Summer Programme 2021

Solutions to exercises from Michael Artin's *Algebra*

Satvik Saha
19MS154

*Indian Institute of Science Education and Research, Kolkata,
Mohanpur, West Bengal, 741246, India.*

Chapter 1

Matrix Operations

1.4 Permutation Matrices

Exercise 1. Consider the permutation p defined by $1 \rightsquigarrow 3$, $2 \rightsquigarrow 1$, $3 \rightsquigarrow 4$, $4 \rightsquigarrow 2$.

(a) Find the associated permutation matrix P.

(b) Write p as a product of transpositions and evaluate the corresponding matrix product.

(c) Compute the sign of p.

Solution.

(a) The column P_i must be the standard basis vector $\boldsymbol{e}_{p(i)}$, so

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

(b) Check that $p = (1,3,4,2) = (1,2)(1,4)(1,3)$. This product is given by

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

(c) Since p is the product of an odd number of transpositions, its sign is -1. This is verified by calculating the determinant

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

Exercise 2. Prove that every permutation matrix is a product of transpositions.

Solution. Note that this is equivalent to stating that any ordered list can be sorted using transpositions.

The statement is trivially true for all 2×2 permutation matrices,

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

the first being the identity and the second being a transposition itself. Suppose that any $n \times n$ permutation matrix is the product of transpositions. Use the fact that for square matrices A and B,

$$\begin{bmatrix} A & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} B & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} AB & 0 \\ 0 & 1 \end{bmatrix},$$

which means that if a permutation matrix $P = E_1 E_2 \dots E_k$, then

$$\begin{bmatrix} P & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} E_1 & 0 \\ 0 & 1 \end{bmatrix} \cdots \begin{bmatrix} E_k & 0 \\ 0 & 1 \end{bmatrix}.$$

Now let Q be an arbitrary $(n+1) \times (n+1)$ permutation matrix. Let j be the index of the row of Q which is precisely $(0 \dots 0\, 1)$, and let E be the transposition matrix which interchanges the rows $j \leftrightarrow n+1$. Then,

$$EQ = \begin{bmatrix} Q' & 0 \\ 0 & 1 \end{bmatrix},$$

where Q' is an $n \times n$ permutation matrix. This is because Q' has exactly one 1 in each row and column, the remaining elements being 0. Multiply both sides by E, and use the fact that $E^2 = \mathbb{I}$. Now, Q' is a product of transpositions $E_1 \dots E_k$, so we finally have

$$Q = E \begin{bmatrix} Q' & 0 \\ 0 & 1 \end{bmatrix} = E \begin{bmatrix} E_1 & 0 \\ 0 & 1 \end{bmatrix} \cdots \begin{bmatrix} E_k & 0 \\ 0 & 1 \end{bmatrix}.$$

Exercise 3. Prove that every matrix with a single 1 in each row and a single 1 in each column, the other entries being zero, is a permutation matrix.

Solution. Note that each column of such a matrix P must be a distinct standard basis vector $\boldsymbol{e}_k$, and we claim that this matrix represents the permutation p defined as $p(j) = k$, where $P_j = \boldsymbol{e}_k$ is the j^{th} column of P. Now, p is a bijection because every column j has one and exactly one 1 in the k^{th} row. This justifies that p is indeed a permutation. When P acts on a column vector $\boldsymbol{x}$, we have

$$P\boldsymbol{x} = P_1 x_1 + P_2 x_2 + \cdots + P_n x_n = \boldsymbol{e}_{p(1)} x_1 + \boldsymbol{e}_{p(2)} x_2 + \cdots + \boldsymbol{e}_{p(n)} x_n.$$

This means that

$$P \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_{p^{-1}(1)} \\ x_{p^{-1}(2)} \\ \vdots \\ x_{p^{-1}(n)} \end{bmatrix}.$$

Exercise 4. Let p be a permutation. Prove that $\operatorname{sign} p = \operatorname{sign} p^{-1}$.

