Chapter 7
Matrix Groups

Among the most important examples of groups are groups of matrices. In this chapter, matrix groups are defined and a number of standard examples are discussed.

7.1 Introduction to Matrix Groups

Let $A = \left(a_{ij}\right)_{m\times n}$ and $B = \left(b_{ij}\right)_{n\times p}$. We recall the *product* AB is the matrix $C = \left(a_{ij}\right)_{m\times p}$, where

$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}.$$

Theorem 7.1 *Matrix multiplication is associative.*

Proof Let $A = \left(a_{ij}\right)_{m\times n}$, $B = \left(b_{ij}\right)_{n\times p}$ and $C = \left(c_{ij}\right)_{p\times q}$ be three arbitrary matrices. Suppose that $AB = D = \left(d_{ij}\right)_{m\times p}$ and $BC = E = \left(e_{ij}\right)_{n\times q}$. Moreover, suppose that

$$(AB)C = DC = X = \left(x_{ij}\right)_{m\times q} \text{ and } A(BC) = AE = Y = \left(y_{ij}\right)_{m\times q}.$$

Then, we obtain

$$\begin{aligned} x_{ij} &= \sum_{l=1}^{p} d_{il}c_{lj} \\ &= \sum_{l=1}^{p} \left(\sum_{k=1}^{n} a_{ik}b_{kj} \right) c_{lj} \end{aligned}$$

B. Davvaz, *A First Course in Group Theory*,
https://doi.org/10.1007/978-981-16-6365-9_7

$$
\begin{aligned}
&= \sum_{l=1}^{p} \left(a_{i1}b_{1l} + a_{i2}b_{2l} + \cdots + a_{in}b_{nl}\right)c_{lj} \\
&= a_{i1}\sum_{l=1}^{p} b_{1l}c_{lj} + a_{i2}\sum_{l=1}^{p} b_{2l}c_{lj} + \cdots + a_{in}\sum_{l=1}^{p} b_{nl}c_{lj} \\
&= \sum_{k=1}^{n} a_{ik}\left(\sum_{l=1}^{p} b_{il}c_{lj}\right) \\
&= \sum_{k=1}^{n} a_{ik}e_{kj} \\
&= y_{ij}.
\end{aligned}
$$

Therefore, we deduce that $X = Y$. ■

A matrix A with n rows and n columns is called a *square matrix* of order n. One example of a square matrix is the $n \times n$ *identity matrix* I_n. This is the matrix $I_n = \left(\delta_{ij}\right)_{n\times n}$ defined by

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases}$$

Definition 7.2 Let A be an $n \times n$ matrix over $\mathbb{F}$. An $n \times n$ matrix B such that $BA = I_n$ is called a *left inverse* of A. Similarly, an $n \times n$ matrix B such that $AB = I_n$ is called a *right inverse* of A. If $AB = BA = I_n$, then B is called an *inverse* of A and A is said to be *invertible*.

Lemma 7.3 *If a matrix A has a left inverse B and a right inverse C, then $B = C$.*

Proof Suppose that $BA = I_n$ and $AC = I_n$. Then, we obtain

$$B = BI_n = B(AC) = (BA)C = I_nC = C$$

as desired. ■

Thus, if A has a left and a right inverse, then A is invertible and has a unique inverse, which we shall denote by A^{-1}.

Theorem 7.4 *Let A and B be two matrices of the same size over $\mathbb{F}$.*

(1) If A is invertible, so is A^{-1} and $(A^{-1})^{-1} = A$.
(2) If both A and B are invertible, so is AB and $(AB)^{-1} = B^{-1}A^{-1}$.

Proof (1) It is evident from the symmetry of the definition.
(2) We have

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AI_nA^{-1} = AA^{-1} = I_n.$$

Similarly, we obtain $(B^{-1}A^{-1})(AB) = I_n$. This completes our proof. ■

Theorem 7.5 *If*

$$GL_n(\mathbb{F}) = \{A \in M_{n\times n} \mid A \text{ is invertible}\},$$

then $GL_n(\mathbb{F})$ is a group under the multiplication of matrices.

Proof The proof follows from Theorem 7.1, the definition of identity matrix and Theorem 7.4. ■

$GL_n(\mathbb{F})$ is called the *general linear group*.

Definition 7.6 The following three operations on a matrix are called *elementary row operations*:

(1) Multiply a row through by a non-zero constant.
(2) Interchange two rows.
(3) Add a multiple of one row to another row.

Indeed, an elementary row operation is a special function f which associate with each matrix $A = \left(a_{ij}\right)_{m\times n}$ a matrix $f(A) = \left(b_{ij}\right)_{m\times n}$ such that

(1) $b_{ij} = a_{ij}$ if $i \neq k$, and $b_{kj} = ca_{kj}$,
(2) $b_{ij} = a_{ij}$ if i is different from both k and l, and $b_{kj} = a_{lj}$, $b_{lj} = a_{kj}$,
(3) $b_{ij} = a_{ij}$ if $i \neq k$, and $b_{kj} = a_{kj} + ca_{lj}$.

Lemma 7.7 *To each elementary row operation f there corresponds an elementary row operation g, of the same type as f, such that $f\big(g(A)\big) = g\big(f(A)\big) = A$, for each A.*

Proof (1) Suppose that f is the operation which multiplies the kth row of a matrix by the non-zero scalar c. Let g be the operation which multiplies row k by c^{-1}.

(2) Suppose that f is the operation which replaces row k by row k plus c times row l, where $k \neq l$. Let g be the operation which replaces row k by row k plus $-c$ times row l.

(3) If f interchanges rows k and l, let $g = f$.

In each of the above cases, we clearly have $f\big(g(A)\big) = g\big(f(A)\big) = A$, for each A. ■

An $n \times n$ matrix is called an *elementary matrix* if it can be obtained from the identity matrix I_n by performing a single elementary row operation.

Lemma 7.8 *If the elementary matrix E results from performing a certain elementary row operation f on I_n and if A is an $m \times n$ matrix, then the product EA is the matrix that results this same row elementary operation is performed on A, i.e., $f(a) = EA$.*

Proof It is straightforward by considering the three types of elementary row operations. ■

Example 7.9 Consider the matrix

$$\begin{bmatrix} 1 & 5 & 4 & 1 \\ 3 & 2 & 3 & 0 \\ 1 & -1 & 3 & 1 \end{bmatrix}$$

and consider the elementary matrix

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 4 & 0 & 1 \end{bmatrix}$$

which results from adding 4 times the first row of I_3 to the third row. The product EA is

$$\begin{bmatrix} 1 & 5 & 4 & 1 \\ 3 & 2 & 3 & 0 \\ 5 & 19 & 19 & 5 \end{bmatrix}.$$

Theorem 7.10 *Each elementary matrix belongs to* $GL_n(\mathbb{F})$.

Proof If A is an $n \times n$ elementary matrix, then A results from performing some row operation on I_n. Let B be the $n \times n$ matrix that results when the inverse operation is performed on I_n. Applying Lemma 7.7 and using the fact that inverse row operations cancel the effect of each other, it follows that $AB = I_n$ and $BA = I_n$. So, the elementary matrix B is the inverse of A. This yields that $A \in GL_n(\mathbb{F})$. ■

Definition 7.11 An $m \times n$ matrix R is called *row-reduced echelon matrix* if

(1) the first non-zero entry in each non-zero row of R is equal to 1,
(2) each column of R which contains the leading non-zero entry of some row has all its other entries 0,
(3) every row of R which has all its entries 0 occurs below every row which has a non-zero entry,
(4) if rows $1, \ldots, r$ are the non-zero rows of R, and if the leading non-zero entry of row i occurs in column k_i, $i = 1, \ldots, r$, then $k_1 < k_2 < \cdots < k_r$.

Example 7.12 (1) One example of row-reduced echelon matrix is I_n.
(2) The matrix

$$\begin{bmatrix} 0 & 1 & 2 & 0 & -3 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

a row-reduced echelon matrix, too.

Matrices that can be obtained from one another by a finite sequence of elementary row operations are said to be *row equivalent*. We left to the readers to show that each matrix is row equivalent to a unique row-reduced echelon matrix.

If one defines an *elementary column operation* and *column equivalent* in a manner analogous to that of elementary row operation and row equivalent, it is clear that each $m \times n$ matrix will be column equivalent to a column-reduced echelon matrix. Also, each elementary column operation will be of the form $A \to AE$, where E is an $n \times n$ elementary matrix, and so on.

Definition 7.13 A *diagonal matrix* is a square matrix in which the entries outside main diagonal are all zero. A diagonal matrix whose diagonal elements all contains the same scalar c is called a *scalar matrix*.

Corollary 7.14 *The set of all $n \times n$ invertible diagonal matrices is a subgroup of $GL_n(\mathbb{F})$.*

Definition 7.15 A matrix $A = \left(a_{ij}\right)_{n\times n}$ is called

(1) an *upper triangular* if $a_{ij} = 0$ for $i > j$.
(2) a *lower triangular* if $a_{ij} = 0$ for $i < j$.

Let $UT_n(\mathbb{F})$ be the set of upper triangular matrices such that all entries on the diagonal are non-zero.

Theorem 7.16 *$UT_n(\mathbb{F})$ is a subgroup of $GL_n(\mathbb{F})$.*

Proof Suppose that A and B are two arbitrary elements of $UT_n(\mathbb{F})$. Let $C = \left(c_{ij}\right)_{n\times n} = AB$. If $i > j$, then

$$c_{ij} = \sum_k a_{ik}b_{kj} = 0 \text{ and } c_{ii} = a_{ii}b_{ii} \neq 0,$$

which shows that $C \in UT_n(\mathbb{F})$.

Now, we notice that

(1) The elementary matrix corresponding to multiply the ith row by a non-zero constant c is the matrix with ones on the diagonal, expect in the ith row, which instead has a c, and zero everywhere else.
(2) If we consider a replacement that add a multiple of the ith row to jth row, where $i < j$, then this matrix has non-zero entries only along the diagonal and in the ijth entry, which is above the diagonal. Therefore, all entries bellow the diagonal are zero.

The above row operations are sufficient to row reduce A to I_n. Since A is upper triangular, it follows that there exists a sequence of row operations of the type described in the above that transforms A into I_n. Consequently, there is a sequence of upper triangular elementary matrices $E_1, E_2, \ldots, E_k$ such that $E_k \ldots E_2E_1A = I_n$, or equivalently

$$A^{-1} = E_k \dots E_2 E_1. \tag{7.1}$$

The right side of (7.1) is a product of upper triangular matrices. Hence, the result of this product is also an upper triangular matrix. Hence, we conclude that the inverse of A and lies in $UT_n(\mathbb{F})$. This completes our proof. ■

Let $LT_n(\mathbb{F})$ be the set of lower triangular matrices such that all entries on the diagonal are non-zero.

