{b}_2) x_2 + h_2 \vec{b}_2^*(\vec{b}_3) x_3 \\ &\quad + h_3 \vec{b}_3^*(\vec{b}_1) x_1 + h_3 \vec{b}_3^*(\vec{b}_2) x_2 + h_3 \vec{b}_3^*(\vec{b}_3) x_3 \end{aligned} \quad (2.60)$$

$$= [h_1 \quad h_2 \quad h_3] \begin{bmatrix} \vec{b}_1^*(\vec{b}_1) & \vec{b}_1^*(\vec{b}_2) & \vec{b}_1^*(\vec{b}_3) \\ \vec{b}_2^*(\vec{b}_1) & \vec{b}_2^*(\vec{b}_2) & \vec{b}_2^*(\vec{b}_3) \\ \vec{b}_3^*(\vec{b}_1) & \vec{b}_3^*(\vec{b}_2) & \vec{b}_3^*(\vec{b}_3) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad (2.61)$$

$$= [h_1 \quad h_2 \quad h_3] \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad (2.62)$$

$$= [h_1, h_2, h_3] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad (2.63)$$

$$= \vec{h}_{\beta^*}^\top \vec{x}_\beta \quad (2.64)$$

The value of $\vec{h} \in L^*$ on $\vec{x} \in L$ is obtained by multiplying \vec{x}_β by the transpose of \vec{h}_{β^*} from the left.

Notice that the middle matrix on the right in Equation 2.61 evaluates into the identity. This is the consequence of using the pair of a basis and its dual basis. The formula 2.64 can be generalized to the situation when bases are not dual by evaluating the middle matrix accordingly. In general

$$\vec{h}(\vec{x}) = \vec{h}_{\bar{\beta}}^\top [\vec{b}_i(\vec{b}_j)] \vec{x}_\beta \quad (2.65)$$

where matrix $[\vec{b}_i(\vec{b}_j)]$ is constructed from the respective bases $\beta, \bar{\beta}$ of L and L^* .

§ 11 Changing the basis in a linear space and in its dual Let us now look at what happens with coordinates of vectors of L^* when passing from the dual basis β^* to the dual basis β'^* induced by passing from a basis β to a basis β' in L . Consider vector $\vec{x} \in L$ and a linear function $\vec{h} \in L^*$ and their coordinates $\vec{x}_\beta, \vec{x}_{\beta'}$ and $\vec{h}_{\beta^*}, \vec{h}_{\beta'^*}$ w.r.t. the respective bases. Introduce further matrix \mathbf{A} transforming coordinates of vectors in L as

$$\vec{x}_{\beta'} = \mathbf{A} \vec{x}_\beta \quad (2.66)$$

when passing from β to β' .

Basis β^* is the dual basis to β and basis β'^* is the dual basis to β' and therefore

$$\vec{h}_{\beta^*}^\top \vec{x}_\beta = \vec{h}(\vec{x}) = \vec{h}_{\beta'^*}^\top \vec{x}_{\beta'} \quad (2.67)$$

for all $\vec{x} \in L$ and all $\vec{h} \in L^*$. Hence

$$\vec{h}_{\beta^*}^\top \vec{x}_\beta = \vec{h}_{\beta'^*}^\top \mathbf{A} \vec{x}_\beta \quad (2.68)$$

for all $\vec{x} \in L$ and therefore

$$\vec{h}_{\beta^*}^\top = \vec{h}_{\beta'^*}^\top \mathbf{A} \quad (2.69)$$

or equivalently

$$\vec{h}_{\beta^*} = \mathbf{A}^\top \vec{h}_{\beta'^*} \quad (2.70)$$

Let us now see what is the meaning of the rows of matrix \mathbf{A} . It becomes clear from Equation 2.69 that the columns of matrix \mathbf{A}^\top can be viewed as vectors of coordinates of basic vectors of $\beta'^* = [\vec{b}'_1, \vec{b}'_2, \vec{b}'_3]$ in the basis $\beta^* = [\vec{b}_1, \vec{b}_2, \vec{b}_3]$ and therefore

$$\mathbf{A} = \begin{bmatrix} \text{---} \vec{b}'_{1\beta^*}{}^\top \text{---} \\ \text{---} \vec{b}'_{2\beta^*}{}^\top \text{---} \\ \text{---} \vec{b}'_{3\beta^*}{}^\top \text{---} \end{bmatrix} \quad (2.71)$$

which means that the rows of \mathbf{A} are coordinates of the dual basis of the primed dual space in the dual basis of the non-primed dual space.

Finally notice that we can also write

$$\vec{h}_{\beta'^*} = \mathbf{A}^{-\top} \vec{h}_{\beta^*} \quad (2.72)$$

which is formally identical with Equation 2.16.

§ 12 When do coordinates transform the same way in a basis and in its dual basis

It is natural to ask when it happens that the coordinates of linear functions in L^* w.r.t. the dual basis β^* transform the same way as the coordinates of vectors of L w.r.t. the original basis β , i.e.

$$\vec{x}_{\beta'} = \mathbf{A} \vec{x}_\beta \quad (2.73)$$

$$\vec{h}_{\beta'^*} = \mathbf{A} \vec{h}_{\beta^*} \quad (2.74)$$

for all $\vec{x} \in L$ and all $\vec{h} \in L^*$. Considering Equation 2.72, we get

$$\mathbf{A} = \mathbf{A}^{-\top} \quad (2.75)$$

$$\mathbf{A}^\top \mathbf{A} = \mathbf{I} \quad (2.76)$$

Notice that this is, for instance, satisfied when \mathbf{A} is a rotation [4]. In such a case, one often does not anymore distinguish between vectors of L and L^* because they behave the same way and it is hence possible to represent linear functions from L^* by vectors of L .

§ 13 Coordinates of the basis dual to a general basis We denote the standard basis in \mathbb{R}^3 by σ and its dual (standard) basis in \mathbb{R}^{3^*} by σ^* . Now, we can further establish another basis $\gamma = [\vec{c}_1 \ \vec{c}_2 \ \vec{c}_3]$ in \mathbb{R}^3 and its dual basis $\gamma^* = [\vec{c}_1^* \ \vec{c}_2^* \ \vec{c}_3^*]$ in \mathbb{R}^{3^*} . We would like to find the coordinates $\gamma_{\sigma^*}^* = [\vec{c}_{1\sigma^*}^* \ \vec{c}_{2\sigma^*}^* \ \vec{c}_{3\sigma^*}^*]$ of vectors of γ^* w.r.t. σ^* as a function of coordinates $\gamma_\sigma = [\vec{c}_{1\sigma} \ \vec{c}_{2\sigma} \ \vec{c}_{3\sigma}]$ of vectors of γ w.r.t. σ .

Considering Equations 2.57 and 2.64, we are getting

$$\vec{c}_{i\sigma^*}^{\top} \vec{c}_{j\sigma} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \quad \text{for } i, j = 1, 2, 3 \quad (2.77)$$

which can be rewritten in a matrix form as

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \vec{c}_{1\sigma^*}^\top \\ \vec{c}_{2\sigma^*}^\top \\ \vec{c}_{3\sigma^*}^\top \end{bmatrix} [\vec{c}_{1\sigma} \quad \vec{c}_{2\sigma} \quad \vec{c}_{3\sigma}] = \gamma_{\sigma^*}^{\star\top} \gamma_\sigma \quad (2.78)$$

and therefore

$$\gamma_{\sigma^*}^\star = \gamma_\sigma^{-\top} \quad (2.79)$$

§ 14 Remark on higher dimensions We have introduced the dual space and the dual basis in a three-dimensional linear space. The definition of the dual space is exactly the same for any linear space. The definition of the dual basis is the same for all finite-dimensional linear spaces [3]. For any n -dimensional linear space L and its basis β , we get the corresponding n -dimensional dual space L^* with the dual basis β^* .

2.5 Operations with matrices

Matrices are a powerful tool which can be used in many ways. Here we review a few useful rules for matrix manipulation. The rules are often studied in multi-linear algebra and tensor calculus. We shall not review the theory of multi-linear algebra but will look at the rules from a phenomenological point of view. They are useful identities making an effective manipulation and concise notation possible.

§ 15 Kronecker product Let A be a $k \times l$ matrix and B be a $m \times n$ matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1l} \\ a_{21} & a_{22} & \cdots & a_{2l} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kl} \end{bmatrix} \in \mathbb{R}^{k \times l} \quad \text{and} \quad B \in \mathbb{R}^{m \times n} \quad (2.80)$$

then $km \times ln$ matrix

$$C = A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1l}B \\ a_{21}B & a_{22}B & \cdots & a_{2l}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1}B & a_{k2}B & \cdots & a_{kl}B \end{bmatrix} \quad (2.81)$$

is the matrix of the *Kronecker product of matrices* A, B (in this order).

Notice that this product is associative, i.e. $(A \otimes B) \otimes C = A \otimes (B \otimes C)$, but it is not commutative, i.e. $A \otimes B \neq B \otimes A$ in general. There holds a useful identity $(A \otimes B)^\top = A^\top \otimes B^\top$.

§ 16 Matrix vectorization Let A be an $m \times n$ matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \in \mathbb{R}^{m \times n} \quad (2.82)$$

We define operator $v(\cdot): \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{mn}$ which reshapes an $m \times n$ matrix \mathbf{A} into a $mn \times 1$ matrix (i.e. into a vector) by stacking columns of \mathbf{A} one above another

$$v(\mathbf{A}) = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \\ a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \\ a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} = \quad (2.83)$$

Let us study the relationship between $v(\mathbf{A})$ and $v(\mathbf{A}^\top)$. We see that vector $v(\mathbf{A}^\top)$ contains permuted elements of $v(\mathbf{A})$ and therefore we can construct permutation matrices [4] $\mathbb{T}_{m \times n}$ and $\mathbb{T}_{n \times m}$ such that

$$\begin{aligned} v(\mathbf{A}^\top) &= \mathbb{T}_{m \times n} v(\mathbf{A}) \\ v(\mathbf{A}) &= \mathbb{T}_{n \times m} v(\mathbf{A}^\top) \end{aligned}$$

We see that there holds

$$\mathbb{T}_{n \times m} \mathbb{T}_{m \times n} v(\mathbf{A}) = \mathbb{T}_{n \times m} v(\mathbf{A}^\top) = v(\mathbf{A}) \quad (2.84)$$

for every $m \times n$ matrix \mathbf{A} . Hence

$$\mathbb{T}_{n \times m} = \mathbb{T}_{m \times n}^{-1} \quad (2.85)$$

Consider a permutation \mathbb{T} . It has exactly one unit element in each row and in each column. Consider the i -th row with 1 in the j -th column. This row sends the j -th element of an input vector to the i -th element of the output vector. The i -th column of the transpose of \mathbb{T} has 1 in the j -th row. It is the only non-zero element in that row and therefore the j -th row of \mathbb{T}^\top sends the i -th element of an input vector to the j -th element of the output vector. We see that \mathbb{T}^\top is the inverse of \mathbb{T} , i.e. permutation matrices are orthogonal. We see that

$$\mathbb{T}_{m \times n}^{-1} = \mathbb{T}_{m \times n}^\top \quad (2.86)$$

and hence conclude

$$\mathbb{T}_{n \times m} = \mathbb{T}_{m \times n}^\top \quad (2.87)$$

We also write $v(\mathbf{A}) = \mathbb{T}_{m \times n}^\top v(\mathbf{A}^\top)$.

§ 17 From matrix equations to linear systems Kronecker product of matrices and matrix vectorization can be used to manipulate matrix equations in order to

get systems of linear equations in the standard matrix form $\mathbf{A} \mathbf{x} = \mathbf{b}$. Consider, for instance, matrix equation

$$\mathbf{A} \mathbf{X} \mathbf{B} = \mathbf{C} \quad (2.88)$$

with matrices $\mathbf{A} \in \mathbb{R}^{m \times k}$, $\mathbf{X} \in \mathbb{R}^{k \times l}$, $\mathbf{B} \in \mathbb{R}^{l \times n}$, $\mathbf{C} \in \mathbb{R}^{m \times n}$. It can be verified by direct computation that

$$v(\mathbf{A} \mathbf{X} \mathbf{B}) = (\mathbf{B}^\top \otimes \mathbf{A}) v(\mathbf{X}) \quad (2.89)$$

This is useful when matrices \mathbf{A} , \mathbf{B} and \mathbf{C} are known and we use Equation 2.88 to compute \mathbf{X} . Notice that matrix Equation 2.88 is actually equivalent to mn scalar linear equations in kl unknown elements of \mathbf{X} . Therefore, we should be able to write it in the standard form, e.g., as

$$\mathbf{M} v(\mathbf{X}) = v(\mathbf{C}) \quad (2.90)$$

with some $\mathbf{M} \in \mathbb{R}^{(mn) \times (kl)}$. We can use Equation 2.89 to get $\mathbf{M} = \mathbf{B}^\top \otimes \mathbf{A}$ which yields the linear system

$$v(\mathbf{A} \mathbf{X} \mathbf{B}) = v(\mathbf{C}) \quad (2.91)$$

$$(\mathbf{B}^\top \otimes \mathbf{A}) v(\mathbf{X}) = v(\mathbf{C}) \quad (2.92)$$

for unknown $v(\mathbf{X})$, which is in the standard form.

Let us next consider two variations of Equation 2.88. First consider matrix equation

$$\mathbf{A} \mathbf{X} \mathbf{B} = \mathbf{X} \quad (2.93)$$

Here unknowns \mathbf{X} appear on both sides but we are still getting a linear system of the form

$$(\mathbf{B}^\top \otimes \mathbf{A} - \mathbf{I}) v(\mathbf{X}) = 0 \quad (2.94)$$

where \mathbf{I} is the $(mn) \times (kl)$ identity matrix.

Next, we add yet another constraints: $\mathbf{X}^\top = \mathbf{X}$, i.e. matrix \mathbf{X} is symmetric, to get

$$\mathbf{A} \mathbf{X} \mathbf{B} = \mathbf{X} \quad \text{and} \quad \mathbf{X}^\top = \mathbf{X} \quad (2.95)$$

which can be rewritten in the vectorized form as

$$(\mathbf{B}^\top \otimes \mathbf{A} - \mathbf{I}) v(\mathbf{X}) = 0 \quad \text{and} \quad (\mathbb{T}_{m \times n} - \mathbf{I}) v(\mathbf{X}) = 0 \quad (2.96)$$

and combined it into a single linear system

$$\begin{bmatrix} \mathbb{T}_{m \times n} - \mathbf{I} \\ \mathbf{B}^\top \otimes \mathbf{A} - \mathbf{I} \end{bmatrix} v(\mathbf{X}) = 0 \quad (2.97)$$

3 Affine space

Let us study the affine space, an important structure underlying geometry and its algebraic representation. The affine space is closely connected to the linear space. The connection is so intimate that the two spaces are sometimes not even distinguished. Consider, for instance, function $f: \mathbb{R} \rightarrow \mathbb{R}$ with non-zero $a, b \in \mathbb{R}$

$$f(x) = ax + b \tag{3.1}$$

It is often called “linear” but it is not a *linear function* [5, 6, 4] since for every $\alpha \in \mathbb{R}$ there holds

$$f(\alpha x) = \alpha ax + b \neq \alpha(ax + b) = \alpha f(x) \tag{3.2}$$

In fact, f is an *affine function*, which becomes a linear function only for $b = 0$.

In geometry, we need to be very precise and we have to clearly distinguish affine from linear. Let us therefore first review the very basics of linear spaces, and in particular their relationship to geometry, and then move to the notion of affine spaces.

3.1 Vectors

Let us start with geometric vectors and study the rules of their manipulation.

Figure 3.1(a) shows the space of points P , which we live in and intuitively understand. We know what is an oriented line segment, which we also call a *marked ruler* (or just a ruler). A marked ruler is oriented from its origin towards its end, which is actually a mark (represented by an arrow in Figure 3.1(b)) on a thought infinite ruler, Figure 3.1(b). We assume that we are able to align the ruler with any

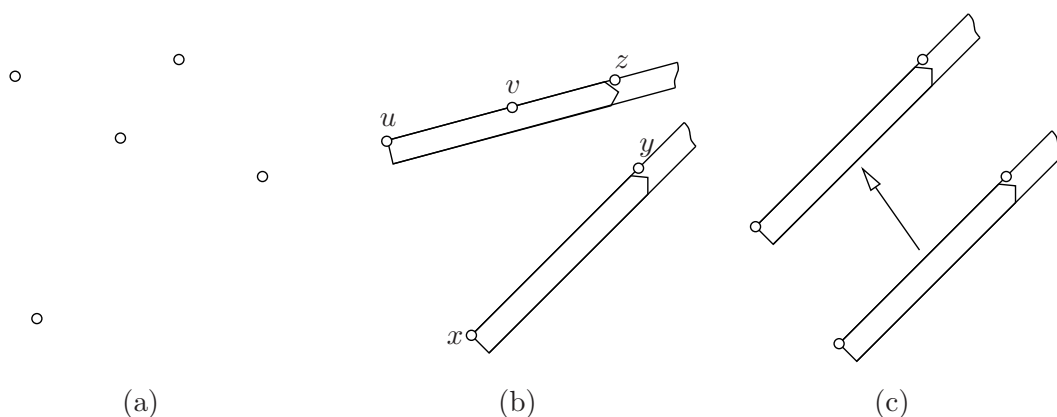


Figure 3.1: (a) The space around us consists of points. Rulers (marked oriented line segments) can be aligned (b) and translated (c) and thus used to transfer, but not measure, distances.

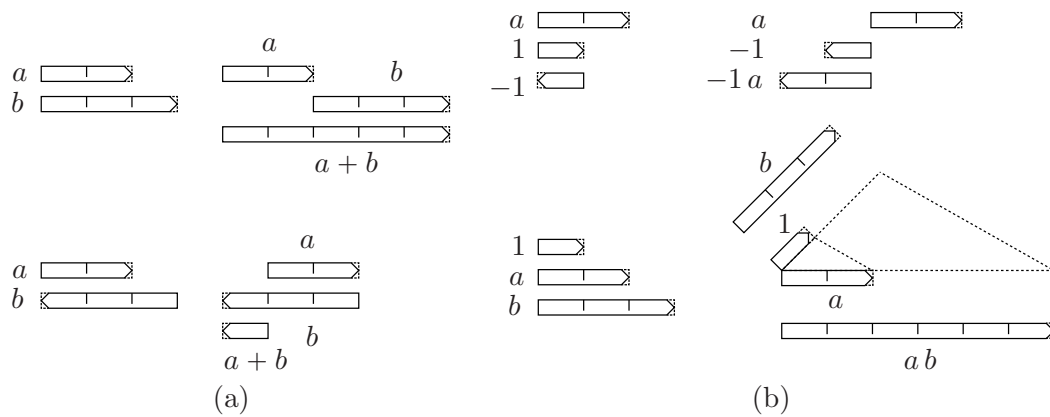


Figure 3.2: Scalars are represented by oriented rulers. They can be added (a) and multiplied (b) purely geometrically by translating and aligning rulers. Notice that we need to single out a unit scalar “1” to perform geometric multiplication.

pair of points x, y , so that the ruler begins in x and a mark is made at the point y . We also know how to align a marked ruler with any pair of distinct points u, v such that the ruler begins in u and aligns with the line connecting u and v in the direction towards point v . The mark on so aligned ruler determines another point, call it z , which is collinear with points u, v . We know how to translate, Figure 3.1(c), a ruler in this space.

To define geometric vectors, we need to first define geometric scalars.

3.1.1 Geometric scalars

Geometric scalars S are horizontal oriented rulers. The ruler, which has its origin identical with its end is called 0. Geometric scalars are equipped with two geometric operations, addition $a + b$ and multiplication ab , defined for every two elements $a, b \in S$.

Figure 3.2(a) shows addition $a + b$. We translate ruler b to align origin of b with the end of a and obtain ruler $a + b$.

Figure 3.2(b) shows multiplication ab . To perform multiplication, we choose a unit ruler “1” and construct its additive inverse -1 using $1 + (-1) = 0$. This introduces orientation to scalars. Scalars aiming to the same side as 1 are *positive* and scalars aiming to the same side as -1 are *negative*. Scalar 0 is neither positive, nor negative. Next we define multiplication by -1 such that $-1 a = -a$, i.e. -1 times a equals the additive inverse of a . Finally, we define multiplication of non-negative (i.e. positive and zero) rulers a, b as follows. We align a with 1 such that origins of 1 and a coincide and such that the rulers contain an acute non-zero angle. We align b with 1 and construct ruler ab by a translation, e.g. as shown in Figure 3.2(b)¹.

All constructions used were purely geometrical and were performed with real rulers. We can verify that so defined addition and multiplication of geometric scalars

¹Notice that ab is well defined since it is the same for all non-zero angles contained by a and 1.