Solution. This follows directly from the fact that $\det P^{-1} = 1/\det P$, and that $\det P = \pm 1$ so $\det P^{-1} = \det P$.

Exercise 5. Prove that the transpose of a permutation matrix P is its inverse.

Solution. Recall that $\det P = \pm 1$, so P is invertible. Write the permutation matrix P in terms of its columns,

$$P = \begin{bmatrix} | & | & \cdots & | \\ \boldsymbol{e}_{p(1)} & \boldsymbol{e}_{p(2)} & \cdots & \boldsymbol{e}_{p(n)} \\ | & | & \cdots & | \end{bmatrix},$$

where p represents the corresponding permutation. Now note that the transpose can be written as

$$P = \begin{bmatrix} — \boldsymbol{e}_{p(1)}^t — \\ — \boldsymbol{e}_{p(2)}^t — \\ \vdots \\ — \boldsymbol{e}_{p(n)}^t — \end{bmatrix}.$$

Therefore, the ij^{th} element of the product P^tP is given by $\boldsymbol{e}_{p(i)}^t \boldsymbol{e}_{p(j)} = \delta_{p(i)p(j)} = \delta_{ij}$, meaning that $P^tP = \mathbb{I}$. We have used the fact that p is a bijection, so $p(i) = p(j)$ if and only if $i = j$. Thus, $P^{-1} = P^t$.

1.5 Cramer's Rule

Exercise 3. Let A be an $n \times n$ matrix with integer entries a_{ij}. Prove that A^{-1} has integer entries if and only if $\det A = \pm 1$.

Solution. First, suppose that $\det A = \pm 1$. If the entries of A^{-1} are b_{ij}, use

$$A^{-1} = \frac{1}{\det A} \operatorname{adj} A$$

to conclude that

$$b_{ij} = \frac{1}{\det A} (-1)^{i+j} \det A_{ji}.$$

Note that A_{ji} contains integer entries, hence its determinant must also be an integer via the complete expansion. Putting $\det A = \pm 1$ means that b_{ij} is always an integer.

Now suppose that A^{-1} has integer entries. Use $\det A = 1/\det A^{-1}$. Now both A and A^{-1} have integer entries, hence integer determinants, with $|\det A^{-1}| \geq 1$. This forces $\det A = \pm 1$.

Miscellaneous Problems

Exercise 2. Find a representation of the complex numbers by real 2 x 2 matrices which is compatible with addition and multiplication.

Solution. Consider the representation

$$z = a + ib \equiv \begin{bmatrix} a & -b \\ b & a \end{bmatrix}.$$

Now, if $z = a + ib$, $w = c + id$, we have addition defined as

$$z + w \equiv \begin{bmatrix} a & -b \\ b & a \end{bmatrix} + \begin{bmatrix} c & -d \\ d & c \end{bmatrix} = \begin{bmatrix} a+c & -b-d \\ b+d & a+c \end{bmatrix} \equiv (a+c) + i(b+d),$$

and multiplication as

$$zw \equiv \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \begin{bmatrix} c & -d \\ d & c \end{bmatrix} = \begin{bmatrix} ac-bd & -ad-bc \\ ad+bc & ac-bd \end{bmatrix} \equiv (ac-bd) + i(ad+bc).$$

Finally,

$$|z|^2 = z\overline{z} = a^2 + b^2 = \det \begin{bmatrix} a & -b \\ b & a \end{bmatrix}.$$

Exercise 3. Find the Vandermonde determinant

$$\det A_n = \det \begin{bmatrix} 1 & 1 & \cdots & 1 \\ a_1 & a_2 & \cdots & a_n \\ a_1^2 & a_2^2 & \cdots & a_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1^{n-1} & a_2^{n-1} & \cdots & a_n^{n-1} \end{bmatrix}.$$

Solution. First look at the 2×2 case,

$$\det A_2 = \det \begin{bmatrix} 1 & 1 \\ a_1 & a_2 \end{bmatrix} = a_2 - a_1.$$