Theorem 7.17 *$LT_n(\mathbb{F})$ is a subgroup of $GL_n(\mathbb{F})$.*

Proof The proof is similar to the proof of Theorem 7.16. ■

Theorem 7.18 *If A is $n \times n$ matrix, then the following are equivalent:*

(1) *$A \in GL_n(\mathbb{F})$,*
(2) *A is row equivalent to I_n,*
(3) *A is a product of elementary matrices.*

Proof Suppose that R is a row-reduced echelon matrix which is row equivalent to A. Then, by Lemma 7.8, there exist elementary matrices $E_1, \dots, E_k$ such that

$$R = E_k \dots E_1 A$$

Since each E_i is invertible, it follows that

$$A = E_1^{-1} \dots E_k^{-1} R.$$

Since products of invertible matrices are invertible, we conclude that A is invertible if and only if R is invertible. Since R is a square row-reduced echelon matrix, it follows that R is invertible if and only if each row of R contains non-zero entry, that is, if and only if $R = I_n$. We have now shown that

$$A \text{ is invertible} \Leftrightarrow R = I,$$

and if $R = I$ then $A = E_1^{-1} \dots E_k^{-1}$. This proves that (1), (2) and (3) are equivalent statements about A. ■

Corollary 7.19 *The general linear group $GL_n(\mathbb{F})$ is generated by elementary matrices.*

Corollary 7.20 *Let A be an invertible matrix. If a sequence of elementary row operations reduces A to the identity, then that same sequence of operations when applied to I_n yields A^{-1}.*

Example 7.21 We want to find the inverse of

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 3 & 4 & 5 \\ 2 & 6 & -1 \end{bmatrix},$$

if exists. If we are able to convert matrix A to the identity matrix by using elementary row operations, i.e.,

$$[A \mid I_n] \longrightarrow [I_n \mid B],$$

then A is invertible and $A^{-1} = B$. The computation can be carried out as follows:

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 3 & 4 & 5 & 0 & 1 & 0 \\ 2 & 6 & -1 & 0 & 0 & 1 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & -1 & -3 & 1 & 0 \\ 2 & 6 & -1 & 0 & 0 & 1 \end{array}\right]$$

$$\longrightarrow \left[\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & -1 & -3 & 1 & 0 \\ 0 & 4 & -5 & -2 & 0 & 1 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 3 & 4 & -1 & 0 \\ 0 & 1 & -1 & -3 & 1 & 0 \\ 0 & 4 & -5 & -2 & 0 & 1 \end{array}\right]$$

$$\longrightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 3 & 4 & -1 & 0 \\ 0 & 1 & -1 & -3 & 1 & 0 \\ 0 & 0 & -1 & 10 & -4 & 1 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 34 & -13 & 3 \\ 0 & 1 & -1 & -3 & 1 & 0 \\ 0 & 0 & -1 & 10 & -4 & 1 \end{array}\right]$$

$$\longrightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 34 & -13 & 3 \\ 0 & 1 & -1 & -3 & 1 & 0 \\ 0 & 0 & 1 & -10 & 4 & -1 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 34 & -13 & 3 \\ 0 & 1 & 0 & -13 & 5 & -1 \\ 0 & 0 & 1 & -10 & 4 & -1 \end{array}\right].$$

Therefore, we get

$$A^{-1} = \begin{bmatrix} 34 & -13 & 3 \\ -13 & 5 & -1 \\ -10 & 4 & -1 \end{bmatrix}.$$

Definition 7.22 The *determinant* of a matrix A, written $\det(A)$, is a certain number associated to A. If $A = (a_{ij})_{n\times n}$, then determinant of A is defined by the following:

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i(i\sigma)}.$$

Theorem 7.23 *If $n = 1$, then $\det(A) = a_{11}$. If $n > 1$, then*

$$\det(A) = \sum_{j=1}^{n} a_{ij} A_{ij}, \quad i = 1, \ldots, n \tag{7.2}$$

and

$$\det(A) = \sum_{i=1}^{n} a_{ij} A_{ij}, \quad j = 1, \ldots, n, \tag{7.3}$$

such that A_{ij} is the ij-cofactor associated with A, i.e.,

$$A_{ij} = (-1)^{i+j} \det(M_{ij}),$$

where M_{ij} is $n-1) \times (n-1)$ matrix obtained from A by removing its ith row and jth column. The M_{ij}'s are called the minors of A.

Proof The proof follows from the expansion of the right sides of (7.2), (7.3) and using the definition of determinant. ■

Formula (7.2) is called the *expansion of the determinant through the ith row* and Formula (7.3) is called the *expansion of the determinant through the jth column*. An alternative notation for the determinant of a matrix A is $|A|$ rather than $\det(A)$.

Example 7.24 Evaluate $\det(A)$, where

$$A = \begin{bmatrix} 2 & 3 & 3 \\ 4 & 5 & -1 \\ 1 & 2 & 3 \end{bmatrix}.$$

We have

$$\begin{aligned} \det(A) &= 2\begin{vmatrix} 5 & -1 \\ 2 & 3 \end{vmatrix} - 3\begin{vmatrix} 4 & -1 \\ 1 & 3 \end{vmatrix} + 3\begin{vmatrix} 4 & 5 \\ 1 & 2 \end{vmatrix} \\ &= 2\big(5 \cdot 3 - (-1) \cdot 2\big) - 3\big(4 \cdot 3 - (-1) \cdot 1\big) + 3\big(4 \cdot 2 - 5 \cdot 1) = 4. \end{aligned}$$

Lemma 7.25 *Let A be any $n \times n$ matrix.*

(1) If B is the matrix that results when a single row of A is multiplied by a constant c, then $\det(B) = c \det(A)$.
(2) If B is the matrix that results when two rows of A are interchanged, then $\det(B) = -\det(A)$.
(3) If B is the matrix that results when a multiple of one row of A is added to another row, then $\det(B) = \det(A)$.

Proof It is straightforward. ■

Lemma 7.26 *If E is an $n \times n$ elementary matrix and A is any $n \times n$ matrix, then*

$$\det(EA) = \det(E) \det(A).$$

Proof We consider the three different types of elementary row operation.

(1) Let E be the elementary matrix obtained from I_n by multiplying the entries of some row of I_n by a non-zero scalar c. Then, we can row reduce E to I_n by multiplying the same row by $\frac{1}{c}$. Hence, we obtain $\det(I_n) = \frac{1}{c} \det(E)$, or equivalently $\det(E) = c$. Since EA is the matrix resulting from multiplying the entries of a row of A by c, it follows that $\det(EA) = c \det(A) = \det(E) \det(A)$.

(2) Let E be the elementary matrix obtained from I_n by interchanging two rows

of I_n. Then, $\det(E) = -\det(I_n) = -1$. Since EA is the matrix resulting from interchanging the corresponding two rows of A, it follows that $\det(EA) = (-1)\det(A)$, and so $\det(EA) = \det(E)\det(A)$.

(3) Let E be the elementary matrix obtained from I_n by adding a multiple of one row of I_n to another row of I_n. Since row operations of this type do not change the determinant, it follows that $\det(E) = \det(I_n) = 1$. Since EA is the result of adding a multiple of a row of A to another row of A, it follows that $\det(EA) = 1 \cdot \det(A) = \det(E)\det(A)$.

So, no matter what type of elementary matrix E is, we have shown that $\det(EA) = \det(E)\det(A)$. Now, we are done. ■

Theorem 7.27 *A square matrix A is invertible if and only if* $\det(A) \neq 0$.

Proof Suppose that $E_1, \ldots, E_k$ are elementary matrices which place A in reduced row echelon matrix R. Then, $R = E_k \ldots E_1 A$. By Lemma 7.26, we have $\det(R) = \det(E_k) \ldots \det(E_1)\det(A)$. Since the determinant of an elementary matrix is non-zero, it follows that $\det(A)$ and $\det(R)$ are either both zero or both non-zero.

Now, if A is invertible, then the reduced row echelon form of A is I_n. In this case, $\det(R) = \det(I_n) = 1$ which implies that $\det(A) \neq 0$.

Conversely, if $\det(A) \neq 0$, then $\det(R) \neq 0$ which implies that R can not have a row of all zeros. This implies that $R = I_n$ and so A is invertible. ■

Corollary 7.28 *We have*

$$GL_n(\mathbb{F}) = \{A \in M_{n\times n} \mid \det(A) \neq 0\}.$$

Theorem 7.29 *A square matrix $A = (a_{ij})_{n\times n}$ is invertible if and only if $AX = 0$, i.e.,*

$$\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

has only the trivial solution.

Proof Suppose that A is invertible and X_0 be any solution to $AX = 0$. Thus, $AX_0 = 0$. Multiplying both sides of this equation by A^{-1} gives $A^{-1}(AX_0) = A^{-1}0$, or equivalently $(A^{-1}A)X_0 = 0$. This implies that $X_0 = 0$. Thus, $AX = 0$ has only the trivial solution.

Conversely, if $AX = 0$ has only the trivial solution, then A must be row equivalent to a reduced row echelon matrix R with n leading 1s. Hence, $R = I_n$. Now, by Theorem 7.18 our proof completes. ■

Definition 7.30 Let A be a square matrix and $V_1, V_2, \ldots, V_n$ be the columns of A. We say that $V_1, V_2, \ldots, V_n$ are *linearly independent* if and only if the only solution for

$$x_1V_1 + x_2V_2 + \cdots + x_nV_n = 0 \ \text{(with } x_i \in \mathbb{F})$$

is $x_1 = x_2 = \cdots = x_n = 0$. Note that the equation $x_1V_1 + x_2V_2 + \cdots + x_nV_n = 0$ can be rewrite as $AX = 0$.

Corollary 7.31 *A square matrix A is invertible if and only if the columns of A are linearly independent.*

Proof The proof results from Theorem 7.29 and Definition 7.30. ■

Theorem 7.32 *If $\mathbb{F}$ is a finite field with q elements, then*

$$|GL_n(\mathbb{F})| = \prod_{k=0}^{n-1}(q^n - q^k).$$

Proof We count $n \times n$ matrices whose columns are linearly independent. We can do this by building up a matrix from scratch. The first column can be anything other than the zero column, so there exist $q^n - 1$ possibilities. The second column must be linearly independent from the first column, which is to say that it must not be a multiple of the first column. Since there exist q multiples of the first column, it follows that there exist $q^n - q$ possibilities for the second column. In general, the ith column must be linearly independent from the first $i - 1$ columns, which means that it can not be a linear combination of the first $i - 1$ columns. Since there exist q^{i-1} linear combinations of the first $i - 1$ columns, it follows that there exist $q^n - q^{i-1}$ possibilities for the ith column. Once we build the entire matrix this way, we know that the column are all linearly independent by choice. Moreover, we can build any $n \times n$ matrix whose columns are linearly independent in this way. Consequently, there exist

$$(q^n - 1)(q^n - q)\ldots(q^n - q^{n-1})$$

matrices. ■

Theorem 7.33 *The center of $GL_n(\mathbb{F})$ is*

$$Z\big(GL_n(\mathbb{F})\big) = \{aI_n \mid a \in \mathbb{F}^*\}.$$

This means that the center of general linear group is the subgroup comprising scalar matrices.

Proof We suppose that $n > 1$. Each elements of $Z\big(GL_n(\mathbb{F})$ must commute with every elements of $GL_n(\mathbb{F})$. In particular must commute with elementary matrices. We complete the our proof in four steps.

Step 1: Each element of $Z\big(GL_n(\mathbb{F})$ commutes with off-diagonal matrix units. Indeed, assume that $c \in \mathbb{F}^*$ and $e_{ij}(c)$ is a matrix with c in ijth entry and zeroes elsewhere. Specially, $e_{ij}(1)$ is called the ijth *matrix unit*. Now, we define $E_{ij}(c) = I_n + e_{ij}(c)$. Then, we have

(1) Since $E_{ij}(-c)$ is the inverse of $E_{ij}(c)$, it follows that $E_{ij}(c) \in GL_n(\mathbb{F})$.
(2) Any matrix that commutes with $E_{ij}(1)$ must commute with $e_{ij}(1)$

Step 2: If a matrix commutes with off-diagonal matrix units, then it is a diagonal matrix. Indeed, let A be a matrix such that $a_{ji} \neq 0$ for some $i \neq j$ and let $B = e_{ij}(1)$. Then, jjth entry of AB is non-zero, while the jjth entry of BA is zero. Consequently, any matrix commutes with all of off-diagonal matrix units $e_{ij}(1)$ can not have any off-diagonal entries.

Step 3: A *permutation matrix* is a matrix obtained by permuting the rows of I_n according to some permutation of the numbers 1 to n. A permutation matrix is invertible. Let A be a diagonal matrix with $a_{ij} \neq a_{jj}$ and let B be a permutation matrix obtained by permuting the ith and jth rows of I_n. Then, A does not commute with B. Therefore, we conclude that any diagonal matrix that commutes with all permutation matrices is scalar.