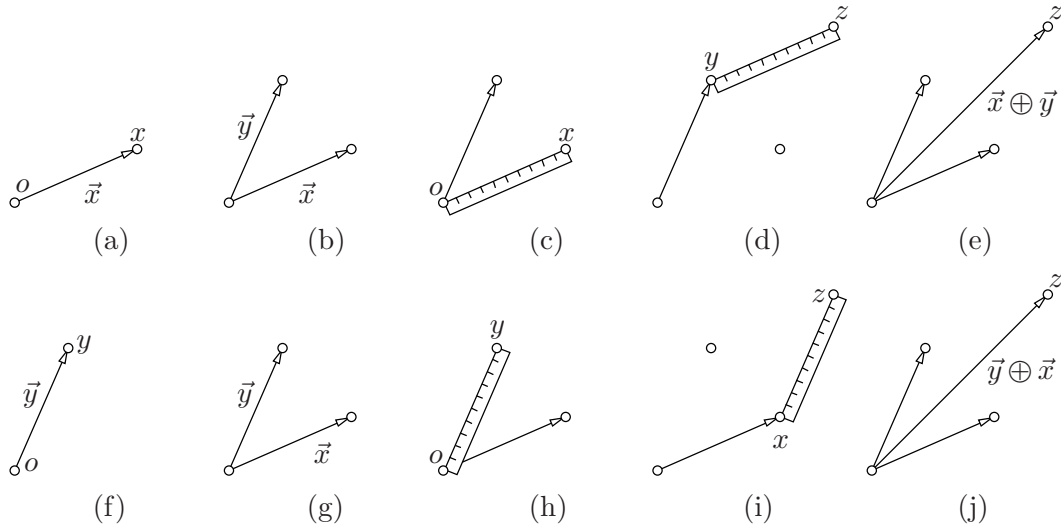


Figure 3.3: Bound vectors are (ordered) pairs of points (o, x) , i.e. arrows $\vec{x} = (o, x)$. Addition of the bound vectors \vec{x}, \vec{y} is realized by parallel transport (using a ruler). We see that the result is the same whether we add \vec{x} to \vec{y} or \vec{y} to \vec{x} . Addition is commutative.

satisfy all rules of addition and multiplication of real numbers. Geometric scalars form a field [10, 13] w.r.t. to $a + b$ and ab .

3.1.2 Geometric vectors

Ordered pairs of points, such as (x, y) in Figure 3.3(a), are called *geometric vectors* and denoted as \vec{xy} , i.e. $\vec{xy} = (x, y)$. Symbol \vec{xy} is often replaced by a simpler one, e.g. by \vec{a} . The set of all geometric vectors is denoted by A .

3.1.3 Bound vectors

Let us now choose one point o and consider all pairs (o, x) , where x can be any point, Figure 3.3(a). We obtain a subset A_o of A , which we call *geometric vectors bound to o* , or just *bound vectors* when it is clear to which point they are bound. We will write $\vec{x} = (o, x)$. Figure 3.3(f) shows another bound vector \vec{y} . The pair (o, o) is special. It will be called the *zero bound vector* and denoted by $\vec{0}$. We will introduce two operations \oplus, \ominus with bound vectors.

First we define *addition of bound vectors* $\oplus: A_o \times A_o \rightarrow A_o$. Let us add vector \vec{x} to \vec{y} as shown on Figure 3.3(b). We take a ruler and align it with \vec{x} , Figure 3.3(c). Then we translate the ruler to align its begin with point y , Figure 3.3(d). The end of the ruler determines point z . We define a new bound vector, which we denote $\vec{x} \oplus \vec{y}$, as the pair (o, z) , Figure 3.3(e). Figures 3.3(f-j) demonstrate that addition gives the same result when we exchange (*commute*) vectors \vec{x} and \vec{y} , i.e. $\vec{x} \oplus \vec{y} = \vec{y} \oplus \vec{x}$. We notice that for every point x , there is exactly one point x' such that $(o, x) \oplus (o, x') = (o, o)$, i.e. $\vec{x} \oplus \vec{x}' = \vec{0}$. Bound vector \vec{x}' is the *inverse* to \vec{x} and is denoted as $-\vec{x}$. Bound vectors are invertible w.r.t. operation \oplus . Finally, we see that $(o, x) \oplus (o, o) = (o, x)$, i.e. $\vec{x} \oplus \vec{0} = \vec{x}$. Vector $\vec{0}$ is the *identity element* of the

operation \oplus . Clearly, operation \oplus behaves exactly as addition of scalars – it is a commutative group [10, 13].

Secondly, we define the *multiplication of a bound vector by a geometric scalar* $\odot: S \times A_o \rightarrow A_o$, where S are geometric scalars and A_o are bound vectors. Operation \odot is a mapping which takes a geometric scalar (a ruler) and a bound vector and delivers another bound vector.

Figure 3.4 shows that to multiply a bound vector $\vec{x} = (o, x)$ by a geometric scalar a , we consider the ruler b whose origin can be aligned with o and end with x . We multiply scalars a and b to obtain scalar ab and align ab with \vec{x} such that the origin of ab coincides with o and ab extends along the line passing through \vec{x} . We obtain end point y of so placed ab and construct the resulting vector $\vec{y} = a \odot \vec{x} = (o, y)$.

We notice that addition \oplus and multiplication \odot of horizontal bound vectors coincides exactly with addition and multiplication of scalars.

3.2 Linear space

We can verify that for every two geometric scalars $a, b \in S$ and every three bound vectors $\vec{x}, \vec{y}, \vec{z} \in A_o$ with their respective operations, there holds the following eight rules

$$\vec{x} \oplus (\vec{y} \oplus \vec{z}) = (\vec{x} \oplus \vec{y}) \oplus \vec{z} \quad (3.3)$$

$$\vec{x} \oplus \vec{y} = \vec{y} \oplus \vec{x} \quad (3.4)$$

$$\vec{x} \oplus \vec{0} = \vec{x} \quad (3.5)$$

$$\vec{x} \oplus -\vec{x} = \vec{0} \quad (3.6)$$

$$1 \odot \vec{x} = \vec{x} \quad (3.7)$$

$$(ab) \odot \vec{x} = a \odot (b \odot \vec{x}) \quad (3.8)$$

$$a \odot (\vec{x} \oplus \vec{y}) = (a \odot \vec{x}) \oplus (a \odot \vec{y}) \quad (3.9)$$

$$(a + b) \odot \vec{x} = (a \odot \vec{x}) \oplus (b \odot \vec{x}) \quad (3.10)$$

These rules are known as axioms of a *linear space* [5, 6, 3]. Bound vectors are one particular model of the linear space. There are many other very useful models, e.g. n -tuples of real or rational numbers for any natural n , polynomials, series of real numbers and real functions. We will give some particularly simple examples useful in geometry later.

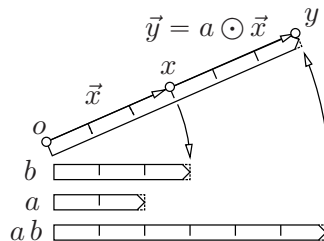


Figure 3.4: Multiplication of the bound vector \vec{x} by a geometric scalar a is realized by aligning rulers to vectors and multiplication of geometric scalars.

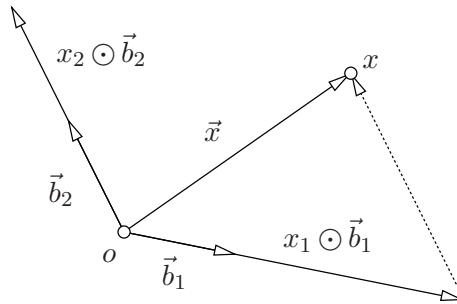


Figure 3.5: Coordinates are the unique scalars that combine independent basic vectors \vec{b}_1, \vec{b}_2 into \vec{x} .

The next concept we will introduce are *coordinates of bound vectors*. To illustrate this concept, we will work in a plane. Figure 3.5 shows two non-collinear bound vectors \vec{b}_1, \vec{b}_2 , which we call *basis*, and another bound vector \vec{x} . We see that there is only one way how to choose scalars x_1 and x_2 such that vectors $x_1 \odot \vec{b}_1$ and $x_2 \odot \vec{b}_2$ add to \vec{x} , i.e.

$$\vec{x} = x_1 \odot \vec{b}_1 \oplus x_2 \odot \vec{b}_2 \quad (3.11)$$

Scalars x_1, x_2 are *coordinates* of \vec{x} in (ordered) basis $[\vec{b}_1, \vec{b}_2]$.

3.3 Free vectors

We can choose any point from A to construct bound vectors and all such choices will lead to the same manipulation of bound vector and to the same axioms of a linear space. Figure 3.6 shows two such choices for points o and o' .

We take bound vectors $\vec{b}_1 = (o, b_1), \vec{b}_2 = (o, b_2), \vec{x} = (o, x)$ at o and construct bound vectors $\vec{b}'_1 = (o', b'_1), \vec{b}'_2 = (o', b'_2), \vec{x}' = (o', x')$ at o' by translating x to x' , b_1 to b'_1 and b_2 to b'_2 by the same translation. Coordinates of \vec{x} w.r.t. $[\vec{b}_1, \vec{b}_2]$ are equal to coordinates of \vec{x}' w.r.t. $[\vec{b}'_1, \vec{b}'_2]$. This interesting property allows us to construct another model of a linear space, which plays an important role in geometry.

Let us now consider the set of all geometric vectors A . Figure 3.7(a) shows an example of a few points and a few geometric vectors. Let us *partition* [1] the set A of

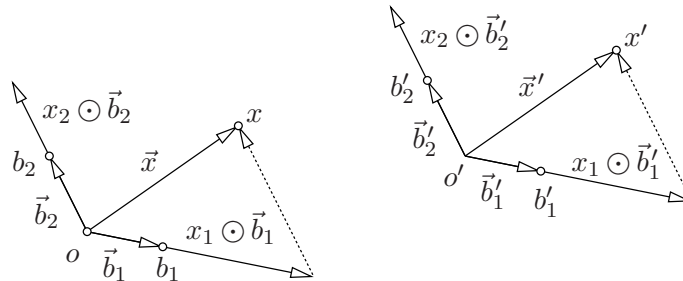


Figure 3.6: Two sets of bound vectors A_o and $A_{o'}$. Coordinates of \vec{x} w.r.t. $[\vec{b}_1, \vec{b}_2]$ are equal to coordinates of \vec{x}' w.r.t. $[\vec{b}'_1, \vec{b}'_2]$.

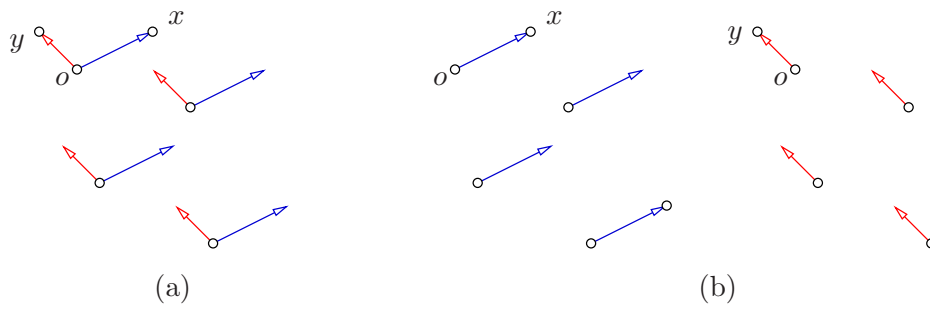


Figure 3.7: The set A of all geometric vectors (a) can be partitioned into subsets which are called free vectors. Two free vectors $A_{(o,x)}$ and $A_{(o,y)}$, i.e. subsets of A , are shown in (b).

geometric vectors into disjoint subsets $A_{(o,x)}$ such that we choose one bound vector (o, x) and put to $A_{(o,x)}$ all geometric vectors that can be obtained by a translation of (o, x) . Figure 3.7(b) shows two such partitions $A_{(o,x)}$, $A_{(o,y)}$. It is clear that $A_{(o,x)} \cap A_{(o,x')} = \emptyset$ for $x \neq x'$ and that every geometric vector is in some (and in exactly one) subset $A_{(o,x)}$.

Two geometric vectors (o, x) and (o', x') form two subsets $A_{(o,x)}$, $A_{(o',x')}$ which are equal if and only if (o', x') is related by a translation to (o, x) .

“To be related by a translation” is an equivalence relation [1]. All geometric vectors in $A_{(o,x)}$ are equivalent to (o, x) .

There are as many sets in the partition as there are bound vectors at a point. We can define the partition by geometric vectors bound to any point o because if we choose another point o' , then for every point x , there is exactly one point x' such that (o, x) can be translated to (o', x') .

We denote the set of subsets $A_{(o,x)}$ by V . Let us see that we can equip set V with a meaningful addition $\boxplus: V \times V \rightarrow V$ and multiplication $\boxtimes: S \times V \rightarrow V$ by geometrical scalars S such that it will become a model of the linear space. Elements of V will be called *free vectors*.

We define the sum of $\vec{x} = A_{(o,x)}$ and $\vec{y} = A_{(o,y)}$, i.e. $\vec{z} = \vec{x} \boxplus \vec{y}$ is the set $A_{(o,x) \boxplus (o,y)}$. Multiplication of $\vec{x} = A_{(o,x)}$ by geometrical scalar a is defined analogically, i.e. $a \boxtimes \vec{x}$ equals the set $A_{a \odot (o,x)}$. We see that the result of \boxplus and \boxtimes does not depend on the choice of o . We have constructed the linear space V of free vectors.

§ 18 Why so many vectors? In the literature, e.g. in [3, 4, 7], linear spaces are often treated purely axiomatically and their geometrical models based on geometrical scalars and vectors are not studied in detail. This is a good approach for a pure mathematician but in engineering we use the geometrical model to study the space we live in. In particular, we wish to appreciate that good understanding of the geometry of the space around us calls for using bound as well as free vectors.

3.4 Affine space

We saw that bound vectors and free vectors were (models of) a linear space. On the other hand, we see that the set of geometric vectors A is not (a model of) a linear

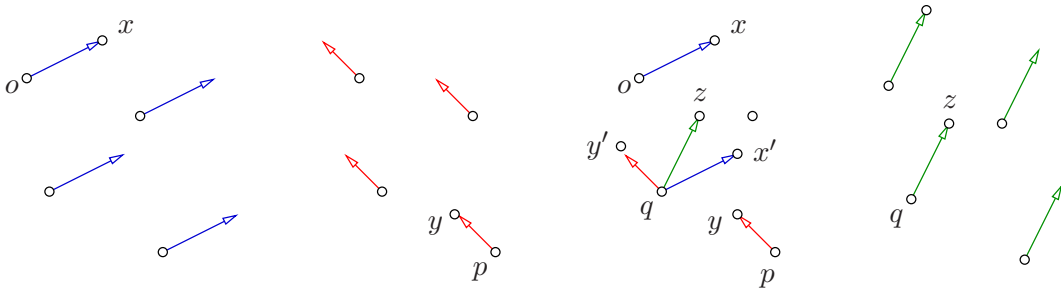


Figure 3.8: Free vector $A_{(o,x)}$ is added to free vector $A_{(p,y)}$ by translating (o,x) to (q,x') and (p,y) to (q,y') , adding bound vectors $(q,z) = (q,x') \oplus (q,y')$ and setting $A_{(o,x)} \boxplus A_{(p,y)} = A_{(q,z)}$

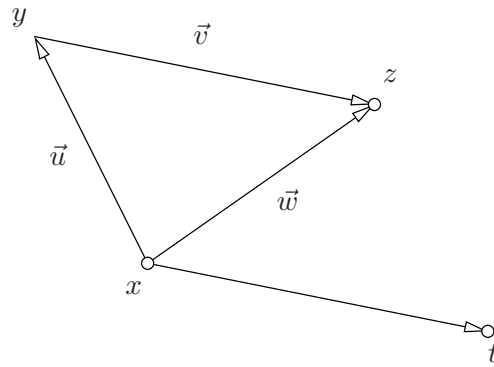


Figure 3.9: Free vectors \vec{u} , \vec{v} and \vec{w} defined by three points x , y and z satisfy triangle identity $\vec{u} \boxplus \vec{v} = \vec{w}$.

space because we do not know how to meaningfully add (by translation) geometric vectors which are not bound to the same point. The set of geometric vectors is an *affine space*.

The affine space connects points, geometric scalars, bound geometric vectors and free vectors in a natural way.

Two points x and y , in this order, give one geometric vector (x,y) , which determines exactly one free vector $\vec{v} = A_{(x,y)}$. We define function $\varphi: A \rightarrow V$, which assigns to two points $x,y \in P$ their corresponding free vector $\varphi(x,y) = A_{(x,y)}$.

Consider a point $a \in P$ and a free vector $\vec{x} \in V$. There is exactly one geometric vector (a,x) , with a at the first position, in the free vector \vec{x} . Therefore, point a and free vector \vec{x} uniquely define point x . We define function $\# : P \times V \rightarrow P$, which takes a point and a free vector and delivers another point. We write $a\#\vec{x} = x$ and require $\vec{x} = \varphi(a,x)$.

Consider three points $x,y,z \in P$, Figure 3.9. We can produce three free vectors $\vec{u} = \varphi(x,y) = A_{(x,y)}$, $\vec{v} = \varphi(y,z) = A_{(y,z)}$, $\vec{w} = \varphi(x,z) = A_{(x,z)}$. Let us investigate the sum $\vec{x} \boxplus \vec{y}$. Chose the representatives of the free vectors, such that they are all bound to x , i.e. bound vectors $(x,y) \in A_{x,y}$, $(x,t) \in A_{(y,z)}$ and $(x,z) \in A_{(x,z)}$. Notice that we could choose the pairs of original points to represent the first and the third free vector but we had to introduce a new pair of points, (x,t) , to represent

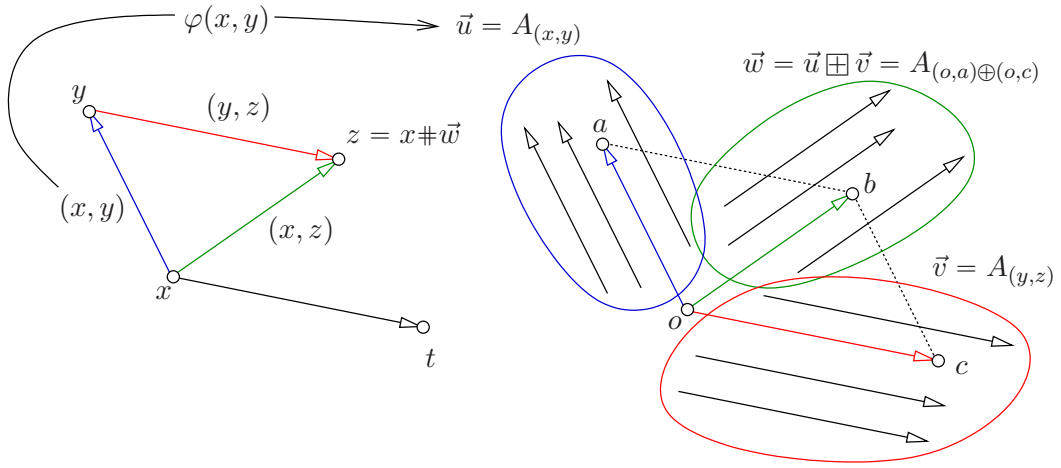


Figure 3.10: Affine space (P, L, φ) , its geometric vectors $(x, y) \in A = P \times P$ and free vector space L and the canonical assignment of pairs of points (x, y) to the free vector $A_{(x,y)}$. Operations \oplus, \boxplus , combining vectors with vectors, and $\#$, combining points with vectors, are illustrated.

the second free vector. Clearly, there holds $(x, y) \oplus (x, t) = (x, z)$. We now see, Figure 3.9, that (y, z) is related to (x, t) by a translation and therefore

$$\vec{u} \boxplus \vec{v} = A_{(x,y)} \boxplus A_{(y,z)} = A_{(x,y)} \boxplus A_{(x,t)} = A_{(x,y) \oplus (x,t)} = A_{(x,z)} = \vec{w} \quad (3.12)$$

Figure 3.10 shows the operations explained above in Figure 3.9 but realized using the vectors bound to another point o .

The above rules are known as *axioms of affine space* and can be used to define even more general affine spaces.

§ 19 Remark on notation We were carefully distinguishing operations $(+, \cdot)$ over scalars, (\oplus, \odot) over bound vectors, (\boxplus, \boxminus) over free vectors, and function $\#$ combining points and free vectors. This is very correct but rarely used. Often, only the symbols introduced for geometric scalars are used for all operations, i.e.

$$+ \equiv +, \oplus, \boxplus, \# \quad (3.13)$$

$$\equiv \cdot, \odot, \boxminus \quad (3.14)$$

§ 20 Affine space Triple (P, L, φ) with a set of points P , linear space (L, \boxplus, \boxminus) (over some field of scalars) and a function $\varphi: P \times P \rightarrow L$, is an affine space when

A1 $\varphi(x, z) = \varphi(x, y) \boxplus \varphi(y, z)$ for every three points $x, y, z \in P$

A2 for every $o \in P$, the function $\varphi_o: P \rightarrow L$, defined by $\varphi_o(x) = \varphi(o, x)$ for all $x \in P$ is a bijection [1].