Now, look at the $n \times n$ case. Perform the row operations $R_k \to R_k - a_1 R_{k-1}$ for all rows $k = 2, \ldots, n$. This leaves the determinant unchanged, so

$$\det A_2 = \det \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & a_2 - a_1 & a_3 - a_1 & \cdots & a_n - a_1 \\ 0 & a_2(a_2 - a_1) & a_3(a_3 - a_1) & \cdots & a_n(a_n - a_1) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & a_2^{n-2}(a_2 - a_1) & a_3^{n-2}(a_3 - a_1) & \cdots & a_n^{n-2}(a_n - a_1) \end{bmatrix}.$$

Using expansion by minors on the first column, we have

$$\det A_2 = \det \begin{bmatrix} a_2 - a_1 & a_3 - a_1 & \cdots & a_n - a_1 \\ a_2(a_2 - a_1) & a_3(a_3 - a_1) & \cdots & a_n(a_n - a_1) \\ \vdots & \vdots & \ddots & \vdots \\ a_2^{n-2}(a_2 - a_1) & a_3^{n-2}(a_3 - a_1) & \cdots & a_n^{n-2}(a_n - a_1) \end{bmatrix}.$$

Factoring out $a_j - a_1$ from each j^{th} column gives

$$\det A_n = \prod_{j=2}^{n} (a_j - a_1) \times \det \begin{bmatrix} 1 & 1 & \cdots & 1 \\ a_2 & a_3 & \cdots & a_n \\ \vdots & \vdots & \ddots & \vdots \\ a_2^{n-2} & a_3^{n-2} & \cdots & a_n^{n-2} \end{bmatrix}$$

Continuing in this fashion, we get

$$\det A_n = \prod_{j=2}^{n} (a_j - a_1) \times \prod_{j=3}^{n} (a_j - a_2) \times \cdots \times (a_{n-1} - a_n).$$

This can be written down concisely as

$$\det A_n = \prod_{1 \le i < j \le n}^{n} (a_j - a_i)$$

Exercise 4. Consider a general system $AX = B$ of m linear equations in n unknowns. If the coefficient matrix A has a left inverse A', a matrix such that $A'A = \mathbb{I}_n$, then we may try to solve the system as follows.

$$\begin{aligned} AX &= B, \\ A'AX &= A'B \\ X &= A'B. \end{aligned}$$

But when we try to check our work by running the solution backward, we get into trouble:

$$\begin{aligned} X &= A'B \\ AX &= AA'B \\ AX &\stackrel{?}{=} B. \end{aligned}$$

We seem to want A' to be a right inverse: $AA' = \mathbb{I}_n$, which isn't what was given. Explain.

Solution. In the case that $m > n$, note that the left inverse is not necessarily unique. An example is

$$\begin{bmatrix} 1 & 0 & a \\ 0 & 1 & b \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbb{I}_2,$$

irrespective of a and b. Hence, $X = A'B$ is not unique, but rather is dependent on our choice of A'. If we had started with

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} p \\ q \\ r \end{bmatrix}$$

then we would have written

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 0 & a \\ 0 & 1 & b \end{bmatrix} \begin{bmatrix} p \\ q \\ r \end{bmatrix} = \begin{bmatrix} p + ar \\ q + br \end{bmatrix} = \begin{bmatrix} p \\ q \end{bmatrix} + r \begin{bmatrix} a \\ b \end{bmatrix}.$$

This means that the given argument is not sufficient to conclude $AA' = \mathbb{I}$.

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & a \\ 0 & 1 & b \end{bmatrix} = \begin{bmatrix} 1 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & 0 \end{bmatrix} \neq \mathbb{I}_3,$$

Note that this system is nonsense for $r \neq 0$ with no solutions, yet the left inverses A' do exist nonetheless. Here, $AX \neq B$ when $r \neq 0$, since

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} p + ar \\ q + ar \end{bmatrix} = \begin{bmatrix} p + ar \\ q + ar \\ 0 \end{bmatrix}.$$