Step 4: Combining the first two steps yields that any matrix in $Z\big(GL_n(\mathbb{F})\big)$ must be diagonal, and the third step then yields that it must be scalar. ■

Corollary 7.34 *For any $n \times n$ matrices A and B,*

$$AB = I_n \;\Leftrightarrow\; BA = I_n.$$

Proof It is enough to prove that $BA = I_n$ implies that $AB = I_n$. Let $BA = I_n$. If $AX = 0$, then $B(AX) = B0 = 0$. Hence, $(BA)X = 0$, or equivalently $I_nX = 0$. This implies that $X = 0$. Now, by Theorem 7.29, we conclude that A is invertible. Since $BA = I_n$, it follows that

$$A(BA)A^{-1} = AI_nA^{-1} = I_n.$$

This implies that $(AB)AA^{-1} = I_n$. Thus, we conclude that $AB = I_n$. ■

Corollary 7.35 *Let A and B be two $n \times n$ matrices. If AB is invertible, then A and B are invertible.*

Proof Suppose that $C = B(AB)^{-1}$ and $D = (AB)^{-1}A$. Then, we obtain

$$\begin{aligned} AC &= A\big(B(AB)^{-1}\big) = (AB)(AB)^{-1} = I_n, \\ DB &= \big((AB)^{-1}A\big)B = (AB)^{-1}(AB) = I_n. \end{aligned}$$

Now, by Corollary 7.34, we deduce that $C = A^{-1}$ and $D = B^{-1}$. ■

The result in Lemma 7.26 can be generalized to any two $n \times n$ matrices.

Theorem 7.36 *If A and B are two $n \times n$ matrices, then*

$$\det(AB) = \det(A)\det(B).$$

Proof We consider two cases:

(1) If A is invertible, then there exist elementary matrices $E_1, \ldots, E_k$ such that $A = E_1 \ldots E_k$. Then, we have

$$\begin{aligned}
\det(AB) &= \det(E_1 \ldots E_k B) \\
&= \det(E_1)\det(E_2 \ldots E_k B) \\
&\vdots \\
&= \det(E_1)\ldots\det(E_k)\det(B) \\
&= \det(E_1 E_2)\det(E_3)\ldots\det(E_k)\det(B) \\
&\vdots \\
&= \det(E_1 \ldots E_k)\det(B) \\
&= \det(A)\det(B).
\end{aligned}$$

(2) If A is not invertible, by Corollary 7.35, AB is not invertible. This yield that $\det(AB) = 0 = \det(A)\det(B)$, and so the theorem holds. ■

Corollary 7.37 *If A is $n \times n$ invertible matrix, then*

$$\det(A^{-1}) = \frac{1}{\det(A)}.$$

Proof Since $AA^{-1} = I_n$, it follows that $\det(AA^{-1}) = \det(I_n)$ and so $\det(A^{-1})\det(A) = 1$. Since $\det(A) \neq 0$, the proof can be completed by dividing through by $\det(A)$. ■

Theorem 7.38 *If*

$$SL_n(\mathbb{F}) = \{A \in GL_n(\mathbb{F}) \mid \det(A) = 1\},$$

then $SL_n(\mathbb{F})$ is a subgroup of $GL_n(\mathbb{F})$.

Proof Let $A, B \in SL_n(\mathbb{F})$ be arbitrary. Since $\det(A) = \det(B) = 1$, by Theorem 7.36 we conclude that $\det(AB) = 1$. This implies that $AB \in SL_n(\mathbb{F})$. Moreover, by Corollary 7.37, $\det(A^{-1}) = 1$, and hence $A^{-1} \in SL_n(\mathbb{F})$. ■

$SL_n(\mathbb{F})$ is called the *special linear group*.

Theorem 7.39 *The center of $SL_n(\mathbb{F})$ is*

$$Z\big(SL_n(\mathbb{F})\big) = \{aI_n \mid a^n = 1\}.$$

Proof The proof results from the definition of $SL_n(\mathbb{F})$ and Theorem 7.33. ■

Theorem 7.40 *The special linear group $SL_2(\mathbb{F})$ is generated by matrices of the form*

$$\begin{bmatrix} 1 & 0 \\ t & 1 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & r \\ 0 & 1 \end{bmatrix}$$

for all $r, s \in \mathbb{F}^*$.

Proof Suppose that

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

is an arbitrary element of $SL_2(\mathbb{F})$. We consider the following three cases:

Case 1: If $b \neq 0$, then

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ (d-1)b^{-1} & 1 \end{bmatrix} \begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ (a-1)b^{-1} & 1 \end{bmatrix}.$$

Case 2: If $c \neq 0$, then

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & (a-1)c^{-1} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ c & 1 \end{bmatrix} \begin{bmatrix} 1 & (d-1)c^{-1} \\ 0 & 1 \end{bmatrix}.$$

Case 3: If $b = c = 0$, then

$$\begin{bmatrix} a & 0 \\ 0 & a^{-1} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ (1-a)a^{-1} & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ a-1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -a^{-1} \\ 0 & 1 \end{bmatrix},$$

and we are done. ■

Let $e_{ij}(\lambda)$ is a matrix with λ in ijth entry and zeroes elsewhere, and let $E_{ij}(\lambda) = I_n + e_{ij}(\lambda)$.

Theorem 7.41 *The special linear group* $SL_n(\mathbb{F})$ *is generated by matrices of the form* $E_{ij}(\lambda)$ *with* $i \neq j$ *and* $\lambda \in \mathbb{F}^*$.

Proof We apply mathematical induction. If $n = 2$, then by Theorem 7.40, we are done. Suppose that the statement is true for each $(n-1) \times (n-1)$ matrix. Let A be any arbitrary matrix in $SL_n(\mathbb{F})$. Multiplying A by $E_{ij}(\lambda)$ on the left or right is an elementary row or column operation:

$$E_{ij}(\lambda)A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{i1} + \lambda a_{j1} & \dots & a_{in} + \lambda a_{jn} \\ \vdots & \vdots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix}$$

and

$$AE_{ij}(\lambda) = \begin{bmatrix} a_{11} & \dots & a_{1j} + \lambda a_{1i} & \dots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \dots & a_{nj} + \lambda a_{ni} & \dots & a_{nn} \end{bmatrix}.$$

These are the only operations we may use. Since $\det(A) = 1$, it follows that some entry of the first column of A is not 0. If $a_{k1} \neq 0$ with $k > 1$, then

$$E_{1k}\left(\frac{1 - a_{11}}{a_{k1}}\right) A = \begin{bmatrix} 1 & \dots \\ \vdots & \ddots \end{bmatrix}. \tag{7.4}$$

If $a_{21}, \dots, a_{n1}$ are all 0, then $a_{11} \neq 0$ and

$$E_{21}\left(\frac{1}{a_{11}}\right) A = \begin{bmatrix} a_{11} & \dots \\ 1 & \dots \\ \vdots & \ddots \end{bmatrix}.$$

Hence, by (7.4) with $k = 2$, we obtain

$$E_{12}(1 - a_{11})E_{21}\left(\frac{1}{a_{11}}\right) A = \begin{bmatrix} 1 & \dots \\ \vdots & \ddots \end{bmatrix}.$$

When we have a matrix with upper left entry 1, multiplying it on the left by $E_{i1}(\lambda)$ for $i \neq 1$ will all λ to the $1i$th entry, hence with a suitable λ, we can make the $i1$th entry of the matrix 0. Consequently, multiplication on the left by suitable matrices of the form $E_{ij}(\lambda)$ produces a block matrix

$$\begin{bmatrix} 1 & * \\ 0 & B \end{bmatrix}$$

whose first column is all 0's except for the upper left entry, which is 1. Multiplying this matrix on the right by $E_{1j}(\lambda)$ for $j \neq 1$ adds λ to the $1j$th entry without changing a column other than the jth column. With a suitable choice of λ we can make the $1j$th entry equal to 0, and carrying this out for $j = 2, \dots, n$ leads to a block matrix

$$\begin{bmatrix} 1 & 0 \\ 0 & A' \end{bmatrix}. \tag{7.5}$$

Since this matrix is in $SL_n(\mathbb{F})$, it follows that $\det(A') = 1$. This implies that $A' \in SL_{n-1}(\mathbb{F})$. By induction hypothesis, A' is a product of elementary matrices $E_{ij}(\lambda)_{(n-1)\times(n-1)}$. We say

$$A' = E_1 E_2 \dots E_r.$$

So, we have

$$\begin{bmatrix} 1 & 0 \\ 0 & A' \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & E_1E_2\dots E_r \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & E_1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & E_2 \end{bmatrix} \cdots \begin{bmatrix} 1 & 0 \\ 0 & E_r \end{bmatrix}.$$

This yields that A is a product of $n \times n$ matrices $E_{ij}(\lambda)$. ■

Definition 7.42 Let A be an $m \times n$ matrix. Then, A^t, the *transpose* of A, is the matrix obtained by interchanging the rows and columns of A. Geometrically, A^t is obtained from A by reflecting across the diagonal of A.

We say A is *symmetric* if $A^t = A$ and A is *skew-symmetric* if $A^t = -A$.

Assuming that the sizes of the matrices are such that the operations can be performed, the transpose operation has the following properties:

(1) $(A^t)^t = A$,
(2) $(A + B)^t = A^t + B^t$,
(3) $(cA)^t = cA^t$, where $c \in \mathbb{F}$,
(4) $(AB)^t = B^tA^t$.

Corollary 7.43 *If A is any square matrix, then* $\det(A) = \det(A^t)$.

Corollary 7.44 *Any orthogonal matrix has determinant equal to* 1 *or* -1.

Theorem 7.45 *If*

$$O_n(\mathbb{F}) = \{A \in GL_n(\mathbb{F}) \mid A^tA = I_n\},$$

then $O_n(\mathbb{F})$ is a subgroup of $GL_n(\mathbb{F})$.

Proof Let A and B be two arbitrary elements of $O_n(\mathbb{F})$. Then, we have $A^tA = I_n$ and $B^tB = I_n$. This implies that

$$(AB)^t(AB) = B^tA^tAB = B^tI_nB = B^tB = I_n.$$

So, we conclude that $AB \in O_n(\mathbb{F})$. Moreover, we have $A^t = A^{-1}$, and so

$$(A^{-1})^tA^{-1} = AA^{-1} = I_n.$$

This implies that $A^{-1} \in O_n(\mathbb{F})$. ■

$O_n(\mathbb{F})$ is called the *orthogonal group*. We can define

$$SO_n(\mathbb{F}) = SL_n(\mathbb{F}) \cap O_n(\mathbb{F}).$$

$SO_n(\mathbb{F})$ is called the *special orthogonal group*.

One can gets a whole of examples by fixing $Q \in M_n(\mathbb{F})$ and defining

$$G_Q(\mathbb{F}) = \{A \in GL_n(\mathbb{F}) \mid A^tQA = Q\}.$$

Theorem 7.46 *$G_Q(\mathbb{F})$ is a subgroup of $GL_n(\mathbb{F})$.*

Proof Suppose that A and B are two arbitrary elements of $G_Q(\mathbb{F})$. Then we have $AB \in G_Q(\mathbb{F})$. Indeed,

$$(AB)^t Q(AB) = B^t A^t QAB = B^t QB = Q.$$

In addition, we have $A^{-1} \in G_Q(\mathbb{F})$. Indeed, when multiplying the equality $Q = A^t QA$ to the right by A^{-1} and to the left by $(A^{-1})^t$, then we obtain $(A^{-1})^t QA^{-1} = Q$. ■

One important case is when $n = 2m$ and we take Q to be the matrix

$$J_m = \begin{bmatrix} 0 & -I_m \\ I_m & 0 \end{bmatrix}$$

Then, $G_Q(\mathbb{F})$ is denoted by $SP_{2m}(\mathbb{F})$ and is called *symplectic group*

Exercises

1. If A is a symmetric matrix, is the matrix A^{-1} symmetric?
2. Show that the non-zero elements of $Mat_{n\times n}(\mathbb{C})$ is not group under multiplication.
3. For what values of a and b the following matrix belong to $GL_4(\mathbb{R})$:

$$\begin{bmatrix} b & 0 & a & 0 \\ 0 & 0 & b & a \\ 0 & a & 0 & b \\ b & a & 0 & 0 \end{bmatrix}.$$

4. The *trace* of a square matrix $A = \left(a_{ij}\right)_{n\times n}$ is

$$tr(A) = \sum_{i=1}^{n} a_{ii}.$$

Let A and B be two square matrices.