Axiom A1 calls for an assignment of pairs of point to vectors. Axiom A2 then makes this assignment such that it is one-to-one when the first argument of φ is fixed.

We can define another function $\#: P \times L \rightarrow P$, defined by $o \# \vec{x} = \varphi_o^{-1}(\vec{x})$, which means $\varphi(o, o \# \vec{x}) = \vec{x}$ for all $\vec{x} \in L$. This function combines points and vectors in a

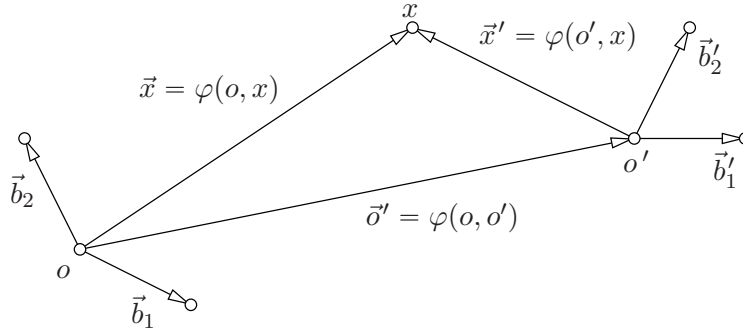


Figure 3.11: Point x is represented in two affine coordinate systems.

way that is very similar to addition and hence is sometimes denoted by $+$ instead of more correct \boxplus .

In our geometrical model of A discussed above, function φ assigned to a pair of points x, y their corresponding free vector $A_{(x,y)}$. Function \boxplus , on the other hand, takes a point x and a free vector \vec{v} and gives another points y such that the bound vector (x, y) is a representative of \vec{v} , i.e. $A_{(x,y)} = \vec{v}$.

3.5 Coordinate system in affine space

We see that function φ assigns the same vector from L to many different pairs of points from P . To represent uniquely points by vectors, we select a point o , called the *origin of affine coordinate system* and represent point $x \in P$ by its *position vector* $\vec{x} = \varphi(o, x)$. In our geometric model of A discussed above, we thus represent point x by bound vector (o, x) or by point o and free vector $A_{(o,x)}$.

To be able to compute with points, we now pass to the representation of points in A by coordinate vectors. We choose a basis $\beta = (\vec{b}_1, \vec{b}_2, \dots)$ in L . That allows us to represent point $x \in P$ by a coordinate vector

$$\vec{x}_\beta = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \end{bmatrix}, \quad \text{such that} \quad \vec{x} = x_1 \vec{b}_1 + x_2 \vec{b}_2 + \dots \quad (3.15)$$

The pair (o, β) , where $o \in P$ and β is a basis of L is called an *affine coordinate system* (often shortly called just coordinate system) of affine space (P, L, φ) .

Let us now study what happens when we choose another point o' and another basis $\beta' = (\vec{b}'_1, \vec{b}'_2, \dots)$ to represent $x \in P$ by coordinate vectors, Figure 3.11. Point x is represented twice: by coordinate vector $\vec{x}_\beta = \varphi(o, x)_\beta = A_{(o,x)\beta}$ and by coordinate vector $\vec{x}'_{\beta'} = \varphi(o', x)_{\beta'} = A_{(o',x)\beta'}$.

To get the relationship between the coordinate vectors \vec{x}_β and $\vec{x}'_{\beta'}$, we employ the triangle equality

$$\varphi(o, x) = \varphi(o, o') \boxplus \varphi(o', x) \quad (3.16)$$

$$\vec{x} = \vec{o}' \boxplus \vec{x}' \quad (3.17)$$

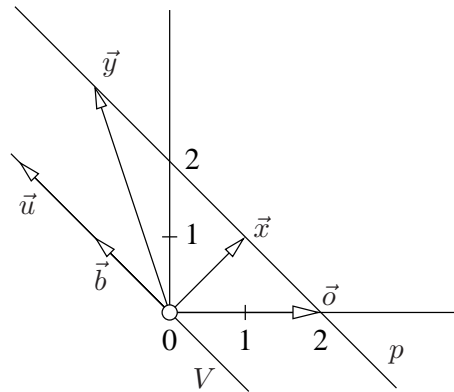


Figure 3.12: Affine space (P, V, φ) of solutions to a linear system is the set of vectors representing points on line p . In coordinate system (\vec{o}, \vec{u}) , vector \vec{x} has coordinate 1. The subspace V of solutions to the associated homogeneous system is the associated linear space. Function φ assigns to two points \vec{o}, \vec{x} the vector $\vec{u} = \vec{y} - \vec{x}$.

which we can write in basis β as (notice that we replace \boxplus by $+$ to emphasize that we are adding coordinate vectors)

$$\vec{x}_\beta = \vec{x}'_\beta + \vec{o}'_\beta \quad (3.18)$$

and use the matrix A transforming coordinates of vectors from basis β' to β to get the desired relationship

$$\vec{x}_\beta = A \vec{x}'_{\beta'} + \vec{o}'_\beta \quad (3.19)$$

Columns of A correspond to coordinate vectors $\vec{b}'_{1\beta}, \vec{b}'_{2\beta}, \dots$. When presented with a situation in a real affine space, we can measure those coordinates by a ruler on a particular representation of L by geometrical vectors bound to, e.g., point o .

3.6 An example of affine space

Let us now present an important example of affine space.

3.6.1 Affine space of solutions of a system of linear equations

When looking at the following system of linear equations in \mathbb{R}^2

$$\begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} \vec{x} = \begin{bmatrix} 2 \\ -2 \end{bmatrix} \quad (3.20)$$

we immediately see that there is an infinite number of solutions. They can be written as

$$\vec{x} = \begin{bmatrix} 2 \\ 0 \end{bmatrix} + \tau \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \tau \in \mathbb{R} \quad (3.21)$$

or as a sum of a particular solution $[2, 0]^\top$ and the set of solutions $\vec{v} = \tau[-1, 1]^\top$ of the accompanied homogeneous system

$$\begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} \vec{v} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (3.22)$$

Figure 3.12 shows that the affine space (P, V, φ) of solutions to the linear system (3.20) is the set of vectors representing points on line p . The subspace V of solutions to the accompanied homogeneous system (3.22) is the linear space associated to A by function φ , which assigns to two points $\vec{x}, \vec{y} \in A$ the vector $\vec{u} = \vec{y} - \vec{x} \in V$. If we choose $\vec{o} = [2, 0]^\top$ as the origin in A and vector $\vec{b} = \varphi(\vec{o}, \vec{x}) = \vec{x} - \vec{o}$ as the basis of V , vector \vec{x} has coordinate 1.

We see that, in this example, points of A are actually vectors of \mathbb{R}^2 , which are the solution to the system (3.20). The vectors of V are the vectors of \mathbb{R}^2 , which are solutions to the associated homogeneous linear system (3.22).

4 Motion

Let us introduce a mathematical model of rigid motion in three-dimensional Euclidean space. The important property of motion is that it only relocates objects without changing their shape. Distances between points on rigidly moving objects remain unchanged.

4.1 Change of position vector coordinates induced by motion

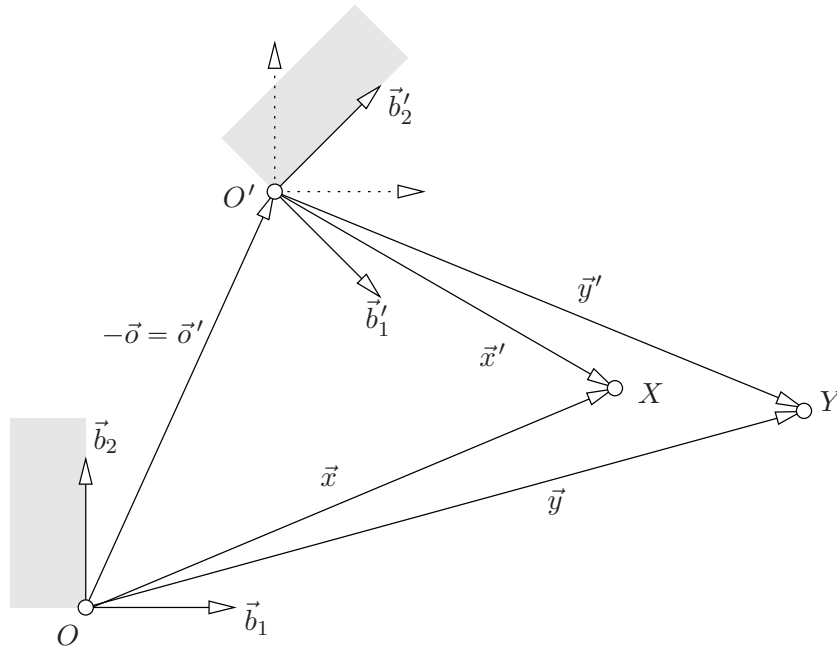


Figure 4.1: Representation of rigid motion.

§ 21 Alias representation of motion¹. Figure 4.1 illustrates a model of (rigid) motion using coordinate systems, points and their position vectors. A coordinate system (O, β) with origin O and basis β is attached to a moving rigid body. As the body moves to a new position, a new coordinate system (O', β') is constructed. Assume a point X in a general position w.r.t. the body, which is represented in the coordinate system (O, β) by its position vector \vec{x} . The same point X is represented in the coordinate system (O', β') by its position vector \vec{x}' . The motion induces a mapping $\vec{x}'_{\beta'} \mapsto \vec{x}_{\beta}$. Such a mapping also determines the motion itself and provides its convenient mathematical model.

¹The terms *alias* and *alibi* were introduced in the classical monograph [13].

Let us derive the formula for the mapping $\vec{x}'_{\beta'} \mapsto \vec{x}_{\beta}$ between the coordinates $\vec{x}'_{\beta'}$ of vector \vec{x}' and coordinates \vec{x}_{β} of vector \vec{x} . Consider the following equations:

$$\vec{x} = \vec{x}' + \vec{o}' \quad (4.1)$$

$$\vec{x}_{\beta} = \vec{x}'_{\beta'} + \vec{o}'_{\beta} \quad (4.2)$$

$$\vec{x}_{\beta} = \begin{bmatrix} \vec{b}'_{1\beta} & \vec{b}'_{2\beta} & \vec{b}'_{3\beta} \end{bmatrix} \vec{x}'_{\beta'} + \vec{o}'_{\beta} \quad (4.3)$$

$$\vec{x}_{\beta} = \mathbf{R} \vec{x}'_{\beta'} + \vec{o}'_{\beta} \quad (4.4)$$

$$\vec{x}'_{\beta'} = \mathbf{R}^{-1} (\vec{x}_{\beta} - \vec{o}'_{\beta}) \quad (4.5)$$

Vector \vec{x} is the sum of vectors \vec{x}' and \vec{o}' , Equation 4.1. We can express all vectors in (the same) basis β , Equation 4.2. To pass to the basis β' we introduce matrix $\mathbf{R} = \begin{bmatrix} \vec{b}'_{1\beta} & \vec{b}'_{2\beta} & \vec{b}'_{3\beta} \end{bmatrix}$, which transforms the coordinates of vectors from β' to β , Equation 4.4. Columns of matrix \mathbf{R} are coordinates $\vec{b}'_{1\beta}, \vec{b}'_{2\beta}, \vec{b}'_{3\beta}$ of basic vectors $\vec{b}'_1, \vec{b}'_2, \vec{b}'_3$ of basis β' in basis β .

§ 22 Alibi representation of motion. An alternative model of (rigid) motion can be developed from the relationship between the points X and Y and their position vectors in Figure 4.1. The point Y is obtained by moving point X altogether with the moving object. It means that the coordinates $\vec{y}'_{\beta'}$ of the position vector \vec{y}' of Y in the coordinate system (O', β') equal the coordinates \vec{x}_{β} of the position vector \vec{x} of X in the coordinate system (O, β) , i.e.

$$\vec{y}'_{\beta'} = \vec{x}_{\beta}$$

$$\vec{y}_{\beta'} + \vec{o}_{\beta'} = \vec{x}_{\beta}$$

$$\mathbf{R}^{-1} (\vec{y}_{\beta} + \vec{o}_{\beta}) = \vec{x}_{\beta}$$

$$\vec{y}_{\beta} = \mathbf{R} \vec{x}_{\beta} - \vec{o}_{\beta} \quad (4.6)$$

$$\vec{y}_{\beta} = \mathbf{R} \vec{x}_{\beta} + \vec{o}'_{\beta} \quad (4.7)$$

Equation 4.6 describes how is the point X moved to point Y w.r.t. the coordinate system (O, β) .

4.2 Rotation matrix

Motion that leaves at least one point fixed is called rotation. Choosing such a fixed point as the origin leads to $O = O'$ and hence to $\vec{o} = \vec{0}$. The motion is then fully described by matrix \mathbf{R} , which is called *rotation matrix*.

§ 23 Two-dimensional rotation. To understand the matrix \mathbf{R} , we shall start with an experiment in two-dimensional plane. Imagine a right-angled triangle ruler as shown in Figure 4.2(a) with arms of equal length and let us define a coordinate system as in the figure. Next, rotate the triangle ruler around its tip, i.e. around the origin O of the coordinate system. We know, and we can verify it by direct physical measurement, that thanks to the symmetry of the situation, the parallelograms through the tips of \vec{b}'_1 and \vec{b}'_2 and along \vec{b}_1 and \vec{b}_2 will be rotated by 90 degrees. We

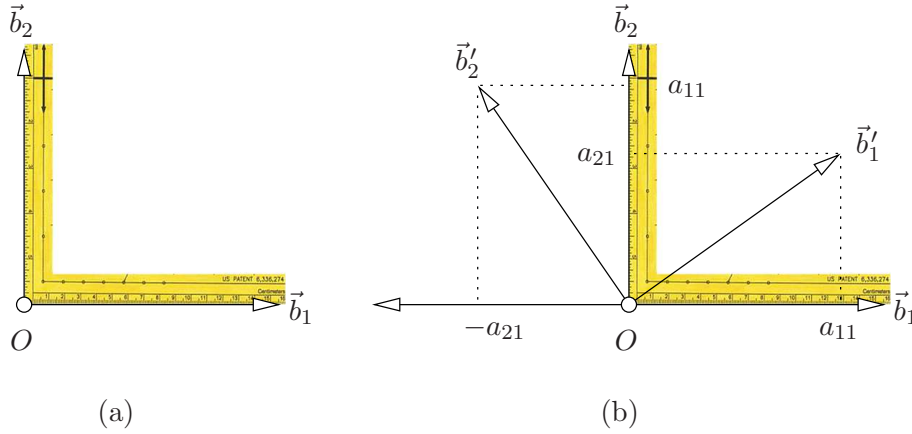


Figure 4.2: Rotation in two-dimensional space.

see that

$$\vec{b}'_1 = a_{11} \vec{b}_1 + a_{21} \vec{b}_2 \quad (4.8)$$

$$\vec{b}'_2 = -a_{21} \vec{b}_1 + a_{11} \vec{b}_2 \quad (4.9)$$

for some real numbers a_{11} and a_{21} . By comparing it with Equation 4.3, we conclude that

$$\mathbf{R} = \begin{bmatrix} a_{11} & -a_{21} \\ a_{21} & a_{11} \end{bmatrix} \quad (4.10)$$

We immediately see that

$$\mathbf{R}^\top \mathbf{R} = \begin{bmatrix} a_{11} & a_{21} \\ -a_{21} & a_{11} \end{bmatrix} \begin{bmatrix} a_{11} & -a_{21} \\ a_{21} & a_{11} \end{bmatrix} = \begin{bmatrix} a_{11}^2 + a_{21}^2 & 0 \\ 0 & a_{11}^2 + a_{21}^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (4.11)$$

since $(a_{11}^2 + a_{21}^2)$ is the squared length of the basic vector of b_1 , which is one. We derived an interesting result

$$\mathbf{R}^{-1} = \mathbf{R}^\top \quad (4.12)$$

$$\mathbf{R} = \mathbf{R}^{-\top} \quad (4.13)$$

Next important observation is that for coordinates \vec{x}_β and $\vec{x}'_{\beta'}$, related by a rotation, there holds

$$(x')^2 + (y')^2 = \vec{x}'_{\beta'}^\top \vec{x}'_{\beta'} = (\mathbf{R} \vec{x}_\beta)^\top \mathbf{R} \vec{x}_\beta = \vec{x}_\beta^\top (\mathbf{R}^\top \mathbf{R}) \vec{x}_\beta = \vec{x}_\beta^\top \vec{x}_\beta = x^2 + y^2 \quad (4.14)$$

Now, if the basis β was constructed as in Figure 4.2, in which case it is called an *orthonormal basis*, then the parallelogram used to measure coordinates x, y of \vec{x} is a rectangle and hence $x^2 + y^2$ is the squared length of \vec{x} by the Pythagoras theorem. If β' is related by rotation, then also $(x')^2 + (y')^2$ is the squared length of \vec{x} , again thanks to the Pythagoras theorem.

We see that $\vec{x}_\beta^\top \vec{x}_\beta$ is the squared length of \vec{x} when β is orthonormal and that this length is preserved by computing it in the same way from the new coordinates of \vec{x} in the new coordinate system after motion. The change of coordinates induced by motion is modeled by rotation matrix \mathbf{R} , which has the desired property $\mathbf{R}^\top \mathbf{R} = \mathbf{I}$, when the bases β, β' are both orthonormal.

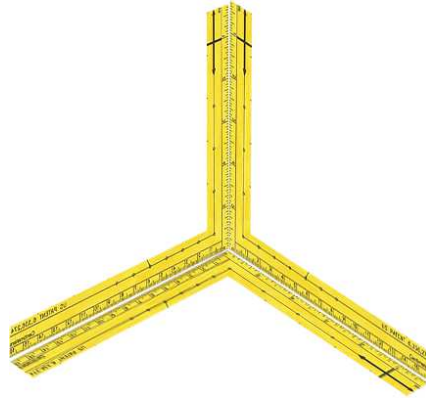


Figure 4.3: A three-dimensional coordinate system.

§ 24 Three-dimensional rotation. Let us now consider three dimensions. It would be possible to generalize Figure 4.2 to three dimensions, construct orthonormal bases and use rectangular parallelograms to establish the relationship between elements of \mathbb{R} in three dimensions. However, the figure and the derivations would become much more complicated.

We shall follow a more intuitive path instead. Consider that we have found that with two-dimensional orthonormal bases, the lengths of vectors could be computed by the Pythagoras theorem since the parallelograms determining the coordinates were rectangular. To achieve this in three dimensions, we need (and can!) use bases consisting from three orthogonal vectors. Then, again, the parallelograms will be rectangular and hence the Pythagoras theorem for three dimensions can be used analogically as in two dimensions, Figure 4.3.

Considering orthonormal bases β, β' , we require the following to hold for all vectors \vec{x} with $\vec{x}_\beta = [x \ y \ z]^\top$ and $\vec{x}'_{\beta'} = [x' \ y' \ z']^\top$

$$\begin{aligned}
 (x')^2 + (y')^2 + (z')^2 &= x^2 + y^2 + z^2 \\
 \vec{x}'_{\beta'}{}^\top \vec{x}'_{\beta'} &= \vec{x}_\beta{}^\top \vec{x}_\beta \\
 (\mathbf{R} \vec{x}_\beta)^\top \mathbf{R} \vec{x}_\beta &= \vec{x}_\beta{}^\top \vec{x}_\beta \\
 \vec{x}_\beta{}^\top (\mathbf{R}^\top \mathbf{R}) \vec{x}_\beta &= \vec{x}_\beta{}^\top \vec{x}_\beta \\
 \vec{x}_\beta{}^\top \mathbf{C} \vec{x}_\beta &= \vec{x}_\beta{}^\top \vec{x}_\beta
 \end{aligned} \tag{4.15}$$

Equation 4.15 must hold for all vectors \vec{x} and hence also for special vectors such as those with coordinates

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \tag{4.16}$$

Let us see what that implies, e.g., for the first vector

$$[1 \ 0 \ 0] \mathbf{C} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = 1 \tag{4.17}$$

$$c_{11} = 1 \tag{4.18}$$

Taking the second and the third vector leads similarly to $c_{22} = c_{33} = 1$. Now, let's try the fourth vector

$$\begin{bmatrix} 1 & 1 & 0 \end{bmatrix} \mathbf{C} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = 2 \quad (4.19)$$

$$1 + c_{12} + c_{21} + 1 = 2 \quad (4.20)$$

$$c_{12} + c_{21} = 0 \quad (4.21)$$

Again, taking the fifth and the sixth vector leads to $c_{13} + c_{31} = c_{23} + c_{32} = 0$. This brings us to the following form of \mathbf{C}

$$\mathbf{C} = \begin{bmatrix} 1 & c_{12} & c_{13} \\ -c_{12} & 1 & c_{23} \\ -c_{13} & -c_{23} & 1 \end{bmatrix} \quad (4.22)$$

Moreover, we see that \mathbf{C} is symmetric since

$$\mathbf{C}^\top = (\mathbf{R}^\top \mathbf{R})^\top = \mathbf{R}^\top \mathbf{R} = \mathbf{C} \quad (4.23)$$

which leads to $-c_{12} = c_{12}$, $-c_{13} = c_{13}$ and $-c_{23} = c_{23}$, i.e. $c_{12} = c_{13} = c_{23} = 0$ and allows us to conclude that

$$\mathbf{R}^\top \mathbf{R} = \mathbf{C} = \mathbf{I} \quad (4.24)$$

Interestingly, not all matrices \mathbf{R} satisfying Equation 4.24 represent motions in three-dimensional space.