In the case that $m < n$, A has no left inverse. Label the columns of A as A_i. Demanding $A'A = \mathbb{I}_n$ means that we want $A'A_i = \boldsymbol{e}_i$ for all $i = 1, \ldots, n$. Since the $m \times n$ matrix A has more columns than rows, its columns must be linearly dependent, so without loss of generality, write the first column A_1 as a non-trivial linear combination of the rest,

$$A_1 = a_2 A_2 + a_3 A_3 + \cdots + a_n A_n.$$

Multiplying by A' gives

$$A'A_1 = \boldsymbol{e}_1 = a_2 \boldsymbol{e}_2 + a_3 \boldsymbol{e}_3 + \cdots + a_n \boldsymbol{e}_n,$$

which is a contradiction since the basis vectors $\{\boldsymbol{e}_i\}$ are linearly independent.

In the case $m = n$, it is indeed true that A' is also a right inverse of A. Note that if A'' is a right inverse of A with $AA'' = \mathbb{I}_n$, then

$$A' = A'\mathbb{I}_n = A'(AA'') = (A'A)A'' = \mathbb{I}_n A'' = A''.$$

To justify that A'' exists, note that $A'A = \mathbb{I}_n$ gives $\det A' \det A = 1$, so $\det A \neq 0$. Thus, A has full rank and its range must be the full n dimensional vector space of column vectors. Multiplying by A, we have $AA'A = A$ or $(AA' - \mathbb{I}_n)A = 0$. Recall that the range of A is the entire vector space, so $(AA' - \mathbb{I}_n)\boldsymbol{x} = \mathbf{0}$ for all possible column vectors $\boldsymbol{x}$. This forces $AA' - \mathbb{I}_n = 0$, or $AA' = \mathbb{I}_n$.

Exercise 5.

(a) Let A be a real 2×2 matrix, and let A_1, A_2 be the rows of A. Let P be the parallelogram whose vertices are 0, A_1, A_2, $A_1 + A_2$. Prove that the area of P is the absolute value of the determinant $\det A$ by comparing the effect of and elementary row operation on the area and on $\det A$.

(b) Prove an analogous result for $n \times n$ matrices.

Solution.

(a) First note that

$$\det \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 1,$$

which is consistent with the fact that the area of a unit square is 1. Now let a_{ij} be the elements of A. Perform the row operation which multiplies the top row by a_{11}, i.e. $R_1 \to a_{11}R_1$. We have

$$\det \begin{bmatrix} a_{11} & 0 \\ 0 & 1 \end{bmatrix} = a_{11}.$$

Now perform $R_1 \to R_1 + a_{12}R_2$. This gives

$$\det \begin{bmatrix} a_{11} & a_{12} \\ 0 & 1 \end{bmatrix} = a_{11}.$$

Next, perform $R_2 \to (a_{11}a_{22} - a_{12}a_{21})R_2$. This gives

$$\det \begin{bmatrix} a_{11} & a_{12} \\ 0 & a_{11}a_{22} - a_{12}a_{21} \end{bmatrix} = (a_{11}a_{22} - a_{12}a_{21})a_{11}.$$

Next, perform $R_2 \to R_2 + a_{21}R_1$. This gives

$$\det \begin{bmatrix} a_{11} & a_{12} \\ a_{11}a_{21} & a_{11}a_{22} \end{bmatrix} = (a_{11}a_{22} - a_{12}a_{21})a_{11}.$$

Finally, perform $R_2 \to R_2/a_{11}$. This gives

$$\det A = \det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

Note that if $a_{11} = 0$, we could have interchanged the roles of a_{11} and a_{21} at the beginning by interchanging the rows of A. This would have given the same result, up to a sign which we are not interested in. If both a_{11} and a_{22} are zero, note that the two rows are linearly dependent, with one being a multiple of the other, so the parallelogram they form has zero area. Thus, we have shown that any matrix A representing a parallelogram with non-zero area can be obtained from the identity matrix $\mathbb{I}_2$ by performing elementary row operations.