(a) Prove that $tr(AB) = tr(BA)$;
(b) If B is invertible, show that the formula $tr(B^{-1}AB) = tr(A)$.

5. Let A be a 2×2 matrix over a field $\mathbb{F}$. Prove that $\det(I_2 + A) = 1 + \det(A)$ if and only if $tr(A) = 0$.
6. Let A be an $n \times n$ matrix over a field $\mathbb{F}$. Prove that there are at most n distinct scalars $c \in \mathbb{F}$ such that $\det(cI_2 - A) == 0$.

7. If A and B are invertible matrices and the involved partitioned products are defined, show that

$$\begin{bmatrix} A & B \\ C & 0 \end{bmatrix}^{-1} = \begin{bmatrix} 0 & C^{-1} \\ B^{-1} & -B^{-1}AC^{-1} \end{bmatrix}.$$

8. Show that the non-zero elements of $Mat_{n\times n}(\mathbb{C})$ is not group under multiplication.
9. Let $A = \left(a_{ij}\right)_{n\times n}$ be a square matrix with $a_{ij} = \max\{i,\ j\}$. Compute $\det(A)$.
10. Let $A = \left(a_{ij}\right)_{n\times n}$ be a square matrix with $a_{ij} = 1/(i + j)$. Show that A is invertible.
11. Show that the set

$$G = \left\{ \begin{bmatrix} \cosh x & \sinh x \\ \sinh x & \cosh x \end{bmatrix} \mid x \in \mathbb{R} \right\},$$

in which sinh and cosh are hyperbolic functions, is a group under multiplication of matrices.
12. Show by example that if $AC = BC$, then it does not follow that $A = B$. However, show that if C is invertible the conclusion $A = B$ is valid.
13. Show that $SL_2(\mathbb{C}) = SP_2(\mathbb{C})$.
14. Does the set $\{A \in GL_2(\mathbb{R}) \mid A^2 = I_2\}$ form a subgroup of $GL_2(\mathbb{R})$?
15. Let A be a 2×2 matrix over a field $\mathbb{F}$, and suppose that $A^2 = 0$. Show that for each scalar c that $\det(cI_2 - A) = c^2$.
16. Determine the *one-dimensional Lorentz group*. That is, all 2×2 matrices A such that

$$A^t \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

17. Determine $GL_2(\mathbb{Z}_2)$, $SL_2(\mathbb{Z}_2)$, $O_2(\mathbb{Z}_2)$, $SO_2(\mathbb{Z}_2)$, and $SP_2(\mathbb{Z}_2)$.
18. (a) Let G be the group of all 2×2 matrices $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$, where $ad - bc \neq 0$ and a, b, c, d are integers modulo 3. Show that $|G| = 48$;
    (b) If we modify the example of G in part (a) by insisting that $ad - bc = 1$, then what is $|G|$?
19. (a) Let G be the group of all 2×2 matrices $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$, where a, b, c, d are integers modulo p, p is a prime number, such that $ad - bc \neq 0$; indeed, $G = GL_2(\mathbb{Z}_p)$. What is $|G|$?
    (b) Let H be the subgroup of G of part (a) defined by

$$\left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \mid ad - bc = 1 \right\};$$

indeed, $H = SL_2(\mathbb{Z}_p)$. What is $|H|$?
20. Let

$$H = \left\{ \begin{bmatrix} 1 & a & b \\ 1 & 1 & c \\ 0 & 0 & 1 \end{bmatrix} \mid a, b, c \in \mathbb{Z}_3 \right\}.$$

(a) Show that H is a subgroup of $SL_3(\mathbb{Z}_3)$;
(b) How many elements does H have?
(c) Find three subgroups of H of order 9. Are these subgroups of order 9 abelian? Is H abelian?

21. Find the centralizer of $\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ in $GL_2(\mathbb{R})$.
22. Let $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix}$ be two elements of $SL_2(\mathbb{Q})$. Show that $o(A) = 4$, $o(B) = 3$, but AB has infinite order.
23. Let $A \in Mat_{2\times 2}(\mathbb{C})$. Show that there is no matrix solution $B \in Mat_{2\times 2}(\mathbb{C})$ to $AB - BA = I_2$. What can you say about the same problem with $A, B \in Mat_{n\times n}(\mathbb{C})$?
24. Let $\lambda \in \mathbb{F}$. Prove that the matrix $E_{ij}(\lambda)$ is of order p if and only if the characteristic of $\mathbb{F}$ is p.

7.2 More About Vectors in $\mathbb{R}^n$

Vectors are the way we represent points in two, three, or n dimensional space. A *vector* can be considered as an object which has a length and a direction, or it can be thought of as a point in space, with coordinates representing that point. In terms of coordinates, we use the notation

$$X = \begin{bmatrix} x \\ y \end{bmatrix}$$

for a column vector. This vector denotes the point in the xy-plane which is x units in the horizontal direction and y units in the vertical direction from the origin. In n dimensional space, we represent a vector by

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Let $X = [x_1\ x_2\ \ldots\ x_n]^t$ and $Y = [y_1\ y_2\ \ldots\ y_n]^t$ be two vectors in $\mathbb{R}^n$. We define the *inner product* $\langle X, Y \rangle$ by

$$\langle X, Y \rangle = x_1y_1 + x_2y_2 + \cdots x_ny_n.$$

The *length* of X is given by

$$\|X\| = \sqrt{X \cdot X} = \sqrt{x_1^2 + x_2^2 + \cdots x_n^2}.$$

The vector $U = \frac{X}{\|X\|}$ is called the *unit vector* in the X-direction. The angle θ between two vectors X and Y is calculated by

$$\langle X, Y\rangle = \|X\|\,\|Y\| \cos\theta,$$

where $0 \le \theta \le \pi$. Two vector X and Y are said to be *orthogonal* if $\langle X, Y\rangle = 0$.

Definition 7.47 Non-zero vectors $X_1, X_2, \ldots, X_k$ in $\mathbb{R}^n$ form an *orthogonal set* if they are orthogonal to each other, i.e., $\langle X_i, X_j\rangle = 0$ for all $i \neq j$. An *orthonormal set* is an orthogonal set with the additional property that all vectors are unit, i.e., $\|X_i\| = 1$ for all $i = 1, \ldots, k$.

Lemma 7.48 *Any orthogonal set S is linearly independent.*

Proof Let $X_1, X_2, \ldots, X_k$ be distinct vectors in S and

$$Y = c_1X_1 + c_2X_2 + \cdots + c_kX_k.$$

Then, we have

$$\langle Y, X_j\rangle = \langle \sum_{i=1}^{k} c_iX_j, X_i\rangle = c_j\langle X_j, X_j\rangle = c_j\|X_j\|^2.$$

Since $\langle X_j, X_j\rangle \neq 0$, it follows that

$$c_j = \frac{\langle Y, X_j\rangle}{\|X_j\|^2},\ 1 \le j \le k.$$

Thus, when $Y = 0$, each $c_j = 0$. ■

Gram-Schmidt orthogonalization process: Let $X_1, X_2, \ldots, X_k$ be any independent vectors. Then one may construct orthogonal vectors $V_1, V_2, \ldots, V_k$ as follows:

$$\begin{aligned}
V_1 &= X_1,\\
V_2 &= X_2 - \frac{\langle X_2, V_1\rangle}{\|V_1\|^2}V_1,\\
&\vdots\\
V_j &= X_j - \sum_{i=1}^{j-1}\frac{\langle X_j, V_i\rangle}{\|V_i\|^2}V_i,\\
&\vdots
\end{aligned}$$

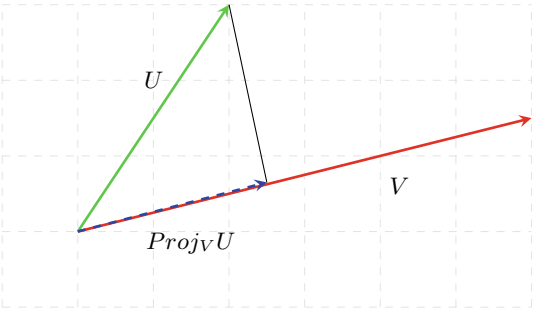


Fig. 7.1 The projection of U onto V

Then, $V_1, V_2, \ldots, V_k$ are orthogonal vectors.

The *vector projection* of a vector U on a non-zero vector V is the orthogonal projection of U on a straight line parallel to V, see Fig. 7.1.

Lemma 7.49 *The vector projection of a vector U on a non-zero vector V is equal to*

$$Proj_V U = \frac{VV^t}{\|V\|^2}U.$$

Proof The projection vector of U onto V is a scalar multiple of V, i.e., $Proj_V U = cV$. Since the vector $U - cV$ is perpendicular to V, it follows that $\langle V, U - cV\rangle = 0$. Hence, we have $\langle V, U\rangle - c\langle V, V\rangle = 0$. This implies that

$$c = \frac{\langle V, U\rangle}{\langle V, V\rangle}.$$

Thus, we conclude that

$$Proj_V U = \frac{\langle V, U\rangle}{\langle V, V\rangle}V. \tag{7.6}$$

When we multiply a vector by a scalar, it does not matter whether we put the scalar before or after the vector. So, by (7.6), we get

$$Proj_V U = V\frac{\langle V, U\rangle}{\langle V, V\rangle}.$$

Since $\langle V, U\rangle = V^t U$ and $\langle V, V\rangle = \|V\|^2$, it follows that

$$Proj_V U = V\frac{V^t U}{\|V\|^2} = \frac{VV^t}{\|V\|^2}U,$$

as desired. ■

If we define

$$P = \frac{VV^t}{\|V\|^2},$$

then the projection formula becomes $Proj_V U = PU$. The matrix P is called the *projection matrix*. So, we can project any vector onto the vector V by multiplying by the matrix P.

Exercises

1. (a) Suppose that $\theta \in \mathbb{R}$. Show that

$$(\cos\theta, \sin\theta), \quad (-\sin\theta, \cos\theta)$$

and

$$(\cos\theta, \sin\theta), \quad (\sin\theta, -\cos\theta)$$

are orthonormal bases of $\mathbb{R}^2$;

(b) Show that each orthonormal basis of $\mathbb{R}^2$ is of the form given by one of the two possibilities of part (a).

2. Suppose that $X, Y \in \mathbb{R}^n$. Prove that $\langle X, Y\rangle = 0$ if and only if

$$\|X\| \leq \|X + aY\|,$$

for all $a \in \mathbb{R}$.

3. Find vectors $X, Y \in \mathbb{R}^2$ such that X is a scalar multiple of $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$, Y is orthogonal to $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$, and $\begin{bmatrix} 1 \\ 2 \end{bmatrix} = X + Y$.

4. Apply Gram-Schmidt orthogonalization process to the following vectors in $\mathbb{R}^3$:

$$\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 8 \\ 1 \\ -6 \end{bmatrix} \text{ and } \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

5. Find the angle between a diagonal of a cube and one of its edges.

6. If $C = \|A\|B + \|B\|A$, where A, B, and C are all non-zero vectors, show that C bisects the angle between A and B.

7.3 Rotation Groups

A rotation matrix describes the rotation of an object. In this section, we derive the matrix for a general rotation.