Consider, e.g., matrix

$$\mathbf{S} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad (4.25)$$

Matrix \mathbf{S} does not correspond to any rotation of the space since it keeps the plane xy fixed and reflects all other points w.r.t. this xy plane. We see that some matrices satisfying Equation 4.24 are rotations but there are also some such matrices that are not rotations. Can we somehow distinguish them?

Notice that $|\mathbf{S}| = -1$ while $|\mathbf{I}| = 1$. It might be therefore interesting to study the determinant of \mathbf{C} in general. Consider that

$$1 = |\mathbf{I}| = |(\mathbf{R}^\top \mathbf{R})| = |\mathbf{R}^\top| |\mathbf{R}| = |\mathbf{R}| |\mathbf{R}| = (|\mathbf{R}|)^2 \quad (4.26)$$

which gives that $|\mathbf{R}| = \pm 1$. We see that the sign of the determinant splits all matrices satisfying Equation 4.24 into two groups – rotations, which have a positive determinant, and reflections, which have a negative determinant. The product of any two rotations will again be a rotation, the product of a rotation and a reflection will be a reflection and the product of two reflections will be a rotation.

To summarize, rotation in three-dimensional space is represented by a 3×3 matrix \mathbf{R} with $\mathbf{R}^\top \mathbf{R} = \mathbf{I}$ and $|\mathbf{R}| = 1$. The set of all such matrices, and at the same time also the corresponding rotations, will be called $SO(3)$, for *special orthonormal three-dimensional*. Two-dimensional rotations will be analogically denoted as $SO(2)$.

4.3 Coordinate vectors

We see that the matrix \mathbf{R} induced by motion has the property that coordinates and the basic vectors are transformed in the same way. This is particularly useful observation when β is formed by the standard basis, i.e.

$$\beta = \left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right) \quad (4.27)$$

For a rotation matrix \mathbf{R} , Equation 2.16 becomes

$$\begin{bmatrix} \vec{b}'_1 \\ \vec{b}'_2 \\ \vec{b}'_3 \end{bmatrix} = \mathbf{R} \begin{bmatrix} \vec{b}_1 \\ \vec{b}_2 \\ \vec{b}_3 \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} \vec{b}_1 \\ \vec{b}_2 \\ \vec{b}_3 \end{bmatrix} = \begin{bmatrix} r_{11} \vec{b}_1 + r_{12} \vec{b}_2 + r_{13} \vec{b}_3 \\ r_{21} \vec{b}_1 + r_{22} \vec{b}_2 + r_{23} \vec{b}_3 \\ r_{31} \vec{b}_1 + r_{32} \vec{b}_2 + r_{33} \vec{b}_3 \end{bmatrix} \quad (4.28)$$

and hence

$$\vec{b}'_1 = r_{11} \vec{b}_1 + r_{12} \vec{b}_2 + r_{13} \vec{b}_3 = r_{11} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + r_{12} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + r_{13} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} r_{11} \\ r_{12} \\ r_{13} \end{bmatrix} \quad (4.29)$$

and similarly for \vec{b}'_2 and \vec{b}'_3 . We conclude that

$$\begin{bmatrix} \vec{b}'_1 & \vec{b}'_2 & \vec{b}'_3 \end{bmatrix} = \begin{bmatrix} r_{11} & r_{21} & r_{31} \\ r_{12} & r_{22} & r_{32} \\ r_{13} & r_{23} & r_{33} \end{bmatrix} = \mathbf{R}^\top \quad (4.30)$$

This also corresponds to solving for \mathbf{R} in Equation 2.14 with $\mathbf{A} = \mathbf{R}$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \vec{b}'_1 & \vec{b}'_2 & \vec{b}'_3 \end{bmatrix} \mathbf{R} \quad (4.31)$$

5 Rotation

5.1 Properties of rotation matrix

Let us study additional properties of the rotation matrix in three-dimensional space.

5.1.1 Inverse of R

Let

$$\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \quad (5.1)$$

be a rotation matrix. We can find the inverse of \mathbf{R} by evaluating its adjugate matrix [4] and use $\mathbf{R}^{-1} = \mathbf{R}^\top$ and $|\mathbf{R}| = 1$

$$\mathbf{R}^{-1} = \frac{1}{|\mathbf{R}|} \text{Adj}(\mathbf{R}) \quad (5.2)$$

$$\mathbf{R}^\top = \text{Adj}(\mathbf{R}) \quad (5.3)$$

$$= \begin{bmatrix} r_{22} r_{33} - r_{23} r_{32} & r_{13} r_{32} - r_{12} r_{33} & r_{12} r_{23} - r_{13} r_{22} \\ r_{23} r_{31} - r_{21} r_{33} & r_{11} r_{33} - r_{13} r_{31} & r_{13} r_{21} - r_{11} r_{23} \\ r_{21} r_{32} - r_{22} r_{31} & r_{12} r_{31} - r_{11} r_{32} & r_{11} r_{22} - r_{12} r_{21} \end{bmatrix} \quad (5.4)$$

which also gives an alternative expression of

$$\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} = \begin{bmatrix} r_{22} r_{33} - r_{23} r_{32} & r_{23} r_{31} - r_{21} r_{33} & r_{21} r_{32} - r_{22} r_{31} \\ r_{13} r_{32} - r_{12} r_{33} & r_{11} r_{33} - r_{13} r_{31} & r_{12} r_{31} - r_{11} r_{32} \\ r_{12} r_{23} - r_{13} r_{22} & r_{13} r_{21} - r_{11} r_{23} & r_{11} r_{22} - r_{12} r_{21} \end{bmatrix} \quad (5.5)$$

5.1.2 Eigenvalues of R

Let \mathbf{R} be a rotation matrix. Then for every $\vec{v} \in \mathbb{C}^3$

$$(\mathbf{R} \vec{v})^\top \mathbf{R} \vec{v} = \vec{v}^\top \mathbf{R}^\top \mathbf{R} \vec{v} = \vec{v}^\top (\mathbf{R}^\top \mathbf{R}) \vec{v} = \vec{v}^\top \vec{v} \quad (5.6)$$

we see that for all $\vec{v} \in \mathbb{C}^3$ and $\lambda \in \mathbb{C}$ such that

$$\mathbf{R} \vec{v} = \lambda \vec{v} \quad (5.7)$$

there holds¹

$$(\lambda \vec{v})^\top (\lambda \vec{v}) = (\vec{v}^\top \vec{v}) \quad (5.8)$$

$$\bar{\lambda} \lambda (\vec{v}^\top \vec{v}) = (\vec{v}^\top \vec{v}) \quad (5.9)$$

$$|\lambda|^2 (\vec{v}^\top \vec{v}) = (\vec{v}^\top \vec{v}) \quad (5.10)$$

and hence $|\lambda|^2 = 1$ for all $\vec{v} \neq \vec{0}$. We conclude that the absolute value of eigenvalues of \mathbf{R} is one.

Next, by looking at the characteristic polynomial of \mathbf{R}

$$p(\lambda) = |(\lambda \mathbf{I} - \mathbf{R})| = \left| \begin{pmatrix} \lambda - r_{11} & -r_{12} & -r_{13} \\ -r_{21} & \lambda - r_{22} & -r_{23} \\ -r_{31} & -r_{32} & \lambda - r_{33} \end{pmatrix} \right| \quad (5.11)$$

$$= \lambda^3 - (r_{11} + r_{22} + r_{33}) \lambda^2 + (r_{11} r_{22} - r_{21} r_{12} + r_{11} r_{33} - r_{31} r_{13} + r_{22} r_{33} - r_{23} r_{32}) \lambda \quad (5.12)$$

$$+ r_{11} (r_{23} r_{32} - r_{22} r_{33}) - r_{21} (r_{32} r_{13} - r_{12} r_{33}) + r_{31} (r_{13} r_{22} - r_{12} r_{23}) \quad (5.13)$$

$$= \lambda^3 - (r_{11} + r_{22} + r_{33}) \lambda^2 + (r_{33} + r_{22} + r_{11}) \lambda - |\mathbf{R}| \quad (5.14)$$

$$= \lambda^3 - \text{trace} \mathbf{R} (\lambda^2 - \lambda) - 1 \quad (5.15)$$

we conclude that 1 is always an eigenvalue of \mathbf{R} . Notice that we have used identities in Equation 5.5 to pass from Equation 5.12 to Equation 5.13².

Let us denote the eigenvalues as $\lambda_1 = 1$, $\lambda_2 = x + yi$ and $\lambda_3 = x - yi$ with real x, y . It follows from the above that $x^2 + y^2 = 1$. We see that there is either one real or three real solutions since if $y = 0$, then $x^2 = 1$ and hence $\lambda_2 = \lambda_3 = \pm 1$. We conclude that we can encounter only two situations when all eigenvalues are real. Either $\lambda_1 = \lambda_2 = \lambda_3 = 1$, or $\lambda_1 = 1$ and $\lambda_2 = \lambda_3 = -1$.

5.1.3 Eigenvectors of \mathbf{R} .

Let us now look at eigenvectors of \mathbf{R} and let's first investigate the situation when all eigenvalues of \mathbf{R} are real.

§ 25 $\lambda_1 = \lambda_2 = \lambda_3 = 1$: Let $\lambda_1 = \lambda_2 = \lambda_3 = 1$. Then $p(\lambda) = (\lambda - 1)^3 = \lambda^3 - 3\lambda^2 + 3\lambda - 1$. It means that $r_{11} + r_{22} + r_{33} = 3$ and since $r_{11} \leq 1$, $r_{22} \leq 1$, $r_{33} \leq 1$, it leads to $r_{11} = r_{22} = r_{33} = 1$, which implies $\mathbf{R} = \mathbf{I}$. Then $\mathbf{I} - \mathbf{R} = \mathbf{0}$ and all non-zero vectors of \mathbb{R}^3 are eigenvectors of \mathbf{R} . Notice that rank of $\mathbf{R} - \mathbf{I}$ is zero in this case.

¹We will use the *conjugate transpose* [4] on vectors with complex coordinates. It means, e.g., that

$$\begin{bmatrix} a_{11} + b_{11} i & a_{12} + b_{12} i \\ a_{21} + b_{21} i & a_{22} + b_{22} i \end{bmatrix}^\top = \begin{bmatrix} a_{11} - b_{11} i & a_{21} - b_{21} i \\ a_{12} - b_{12} i & a_{22} - b_{22} i \end{bmatrix}$$

for all $a_{11}, a_{12}, a_{21}, a_{22}, b_{11}, b_{12}, b_{21}, b_{22} \in \mathbb{R}$. Also recall [2] that $\overline{a\bar{b}} = a$ for all $a, b \in \mathbb{R}$.

²Alternatively. It follows from the Fundamental theorem of algebra [6] the $p(\lambda) = 0$ has always a solution in \mathbb{C} and since coefficients of $p(\lambda)$ are all real, the solutions must come in complex conjugated pairs. The degree of $p(\lambda)$ is three and thus at least one solution must be real and hence equal to ± 1 . Now, since $p(0) = -|\mathbf{R}| = -1$, $\lim_{\lambda \rightarrow \infty} p(\lambda) = \infty$, and $p(\lambda)$ is a continuous function, it must (by the mean value theorem [2]) cross the positive side of the real axis and hence one of its eigenvalues has to be equal to one.

Next consider $\lambda_1 = 1$ and $\lambda_2 = \lambda_3 = -1$. The eigenvectors \vec{v} corresponding to $\lambda_2 = \lambda_3 = -1$ are solutions to

$$\mathbf{R} \vec{v} = -\vec{v} \quad (5.16)$$

There is always at least one one-dimensional space of such vectors. We also see that there is a rotation matrix

$$\mathbf{R} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad (5.17)$$

with eigenvectors

$$r \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, r \neq 0, \quad \text{and} \quad s \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, s^2 + t^2 \neq 0, \quad (5.18)$$

which means that there is a one-dimensional space of eigenvectors corresponding to 1 and a two-dimensional space of eigenvectors corresponding to -1 . Notice that rank of $\mathbf{R} - \mathbf{I}$ is two here.

§ 26 $\lambda_1 = 1, \lambda_2 = \lambda_3 = -1$: How does the situation look for a general \mathbf{R} with eigenvalues $1, -1, -1$? Consider an eigenvector \vec{v}_1 corresponding to 1 and an eigenvector \vec{v}_2 corresponding to -1 . They are linearly independent. Otherwise there has to be $s \in \mathbb{R}$ such that $\vec{v}_2 = s \vec{v}_1 \neq 0$ and then

$$\vec{v}_2 = s \vec{v}_1 \quad (5.19)$$

$$\mathbf{R} \vec{v}_2 = s \mathbf{R} \vec{v}_1 \quad (5.20)$$

$$-\vec{v}_2 = s \vec{v}_1 \quad (5.21)$$

leading to $s = -s$ and therefore $s = 0$ which contradicts $\vec{v}_2 \neq 0$. Now, let us look at vectors $\vec{v}_3 \in \mathbb{R}^3$ defined by

$$\begin{bmatrix} \vec{v}_1^\top \\ \vec{v}_2^\top \end{bmatrix} \vec{v}_3 = 0 \quad (5.22)$$

The above linear system has a one-dimensional space of solutions since the rows of its matrix are independent. Chose a fixed solution $\vec{v}_3 \neq 0$. Then

$$\begin{bmatrix} \vec{v}_1^\top \\ \vec{v}_2^\top \end{bmatrix} \mathbf{R}^\top \vec{v}_3 = \begin{bmatrix} \vec{v}_1^\top \mathbf{R}^\top \\ \vec{v}_2^\top \mathbf{R}^\top \end{bmatrix} \vec{v}_3 = \begin{bmatrix} \vec{v}_1^\top \\ -\vec{v}_2^\top \end{bmatrix} \vec{v}_3 = 0 \quad (5.23)$$

We see that $\mathbf{R}^\top \vec{v}_3$ and \vec{v}_3 are in the same one-dimensional space, i.e. they are linearly dependent and we can write

$$\mathbf{R}^\top \vec{v}_3 = s \vec{v}_3 \quad (5.24)$$

for some non-zero $s \in \mathbb{C}$. Multiplying equation 5.24 by \mathbf{R} from the left and dividing both sides by s gives

$$\frac{1}{s} \vec{v}_3 = \mathbf{R} \vec{v}_3 \quad (5.25)$$

Clearly, \vec{v}_3 is an eigenvector of \mathbf{R} . Since it is not a multiple of \vec{v}_1 , it must correspond to eigenvalue -1 . Moreover, $\vec{v}_2^\top \vec{v}_3 = 0$ and hence they are linearly independent.

We have shown that if -1 is an eigenvalue of \mathbf{R} , then there are always at least two linearly independent vectors corresponding to the eigenvalue -1 and therefore there is a two-dimensional space of eigenvectors corresponding to -1 . Notice that the rank of $\mathbf{R} - \mathbf{I}$ is two in this case since the two-dimensional subspace corresponding to -1 can be complemented only by a one-dimensional subspace corresponding to 1 to avoid intersecting the subspaces in a non-zero vector.

§ 27 General $\lambda_1, \lambda_2, \lambda_3$: Finally, let us look at arbitrary (even non-real) eigenvalues. Assume $\lambda = x + yi$ for real x, y . Then we have

$$\mathbf{R} \vec{v} = (x + yi) \vec{v} \quad (5.26)$$

If $y \neq 0$, vector \vec{v} must be non-real since otherwise we would have a real vector on the left and a non-real vector on the right. Furthermore, the eigenvalues are pairwise distinct and hence there are three one-dimensional subspaces of eigenvectors (we now understand the space as \mathbb{C}^3 over \mathbb{C}). In particular, there is exactly one one-dimensional subspace corresponding to eigenvalue 1 . The rank of $\mathbf{R} - \mathbf{I}$ is two.

Let \vec{v} be an eigenvector of a rotation matrix \mathbf{R} . Then

$$\mathbf{R} \vec{v} = (x + yi) \vec{v} \quad (5.27)$$

$$\mathbf{R}^\top \vec{v} = (x + yi) \mathbf{R}^\top \vec{v} \quad (5.28)$$

$$\vec{v} = (x + yi) \mathbf{R}^\top \vec{v} \quad (5.29)$$

$$\frac{1}{(x + yi)} \vec{v} = \mathbf{R}^\top \vec{v} \quad (5.30)$$

$$(x - yi) \vec{v} = \mathbf{R}^\top \vec{v} \quad (5.31)$$

We see that the eigenvector \vec{v} of \mathbf{R} corresponding to eigenvalue $x + yi$ is the eigenvector of \mathbf{R}^\top corresponding to eigenvalue $x - yi$ and vice versa. Thus there is the following interesting correspondence between eigenvalues and eigenvectors of \mathbf{R} and \mathbf{R}^\top . Considering eigenvalue-eigenvector pairs $(1, \vec{v}_1), (x + yi, \vec{v}_2), (x - yi, \vec{v}_3)$ of \mathbf{R} we have $(1, \vec{v}_1), (x - yi, \vec{v}_2), (x + yi, \vec{v}_3)$ pairs of \mathbf{R}^\top , respectively.

§ 28 Orthogonality of eigenvectors The next question to ask is what are the angles between eigenvectors of \mathbf{R} ? We will consider pairs $(\lambda_1 = 1, \vec{v}_1), (\lambda_2 = x + yi, \vec{v}_2), (\lambda_3 = x - yi, \vec{v}_3)$ of eigenvectors associated with their respective eigenvalues. For instance, vector \vec{v}_1 denotes an eigenvector associated with eigenvalue 1 .

If all eigenvalues are equal to 1 , i.e. $\mathbf{R} = \mathbf{I}$, then all non-zero vectors of \mathbb{R}^3 are eigenvectors of \mathbf{R} and hence we can always find two eigenvectors containing a given angle. In particular, we can choose three mutually orthogonal eigenvectors.

If $\lambda_1 = 1$ and $\lambda_2 = \lambda_3 = -1$, then we have seen that every \vec{v}_1 is perpendicular to \vec{v}_2 and \vec{v}_3 and that \vec{v}_2 and \vec{v}_3 can be any two non-zero vectors in a two-dimensional subspace of \mathbb{R}^3 , which is orthogonal to \vec{v}_1 . Therefore, for every angle, there are \vec{v}_2 and \vec{v}_3 which contain it. In particular, it is possible to choose \vec{v}_2 to be orthogonal to \vec{v}_3 and hence there are three mutually orthogonal eigenvectors.

Finally, if λ_2, λ_3 are non-real, i.e. $y \neq 0$, we have three mutually distinct eigenvalues and hence there are exactly three one-dimensional subspaces (each without the zero vector) of eigenvectors. If two eigenvectors are from the same subspace, then they are linearly dependent and hence they contain the zero angle.

Let us now evaluate $\vec{v}_1^\top \vec{v}_2$

$$\vec{v}_1^\top \vec{v}_2 = \vec{v}_1^\top \mathbf{R}^\top \mathbf{R} \vec{v}_2 = \vec{v}_1^\top (x + yi) \vec{v}_2 = (x + yi) \vec{v}_1^\top \vec{v}_2 \quad (5.32)$$

We conclude that either $(x + yi) = 1$ or $\vec{v}_1^\top \vec{v}_2 = 0$. Since the former can't be the case as $y \neq 0$, the latter must hold true. We see that \vec{v}_1 is orthogonal to \vec{v}_2 . We can show that \vec{v}_1 is orthogonal to \vec{v}_3 exactly in the same way.