Theorem 7.50 *Any* 2×2 *orthogonal matrix with entries in* $\mathbb{R}$ *and determinant* 1 *has the form*

$$\text{Rot}(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix},$$

and represents a rotation in $\mathbb{R}^2$ *by an angle* θ *counterclockwise, with center the origin.*

Proof Suppose that

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

be an orthogonal matrix with entries in $\mathbb{R}$. Then, we have

$$A^t A = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

So, both columns of A are mutually orthogonal unit vectors. The only two unit vectors orthogonal to $\begin{bmatrix} a \\ c \end{bmatrix}$ are $\begin{bmatrix} -c \\ a \end{bmatrix}$ and $\begin{bmatrix} c \\ -a \end{bmatrix}$. Hence, one of these must be $\begin{bmatrix} b \\ d \end{bmatrix}$. Consequently, there exist two possibilities for A:

$$\begin{bmatrix} a & -c \\ c & a \end{bmatrix} \text{ or } \begin{bmatrix} a & c \\ c & -a \end{bmatrix}.$$

Since $a^2 + b^2 = 1$, it follows that the first matrix has determinant 1 and the second matrix has determinant -1. If we consider the first matrix, then there exists a unique angle θ such that $\cos\theta = a$ and $\sin\theta = b$. Hence, we denote the resulting matrix A by

$$\text{Rot}(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix},$$

to emphasis that it is a function of θ. Now, we obtain

$$\text{Rot}(\theta) \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{bmatrix}.$$

The right side is the image of the vector $\begin{bmatrix} x \\ y \end{bmatrix}$ under a rotation about the origin by angle θ. ■

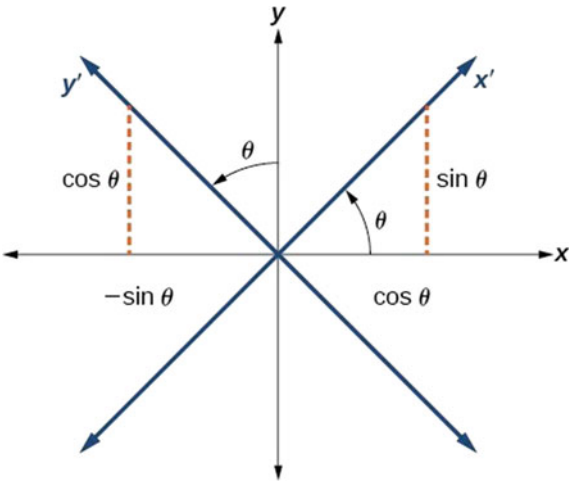


Fig. 7.2 The x- and y-axes and the resulting x'- and y'-axes formed by a rotation through an angle θ

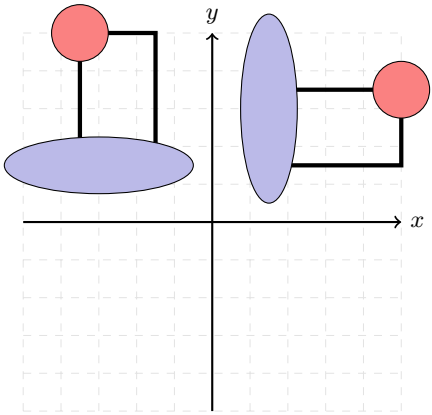


Fig. 7.3 Rotation with $\theta = \pi/2$

Indeed, according to Theorem 7.50, if the x and y axes are rotated through an angle θ, then every point on the plane may be thought of as having two representations:

(1) (x, y) on the xy plane with the original x-axis and y-axis.
(2) (x', y') on the new plane defined by the new, rotated axes, called the x'-axis and y'-axis, see Fig. 7.2.

Example 7.51 In particular, the matrix

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

is used for $\theta = \pi/2$, see Fig. 7.3.

The matrix

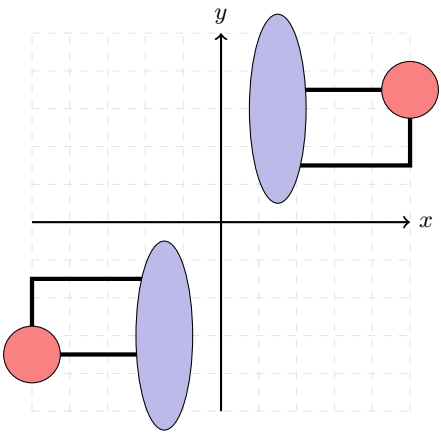


Fig. 7.4 Rotation with $\theta = \pi$

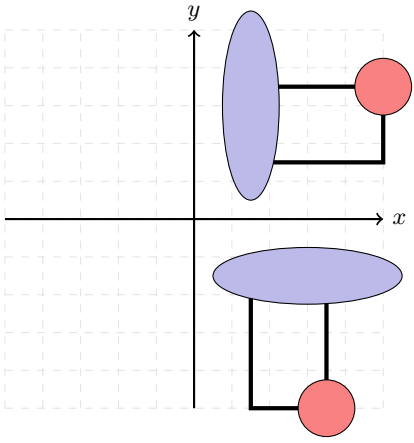


Fig. 7.5 Rotation with $\theta = 3\pi/2$

$$\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

is used for $\theta = \pi$, see Fig. 7.4.

The matrix

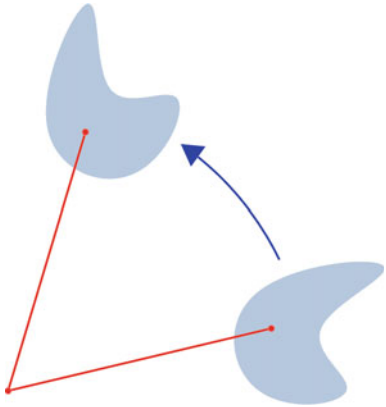
$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

is used for $\theta = 3\pi/2$, see Fig. 7.5.

Lemma 7.52 *If A is an $n \times n$ orthogonal matrix with $\det(A) = 1$, then*

$$\det(A - I_n) = (-1)^n \det(A - I_n).$$

Fig. 7.6 A rotation in $\mathbb{R}^3$



Proof We have

$$\begin{aligned}\det(A - I_n) &= \det(A^t)\det(A - I_n)\\ &= \det\left(A^t(A - I_n)\right)\\ &= \det(I_n - A^t)\\ &= \det\left((I_n - A)^t\right)\\ &= \det(I_n - A)\\ &= (-1)^n \det(A - I_n),\end{aligned}$$

as desired. ■

Corollary 7.53 *If A is a 3×3 orthogonal matrix with $\det(A) = 1$, then there exists a non-zero column vector $X \in \mathbb{R}^3$ such that $AX = X$.*

Proof In Lemma 7.52, let $n = 3$. Then, we deduce that $\det(A - I_n) = 0$. This yields that $A - I_n$ is not invertible. So, $(A - I_n)X = 0$ has a non-zero solution. Consequently, there exists a non-zero column vector $X \in \mathbb{R}^3$ such that $AX = X$. ■

Definition 7.54 A rotation in $\mathbb{R}^3$ is a matrix $R(e_r, \theta)$ determined by a unit vector $e_R \in \mathbb{R}^3$ and an angle θ. More precisely, $R(e_R, \theta)$ has the line through e_R as the axis and the plane perpendicular to the line is rotated by the angle θ in the counterclockwise direction (see Fig. 7.6).

Corollary 7.53 can be used to show that any 3×3 orthogonal matrix with determinant 1 represents a rotation in $\mathbb{R}^3$ about an axis passing through the origin.

Theorem 7.55 *Each rotation matrix in $\mathbb{R}^3$ lies in $SO_3(\mathbb{R})$.*

Proof Suppose that $R(e_R, \theta)$ is a rotation in $\mathbb{R}^3$, where e_R is the unit vector in the direction of the axis of rotation. By Gram-Schmidt orthogonalization process, we can find two more vectors. So, we can consider an orthonormal set $\{V_1, V_2, V_3\}$ of $\mathbb{R}^3$ such that $V_3 = e_R$. Therefore, the matrix $[V_1\ V_2\ V_3]$ is orthogonal. Now, by referring to Theorem 7.50, this rotation described by

$$R = R(e_R, \theta) = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Since $R^t R = I_3$ and $\det(R) = 1$, we conclude that $R \in SO_3(\mathbb{R})$. ■

Theorem 7.56 *Each matrix in $SO_3(\mathbb{R})$ is a rotation.*

Proof Suppose that $R \in SO_3(\mathbb{R})$ is an arbitrary element. By Lemma 7.53, there exists a column unit vector V_3 such that $RV_3 = V_3$. By Gram-Schmidt orthogonalization process, we can extend this to an orthonormal set $\{V_1, V_2, V_3\}$. Then, the matrix $A = [V_1\ V_2\ V_3] \in SO_3(\mathbb{R})$. Hence, we conclude that $RA \in SO_3(\mathbb{R})$. We can write

$$\begin{aligned} RV_1 &= aV_1 + bV_2, \\ RV_2 &= cV_1 + dV_2, \\ RV_3 &= V_3. \end{aligned}$$

Since

$$A^{-1}RA = \begin{bmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & 1 \end{bmatrix} \in SO_3(\mathbb{R}),$$

it follows that

$$A^{-1}RA = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in SO_2(\mathbb{R}),$$

which means that it is a planer rotation matrix $\mathrm{Rot}(\theta)$. Consequently, we have $R = R(V_3, \theta)$. ■

Corollary 7.57 *The set of rotation in $\mathbb{R}^3$ can be identified with $SO_3(\mathbb{R})$.*

Exercises

1. For any rotation matrix R, show that $R^t = R^{-1}$.
2. Use standard trigonometric identities to verify that

$$\mathrm{Rot}(\theta)\mathrm{Rot}(\phi) = \mathrm{Rot}(\theta + \phi).$$

   Deduce that $\mathrm{Rot}(2\theta) = \mathrm{Rot}(\theta)^2$ and $\mathrm{Rot}(-\theta) = \mathrm{Rot}(\theta)^{-1}$.
3. Use rotation matrix to find the image of the vector $\begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix}$ if it is rotated
   (a) 30° about the x-axis;
   (b) -30° about the x-axis;

(c) 45° about the y-axis;
(d) $-45°$ about the y-axis;
(e) 90° about the z-axis;
(f) $-90°$ about the z-axis.

7.4 Reflections in $\mathbb{R}^2$ and $\mathbb{R}^3$

Reflection matrix is the matrix which can be used to make reflection transformation of a figure.

Theorem 7.58 *A reflection across the line $y = mx$ is performed by matrix*

$$\text{Ref}(\theta) = \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix},$$

in terms of an angle θ that mirror makes with the positive x-axis (Fig. 7.7).

Proof The formula for the distance of $P(x, y)$ from the line $mx - y = 0$ is

$$d = \frac{mx - y}{\sqrt{1 + m^2}},$$

where $m = \tan\theta$. Hence, we have

$$d = \frac{x\tan\theta - y}{\sqrt{1 + \tan^2\theta}} = \frac{x\tan\theta - y}{\sec\theta} = x\sin\theta - y\cos\theta.$$

So, we conclude that

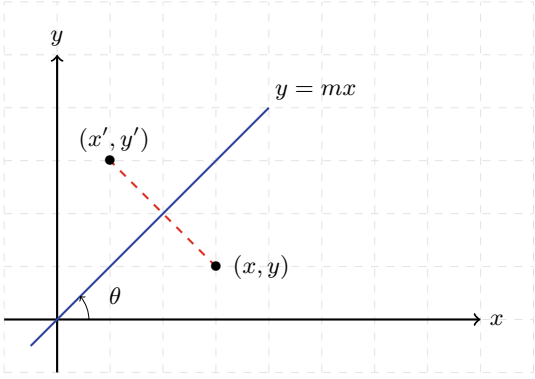


Fig. 7.7 A reflection across the line $y = mx$

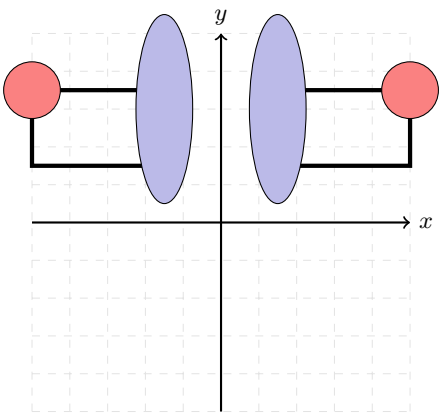


Fig. 7.8 Reflection with $\theta = \pi/2$

$$\begin{aligned} x' &= x - 2d\sin\theta = x - 2\sin\theta(x\sin\theta - y\cos\theta) \\ &= x(1 - 2\sin^2\theta) + 2y\sin\theta\cos\theta = x\cos 2\theta + y\sin 2\theta \end{aligned}$$

and

$$\begin{aligned} y' &= y + 2d\cos\theta = y + 2\cos\theta(x\sin\theta - y\cos\theta) \\ &= 2x\sin\theta\cos\theta + y(1 - 2\cos^2\theta) = x\sin 2\theta - y\cos 2\theta. \end{aligned}$$

Consequently, we obtain

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix},$$

and we are done. ■

Example 7.59 In particular, the matrix

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

is used for $\theta = \pi/2$, see Fig. 7.8. The matrix

$$\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

is used for $\theta = \pi$, see Fig. 7.9.

Corollary 7.60 *The set of reflections in* $\mathbb{R}^2$ *is exactly the set of elements in* $O_2(\mathbb{R})$ *that do not belong to* $SO_2(\mathbb{R})$.

Definition 7.61 A *reflection* in $O_3(\mathbb{R})$ is a matrix A whose effect is to send every column vector of $\mathbb{R}^3$ to its mirror image with respect to a plane S containing the

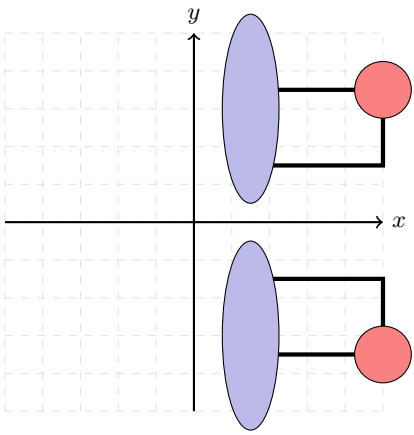


Fig. 7.9 Reflection with $\theta = \pi$

origin. More precisely, suppose that

$$S = \{ax + by + cz = 0 \mid a, b, c \in \mathbb{R}\}$$

is a plane through the origin. We say a matrix A makes a *reflection across* S if

(1) $AX = X$ for all vectors X in S,
(2) $AX = -X$ for all vectors X perpendicular to S.

If N is chosen to be a unit column vector perpendicular to S, then

$$AX = X - 2\langle X, N\rangle N, \tag{7.7}$$

for all $X \in \mathbb{R}^3$.

According to (7.7), we obtain

$$\begin{aligned} AX &= X - 2\langle X, N\rangle N \\ &= X - 2N\langle X, N\rangle \\ &= X - 2N(N^t X) \\ &= X - 2(NN^t)X \\ &= (I_n - 2NN^t)X. \end{aligned}$$

The matrix $(I_n - 2NN^t)$ is called *reflection matrix* for the plane S, and is also sometimes called a *Householder matrix*.

Example 7.62 Let $X^t = [x_1 \ x_2 \ x_3]$. Table 7.1 describes some reflections of the column vector X^t across some planes.

Example 7.63 We want to compute the reflection of the column vector

Table 7.1 Some reflections of the column vector X^t across some planes

Operator	Equations defining the image	Reflection matrix
Reflection across the xy plane	$x_1' = x + 0y + 0z$ $x_2' = 0x + y + 0z$ $x_3' = 0x + 0y - z$	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$
Reflection across the xz plane	$x_1' = x + 0y + 0z$ $x_2' = 0x - y + 0z$ $x_3' = 0x + 0y + z$	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$
Reflection across the yz plane	$x_1' = -x + 0y + 0z$ $x_2' = 0x + y + 0z$ $x_3' = 0x + 0y + z$	$\begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

$$X = \begin{bmatrix} -1 \\ 2 \\ -2 \end{bmatrix}$$

across the plane $2x - y + 3z = 0$. The column vector

$$V = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}$$

is normal to the plane and $\|V\|^2 = \langle V, V\rangle = 2^2 + (-1)^2 + 3^2 = \sqrt{14}$. So, a unit normal vector is

$$N = \frac{V}{\|V\|} = \frac{1}{\sqrt{14}} \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}.$$

Consequently, the reflection matrix is equal to

$$\begin{aligned} I_n - 2NN^t &= I_n - \frac{1}{7} VV^t \\ &= I_n - \frac{1}{7} \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} [2 \quad -1 \quad 3] \\ &= \begin{bmatrix} \frac{3}{7} & \frac{2}{7} & \frac{-6}{7} \\ \frac{2}{7} & \frac{6}{7} & \frac{3}{7} \\ \frac{-6}{7} & \frac{3}{7} & \frac{-2}{7} \end{bmatrix}. \end{aligned}$$

Theorem 7.64 *A reflection across a plane S can be performed by the matrix*

$$\text{Ref}(\theta) = \begin{bmatrix} \cos 2\theta & \sin 2\theta & 0 \\ \sin 2\theta & -\cos 2\theta & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Proof Let V_1 be a unit vector in $\mathbb{R}^3$, V_2 be a unit vector orthogonal to V_1, and let $V_3 = U \times V$. Then, (V_1, V_2, V_3) is a orthogonal triple. Suppose that $A = \text{Ref}(\theta)$ is a reflection in $\mathbb{R}^3$ across the plane S through the origin whose unit normal column vector is N, and orthogonal to V_3. We can find an angle θ such that

$$N = -\sin\theta V_1 + \cos\theta V_2.$$

Using Eq. (7.7), we obtain

$$\begin{aligned} AV_1 &= (1 - 2\sin^2\theta)V_1 + 2\sin\theta\cos\theta V_2, \\ AV_2 &= 2\sin\theta\cos\theta V_1 + (1 - 2\cos^2\theta)V_2, \\ AV_3 &= V_3. \end{aligned}$$

That is,

$$\begin{aligned} AV_1 &= \cos 2\theta V_1 + \sin 2\theta V_2, \\ AV_2 &= \sin 2\theta V_1 - \cos 2\theta V_2, \\ AV_3 &= V_3. \end{aligned}$$

Therefore, we obtain

$$A = \text{Ref}(\theta) = \begin{bmatrix} \cos 2\theta & \sin 2\theta & 0 \\ \sin 2\theta & -\cos 2\theta & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

as required. ■

Exercises

1. Use standard trigonometric identities to verify that
   (a) $\text{Ref}(\theta)\text{Ref}(\phi) = \text{Rot}(2(\theta - \phi))$,
   (b) $\text{Rot}(\theta)\text{Ref}(\phi) = \text{Ref}(\phi + \frac{1}{2}\theta)$,
   (c) $\text{Ref}(\phi)\text{Rot}(\theta) = \text{Ref}(\phi - \frac{1}{2}\theta)$.
2. Find a matrix which represents a reflection in the line $y = 2x$.
3. Prove that each element of $O_3(\mathbb{R})$ can be expressed as a product of at most three reflections.

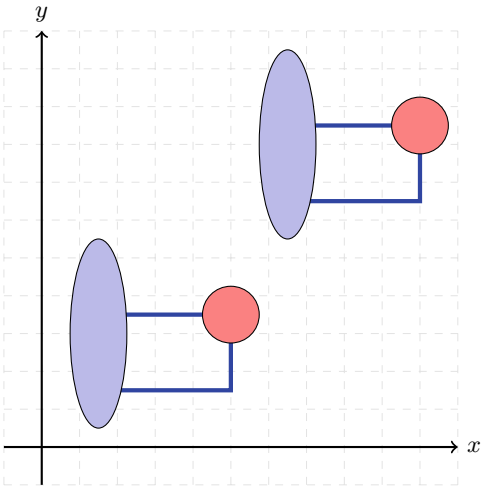


Fig. 7.10 A translation in $\mathbb{R}^2$

4. Find the matrix which induces reflection with respect to given vector $[a\ b\ c]^t$ in $\mathbb{R}^3$. Check your result for the case $[1\ 0\ 0]^t$.

7.5 Translation and Scaling Matrices

Suppose that $\mathcal{T}_n$ is the set of all translations of a fixed point $(x_1, x_2, \ldots, x_n)$ in $\mathbb{R}^n$. If

$$A = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \in \mathcal{T}_n.$$

then

$$\begin{bmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}.$$

$\mathcal{T}_n$ with addition of columns matrices as the composition rule is a group. This group is called the *translation group*.

Example 7.65 A translation is shown in Fig. 7.10 by $A = \begin{bmatrix} 5 \\ 5 \end{bmatrix} \in \mathcal{T}_2$.

Example 7.66 The set of all translations parallel to the x-axis, i.e., all translations of the form

$$\left\{ \begin{bmatrix} a \\ 0 \end{bmatrix} \mid a \in \mathbb{R} \right\}$$

is a subgroup of $\mathcal{T}_2$.

Example 7.67 The set of all translations parallel to the y axis, i.e., all translations of the form

$$\left\{ \begin{bmatrix} 0 \\ b \end{bmatrix} \mid b \in \mathbb{R} \right\}$$

is a subgroup of $\mathcal{T}_2$.

A scaling can be represent by a *scaling matrix*. To scale an object by a vector $V^t = [c_1\ c_2\ \ldots\ c_n]$ ($c_1 = \cdots = c_n \neq 0$), each point X would need to be multiplied by the following diagonal matrix:

$$A = \begin{bmatrix} c_1 & 0 & \ldots & 0 \\ 0 & c_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & c_n \end{bmatrix}$$

As shown below, the multiplication give the expected result:

$$AX = \begin{bmatrix} c_1 & 0 & \ldots & 0 \\ 0 & c_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & c_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} c_1 x_1 \\ c_2 x_2 \\ \vdots \\ c_n x_n \end{bmatrix}.$$

Remark 7.68 The set of all diagonal matrices with non-zero determinants in $GL_n(\mathbb{R})$ forms a subgroup.

Such a scaling changes the diameter of an object by a factor between the scale factors, the area by a factor between the smallest and the largest product of two scale factors, and the volume by the product of all three.

The scaling is uniform if and only if the scaling factors are equal ($v_1 = \cdots = v_n$). If all except one of the scale factors are equal to 1, we have directional scaling.

In the case where $v_1 = \cdots = v_n = c$, scaling changes the area of any surface by a factor of c^2 (see Fig. 7.11) and the volume of any solid object by a factor of c^3 (see Fig. 7.12).