Let us next consider the angle between eigenvectors \vec{v}_2 and \vec{v}_3

$$\vec{v}_3^\top \vec{v}_2 = \vec{v}_3^\top \mathbf{R}^\top \mathbf{R} \vec{v}_2 = (\vec{v}_3 (x - yi))^\top (x + yi) \vec{v}_2 \quad (5.33)$$

$$= \vec{v}_3^\top (x + yi) (x + yi) \vec{v}_2 \quad (5.34)$$

$$\vec{v}_3^\top \vec{v}_2 = (x^2 + 2xyi - y^2) \vec{v}_3^\top \vec{v}_2 \quad (5.35)$$

We conclude that either $(x^2 + 2xyi - y^2) = 1$ or $\vec{v}_3^\top \vec{v}_2 = 0$. The former implies $xy = 0$ and therefore $x = 0$ since $y \neq 0$ but then $-y^2 = 1$, which is, for a real y , impossible. We see that $\vec{v}_3^\top \vec{v}_2 = 0$, i.e. vectors \vec{v}_2 are orthogonal to vectors \vec{v}_3 .

Clearly, it is always possible to choose three mutually orthogonal eigenvectors. We can further normalize them to unit length and thus obtain an orthonormal basis as non-zero orthogonal vectors are linearly independent. Therefore

$$\mathbf{R} [\vec{v}_1 \ \vec{v}_2 \ \vec{v}_3] = [\vec{v}_1 \ \vec{v}_2 \ \vec{v}_3] \begin{bmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_3 \end{bmatrix} \quad (5.36)$$

$$[\vec{v}_1 \ \vec{v}_2 \ \vec{v}_3]^\top \mathbf{R} [\vec{v}_1 \ \vec{v}_2 \ \vec{v}_3] = \begin{bmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_3 \end{bmatrix} \quad (5.37)$$

Let us further investigate the structure of eigenvectors \vec{v}_2, \vec{v}_3 . We shall show that they are “conjugated”. Let's write $\vec{v}_2 = \vec{u} + \vec{w}i$ with real vectors \vec{u}, \vec{w} . There holds

$$\mathbf{R} \vec{v}_2 = \mathbf{R} (\vec{u} + \vec{w}i) = \mathbf{R} \vec{u} + \mathbf{R} \vec{w}i \quad (5.38)$$

$$(x + yi) \vec{v}_2 = (x + yi) (\vec{u} + \vec{w}i) = x \vec{u} - y \vec{w} + (x \vec{w} + y \vec{u})i \quad (5.39)$$

which implies

$$\mathbf{R} \vec{u} = x \vec{u} - y \vec{w} \quad \text{and} \quad \mathbf{R} \vec{w} = x \vec{w} + y \vec{u} \quad (5.40)$$

Now, let us compare two expressions: $\mathbf{R} (\vec{u} - \vec{w}i)$ and $(x - yi) (\vec{u} - \vec{w}i)$

$$\mathbf{R} (\vec{u} - \vec{w}i) = \mathbf{R} \vec{u} - \mathbf{R} \vec{w}i = x \vec{u} - y \vec{w} - (x \vec{w} + y \vec{u})i \quad (5.41)$$

$$(x - yi) (\vec{u} - \vec{w}i) = x \vec{u} - y \vec{w} - (x \vec{w} + y \vec{u})i \quad (5.42)$$

We see that

$$\mathbf{R} (\vec{u} - \vec{w}i) = (x - yi) (\vec{u} - \vec{w}i) \quad (5.43)$$

which means that $(x - yi, \vec{u} - \vec{w}i)$ are an eigenvalue-eigenvector pair of \mathbf{R} . It is important to understand what has been shown. We have shown that if $\vec{u} + \vec{w}i$ is an eigenvector of \mathbf{R} corresponding to an eigenvalue λ , then the conjugated vector $\vec{u} - \vec{w}i$ is an eigenvector of \mathbf{R} corresponding to eigenvalue, which is conjugated to λ (This does not mean that every two eigenvectors corresponding to $x + yi$ and $x - yi$ must be conjugated).

The conclusion from the previous analysis is that the both non-real eigenvectors of \mathbf{R} are generated by the same two real vectors \vec{u} and \vec{w} . Let us look at the angle between \vec{u} and \vec{w} . Consider that

$$0 = \vec{v}_3^\top \vec{v}_2 = (\vec{u} - \vec{w}i)^\top (\vec{u} + \vec{w}i) = (\vec{u}^\top + \vec{w}^\top i)(\vec{u} + \vec{w}i) \quad (5.44)$$

$$= (\vec{u}^\top \vec{u} - \vec{w}^\top \vec{w}) + (\vec{u}^\top \vec{w} + \vec{w}^\top \vec{u})i \quad (5.45)$$

$$= (\vec{u}^\top \vec{u} - \vec{w}^\top \vec{w}) + 2\vec{w}^\top \vec{u}i \quad (5.46)$$

and therefore

$$\vec{u}^\top \vec{u} = \vec{w}^\top \vec{w} \quad \text{and} \quad \vec{w}^\top \vec{u} = 0 \quad (5.47)$$

which means that vectors \vec{u} and \vec{w} are orthogonal.

Finally, let us consider

$$0 = \vec{v}_1^\top \vec{v}_2 = \vec{v}_1^\top \vec{u} + \vec{v}_1^\top \vec{w}i \quad (5.48)$$

and hence

$$\vec{v}_1^\top \vec{u} = 0 \quad \text{and} \quad \vec{v}_1^\top \vec{w} = 0 \quad (5.49)$$

which means that \vec{u} and \vec{w} are also orthogonal to \vec{v}_1 .

5.1.4 Rotation axis

A one-dimensional subspace generated by an eigenvector \vec{v}_1 of \mathbf{R} corresponding to $\lambda = 1$, is called the *rotation axis* (or axis of rotation) of \mathbf{R} . If $\mathbf{R} = \mathbf{I}$, then there is an infinite number of rotation axes, otherwise there is exactly one. Vectors \vec{v} , which are in a rotation axis of rotation \mathbf{R} , remain unchanged by \mathbf{R} , i.e. $\mathbf{R}\vec{v} = \vec{v}$.

Consider that the eigenvector of \mathbf{R} corresponding to 1 is also an eigenvector of \mathbf{R}^\top since

$$\mathbf{R}\vec{v}_1 = \vec{v}_1 \quad (5.50)$$

$$\mathbf{R}^\top \mathbf{R}\vec{v}_1 = \mathbf{R}^\top \vec{v}_1 \quad (5.51)$$

$$\vec{v}_1 = \mathbf{R}^\top \vec{v}_1 \quad (5.52)$$

It implies

$$(\mathbf{R} - \mathbf{R}^\top)\vec{v}_1 = 0 \quad (5.53)$$

$$\begin{bmatrix} 0 & r_{12} - r_{21} & r_{13} - r_{31} \\ r_{21} - r_{12} & 0 & r_{23} - r_{32} \\ r_{31} - r_{13} & r_{32} - r_{23} & 0 \end{bmatrix} \vec{v}_1 = 0 \quad (5.54)$$

and we see that

$$\begin{bmatrix} 0 & r_{12} - r_{21} & r_{13} - r_{31} \\ r_{21} - r_{12} & 0 & r_{23} - r_{32} \\ r_{31} - r_{13} & r_{32} - r_{23} & 0 \end{bmatrix} \begin{bmatrix} r_{32} - r_{23} \\ r_{13} - r_{31} \\ r_{21} - r_{12} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad (5.55)$$

Clearly, we have a nice formula for an eigenvector corresponding to $\lambda_1 = 1$, when vector $[r_{32} - r_{23} \quad r_{13} - r_{31} \quad r_{21} - r_{12}]^\top$ is non-zero. That is when $\mathbf{R} - \mathbf{R}^\top$ is a non-zero matrix, which is exactly when \mathbf{R} is not symmetric.

Let us now investigate the situation when \mathbf{R} is symmetric. Then, $\mathbf{R} = \mathbf{R}^\top = \mathbf{R}^{-1}$ and therefore

$$\mathbf{R}(\mathbf{R} + \mathbf{I}) = \mathbf{R}\mathbf{R} + \mathbf{R} = \mathbf{I} + \mathbf{R} = \mathbf{R} + \mathbf{I} \quad (5.56)$$

which shows that the non-zero columns of the matrix $\mathbf{R} + \mathbf{I}$ are eigenvectors corresponding to the unit eigenvalue. Clearly, at least one of the columns must be non-zero since otherwise, $\mathbf{R} = -\mathbf{I}$ and $|\mathbf{R}|$ would be minus one, which is impossible for a rotation.

5.1.5 Rotation angle

Rotation angle θ of rotation \mathbf{R} is the angle between a non-zero real vector \vec{x} which is orthogonal to \vec{v}_1 and its image $\mathbf{R}\vec{x}$. There holds

$$\cos \theta = \frac{\vec{x}^\top \mathbf{R} \vec{x}}{\vec{x}^\top \vec{x}} \quad (5.57)$$

Let us set

$$\vec{x} = \vec{u} + \vec{w} \quad (5.58)$$

Clearly, \vec{x} is a real vector which is orthogonal to \vec{v}_1 since both \vec{u} and \vec{w} are. Let's see that it is non-zero. Vector \vec{v}_2 is an eigenvector and thus

$$0 \neq \vec{v}_2^\top \vec{v}_2 = \vec{u}^\top \vec{u} + \vec{w}^\top \vec{w} \quad (5.59)$$

and therefore $\vec{u} \neq \vec{0}$ or $\vec{w} \neq \vec{0}$. Vectors \vec{u} , \vec{w} are orthogonal and therefore their sum can be zero only if they both are zero since otherwise for, e.g., a non-zero \vec{u} we get the following contradiction

$$0 = \vec{u}^\top \vec{0} = \vec{u}^\top (\vec{u} + \vec{v}) = \vec{u}^\top \vec{u} + \vec{u}^\top \vec{v} = \vec{u}^\top \vec{u} \neq 0 \quad (5.60)$$

Let us now evaluate

$$\begin{aligned} \cos \theta &= \frac{\vec{x}^\top \mathbf{R} \vec{x}}{\vec{x}^\top \vec{x}} = \frac{(\vec{u} + \vec{w})^\top \mathbf{R} (\vec{u} + \vec{w})}{(\vec{u} + \vec{w})^\top (\vec{u} + \vec{w})} = \frac{(\vec{u} + \vec{w})^\top (x \vec{u} - y \vec{w} + x \vec{w} + y \vec{u})}{\vec{u}^\top \vec{u} + \vec{w}^\top \vec{w}} \\ &= \frac{x (\vec{u}^\top \vec{u} + \vec{w}^\top \vec{w}) + y (\vec{u}^\top \vec{u} - \vec{w}^\top \vec{w})}{\vec{u}^\top \vec{u} + \vec{w}^\top \vec{w}} \end{aligned} \quad (5.61)$$

$$= x \quad (5.62)$$

We have used equation 5.40 and equation 5.47. We see that the rotation angle is given by the real part of λ_2 (or λ_3). Consider the characteristic equation of \mathbf{R} , Equation 5.12

$$0 = \lambda^3 - \text{trace} \mathbf{R} \lambda^2 + (\mathbf{R}_{11} + \mathbf{R}_{22} + \mathbf{R}_{33}) \lambda - |\mathbf{R}| \quad (5.63)$$

$$= (\lambda - 1)(\lambda - x - yi)(\lambda - x + yi) \quad (5.64)$$

$$= \lambda^3 - (2x + 1)\lambda^2 + (x^2 + 2x + y^2)\lambda - (x^2 + y^2) \quad (5.65)$$

We see that $\text{trace} \mathbf{R} = 2x + 1$ and thus

$$\cos \theta = \frac{1}{2}(\text{trace} \mathbf{R} - 1) \quad (5.66)$$

5.1.6 Matrix $(\mathbf{R} - \mathbf{I})$.

We have seen that $\text{rank}(\mathbf{R} - \mathbf{I}) = 0$ for $\mathbf{R} = \mathbf{I}$ and $\text{rank}(\mathbf{R} - \mathbf{I}) = 2$ for all rotation matrices $\mathbf{R} \neq \mathbf{I}$.

Let us next investigate the relationship between the range and the null space of $(\mathbf{R} - \mathbf{I})$. The null space of $(\mathbf{R} - \mathbf{I})$ is generated by eigenvectors corresponding to 1 since $(\mathbf{R} - \mathbf{I})\vec{v} = \vec{0}$ implies $\mathbf{R}\vec{v} = \vec{v}$.

Now assume that vector \vec{v} is also in the range of $(\mathbf{R} - \mathbf{I})$. Then, there is a vector $\vec{a} \in \mathbb{R}^3$ such that $\vec{v} = (\mathbf{R} - \mathbf{I})\vec{a}$. Let us evaluate the square of the length of \vec{v}

$$\vec{v}^\top \vec{v} = \vec{v}^\top (\mathbf{R} - \mathbf{I}) \vec{a} = (\vec{v}^\top \mathbf{R} - \vec{v}^\top) \vec{a} = (\vec{v}^\top - \vec{v}^\top) \vec{a} = 0 \quad (5.67)$$

which implies $\vec{v} = \vec{0}$. We have used result 5.31 with $x = 1$ and $y = 0$.

We conclude that in this case the range and the null space intersect only in the zero vector.

6 Axis of Motion

We will study motion and show that every motion in three dimensional space has an axis of motion. *Axis of motion* is a line of points that remain in the line after the motion. The existence of such an axis will allow us to decompose every motion into a sequence of a rotation around the axis followed by a translation along the axis as shown in Figure 6.1(a).

§ 29 Algebraic characterization of the axis of motion. Consider Equation 4.7 and denote the motion so defined as $m(\vec{x}_\beta) = \mathbf{R}\vec{x}_\beta + \vec{o}'_\beta$ w.r.t. a fixed coordinate system (O, β) . Now let us study the sets of points that remain fixed by the motion, i.e. sets F such that for all $\vec{x}_\beta \in F$ motion m leaves the $m(\vec{x}_\beta)$ in the set, i.e. $m(\vec{x}_\beta) \in F$. Obviously, complete space and the empty set are fixed sets. How do look other, non-trivial, fixed sets?

A nonempty F contains at least one \vec{x}_β . Then, both $\vec{y}_\beta = m(\vec{x}_\beta)$ and $\vec{z}_\beta = m(\vec{y}_\beta)$ must be in F , see Figure 6.1(b). Let us investigate such fixed points \vec{x}_β for which

$$\vec{z}_\beta - \vec{y}_\beta = \vec{y}_\beta - \vec{x}_\beta \quad (6.1)$$

holds. We do not yet know whether such equality has to necessary hold for points of all fixed sets F but we see that it holds for the identity motion id that leaves all points unchanged, i.e. $id(\vec{x}_\beta) = \vec{x}_\beta$. We will find later that it holds for all motions and all their fixed sets. Consider the following sequence of equatities

$$\begin{aligned} \vec{z}_\beta - \vec{y}_\beta &= \vec{y}_\beta - \vec{x}_\beta \\ \mathbf{R}(\mathbf{R}\vec{x}_\beta + \vec{o}'_\beta) + \vec{o}'_\beta - \mathbf{R}\vec{x}_\beta - \vec{o}'_\beta &= \mathbf{R}\vec{x}_\beta + \vec{o}'_\beta - \vec{x}_\beta \\ \mathbf{R}^2\vec{x}_\beta + \mathbf{R}\vec{o}'_\beta - \mathbf{R}\vec{x}_\beta &= \mathbf{R}\vec{x}_\beta + \vec{o}'_\beta - \vec{x}_\beta \\ \mathbf{R}^2\vec{x}_\beta - 2\mathbf{R}\vec{x}_\beta + \vec{x}_\beta &= -\mathbf{R}\vec{o}'_\beta + \vec{o}'_\beta \\ (\mathbf{R}^2 - 2\mathbf{R} + \mathbf{I})\vec{x}_\beta &= -(\mathbf{R} - \mathbf{I})\vec{o}'_\beta \\ (\mathbf{R} - \mathbf{I})(\mathbf{R} - \mathbf{I})\vec{x}_\beta &= -(\mathbf{R} - \mathbf{I})\vec{o}'_\beta \end{aligned} \quad (6.2)$$

$$(\mathbf{R} - \mathbf{I})((\mathbf{R} - \mathbf{I})\vec{x}_\beta + \vec{o}'_\beta) = 0 \quad (6.3)$$

Equation 6.3 always has a solution. Let us see why.

Recall that $\text{rank}(\mathbf{R} - \mathbf{I})$ is either two or zero. If it is zero, then $\mathbf{R} - \mathbf{I} = 0$ and (i) Equation 6.3 holds for every \vec{x}_β .

Let $\text{rank}(\mathbf{R} - \mathbf{I})$ be two. Vector \vec{o}'_β either is zero or it is not zero. If it is zero, then Equation 6.3 becomes $(\mathbf{R} - \mathbf{I})^2\vec{x}_\beta = 0$, which has (ii) a one-dimensional space of solutions because the null space and the range of $\mathbf{R} - \mathbf{I}$ intersect only in the zero vector for $\mathbf{R} \neq \mathbf{I}$.

Let \vec{o}'_β be non-zero. Vector \vec{o}'_β either is in the span of $\mathbf{R} - \mathbf{I}$ or it is not. If \vec{o}'_β is in the span of $\mathbf{R} - \mathbf{I}$, then $(\mathbf{R} - \mathbf{I})\vec{x}_\beta + \vec{o}'_\beta = 0$ has (iii) one-dimensional affine space of solutions.

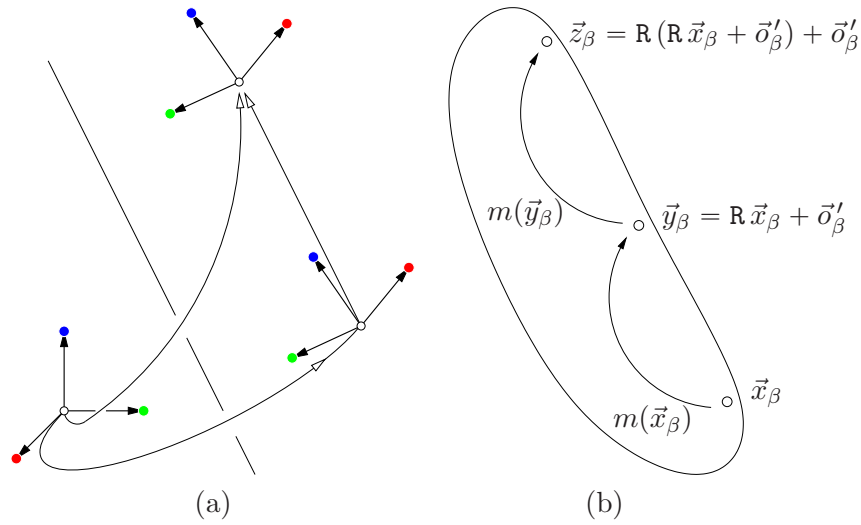


Figure 6.1: Axis of motion.

If \vec{o}'_β is not in the span of $\mathbf{R} - \mathbf{I}$, then $(\mathbf{R} - \mathbf{I})\vec{x}_\beta + \vec{o}'_\beta$ for $\vec{x}_\beta \in \mathbb{R}^3$ generates a vector in all one-dimensional subspaces of \mathbb{R}^3 which are not in the span of $\mathbf{R} - \mathbf{I}$. Therefore, it generates a non-zero vector $\vec{z}_\beta = (\mathbf{R} - \mathbf{I})\vec{y}_\beta + \vec{o}'_\beta$ in the one-dimensional null space of $\mathbf{R} - \mathbf{I}$, because the null space and the span of $(\mathbf{R} - \mathbf{I})$ intersect only in the zero vector for $\mathbf{R} \neq \mathbf{I}$. Equation $(\mathbf{R} - \mathbf{I})\vec{x}_\beta = \vec{z}_\beta - \vec{o}'_\beta$ is satisfied by (iv) a one-dimensional affine set of vectors.

We can conclude that every motion has a fixed line of points for which Equation 6.1 holds. Therefore, every motion has a fixed line of points, every motion has an axis.

§ 30 Geometrical characterization of the axis of motion We now understand the algebraic description of motion. Can we also understand the situation geometrically? Figure 6.2 gives the answer. We shall concentrate on the general situation with $\mathbf{R} \neq \mathbf{I}$ and $\vec{o}'_\beta \neq 0$. The main idea of the figure is that the axis of motion a consists of points that are first rotated away from a by the pure rotation \mathbf{R} around r and then returned back to a by the pure translation \vec{o}'_β .

Figure 6.2 shows axis a of motion, which is parallel to the axis of rotation r and intersects the perpendicular plane σ passing through the origin O at a point P , which is first rotated in σ away from a to P' and then returned back to P'' on a by translation \vec{o}'_β . Point P is determined by the component $\vec{o}'_{\sigma\beta}$ of \vec{o}'_β , which is in the plane σ . Notice that every vector \vec{o}'_β can be written as a sum of its component $\vec{o}'_{r\beta}$ parallel to r and component $\vec{o}'_{\sigma\beta}$ perpendicular to r .