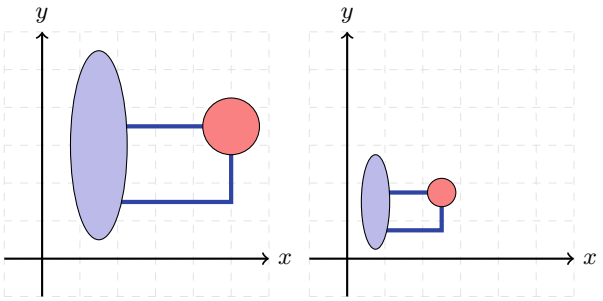


Fig. 7.11 Scaling with $c = 0.5$ in $\mathbb{R}^2$

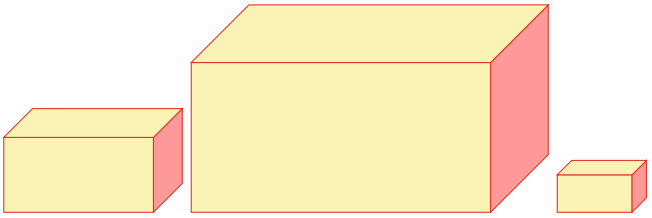


Fig. 7.12 Scaling with $c = 2$ and $c = 0.5$ in $\mathbb{R}^3$

Exercises

1. Prove that every positive rigid motion in $\mathbb{R}^3$ can be obtained by a rotation about a l-axis, followed by a translation along l (this type of motion is called *screw*).
2. Consider the graph $y = e^x$. Suppose that the graph is dilated from the y-axis by a factor of 3, reflected in the x-axis, and then translated 1 unit to the left and 2 units down. What is the equation of the resulting graph?
3. Consider the graph $y = x^2$. Suppose the graph is dilated from the x axis by a factor of 2, and then translated 3 units to the right. What is the equation of the resulting graph?
4. Suppose the graph of $y = f(x)$ is transformed by a dilation of factor k from the x axis, factor $h \neq 0$ from the y axis, and then translated c units to the right and d units up. Show that the resulting graph has equation

$$y = kf\left(\frac{1}{h}x - c\right) + d.$$

5. Two geometric shapes A and B are shown in Fig. 7.13. Shape A has one point at $(3, 4)$ and shape B has one point at $(0, -2)$. Calculate a chain of matrices that, when post-multiplied by the vertices of shape A, will transform all the vertices of shape A into the vertices of shape B, i.e., translate and rotate the shape point

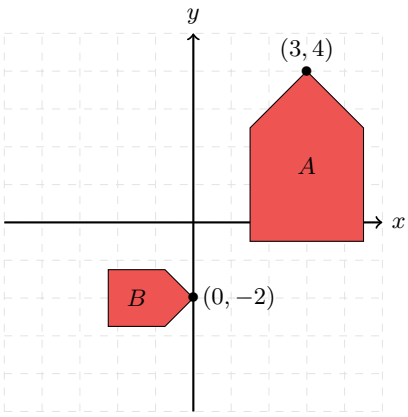


Fig. 7.13 Two geometric shapes A and B

$(3, 4)$ to $(0, -2)$. The transformation must also scale the size of shape A by half to shape B.

7.6 Dihedral Groups

The dihedral groups form an important infinite family of examples of finite groups. They arise as groups of symmetries of the regular n-gons, and they play an important role in group theory, geometry, and chemistry.

Definition 7.69 The *dihedral group* D_n is the group of symmetries of a regular polygon with n vertices.

We may consider this polygon as having vertices on the unit circle, with vertices labelled $1, 2, \ldots, n-1$ starting at $(1, 0)$ and proceeding counterclockwise at angles in multiples of $2\pi/n$ radians. There are two types of symmetries of the regular n-gon, each one giving rise to n elements in the group D_n:

(1) Rotations $R_0, R_1, \ldots, R_{n-1}$, where R_k is rotation of angle $2k\pi/n$.
(2) Reflections $S_0, S_1, \ldots, S_{n-1}$, where S_k is reflection about the line through the origin and making an angle of $k\pi/n$ with the horizontal axis.

The group operation is given by composition of symmetries. Similar to permutations, if x and y are two elements in D_n, then $xy = y \circ x$. That means that xy is the symmetry obtained by applying first x, followed by y.

So, according to our discussion in Sects. 7.3 and 7.4, the elements of D_n can be thought 2×2 matrices, with group operation corresponding to matrix multiplication. Indeed, we have

$$R_k = \begin{bmatrix} \cos\left(\frac{2k\pi}{n}\right) & -\sin\left(\frac{2k\pi}{n}\right) \\ \sin\left(\frac{2k\pi}{n}\right) & \cos\left(\frac{2k\pi}{n}\right) \end{bmatrix},$$

$$S_k = \begin{bmatrix} \cos\left(\frac{2k\pi}{n}\right) & \sin\left(\frac{2k\pi}{n}\right) \\ \sin\left(\frac{2k\pi}{n}\right) & -\cos\left(\frac{2k\pi}{n}\right) \end{bmatrix}.$$

Now, it is not difficult to see that the following relations hold in D_n:

$$\begin{aligned} R_i R_j &= R_{i+j}, \\ R_i S_j &= S_{i+j}, \\ S_i R_j &= S_{i-j}, \\ S_i S_j &= R_{i-j}, \end{aligned}$$

where $0 \leq i, j \leq n-1$, and both $i+j$ and $i-j$ are computed modulo n. The Cayley table for D_n can be readily computed from the above relations. In particular, we see that R_0 is the identity, $R_i^{-1} = R_{n-i}$ and $S_i^{-1} = S_i$.

Example 7.70 In Example 3.39, we investigated D_3, the symmetry group of the equilateral triangle. The matrix representation of D_3 is given by

$$R_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad R_1 = \begin{bmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}, \quad R_2 = \begin{bmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{bmatrix},$$

$$S_0 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad S_1 = \begin{bmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}, \quad S_2 = \begin{bmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}.$$

Example 7.71 In Example 3.40, we investigated D_4, the symmetry group of a square. The matrix representation of D_4 is given by

$$R_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad R_1 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad R_2 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, \quad R_3 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix},$$

$$S_0 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad S_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad S_2 = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \quad S_3 = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}.$$

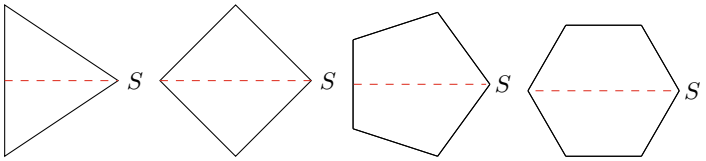


Fig. 7.14 Examples of polygons

Theorem 7.72 *Let R be a counterclockwise rotation of the n-gon by $2\pi/n$ radians and let S be a reflection across a line through a vertex. See examples in the polygons in Fig. 7.14. Then*

(1) The n rotations in D_n are I, R, R^2, ..., R^{n-1};
(2) $o(S) = 2$;
(3) The n reflections in D_n are S, RS, R^2S, ..., $R^{n-1}S$.

Proof (1) The rotations I, R, R^2, ..., R^{n-1} are distinct since R has order n.
(2) Simply consider what applying s twice to each vertex will do to it.
(3) Since I, R, R^2, ..., R^{n-1} are distinct, it follows that S, RS, R^2S, ..., $R^{n-1}S$ are distinct. In addition, for each j, R^jS is not a rotation because if $R^jS = R^i$, then $S = R^{i-j}$, while S is not a rotation. ■

Corollary 7.73 *The dihedral group D_n has $2n$ elements and*

$$D_n = \{I,\ R,\ R^2,\ \ldots,\ R^{n-1},\ S,\ RS,\ R^2S,\ \ldots,\ R^{n-1}S\}.$$

In particular, all elements of D_n with order greater than 2 are powers of R.

Corollary 7.74 *The dihedral group D_n is generated by R and S.*

Proof Since every element of D_n is a product of R and S, by Theorem 4.40 it follows that $\{R,\ S\}$ generate D_n. ■

Theorem 7.75 *In the dihedral group D_n, we have $RS = SR^{-1}$.*

Proof Since RS is a reflection, it follows that $(RS)^2 = I$, or equivalently $RSRS = I$. This implies that $RS = S^{-1}R^{-1}$. Since $o(S) = 2$, it follows that $S = S^{-1}$ and so we obtain $RS = SR^{-1}$. ■

Corollary 7.76 *For each $1 \leq j \leq n$, we have $R^kS = SR^{-k}$.*

Proof It is straightforward by mathematical induction. ■

Corollary 7.77 *The dihedral group D_n is not abelian if $n \geq 3$.*

Theorem 7.78 *For each $n \geq 3$, the center of the dihedral group D_n is*

(1) $Z(D_n) = \{I\}$ if n is odd;

(2) $Z(D_n) = \{I,\ R^{n/2}\}$ *if* n *is even.*

Proof In order to determine the center of D_n it suffices to determine those elements which commute with the generators R and S. Since $n \geq 3$, it follows that $R^{-1} \neq R$.

If $R^k S \in Z(D_n)$ for some $1 \leq k \leq n$, then $R(R^k S) = (R^k S)R$. Then, we must have

$$R^{k+1} S = R(R^k S) = (R^k S)R = R^k (SR) = R^k (R^{-1} S) = R^{k-1} S.$$

This implies that $R^2 = I$ a contradiction. Thus, we conclude that no reflections are in the center of D_n since reflections do not commute with R.

Similarly, if for $1 \leq j \leq n$, $R^j S = SR^j$, then $R^j S = R^{-j} S$. Hence, $R^{2j} = I$. Since R has order n, it follows that

$$R^{2j} = I \ \Leftrightarrow\ n | 2j.$$

If n is odd, then

$$R^{2j} = I \Leftrightarrow\ n | j \ \Leftrightarrow\ j \text{ is multiple of } n \ \Leftrightarrow\ R^j = I.$$

So, if n is odd, then the only rotation that could be in $Z(D_n)$ is I.

If n is even, then

$$R^{2j} = I \ \Leftrightarrow\ n | 2j \ \Leftrightarrow (n/2) | j.$$

Hence, the only choice for j are $j = 0$ and $j = (n/2)$. So, we have

$$R^j = R^0 = I \text{ or } R^j = R^{n/2}.$$

To show that $R^{n/2}$ is in $Z(D_n)$, we check it commutes with every rotations and reflections in D_n. Clearly, $R^{n/2}$ commutes with all rotations, since all rotations are powers of R. Hence, we check $R^{n/2}$ commutes with each reflections. Indeed, we have

$$R^{n/2}(R^i S) = R^{(n/2)+i} S \tag{7.8}$$

and

$$\begin{aligned} (R^i S)R^{n/2} &= R^i (SR^{n/2}) = R^i R^{-n/2} S \\ &= R^i R^{n/2} S = R^{i+(n/2)} S = R^{(n/2)+i} S. \end{aligned} \tag{7.9}$$

From (7.8) and (7.9), we obtain $R^{n/2}(R^i S) = (R^i S)R^{n/2}$.

Therefore, we conclude that

$$Z(D_n) = \begin{cases} \{I\} & \text{if } n \text{ is odd} \\ \{I,\ R^{n/2}\} & \text{if } n \text{ is even.} \end{cases}$$

■

Theorem 7.79 *The conjugacy classes of the dihedral group D_n are as follows:*

(1) If n is odd,

- *the identity element:* $\{I\}$,
- $(n-1)/2$ *conjugacy classes of size 2:* $\{R,\ R^{-1}\}$, $\{R^2,\ R^{-2}\}$, ..., $\{R^{(n-1)/2}, R^{-(n-1)/2}\}$,
- *all the reflections:* $\{R,\ RS,\ R^2S,\ \ldots,\ R^{n-1}S\}$.

(2) If n is even,

- *two conjugacy classes of size* 1: $\{I\}$, $\{R^{n/2}\}$,
- $(n/2)-1$ *conjugacy classes of size 2:* $\{R,\ R^{-1}\}$, $\{R^2,\ R^{-2}\}$, ..., $\{R^{(n/2)-1}, R^{-((n/2)-1)}\}$,
- *the reflection divided into two conjugacy classes:* $\{R^{2k}S \mid 0 \le k \le (n/2)-1\}$ *and* $\{R^{2k+1}S \mid 0 \le k \le (n/2)-1\}$.