§ 31 Motion axis is parallel to rotation axis. Let us verify algebraically that the rotation axis r is parallel to the motion axis a . Consider Equation 6.2, which we can rewrite as

$$(\mathbf{R} - \mathbf{I})^2 \vec{x}_\beta = -(\mathbf{R} - \mathbf{I}) \vec{o}'_\beta \quad (6.4)$$

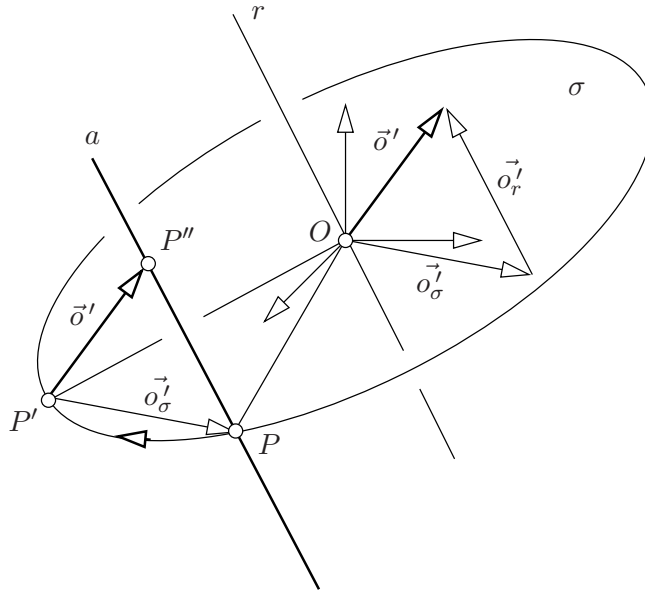


Figure 6.2: Axis a of motion is parallel to the axis of rotation r and intersects the perpendicular plane σ passing through the origin O at a point P , which is first rotated in σ away from a to P' and then returned back to P'' on a by translation \vec{o}' . Point P is determined by the component \vec{o}'_{σ} of \vec{o}' , which is in the plane σ .

Define axis r of motion as the set of points that are left fixed by the pure rotation \mathbf{R} , i.e.

$$(\mathbf{R} - \mathbf{I}) \vec{x}_{\beta} = 0 \quad (6.5)$$

$$\mathbf{R} \vec{x}_{\beta} = \vec{x}_{\beta} \quad (6.6)$$

These are eigenvectors of \mathbf{R} and the zero vector. Take any two solutions $\vec{x}_{1\beta}, \vec{x}_{2\beta}$ of Equation 6.4 and evaluate

$$(\mathbf{R} - \mathbf{I})^2 (\vec{x}_{1\beta} - \vec{x}_{2\beta}) = -(\mathbf{R} - \mathbf{I}) \vec{o}'_{\beta} + (\mathbf{R} - \mathbf{I}) \vec{o}'_{\beta} = 0 \quad (6.7)$$

and thus a non-zero $\vec{x}_{1\beta} - \vec{x}_{2\beta}$ is an eigenvector of \mathbf{R} . We see that the direction vectors of a lie in the subspace of direction vectors of r .

7 Rotation representation and parameterization

We have seen Chapter 5 that rotation can be represented by an orthonormal matrix \mathbf{R} . Matrix \mathbf{R} has nine elements and there are six constraints $\mathbf{R}^\top \mathbf{R} = \mathbf{I}$ and one constraint $|\mathbf{R}| = 1$. Hence, we can view the space of all rotation matrices as a subset of \mathbb{R}^9 . This subset¹ is determined by seven polynomial equations in nine variables. We will next investigate how to describe, i.e. *parameterize*, this set with fewer parameters and fewer constraints.

7.1 Angle-axis representation of rotation

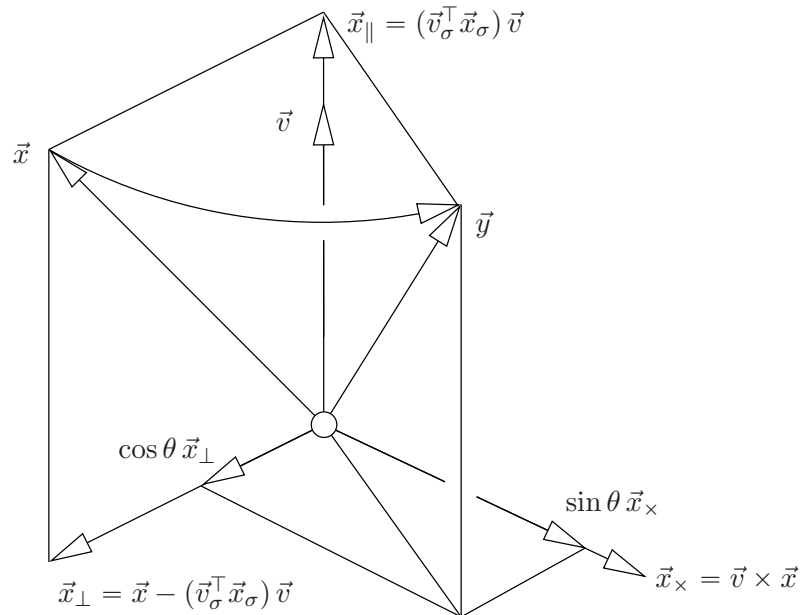


Figure 7.1: Vector \vec{y} is obtained by rotating vector \vec{x} by angle θ around the rotation axis given by unit vector \vec{v} . Vector \vec{y} can be written as a linear combination of an orthogonal basis $[\vec{x} - (\vec{v}^\top \vec{x}_\sigma) \vec{v}, \vec{v} \times \vec{x}, (\vec{v}^\top \vec{x}_\sigma) \vec{v}]$.

We know, Paragraph 5.1.4, that every rotation is determined by a rotation axis and a rotation angle. Let us next give a classical construction of the rotation matrix from an axis and angle.

Figure 7.1 shows how the vector \vec{x} rotates by angle θ around an axis given by a unit vector \vec{v} into vector \vec{y} . To find the relationship between \vec{x} and \vec{y} , we shall construct a special basis of \mathbb{R}^3 . Vector \vec{x} either is, or it is not a multiple of \vec{v} . If it is, then $\vec{y} = \vec{x}$ and $\mathbf{R} = \mathbf{I}$. Let us alternatively consider \vec{x} , which is not a multiple of

¹It is often called algebraic variety in specialized literature [16].

\vec{v} (and hence is not the zero vector!). Further, let us consider the standard basis σ of \mathbb{R}^3 and coordinates of vectors \vec{x}_σ and \vec{v}_σ . We construct three non-zero vectors

$$\vec{x}_{\parallel\sigma} = (\vec{v}_\sigma^\top \vec{x}_\sigma) \vec{v}_\sigma \quad (7.1)$$

$$\vec{x}_{\perp\sigma} = \vec{x} - (\vec{v}_\sigma^\top \vec{x}_\sigma) \vec{v}_\sigma \quad (7.2)$$

$$\vec{x}_{\times\sigma} = \vec{v}_\sigma \times \vec{x}_\sigma \quad (7.3)$$

which are mutually orthogonal and hence form a basis of \mathbb{R}^3 . We may notice that coordinate vectors $\vec{x} \in \mathbb{R}^3$, are actually equal to their coordinates w.r.t. the standard basis σ . Hence we can drop σ index and write

$$\vec{x}_{\parallel} = (\vec{v}^\top \vec{x}) \vec{v} = \vec{v} (\vec{v}^\top \vec{x}) = (\vec{v} \vec{v}^\top) \vec{x} = [\vec{v}]_{\parallel} \vec{x} \quad (7.4)$$

$$\vec{x}_{\perp} = \vec{x} - (\vec{v}^\top \vec{x}) \vec{v} = \vec{x} - (\vec{v} \vec{v}^\top) \vec{x} = (\mathbf{I} - \vec{v} \vec{v}^\top) \vec{x} = [\vec{v}]_{\perp} \vec{x} \quad (7.5)$$

$$\vec{x}_{\times} = \vec{v} \times \vec{x} = [\vec{v}]_{\times} \vec{x} \quad (7.6)$$

We have introduced two new matrices

$$[\vec{v}]_{\parallel} = \vec{v} \vec{v}^\top \quad \text{and} \quad [\vec{v}]_{\perp} = \mathbf{I} - \vec{v} \vec{v}^\top \quad (7.7)$$

Let us next study how the three matrices $[\vec{v}]_{\parallel}$, $[\vec{v}]_{\perp}$, $[\vec{v}]_{\times}$ behave under the transposition and mutual multiplication. We see that the following identities

$$\begin{aligned} [\vec{v}]_{\parallel}^\top &= [\vec{v}]_{\parallel}, & [\vec{v}]_{\parallel} [\vec{v}]_{\parallel} &= [\vec{v}]_{\parallel}, & [\vec{v}]_{\parallel} [\vec{v}]_{\perp} &= \mathbf{0}, & [\vec{v}]_{\parallel} [\vec{v}]_{\times} &= \mathbf{0}, \\ [\vec{v}]_{\perp}^\top &= [\vec{v}]_{\perp}, & [\vec{v}]_{\perp} [\vec{v}]_{\parallel} &= \mathbf{0}, & [\vec{v}]_{\perp} [\vec{v}]_{\perp} &= [\vec{v}]_{\perp}, & [\vec{v}]_{\perp} [\vec{v}]_{\times} &= [\vec{v}]_{\times}, \\ [\vec{v}]_{\times}^\top &= -[\vec{v}]_{\times}, & [\vec{v}]_{\times} [\vec{v}]_{\parallel} &= \mathbf{0}, & [\vec{v}]_{\times} [\vec{v}]_{\perp} &= [\vec{v}]_{\times}, & [\vec{v}]_{\times} [\vec{v}]_{\times} &= -[\vec{v}]_{\perp} \end{aligned} \quad (7.8)$$

hold true. The last identity is obtained as follows

$$[\vec{v}]_{\times} [\vec{v}]_{\times} = \begin{bmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{bmatrix} \begin{bmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{bmatrix} \quad (7.9)$$

$$= \begin{bmatrix} -v_2^2 - v_3^2 & v_1 v_2 & v_1 v_3 \\ v_1 v_2 & -v_1^2 - v_3^2 & v_2 v_3 \\ v_1 v_3 & v_2 v_3 & -v_1^2 - v_2^2 \end{bmatrix} \quad (7.10)$$

$$= \begin{bmatrix} v_1^2 - 1 & v_1 v_2 & v_1 v_3 \\ v_1 v_2 & v_2^2 - 1 & v_2 v_3 \\ v_1 v_3 & v_2 v_3 & v_3^2 - 1 \end{bmatrix} = [\vec{v}]_{\parallel} - \mathbf{I} = -[\vec{v}]_{\perp} \quad (7.11)$$

We can now write \vec{y} in the basis $[\vec{x}_{\parallel}, \vec{x}_{\perp}, \vec{x}_{\times}]$

$$\vec{y} = \vec{x}_{\parallel} + \|\vec{x}_{\perp}\| \cos \theta \frac{\vec{x}_{\perp}}{\|\vec{x}_{\perp}\|} + \|\vec{x}_{\perp}\| \sin \theta \frac{\vec{x}_{\times}}{\|\vec{x}_{\times}\|} \quad (7.12)$$

$$= [\vec{v}]_{\parallel} \vec{x} + \cos \theta [\vec{v}]_{\perp} \vec{x} + \sin \theta [\vec{v}]_{\times} \vec{x} \quad (7.13)$$

$$= ([\vec{v}]_{\parallel} + \cos \theta [\vec{v}]_{\perp} + \sin \theta [\vec{v}]_{\times}) \vec{x} = \mathbf{R} \vec{x} \quad (7.14)$$

We used the fact that $\|\vec{x}_{\perp}\| = \|\vec{x}_{\times}\|$ which follows from

$$\|\vec{x}_{\times}\|^2 = \vec{x}_{\times}^\top \vec{x}_{\times} = \vec{x}^\top [\vec{v}]_{\times}^\top [\vec{v}]_{\times} \vec{x} = \vec{x}^\top (-[\vec{v}]_{\times}^2) \vec{x} = \vec{x}^\top [\vec{v}]_{\perp} \vec{x} \quad (7.15)$$

$$\|\vec{x}_{\perp}\|^2 = \vec{x}_{\perp}^\top \vec{x}_{\perp} = \vec{x}^\top [\vec{v}]_{\perp}^\top [\vec{v}]_{\perp} \vec{x} = \vec{x}^\top [\vec{v}]_{\perp}^2 \vec{x} = \vec{x}^\top [\vec{v}]_{\perp} \vec{x} \quad (7.16)$$

We obtained matrix

$$\mathbf{R} = [\vec{v}]_{\parallel} + \cos \theta [\vec{v}]_{\perp} + \sin \theta [\vec{v}]_{\times} \quad (7.17)$$

Let us check that this indeed is a rotation matrix

$$\begin{aligned} \mathbf{R}^{\top} \mathbf{R} &= \left([\vec{v}]_{\parallel} + \cos \theta [\vec{v}]_{\perp} + \sin \theta [\vec{v}]_{\times} \right)^{\top} \left([\vec{v}]_{\parallel} + \cos \theta [\vec{v}]_{\perp} + \sin \theta [\vec{v}]_{\times} \right) \\ &= \left([\vec{v}]_{\parallel} + \cos \theta [\vec{v}]_{\perp} - \sin \theta [\vec{v}]_{\times} \right) \left([\vec{v}]_{\parallel} + \cos \theta [\vec{v}]_{\perp} + \sin \theta [\vec{v}]_{\times} \right) \\ &= [\vec{v}]_{\parallel} + \cos^2 \theta [\vec{v}]_{\perp} + \sin \theta \cos \theta [\vec{v}]_{\times} - \sin \theta \cos \theta [\vec{v}]_{\times} + \sin^2 \theta [\vec{v}]_{\perp} \\ &= [\vec{v}]_{\parallel} + [\vec{v}]_{\perp} = \mathbf{I} \end{aligned} \quad (7.18)$$

It is sometimes useful to write \mathbf{R} as a function of $[\vec{v}]_{\times}$

$$\mathbf{R} = [\vec{v}]_{\parallel} + \cos \theta [\vec{v}]_{\perp} + \sin \theta [\vec{v}]_{\times} \quad (7.19)$$

$$= [\vec{v}]_{\parallel} + \cos \theta (\mathbf{I} - [\vec{v}]_{\parallel}) + \sin \theta [\vec{v}]_{\times} \quad (7.20)$$

$$= (1 - \cos \theta) [\vec{v}]_{\parallel} + \cos \theta \mathbf{I} + \sin \theta [\vec{v}]_{\times} \quad (7.21)$$

$$= (1 - \cos \theta) ([\vec{v}]_{\times}^2 + \mathbf{I}) + \cos \theta \mathbf{I} + \sin \theta [\vec{v}]_{\times} \quad (7.22)$$

$$= \mathbf{I} + \sin \theta [\vec{v}]_{\times} + (1 - \cos \theta) [\vec{v}]_{\times}^2 \quad (7.23)$$

§ 32 **Angle-axis parameterization** Let us write \mathbf{R} in more detail

$$\mathbf{R} = (1 - \cos \theta) [\vec{v}]_{\parallel} + \cos \theta \mathbf{I} + \sin \theta [\vec{v}]_{\times} \quad (7.24)$$

$$\begin{aligned} &= (1 - \cos \theta) \begin{bmatrix} v_1 v_1 & v_1 v_2 & v_1 v_3 \\ v_2 v_1 & v_2 v_2 & v_2 v_3 \\ v_3 v_1 & v_3 v_2 & v_3 v_3 \end{bmatrix} + \cos \theta \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \sin \theta \begin{bmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} v_1 v_1 (1 - \cos \theta) + \cos \theta & v_1 v_2 (1 - \cos \theta) - v_3 \sin \theta & v_1 v_3 (1 - \cos \theta) + v_2 \sin \theta \\ v_2 v_1 (1 - \cos \theta) + v_3 \sin \theta & v_2 v_2 (1 - \cos \theta) + \cos \theta & v_2 v_3 (1 - \cos \theta) - v_1 \sin \theta \\ v_3 v_1 (1 - \cos \theta) - v_2 \sin \theta & v_3 v_2 (1 - \cos \theta) + v_1 \sin \theta & v_3 v_3 (1 - \cos \theta) + \cos \theta \end{bmatrix} \end{aligned} \quad (7.25)$$

which allows us to parameterize rotation by four numbers

$$\begin{bmatrix} \theta & v_1 & v_2 & v_3 \end{bmatrix}^{\top} \quad \text{with} \quad v_1^2 + v_2^2 + v_3^2 = 1 \quad (7.26)$$

The parameterization uses goniometric functions.

§ 33 **Computing the axis and the angle of rotation from \mathbf{R}** Let us now discuss how to get a unit vector \vec{v} of the axis and the corresponding angle θ of rotation from a rotation matrix \mathbf{R} , such that the pair $[\theta, \vec{v}]$ gives \mathbf{R} by Equation 7.25. To avoid multiple representations due to periodicity of θ , we will confine θ to real interval $(-\pi, \pi]$.

We can get $\cos(\theta)$ from Equation 5.66.

If $\cos \theta = 1$, then $\sin \theta = 0$, and thus $\theta = 0$. Then, $\mathbf{R} = \mathbf{I}$ and any unit vector can be taken as \vec{v} , i.e. all pairs $[0, \vec{v}]$ for unit vector $\vec{v} \in \mathbb{R}^3$ represent \mathbf{I} .

If $\cos \theta = -1$, then $\sin \theta = 0$, and thus $\theta = \pi$. Then \mathbf{R} is a symmetrical matrix and we use Equation 5.56 to get \vec{v}_1 , a non-zero multiple of \vec{v} , i.e. $\vec{v} = \alpha \vec{v}_1$, with real

non-zero α , and therefore $\vec{v}_1/\|\vec{v}_1\| = s \vec{v}$ with $s = \pm 1$. We are getting

$$\mathbf{R} = 2 [\vec{v}]_{\parallel} - \mathbf{I} = 2 \vec{v} \vec{v}^{\top} - \mathbf{I} = 2 s^2 \vec{v} \vec{v}^{\top} - \mathbf{I} = 2 (s \vec{v}) (s \vec{v})^{\top} - \mathbf{I} \quad (7.27)$$

$$= 2 \left(\frac{\vec{v}_1}{\|\vec{v}_1\|} \right) \left(\frac{\vec{v}_1}{\|\vec{v}_1\|} \right)^{\top} - \mathbf{I} = 2 \left(-\frac{\vec{v}_1}{\|\vec{v}_1\|} \right) \left(-\frac{\vec{v}_1}{\|\vec{v}_1\|} \right)^{\top} - \mathbf{I} \quad (7.28)$$

from Equation 7.24 and hence we can form two pairs

$$\left[\pi, +\frac{\vec{v}_1}{\|\vec{v}_1\|} \right], \quad \left[\pi, -\frac{\vec{v}_1}{\|\vec{v}_1\|} \right] \quad (7.29)$$

representing this rotation.

Let's now move to $-1 < \cos \theta < 1$. We construct matrix

$$\begin{aligned} \mathbf{R} - \mathbf{R}^{\top} &= (1 - \cos \theta) [\vec{v}]_{\parallel} + \cos \theta \mathbf{I} + \sin \theta [\vec{v}]_{\times} \\ &\quad - \left((1 - \cos \theta) [\vec{v}]_{\parallel} + \cos \theta \mathbf{I} + \sin \theta [\vec{v}]_{\times} \right)^{\top} \end{aligned} \quad (7.30)$$

$$\begin{aligned} &= (1 - \cos \theta) [\vec{v}]_{\parallel} + \cos \theta \mathbf{I} + \sin \theta [\vec{v}]_{\times} \\ &\quad - \left((1 - \cos \theta) [\vec{v}]_{\parallel} + \cos \theta \mathbf{I} - \sin \theta [\vec{v}]_{\times} \right) \end{aligned} \quad (7.31)$$

$$= 2 \sin \theta [\vec{v}]_{\times} \quad (7.32)$$

which gives

$$\begin{bmatrix} 0 & r_{12} - r_{21} & r_{13} - r_{31} \\ r_{21} - r_{12} & 0 & r_{23} - r_{32} \\ r_{31} - r_{13} & r_{32} - r_{23} & 0 \end{bmatrix} = 2 \sin \theta \begin{bmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{bmatrix} \quad (7.33)$$

and thus

$$\sin \theta \vec{v} = \frac{1}{2} \begin{bmatrix} r_{32} - r_{23} \\ r_{13} - r_{31} \\ r_{21} - r_{12} \end{bmatrix} \quad (7.34)$$

We thus get

$$|\sin \theta| \|\vec{v}\| = |\sin \theta| = \frac{1}{2} \sqrt{(r_{23} - r_{32})^2 + (r_{31} - r_{13})^2 + (r_{12} - r_{21})^2} \quad (7.35)$$

There holds

$$\sin \theta \vec{v} = \sin(-\theta) (-\vec{v}) \quad (7.36)$$

true and hence we define

$$\theta = \arccos \left(\frac{1}{2} (\text{trace}(\mathbf{R}) - 1) \right), \quad \vec{r} = \frac{1}{2} \begin{bmatrix} r_{32} - r_{23} \\ r_{13} - r_{31} \\ r_{21} - r_{12} \end{bmatrix} \quad (7.37)$$

and write two pairs

$$\left[+\theta, +\frac{\vec{r}}{\sin \theta} \right], \quad \left[-\theta, -\frac{\vec{r}}{\sin \theta} \right] \quad (7.38)$$

representing rotation \mathbf{R} .