Proof We know that every element of D_n is of the form R^i or R^iS for some integer i. Thus, in order to determine the conjugacy class of an element X, we compute

$$R^{-i}XR \text{ and } (R^iS)^{-1}X(R^iS).$$

Since $R^{-i}R^jR^i = R^j$ and $(R^iS)^{-1}R^j(R^iS) = R^{-j}$, it follows that the only conjugates of R^j in D_n are R^j and R^{-j}. It is necessary the more computation to be sure nothing more is conjugate as well. Now, we try to find the conjugacy class of S. We have

$$\begin{aligned} R^{-i}SR^i &= R^{n-2i}S, \\ (R^iS)^{-1}S(R^iS) &= R^{2i}S. \end{aligned}$$

Since $1 \le i \le n$, it follows that R^{2i} and R^{n-2i} run through powers of R divisible by 2.

Let n be odd. Since 2 is invertible modulo n, it follows that we can solve the linear congruence equation $2i \equiv k \pmod{n}$ for each i. This yields that

$$\{R^{2i}S \mid i \in \mathbb{Z}\} = \{R^kS \mid k \in \mathbb{Z}\}.$$

Hence, every reflections in D_n is conjugate to S.

Now, let n be even. We only obtain half the reflections as conjugates of S. The other half are conjugate to RS. Indeed, we have

$$\begin{aligned} R^{-i}(RS)R^i &= R^{-2i+1}S, \\ (R^iS)^{-1}(RS)(R^iS) &= R^{2i-1}S. \end{aligned}$$

Since $1 \le i \le n$, this gives us $\{RS,\ R^3S,\ \ldots,\ R^{n-1}S\}$. ■

Example 7.80 *An application of D_4.* One application of D_4 is in the design of a letter-facing machine. Imagine letters entering a conveyor belt to be postmarked.

They are placed on the conveyor belt at random so that two sides are parallel to the belt. Suppose that a postmarker can recognize a stamp in the top right corner of the envelope, on the side facing up. In Fig. 7.15, a sequence of machines is shown that will recognize a stamp on any letter, no matter what position in which the letter starts. The letter P stands for a postmarker. The letters R and S stand for rotating and flipping machines that perform the motions of R and S. The arrows pointing up indicate that if a letter is postmarked, it is taken off the conveyor belt for delivery. If a letter reaches the end, it must not have a stamp. Letter-facing machines like this have been designed. One economic consideration is that R-machines tend to cost more than S-machines. R-machines also tend to damage more letters. Taking these facts into consideration, the reader is invited to design a better letter-facing machine. Assume that R-machines cost \$1000 and S-machines cost \$750. Be sure that all corners of incoming letters will be examined as they go down the conveyor belt.

Exercises

1. Show that D_4 is non-abelian group of order 8, where each element of D_4 is of the form $a^i b^j$, $0 \leq i \leq 3$ and $0 \leq j \leq 1$.
2. Draw the Hasse diagram for subgroups of the dihedral group D_4.
3. In the dihedral group D_n, suppose that R is a rotation and that S is a reflection. Use the fact that RS is also a reflection, together with the fact that reflections have order 2, to show that RSR is the inverse of S.
4. For each of the snowflakes in Fig. 7.16 find the symmetry group and locate the axes of the reflective symmetry.
5. Find the symmetry group of the Iranian architecture in Fig. 7.17.
6. Design a better letter-facing machine. How can you verify that a letter facing machine does indeed check every corner of a letter? Can it be done on paper without actually sending letters through it?
7. Let S be a reflection in the dihedral group D_n and R be a rotation in D_n. Determine
   (a) $C_{D_n}(S)$ when n is odd;
   (b) $C_{D_n}(S)$ when n is even;
   (c) $C_{D_n}(R)$.

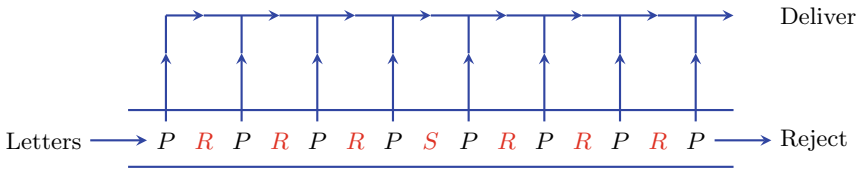


Fig. 7.15 A letter facer

Fig. 7.16 Snowflakes

7.7 Quaternion Group

Consider the case $\mathbb{F} = \mathbb{C}$, the complex numbers, and the set of eight elements

$$Q_8 = \left\{ \pm \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \pm \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, \pm \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \pm \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} \right\}.$$

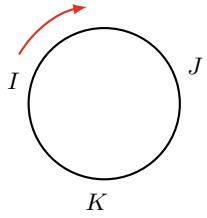
One can use the notation

$$I = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, \; J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \text{ and } K = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}.$$

To avoid confusion, we may show the identity matrix by 1 instead of I_2. Then, we obtain

Fig. 7.17 Iranian architecture

Fig. 7.18 The rules involving I, J, and K in the quaternion group



$$
\begin{aligned}
&I^2 = J^2 = K^2 = -1,\\
&IJ = -JI = K,\\
&JK = -JK = I,\\
&KI = -IK = J.
\end{aligned}
$$

So, $Q_8 = \{\pm 1,\ \pm I,\ \pm J,\ \pm K\}$.

The rules involving I, J and K can be remembered by using Fig. 7.18.

Going clockwise, the product of two consecutive elements is the third one. The same is true for going counterclockwise, except that we obtain the negative of the third element. This group was invented by William Hamilton in 1834. The quaternions are used to describe rotations in three-dimensional space, and they are used in physics. The quaternions can be used to extend the complex numbers in a natural way.

Theorem 7.81 *Q_8 is a subgroup of $GL_2(\mathbb{C})$ and so in particular Q_8 is a group of order* 8.

Proof Clearly, Q_8 is non-empty. By the above relations, it is closed under multiplication. In addition, ± 1 are their own inverse, and all the other elements have an

inverse which is minus themselves (since $I(-I) = 1$, etc.). Therefore, Q_8 is closed under inverse. ■

Q_8 is called *quaternion group*. The quaternion group is not abelian. Its Cayley table is shown below.

$\cdot$	1	-1	I	$-I$	J	$-J$	K	$-K$
1	1	-1	I	$-I$	J	$-J$	K	$-K$
-1	-1	1	$-I$	I	$-J$	J	$-K$	K
I	I	I	-1	1	K	$-K$	$-J$	J
$-I$	$-I$	I	1	-1	$-K$	K	J	$-J$
J	J	$-J$	$-K$	K	-1	1	I	$-I$
$-J$	$-J$	J	K	$-K$	1	-1	$-I$	I
K	K	$-K$	J	$-J$	$-I$	I	-1	1
$-K$	$-K$	K	$-J$	J	I	$-I$	1	-1

Theorem 7.82 *The quaternion group Q_8 is generated by J and K.*

Proof We observe that each element of Q_8 is of the form $J^r K^s$ for some integers r and s. ■

Remark 7.83 Note that I, J and K have order 4 and that any two of them generate the entire group.

The proper subgroups of Q_8 are $\langle I\rangle$, $\langle J\rangle$, $\langle K\rangle$ and $\langle -1\rangle$. We have

$$\langle -1\rangle = \langle I\rangle \cap \langle J\rangle \cap \langle K\rangle$$

and the center of Q_8 is $\langle -1\rangle$.

Exercises

1. Show that Q_8 has exactly one element of order 2 and six elements of order 4.
2. Draw the Hasse diagram for subgroups of the quaternion group Q_8.

7.8 Worked-Out Problems

Problem 7.84 Let n be a positive integer and

$$A = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

Prove that $A^n = I_3$ if and only if $4|n$.

Solution First, we compute A^2, A^3, A^4, etc. We find that

$$A^2 = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix},$$

$$A^3 = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix},$$

$$A^4 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I_3.$$

Now, if $4|n$, then we can write $n = 4q$. Hence, $A^n = A^{4q} = (A^4)^q = I_3^q = I_3$.

Conversely, if $A^n = I_3$, then by the division algorithm, there exist integers q and r such that $n = 4q + r$ with $0 \leq r < 4$. Next, we obtain

$$A^r = A^{n-4q} = A^n(A^{-4})^q = I_3 I_3^q = I_3.$$

Since A, A^2 and A^3 are not equal to I_3, it follows that $r = 0$. Consequently, we obtain $4|n$. ■

Problem 7.85 Let n be a positive integer and

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Compute A^n.

Solution Suppose that

$$B = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Then, we observe that $B^n = 0$, for each $n \geq 3$. Note that $A = I_3 + B$. Since I_3 and B commute, we can use the binomial theorem. So, we can write

$$A^n = (I_3 + B)^n = \sum_{i=0}^{n} \binom{n}{i} I_3^{n-i} B^i.$$

Therefore, we obtain

$$A^n = I_3 + nB + \frac{n(n-1)}{2}B^2 = \begin{bmatrix} 0 & n & \frac{n(n-1)}{2} \\ 0 & 1 & n \\ 0 & 0 & 1 \end{bmatrix},$$

and we are done. ■

7.9 Supplementary Exercises

1. Prove that the elements of $GL_n(\mathbb{R})$ which have integer entries and determinant equal to 1 or -1 form a subgroup of $GL_n(\mathbb{R})$.
2. Let $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix} \in GL_2(\mathbb{Z})$. Show that A has order 4, B has order 3 and AB has infinite order.
3. The matrix
$$\begin{bmatrix} 1 & t_0 & t_0^2 & \dots & t_0^n \\ 1 & t_1 & t_1^2 & \dots & t_1^n \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & t_n & t_n^2 & \dots & t_n^n \end{bmatrix}$$
is called a *Vandermonde matrix*. Show that such a matrix is invertible, when t_0, $t_1, \dots, t_n$ are $n+1$ distinct elements of $\mathbb{C}$.
4. Show by example that there are matrices A and B for which $\lim_{n\to\infty} A^n$ and $\lim_{n\to\infty} B^n$ both exist, but for which $\lim_{n\to\infty} (AB)^n$ does not exist.
5. Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be a matrix over a field $\mathbb{F}$. Then, the set of all matrices of the form $p(A)$, where p is a polynomial over $\mathbb{F}$, is a commutative ring R with identity. If B is a 2×2 matrix over R, the determinant of B is then a 2×2 matrix over $\mathbb{F}$, of the form $p(A)$. Suppose that B is the 2×2 matrix over R as follows:
$$\begin{bmatrix} A - aI_2 & -bI_2 \\ -cI_2 & A - dI_2 \end{bmatrix}.$$
Show that $\det(B) = p(A)$, where
$$p(x) = x^2 - (a+d)x + \det(A), \tag{7.10}$$
and also that $p(A) = 0$. The polynomial p in (7.10) is the characteristic polynomial of A.
6. Given a matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ in $GL_2(\mathbb{Z}_p)$. Consider its characteristic polynomial.

(a) If the roots λ_1 and λ_2 are distinct in $\mathbb{Z}_p$, show that A is conjugate to both $\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$ and $\begin{bmatrix} \lambda_2 & 0 \\ 0 & \lambda_1 \end{bmatrix}$.

(b) If $\lambda_1 = \lambda_2$, show that A is $\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_1 \end{bmatrix}$ or it is conjugate to $\begin{bmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{bmatrix}$ in $GL_2(\mathbb{Z}_p)$.

7. Let

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

be a rotation matrix in $\mathbb{R}^3$. Show that the rotation axis is the vector

$$V = \frac{1}{\sin\theta} \begin{bmatrix} r_{32} - r_{23} \\ r_{13} - r_{31} \\ r_{21} - r_{12} \end{bmatrix}.$$

If the angle of rotation θ is different from π and 0, then

$$\theta = \cos^{-1}\left(\frac{r_{11} + r_{22} + r_{33} - 1}{2}\right).$$

What happens if the rotation angle is very small? Describe a robust method to find the rotation axis (your method should include the cases for $\theta = 0$ and $\theta = \pi$).