We see that all rotations are represented by two pairs of $[\theta, \vec{v}]$ except for the identity, which is represented by an infinite number of pairs.

7.2 Euler vector representation and the exponential map

Let us now discuss another classical and natural representation of rotations. It may seem as only a slight variation of the angle-axis representation but it leads to several interesting connections and properties.

Let us consider the *euler vector* defined as

$$\vec{e} = \theta \vec{v} \quad (7.39)$$

where θ is the rotation angle and \vec{v} is the unit vector representing the rotation axis in the angle-axis representation as in Equation 7.24.

Next, let us recall the very fundamental real functions [2] and their related power series

$$\exp x = \sum_{n=0}^{\infty} \frac{x^n}{n!} \quad (7.40)$$

$$\sin x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} \quad (7.41)$$

$$\cos x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} \quad (7.42)$$

It makes sense to define the exponential function of an $m \times m$ real matrix $\mathbf{A} \in \mathbb{R}^{m \times m}$ as

$$\exp \mathbf{A} = \sum_{n=0}^{\infty} \frac{\mathbf{A}^n}{n!} \quad (7.43)$$

We will now show that the rotation matrix \mathbf{R} corresponding to the angle-axis parameterization $[\theta, \vec{v}]$ can be obtained as

$$\mathbf{R}([\theta, \vec{v}]) = \exp [\vec{e}]_{\times} = \exp [\theta \vec{v}]_{\times} \quad (7.44)$$

The basic tool we have to employ is the relationship between $[\vec{e}]_{\times}^3$ and $[\vec{e}]_{\times}$. It will allow us to pass from the infinite summation of matrix powers to the infinite summation of the powers of the θ and hence to $\sin \theta$ and $\cos \theta$, which will, at the end, give the Rodrigues formula. We write, Equation 7.11,

$$\begin{aligned} [\theta \vec{v}]_{\times}^2 &= \theta^2 (\vec{v} \vec{v}^{\top} - \mathbf{I}) \\ [\theta \vec{v}]_{\times}^3 &= -\theta^2 [\theta \vec{v}]_{\times} \\ [\theta \vec{v}]_{\times}^4 &= -\theta^2 [\theta \vec{v}]_{\times}^2 \\ [\theta \vec{v}]_{\times}^5 &= \theta^4 [\theta \vec{v}]_{\times} \\ [\theta \vec{v}]_{\times}^6 &= \theta^4 [\theta \vec{v}]_{\times}^2 \\ &\vdots \end{aligned} \quad (7.45)$$

and substitute into Equation 7.43 to get

$$\exp [\theta \vec{v}]_{\times} = \sum_{n=0}^{\infty} \frac{[\theta \vec{v}]_{\times}^n}{n!} \quad (7.46)$$

$$= \sum_{n=0}^{\infty} \frac{[\theta \vec{v}]_{\times}^{2n}}{(2n)!} + \sum_{n=0}^{\infty} \frac{[\theta \vec{v}]_{\times}^{2n+1}}{(2n+1)!} \quad (7.47)$$

Let us notice the identities, which are obtained by generalizing Equations 7.45 to an arbitrary power n

$$[\theta \vec{v}]_{\times}^0 = \mathbf{I} \quad (7.48)$$

$$[\theta \vec{v}]_{\times}^{2n} = (-1)^{n-1} \theta^{2(n-1)} [\theta \vec{v}]_{\times}^2 \text{ for } n = 1, \dots \quad (7.49)$$

$$[\theta \vec{v}]_{\times}^{2n+1} = (-1)^n \theta^{2n} [\theta \vec{v}]_{\times} \text{ for } n = 0, \dots \quad (7.50)$$

and substitute them into Equation 7.47 to get

$$\begin{aligned} \exp [\theta \vec{v}]_{\times} &= \mathbf{I} + \left(\sum_{n=1}^{\infty} \frac{(-1)^{n-1} \theta^{2(n-1)}}{(2n)!} \right) [\theta \vec{v}]_{\times}^2 + \left(\sum_{n=0}^{\infty} \frac{(-1)^n \theta^{2n}}{(2n+1)!} \right) [\theta \vec{v}]_{\times} \\ &= \mathbf{I} + \left(\sum_{n=1}^{\infty} \frac{(-1)^{n-1} \theta^{2n}}{(2n)!} \right) [\vec{v}]_{\times}^2 + \left(\sum_{n=0}^{\infty} \frac{(-1)^n \theta^{2n+1}}{(2n+1)!} \right) [\vec{v}]_{\times} \\ &= \mathbf{I} - \left(\sum_{n=0}^{\infty} \frac{(-1)^n \theta^{2n}}{(2n)!} - 1 \right) [\vec{v}]_{\times}^2 + \sin \theta [\vec{v}]_{\times} \\ &= \mathbf{I} - (\cos \theta - 1) [\vec{v}]_{\times}^2 + \sin \theta [\vec{v}]_{\times} \\ &= \mathbf{I} + \sin \theta [\vec{v}]_{\times} + (1 - \cos \theta) [\vec{v}]_{\times}^2 \\ &= \mathbf{I} + \sin \|\vec{e}\| \left[\frac{\vec{e}}{\|\vec{e}\|} \right]_{\times} + (1 - \cos \|\vec{e}\|) \left[\frac{\vec{e}}{\|\vec{e}\|} \right]_{\times}^2 \\ &= \mathbf{R}([\theta, \vec{v}]) \end{aligned} \quad (7.51)$$

by the comparison with Equation 7.23.

7.3 Quaternion representation of rotation

§ 34 Quaternion parameterization We shall now introduce another parameterization of \mathbf{R} by four numbers but this time we will not use goniometric functions but only polynomials. We shall see later that this parameterization has other useful properties.

Let us do a seemingly unnecessary trick. We will pass from θ to $\frac{\theta}{2}$ and introduce

$$\vec{q} = \begin{bmatrix} \cos \frac{\theta}{2} \\ \vec{v} \sin \frac{\theta}{2} \end{bmatrix} = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{bmatrix} = \begin{bmatrix} \cos \frac{\theta}{2} \\ v_1 \sin \frac{\theta}{2} \\ v_2 \sin \frac{\theta}{2} \\ v_3 \sin \frac{\theta}{2} \end{bmatrix} \quad (7.52)$$

There still holds

$$\|\vec{q}\| = q_1^2 + q_2^2 + q_3^2 + q_4^2 = \cos^2 \frac{\theta}{2} + \sin^2 \frac{\theta}{2} v_1^2 + \sin^2 \frac{\theta}{2} v_2^2 + \sin^2 \frac{\theta}{2} v_3^2 = \cos^2 \frac{\theta}{2} + \sin^2 \frac{\theta}{2} = 1 \quad (7.53)$$

We can verify that there holds

$$\sin \theta = 2 \cos \frac{\theta}{2} \sin \frac{\theta}{2} \quad (7.54)$$

$$\cos \theta = 1 - 2 \sin^2 \frac{\theta}{2} \quad (7.55)$$

$$1 - \cos \theta = 2 \sin^2 \frac{\theta}{2} \quad (7.56)$$

true. We can now substitute the above into Equation 7.21 to get

$$\mathbf{R} = \mathbf{I} + \sin \theta [\vec{v}]_{\times} + (1 - \cos \theta) [\vec{v}]_{\times}^2 \quad (7.57)$$

$$= \mathbf{I} + 2 \cos \frac{\theta}{2} \sin \frac{\theta}{2} [\vec{v}]_{\times} + 2 \sin^2 \frac{\theta}{2} [\vec{v}]_{\times}^2 \quad (7.58)$$

$$= \mathbf{I} + 2 \cos \frac{\theta}{2} \left[\sin \frac{\theta}{2} \vec{v} \right]_{\times} + 2 \left[\sin \frac{\theta}{2} \vec{v} \right]_{\times}^2 \quad (7.59)$$

$$= \mathbf{I} + 2 \cos \frac{\theta}{2} \left[\sin \frac{\theta}{2} \vec{v} \right]_{\times} + 2 \left(\left[\sin \frac{\theta}{2} \vec{v} \right]_{\parallel} - \mathbf{I} \right) \quad (7.60)$$

$$= \mathbf{I} + 2 q_1 \left[\begin{bmatrix} q_2 \\ q_3 \\ q_4 \end{bmatrix} \right]_{\times} + 2 \left(\left[\begin{bmatrix} q_2 \\ q_3 \\ q_4 \end{bmatrix} \right]_{\parallel} - \mathbf{I} \right) \quad (7.61)$$

$$= \begin{bmatrix} 1 & -2 q_1 q_4 & 2 q_1 q_3 \\ 2 q_1 q_4 & 1 & -2 q_1 q_2 \\ -2 q_1 q_3 & 2 q_1 q_2 & 1 \end{bmatrix} + \begin{bmatrix} 2 q_2 q_2 - 2 & 2 q_2 q_3 & 2 q_2 q_4 \\ 2 q_3 q_2 & 2 q_3 q_3 - 2 & 2 q_3 q_4 \\ 2 q_4 q_2 & 2 q_4 q_3 & 2 q_4 q_4 - 2 \end{bmatrix} \\ = \begin{bmatrix} q_1^2 + q_2^2 - q_3^2 - q_4^2 & 2 (q_2 q_3 - q_1 q_4) & 2 (q_2 q_4 + q_1 q_3) \\ 2 (q_2 q_3 + q_1 q_4) & q_1^2 - q_2^2 + q_3^2 - q_4^2 & 2 (q_3 q_4 - q_1 q_2) \\ 2 (q_2 q_4 - q_1 q_3) & 2 (q_3 q_4 + q_1 q_2) & q_1^2 - q_2^2 - q_3^2 + q_4^2 \end{bmatrix} \quad (7.62)$$

which uses only second order polynomials in elements of \vec{q} .

§ 35 Computing quaternions from R To get the quaternions representing a rotation matrix \mathbf{R} , we start with Equation 7.59. Let us first confine θ to the real interval $(-\pi, \pi]$ as we did for the angle-axis parameterization.

Matrix \mathbf{R} either is or it is not symmetric.

If \mathbf{R} is symmetric, then either $\sin \theta/2 \vec{v} = \vec{0}$ or $\cos \theta/2 = 0$. If $\sin \theta/2 \vec{v} = \vec{0}$, then $\sin \theta/2 = 0$ since $\|\vec{v}\| = 1$ and thus $\cos \theta/2 = \pm 1$. However, $\cos \theta/2 = -1$ for no $\theta \in (-\pi, \pi]$ and hence $\cos \theta/2 = 1$. This corresponds to $\theta = 0$ and hence to $\mathbf{R} = \mathbf{I}$ which is thus represented by quaternion

$$[1 \ 0 \ 0 \ 0]^{\top} \quad (7.63)$$

If $\cos \theta/2 = 0$, then $\sin \theta/2 = \pm 1$ but $\sin \theta/2 = -1$ for no $\theta \in (-\pi, \pi]$ and hence $\sin \theta/2 = 1$. This corresponds to the rotation the by $\theta = \pi$ around the axis given by unit $\vec{v} = [v_1, v_2, v_3]^{\top}$. This rotation is thus represented by quaternion

$$[0 \ v_1 \ v_2 \ v_3]^{\top} \quad (7.64)$$

Notice that \vec{v} and $-\vec{v}$ generate the same rotation matrix \mathbf{R} and hence every rotation by $\theta = \pi$ is represented by two quaternions.

If \mathbf{R} is not symmetric, then $\mathbf{R} - \mathbf{R}^{\top} \neq 0$ and hence we are getting a useful relationship

$$\mathbf{R} - \mathbf{R}^{\top} = 4 \cos \frac{\theta}{2} \left[\sin \frac{\theta}{2} \vec{v} \right]_{\times} \quad (7.65)$$

and next continue with writing

$$\cos^2 \frac{\theta}{2} = 1 - \sin^2 \frac{\theta}{2} = 1 - \frac{1}{2} (1 - \cos \theta) = 1 - \frac{1}{2} \left(1 - \frac{1}{2} (\text{trace } \mathbf{R} - 1) \right) = \frac{1}{4} (1 + \text{trace } \mathbf{R}) \quad (7.66)$$

using $\text{trace } \mathbf{R}$, and thus

$$q_1 = \cos \frac{\theta}{2} = \frac{s}{2} \sqrt{\text{trace } \mathbf{R} + 1} \quad (7.67)$$

with $s = \pm 1$. We can form equation

$$\begin{bmatrix} 0 & r_{12} - r_{21} & r_{13} - r_{31} \\ r_{21} - r_{12} & 0 & r_{23} - r_{32} \\ r_{31} - r_{13} & r_{32} - r_{23} & 0 \end{bmatrix} = \begin{bmatrix} r_{32} - r_{23} \\ r_{13} - r_{31} \\ r_{21} - r_{12} \end{bmatrix}_{\times} = s \sqrt{\text{trace } \mathbf{R} + 1} \begin{bmatrix} q_2 \\ q_3 \\ q_4 \end{bmatrix}_{\times} \quad (7.68)$$

which gives the following two quaternions

$$\frac{+1}{2\sqrt{\text{trace } \mathbf{R} + 1}} \begin{bmatrix} \text{trace } \mathbf{R} + 1 \\ r_{32} - r_{23} \\ r_{13} - r_{31} \\ r_{21} - r_{12} \end{bmatrix}, \quad \frac{-1}{2\sqrt{\text{trace } \mathbf{R} + 1}} \begin{bmatrix} \text{trace } \mathbf{R} + 1 \\ r_{32} - r_{23} \\ r_{13} - r_{31} \\ r_{21} - r_{12} \end{bmatrix} \quad (7.69)$$

which represent the same rotation as \mathbf{R} .

We see that all rotations are represented by two quaternions \vec{q} and $-\vec{q}$ except for the identity, which is represented by exactly one quaternion.

The quaternion representation of rotation presented above represents every rotation by a finite number of quaternions whereas angle-axis representation allowed for an infinite number of angle-axis pairs to correspond to the identity. Yet, even this still has an “aesthetic flaw” at the identity, which has only one quaternion whereas all other rotations have two quaternions. The “flaw” can be removed by realizing that $\vec{q} = [-1, 0, 0, 0]^T$ also maps to the identity. However, if we look for θ that corresponds to $\cos \theta/2 = -1$ we see that such $\theta/2 = \pm k\pi$ and hence $\theta = \pm 2k\pi$ for $k = 1, 2, \dots$, which are points isolated from $(\pi, \pi]$. Now, if we allow θ to be in interval $(-2\pi, +2\pi]$, then the set

$$\left\{ \begin{bmatrix} \cos \theta/2 \\ \vec{v} \sin \theta/2 \end{bmatrix} \mid \theta \in [-2\pi, +2\pi], \vec{v} \in \mathbb{R}^3, \|\vec{v}\| = 1 \right\} \quad (7.70)$$

of quaternions contains exactly two quaternions for every rotation matrix \mathbf{R} and is obtained by a continuous mapping of a closed interval of angles, which is bounded, times a sphere in \mathbb{R}^3 , which is also closed and bounded.

§ 36 Quaternion composition Consider two rotations represented by \vec{q}_1 and \vec{q}_2 . The respective rotation matrices $\mathbf{R}_1, \mathbf{R}_2$ can be composed into rotation matrix $\mathbf{R}_{21} = \mathbf{R}_2 \mathbf{R}_1$, which can be represented by \vec{q}_{21} . Let us investigate how to obtain \vec{q}_{21} from \vec{q}_1 and \vec{q}_2 . We shall use Equation 7.71 to relate \mathbf{R}_1 to \vec{q}_1 and \mathbf{R}_2 to \vec{q}_1 , then evaluate $\mathbf{R}_{21} = \mathbf{R}_2 \mathbf{R}_1$ and recover \vec{q}_{21} from \mathbf{R}_{21} . We use Equation 7.21 to write

$$\mathbf{R} = 2 \sin^2 \frac{\theta}{2} \vec{v} \vec{v}^T + (2 \cos^2 \frac{\theta}{2} - 1) \mathbf{I} + 2 \cos \frac{\theta}{2} \sin \frac{\theta}{2} [\vec{v}]_{\times} \quad (7.71)$$

and

$$\mathbf{R}_1 = 2 (s_1 \vec{v}_1) (s_1 \vec{v}_1)^T + (2 c_1^2 - 1) \mathbf{I} + 2 c_1 [s_1 \vec{v}_1]_{\times} \quad (7.72)$$

$$\mathbf{R}_2 = 2 (s_2 \vec{v}_2) (s_2 \vec{v}_2)^T + (2 c_2^2 - 1) \mathbf{I} + 2 c_2 [s_2 \vec{v}_2]_{\times} \quad (7.73)$$

$$\mathbf{R}_{21} = 2 (s_{21} \vec{v}_{21}) (s_{21} \vec{v}_{21})^T + (2 c_{21}^2 - 1) \mathbf{I} + 2 c_{21} [s_{21} \vec{v}_{21}]_{\times}$$

with shortcuts

$$c_1 = \cos \frac{\theta_1}{2}, s_1 = \sin \frac{\theta_1}{2}, c_2 = \cos \frac{\theta_2}{2}, s_2 = \sin \frac{\theta_2}{2}, c_{21} = \cos \frac{\theta_{21}}{2}, s_{21} = \sin \frac{\theta_{21}}{2}$$

Let us next assume that both $\mathbf{R}_1, \mathbf{R}_2$ are not identities. Then $\theta_1 \neq 0$ and $\theta_2 \neq 0$ and rotation axes $\vec{v}_1 \neq \vec{0}, \vec{v}_2 \neq \vec{0}$ are well defined. We can now distinguish two cases. Either $\vec{v}_1 = \pm \vec{v}_2$, and then $\vec{v}_{21} = \vec{v}_1 = \pm \vec{v}_2$, or $\vec{v}_1 \neq \pm \vec{v}_2$, and then

$$[\vec{v}_1, \vec{v}_2, \vec{v}_2 \times \vec{v}_1] \quad (7.74)$$

forms a basis of \mathbb{R}^3 . We can thus write

$$\sin \frac{\theta_{21}}{2} \vec{v}_{21} = a_1 \sin \frac{\theta_1}{2} \vec{v}_1 + a_2 \sin \frac{\theta_2}{2} \vec{v}_2 + a_3 (\vec{v}_2 \times \vec{v}_1) \quad (7.75)$$

with coefficients $a_1, a_2, a_3 \in \mathbb{R}$. To find coefficients a_1, a_2, a_3 , we will consider the following special situations:

1. $\vec{v}_1 = \pm \vec{v}_2$ implies $\vec{v}_{21} = \vec{v}_1 = \pm \vec{v}_2$ and $\theta_{21} = \theta_1 \pm \theta_2$ for all real θ_1 and θ_2 .
2. $\vec{v}_2^\top \vec{v}_1 = 0$ and $\theta_1 = \theta_2 = \pi$ implies

$$\mathbf{R}_1 = 2 \vec{v}_1 \vec{v}_1^\top - \mathbf{I} \quad (7.76)$$

$$\mathbf{R}_2 = 2 \vec{v}_2 \vec{v}_2^\top - \mathbf{I} \quad (7.77)$$

$$\mathbf{R}_{21} = (2 \vec{v}_2 \vec{v}_2^\top - \mathbf{I})(2 \vec{v}_1 \vec{v}_1^\top - \mathbf{I}) = \mathbf{I} - 2(\vec{v}_2 \vec{v}_2^\top + \vec{v}_1 \vec{v}_1^\top) \quad (7.78)$$

We see that in the former case we are getting

$$\sin \frac{\theta_{21}}{2} \vec{v}_1 = (a_1 \sin \frac{\theta_1}{2} + a_1 \sin \frac{\theta_2}{2}) \vec{v}_1 \quad \text{for all } \theta_1, \theta_2 \in \mathbb{R} \quad (7.79)$$

which for $\vec{v}_1 \neq \vec{0}$ leads to

$$\sin \frac{\theta_{21}}{2} = a_1 \sin \frac{\theta_1}{2} + a_2 \sin \frac{\theta_2}{2} \quad (7.80)$$

$$\sin \frac{\theta_1 + \theta_2}{2} = a_1 \sin \frac{\theta_1}{2} + a_2 \sin \frac{\theta_2}{2} \quad (7.81)$$

$$\sin \frac{\theta_1}{2} \cos \frac{\theta_2}{2} + \cos \frac{\theta_1}{2} \sin \frac{\theta_2}{2} = a_1 \sin \frac{\theta_1}{2} + a_2 \sin \frac{\theta_2}{2} \quad (7.82)$$

for all $\theta_1, \theta_2 \in \mathbb{R}$. But that means that

$$a_1 = \cos \frac{\theta_2}{2} \quad \text{and} \quad a_2 = \cos \frac{\theta_1}{2} \quad (7.83)$$

In the latter case we find that \vec{v}_{21} is a non-zero multiple of $\vec{v}_2 \times \vec{v}_1$ since

$$\mathbf{R}_{21} (\vec{v}_2 \times \vec{v}_1) = (\mathbf{I} - 2(\vec{v}_2 \vec{v}_2^\top + \vec{v}_1 \vec{v}_1^\top)) (\vec{v}_2 \times \vec{v}_1) \quad (7.84)$$

$$= \vec{v}_2 \times \vec{v}_1 - 2 \vec{v}_2 \vec{v}_2^\top (\vec{v}_2 \times \vec{v}_1) - 2 \vec{v}_1 \vec{v}_1^\top (\vec{v}_2 \times \vec{v}_1) \quad (7.85)$$

$$= \vec{v}_2 \times \vec{v}_1 \quad (7.86)$$

But that means that

$$\sin \frac{\theta_{21}}{2} (\vec{v}_2 \times \vec{v}_1) = a_3 (\vec{v}_2 \times \vec{v}_1) \quad (7.87)$$

and hence for non-zero $\vec{v}_2 \times \vec{v}_1$ we are getting

$$a_3 = \sin \frac{\theta_{21}}{2} \quad (7.88)$$

We get θ_{21} using Equation 5.66 as

$$\cos \theta_{21} = \frac{1}{2}(\text{trace } \mathbf{R} - 1) \quad (7.89)$$

$$= \frac{1}{2}(3 - 2(\|\vec{v}_2\|^2 + \|\vec{v}_1\|^2) - 1) \quad (7.90)$$

$$= \frac{1}{2}(3 - 4 - 1) = -1 \quad (7.91)$$

and hence

$$\theta_{21} = \pm\pi \quad \text{suggests} \quad a_3 = \sin \frac{\pm\pi}{2} = \pm 1 = \pm \sin \frac{\pi}{2} \sin \frac{\pi}{2} = \pm \sin \frac{\theta_1}{2} \sin \frac{\theta_2}{2} \quad (7.92)$$

We can thus hypothesize

$$\sin \frac{\theta_{21}}{2} \vec{v}_{21} = \cos \frac{\theta_2}{2} \left(\sin \frac{\theta_1}{2} \vec{v}_1 \right) + \cos \frac{\theta_1}{2} \left(\sin \frac{\theta_2}{2} \vec{v}_2 \right) + \left(\sin \frac{\theta_2}{2} \vec{v}_2 \right) \times \left(\sin \frac{\theta_1}{2} \vec{v}_1 \right) \quad (7.93)$$

Let's next find $\cos \frac{\theta_{21}}{2}$ consistent with the above hypothesis. We see that

$$\cos^2 \frac{\theta_{21}}{2} = 1 - \sin^2 \frac{\theta_{21}}{2} \quad (7.94)$$

and hence we evaluate

$$\sin^2 \frac{\theta_{21}}{2} = \sin^2 \frac{\theta_{21}}{2} \vec{v}_{21}^\top \vec{v}_{21} = \left(\sin \frac{\theta_{21}}{2} \vec{v}_{21} \right)^\top \left(\sin \frac{\theta_{21}}{2} \vec{v}_{21} \right) \quad (7.95)$$

$$= \cos^2 \frac{\theta_2}{2} \sin^2 \frac{\theta_1}{2} + \cos^2 \frac{\theta_1}{2} \sin^2 \frac{\theta_2}{2} \quad (7.96)$$

$$+ 2 \cos \frac{\theta_2}{2} \cos \frac{\theta_1}{2} \left(\sin \frac{\theta_2}{2} \vec{v}_2 \right)^\top \left(\sin \frac{\theta_1}{2} \vec{v}_1 \right) \quad (7.97)$$

$$+ \left[\left(\sin \frac{\theta_2}{2} \vec{v}_2 \right) \times \left(\sin \frac{\theta_1}{2} \vec{v}_1 \right) \right]^\top \left[\left(\sin \frac{\theta_2}{2} \vec{v}_2 \right) \times \left(\sin \frac{\theta_1}{2} \vec{v}_1 \right) \right] \quad (7.98)$$

We used the fact that \vec{v}_1, \vec{v}_2 are perpendicular to their vector product. To move further, we will use that for every two unit vectors \vec{u}, \vec{v} in \mathbb{R}^3 there holds true

$$(\vec{u} \times \vec{v})^\top (\vec{u} \times \vec{v}) = \|\vec{u} \times \vec{v}\|^2 = \|\vec{u}\|^2 \|\vec{v}\|^2 \sin^2 \angle(\vec{u}, \vec{v}) \quad (7.99)$$

$$= \|\vec{u}\|^2 \|\vec{v}\|^2 (1 - \cos^2 \angle(\vec{u}, \vec{v})) = \|\vec{u}\|^2 \|\vec{v}\|^2 - (\vec{u}^\top \vec{v})^2 \quad (7.100)$$

Applying this to the last summand in Equation 7.98, we get

$$\sin^2 \frac{\theta_{21}}{2} = \cos^2 \frac{\theta_2}{2} \sin^2 \frac{\theta_1}{2} + \cos^2 \frac{\theta_1}{2} \sin^2 \frac{\theta_2}{2} \quad (7.101)$$

$$+ 2 \cos \frac{\theta_2}{2} \cos \frac{\theta_1}{2} \left(\sin \frac{\theta_2}{2} \vec{v}_2 \right)^\top \left(\sin \frac{\theta_1}{2} \vec{v}_1 \right) \quad (7.102)$$

$$+ \sin^2 \frac{\theta_2}{2} \sin^2 \frac{\theta_1}{2} - \left[\left(\sin \frac{\theta_2}{2} \vec{v}_2 \right)^\top \left(\sin \frac{\theta_1}{2} \vec{v}_1 \right) \right]^2 \quad (7.103)$$

$$= \sin^2 \frac{\theta_1}{2} + \cos^2 \frac{\theta_1}{2} \sin^2 \frac{\theta_2}{2} \quad (7.104)$$

$$+ 2 \cos \frac{\theta_2}{2} \cos \frac{\theta_1}{2} \left(\sin \frac{\theta_2}{2} \vec{v}_2 \right)^\top \left(\sin \frac{\theta_1}{2} \vec{v}_1 \right) - \left[\left(\sin \frac{\theta_2}{2} \vec{v}_2 \right)^\top \left(\sin \frac{\theta_1}{2} \vec{v}_1 \right) \right]^2$$

$$= 1 - \cos^2 \frac{\theta_1}{2} \cos^2 \frac{\theta_2}{2} \quad (7.105)$$

$$+ 2 \cos \frac{\theta_2}{2} \cos \frac{\theta_1}{2} \left(\sin \frac{\theta_2}{2} \vec{v}_2 \right)^\top \left(\sin \frac{\theta_1}{2} \vec{v}_1 \right) - \left[\left(\sin \frac{\theta_2}{2} \vec{v}_2 \right)^\top \left(\sin \frac{\theta_1}{2} \vec{v}_1 \right) \right]^2$$

where we used the fact that

$$\sin^2 \frac{\theta_1}{2} + \cos^2 \frac{\theta_1}{2} \sin^2 \frac{\theta_2}{2} = 1 - \cos^2 \frac{\theta_1}{2} + \cos^2 \frac{\theta_1}{2} \sin^2 \frac{\theta_2}{2} \quad (7.106)$$

$$= 1 + \cos^2 \frac{\theta_1}{2} \left(\sin^2 \frac{\theta_2}{2} - 1 \right) = 1 - \cos^2 \frac{\theta_1}{2} \cos^2 \frac{\theta_2}{2}$$

We are thus obtaining

$$\cos^2 \frac{\theta_{21}}{2} = 1 - \sin^2 \frac{\theta_{21}}{2} \quad (7.107)$$

$$= \cos^2 \frac{\theta_1}{2} \cos^2 \frac{\theta_2}{2} \quad (7.108)$$

$$- 2 \cos \frac{\theta_2}{2} \cos \frac{\theta_1}{2} \left(\sin \frac{\theta_2}{2} \vec{v}_2 \right)^\top \left(\sin \frac{\theta_1}{2} \vec{v}_1 \right) + \left[\left(\sin \frac{\theta_2}{2} \vec{v}_2 \right)^\top \left(\sin \frac{\theta_1}{2} \vec{v}_1 \right) \right]^2$$

$$= \left(\cos \frac{\theta_1}{2} \cos \frac{\theta_2}{2} - \left(\sin \frac{\theta_2}{2} \vec{v}_2 \right)^\top \left(\sin \frac{\theta_1}{2} \vec{v}_1 \right) \right)^2 \quad (7.109)$$

Our complete hypothesis will be

$$\sin \frac{\theta_{21}}{2} \vec{v}_{21} = \cos \frac{\theta_2}{2} \left(\sin \frac{\theta_1}{2} \vec{v}_1 \right) + \cos \frac{\theta_1}{2} \left(\sin \frac{\theta_2}{2} \vec{v}_2 \right) + \left(\sin \frac{\theta_2}{2} \vec{v}_2 \right) \times \left(\sin \frac{\theta_1}{2} \vec{v}_1 \right)$$

$$\cos \frac{\theta_{21}}{2} = \cos \frac{\theta_1}{2} \cos \frac{\theta_2}{2} - \left(\sin \frac{\theta_2}{2} \vec{v}_2 \right)^\top \left(\sin \frac{\theta_1}{2} \vec{v}_1 \right) \quad (7.110)$$

To verify this, we will run the following Maple [17] program

```
> restart:
> with(LinearAlgebra):
> E:=IdentityMatrix(3):
```



```

> X_:=proc(u) <<0|-u[3]|u[2]>, <u[3]|0|-u[1]>, <-u[2]|u[1]|0>> end proc:
> v1:=<x1,y1,z1>:
> v2:=<x2,y2,z2>:
> R1:=2*(s1*v1).Transpose(s1*v1)+(2*c1^2-1)*E+2*c1*X_(s1*v1):
> R2:=2*(s2*v2).Transpose(s2*v2)+(2*c2^2-1)*E+2*c2*X_(s2*v2):
> R21:=expand~(R2.R1):
> c21:=c2*c1-Transpose(s2*v2).(s1*v1);

```

$$c21 := c2 c1 - s1 x1 s2 x2 - s1 y1 s2 y2 - s1 z1 s2 z2$$

```

> s21v21:=c2*s1*v1+s2*c1*v2+X_(s2*v2).(s1*v1);

```

$$s21v21 := \begin{bmatrix} c2 s1 x1 + s2 c1 x2 - s2 z2 s1 y1 + s2 y2 s1 z1 \\ c2 s1 y1 + s2 c1 y2 + s2 z2 s1 x1 - s2 x2 s1 z1 \\ c2 s1 z1 + s2 c1 z2 - s2 y2 s1 x1 + s2 x2 s1 y1 \end{bmatrix}$$

```

> RR21:=2*s21v21.Transpose(s21v21)+(2*c21^2-1)*E+2*c21*X_(s21v21):
> simplify(expand~(RR21-R21), [x1^2+y1^2+z1^2=1, x2^2+y2^2+z2^2=1,
                                c1^2+s1^2=1, c2^2+s2^2=1]);

```

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

which verifies that our hypothesis was correct.

Considering two unit quaternions

$$\vec{p} = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \end{bmatrix}, \quad \text{and} \quad \vec{q} = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{bmatrix} \quad (7.111)$$

we can now give their composition as

$$\vec{q}_{21} = \vec{q}\vec{p} = \begin{bmatrix} q_1 p_1 - q_2 p_2 - q_3 p_3 - q_4 p_4 \\ q_1 p_2 + q_2 p_1 + q_3 p_4 - q_4 p_3 \\ q_1 p_3 + q_3 p_1 + q_4 p_2 - q_2 p_4 \\ q_1 p_4 + q_4 p_1 + q_2 p_3 - q_3 p_2 \end{bmatrix} \quad (7.112)$$

$$= \begin{bmatrix} q_1 p_1 - q_2 p_2 - q_3 p_3 - q_4 p_4 \\ q_2 p_1 + q_1 p_2 - q_4 p_3 + q_3 p_4 \\ q_3 p_1 + q_4 p_2 + q_1 p_3 - q_2 p_4 \\ q_4 p_1 - q_3 p_2 + q_2 p_3 + q_1 p_4 \end{bmatrix} \quad (7.113)$$

$$= \begin{bmatrix} q_1 & -q_2 & -q_3 & -q_4 \\ q_2 & q_1 & -q_4 & q_3 \\ q_3 & q_4 & q_1 & -q_2 \\ q_4 & -q_3 & q_2 & q_1 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \end{bmatrix} \quad (7.114)$$

§ 37 Application of quaternions to vectors Consider a rotation by angle θ around an axis with direction \vec{v} represented by a unit quaternion $\vec{q} = [\cos \frac{\theta}{2} \quad \sin \frac{\theta}{2} \vec{v}]$ and a vector $\vec{x} \in \mathbb{R}^3$. To rotate the vector, we may construct the rotation matrix $\mathbf{R}(\vec{q})$ and apply it to the vector \vec{x} as $\mathbf{R}(\vec{q}) \vec{x}$.

Interestingly enough, it is possible to accomplish this in somewhat different and more efficient way by first “embedding” vector \vec{x} into a quaternion

$$\vec{p}(\vec{x}) = \begin{bmatrix} 0 \\ \vec{x} \end{bmatrix} = \begin{bmatrix} 0 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad (7.115)$$

and then composing it with quaternion \vec{q} from both sides

$$\vec{q} \vec{p}(\vec{x}) \vec{q}^{-1} = \begin{bmatrix} \cos \frac{\theta}{2} \\ \sin \frac{\theta}{2} \vec{v} \end{bmatrix} \begin{bmatrix} 0 \\ \vec{x} \end{bmatrix} \begin{bmatrix} \cos \frac{\theta}{2} \\ -\sin \frac{\theta}{2} \vec{v} \end{bmatrix} \quad (7.116)$$

One can verify that the following

$$\begin{bmatrix} 0 \\ \mathbf{R}(\vec{q}) \vec{x} \end{bmatrix} = \vec{q} \vec{p}(\vec{x}) \vec{q}^{-1} \quad (7.117)$$

holds true.

7.4 Topology of the space of rotations

We will next study what happens with the rotated space when the rotation parameterization changes only a little. We understand the geometry of rotations using the angle-axis parameterization and by that also the quaternion parameterization. Unfortunately, it is not easy to visualize the angle-axis parameterization completely because it is four dimensional. To get some intuition about its behaviour, we will first look at the subset of rotations with rotation axes in the xy plane. Figure 7.2

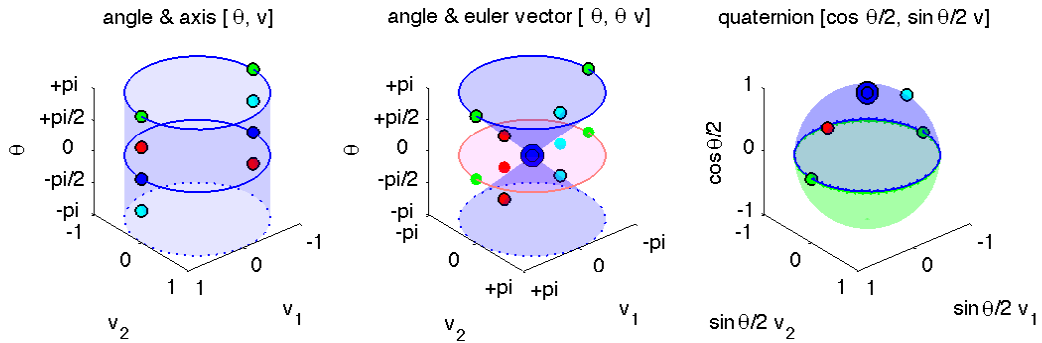


Figure 7.2: Representations of rotations around the rotation axes \vec{v} in xy plane, i.e. $\vec{v} = [v_1, v_2, 0]^\top$. Points in the same color correspond to the same rotation. (Left) The angle-axis representation $[\theta, \vec{v}]^\top$ with $\|\vec{v}\| = 1$. Rotations are represented by vectors that end at the blue cylinder. The opposite points on the top rim of the cylinder represent the same rotation. The top rim of the cylinder is connected to the (open) bottom rim of the cylinder. There is and an infinite number of vectors $[0, \vec{v}]^\top$, which correspond to the blue circle in the middle of the cylinder, that all map to the identity. (Middle) The rotation representation, $[\theta, \theta \vec{v}]^\top$, removes the multiplicities representing the identity. Only vector $[0, 0, 0, 0]^\top$ represents the identity. We see that all the rotations can also be well represented by a two-dimensional disc of radius π shown in red, which corresponds to the euler vector. The opposite points on the boundary of the red disc (shown as the red circle) represent the same rotation by $\pm\pi$. (Right) With the quaternion representation, both the upper and lower parts of the cone of $[\theta, \theta \vec{v}]^\top$ are mapped to the upper (blue) hemisphere of the unit ball. We can add the bottom (green) hemisphere to obtain a complete sphere without a boundary. The opposite points on the complete sphere correspond to the same rotation.

Bibliography

- [1] Paul R. Halmos. *Naive Set Theory*. Springer-Verlag, 1974.
- [2] Walter Rudin. *Principles of Mathematical Analysis*. McGraw-Hill, 1976.
- [3] Paul R. Halmos. *Finite-Dimensional Vector Spaces*. Springer-Verlag, 2000.
- [4] Carl D. Meyer. *Matrix Analysis and Applied Linear Algebra*. SIAM: Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2001.
- [5] Petr Olšák. *Úvod do algebry, zejména lineární*. FEL ČVUT, Praha, 2007.
- [6] Pavel Pták. *Introduction to linear algebra*. Vydavatelství ČVUT, Praha, 2007.
- [7] Gilbert Strang. *Introduction to Linear Algebra*. Wellesley Cambridge, 3rd edition, 2003.
- [8] Seymour Lipschutz. *3,000 Solved Problems in Linear Algebra*. McGraw-Hill, 1st edition, 1989.
- [9] Jim Hefferon. *Linear Algebra*. Virginia Commonwealth University Mathematics, 2009.
- [10] Michael Artin. *Algebra*. Prentice Hall, Upper Saddle River, NJ 07458, USA, 1991.
- [11] Yoshiaki Ueno and Kazuhisa Maehara. An elementary proof of the generalized Laplace expansion formula. *Tokyo Polytechnic University Bulletin*, 25(1):61–63, 2003. http://www.t-kougei.ac.jp/research/bulletin_e/.
- [12] Cyrus Colton MacDuffee. *The Theory of Matrices*. Dover Publications Inc., 2004. <http://books.google.cz/books?id=8hQG7ByP53oC>.
- [13] Saunders MacLane and Garret Birkhoff. *Algebra*. Chelsea Pub. Co., 3rd edition, 1988.
- [14] Richard Hartley and Andrew Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge, 2nd edition, 2003.
- [15] J. A. Grunert. Das pothenotische problem in erweiterter gestalt nebst über seine anwendungen in der geodäsie. *Grunerts Archiv für Mathematik und Physik*, 1:238–248, 1841.
- [16] David Cox, John Little, and Donald O’Shea. *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra*. Springer, 2nd edition, 1996.
- [17] Ltd. Cybernet Systems Co. Maple. <http://www.maplesoft.com/products/maple/>.

- [18] Z. Kúkelová, M. Bujňák, and Tomáš Pajdla. Polynomial eigenvalue solutions to minimal problems in computer vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012.
- [19] D. Nistér. An efficient solution to the five-point relative pose problem. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(6):756–770, June 2004.
- [20] H. Stewenius, C. Engels, and D. Nistér. Recent developments on direct relative orientation. *ISPRS J. of Photogrammetry and Remote Sensing*, 60:284–294, 2006.
- [21] Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, and H. van der Vorst. *Templates for the Solution of Algebraic Eigenvalue Problems*. SIAM, Philadelphia, 2000.

Index

$[n]$, 5
determinant, 5
sign, 6
determinant, 6
inversion, 6
monotonic, 6
permutation, 5

affine coordinate system, 23
affine function, 15
affine space, 21
axioms of linear space, 18
axioms of affine space, 22

basis, 19
bound vector, 17

coordinate linear space, 2
coordinates, 19
cross product, 7

dual basis, 9
dual space, 9

free vector, 20

geometric scalars, 16
geometric vector, 17

Kronecker product, 12

linear function, 15
linear space, 18

marked ruler, 15

origin of affine coordinate system, 23

partition, 19
position vector, 23

standard basis, 2

vector product, 7
vector product, 7

zero bound vector, 17