Now, we consider the effect of these row operations on the area of a parallelogram with legs A_1 and A_2. Note that the operation $A_1 \to kA_1$ for some real scaling factor k has the effect of scaling the area by the same factor k. The operation of interchanging the legs A_1 and A_2 has no effect on the area. The operation $A_1 \to A_1 + kA_2$ also has no effect on the area, because this has the effect of linearly shearing the parallelogram, in a manner parallel to the other leg A_2 which remains fixed. Thus, when we performed our row operations in the square to reach our parallelogram, our area transformed in precisely the same way as the unsigned determinant, which means that

$$\operatorname{area} A_{\|} = |\det A| = |a_{11}a_{22} - a_{12}a_{21}|.$$

(b) We use the fact that any square matrix A with non-zero determinant can be written as the product of row operations acting on the identity matrix $\mathbb{I}_n$, which represents the unit hypercube of hypervolume 1. The Gauss-Jordan elimination algorithm can be used to extract these operations. We see that all scaling operations will scale the hypervolume in the same way, all transpositions have no effect on the hypervolume, and all additions of linear combinations of other rows also have no effect, since they correspond to successive shearing of the hyperparallelopiped along a direction parallel to another leg. Thus, the area of the hypercube transformed in the same way as the unsigned determinant of A, so

$$\operatorname{hypervolume} A_{\|} = |\det A|.$$

Note that we are not interested in matrices with zero determinant, because such a matrix is not of full rank, hence its rows are linearly dependent. Thus, one of the legs of the corresponding hyperparallelopiped can be sheared until it is parallel to another, which immediately gives a zero hypervolume.

Exercise 6. Most invertible matrices can be written as a product $A = LU$ of a lower triangular matrix L and an upper triangular matrix U, where in addition all diagonal entries of U are 1.

(a) Prove uniqueness, that is, prove that there is at most one way to write A as a product.

(b) Explain how to compute L and U when the matrix A is given.

(c) Show that every invertible matrix can be written as a product LPU, where L, U are as above and P is a permutation matrix.

Solution. We first show that the determinant of a triangular matrix is equal to the product of its diagonals. To see this, note that this holds for all 2×2 lower triangular matrices,

$$\det \begin{bmatrix} a & 0 \\ c & d \end{bmatrix} = ad.$$

Next, suppose that this holds for all $n \times n$ lower triangular matrices. Using expansion of minors along the first row and our induction hypothesis on the minor A_{11}, compute

$$\det \begin{bmatrix} a_{11} & 0 & 0 & \cdots & a_{1n} \\ a_{21} & a_{22} & 0 & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{bmatrix} = a_{11} \det A_{11} + 0 + 0 + \cdots + 0 = a_{11}a_{22}a_{33}\cdots a_{nn}.$$

For upper triangular matrices, simply note that $\det U = \det U^t$, and U^t is lower triangular.

Now, we show that the inverse of a triangular matrix is also triangular of the same kind. Note that this holds for all invertible 2×2 matrices, with

$$\begin{bmatrix} a & 0 \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad} \begin{bmatrix} d & 0 \\ -c & a \end{bmatrix}.$$

Next, suppose that an invertible lower triangular matrix L has an inverse L^{-1}, whose columns are labelled $\boldsymbol{x}_j$. Since $LL^{-1} = \mathbb{I}_n$, we want

$$L\boldsymbol{x}_j = \boldsymbol{e}_j.$$

We claim that $(x_j)_i = 0$ for all $i < j$. To see this, note that the first $j-1$ rows expand to

$$\begin{aligned} 0 &= L_{11}x_{j1} \\ 0 &= L_{12}x_{j1} + L_{22}x_{j2} \\ \vdots \quad &\vdots \\ 0 &= L_{j-1,1}x_{j,j-1} + \cdots + L_{j-1,j-1}x_{j,j-1} \end{aligned}$$

All L_{ij} with $i < j$ are zero, and L_{ii} are non-zero since L is invertible hence $\det L \neq 0$. Thus, the first equation gives $x_{j1} = 0$, which when plugged into the second gives $x_{j2} = 0$, and so on up to $x_{j,j-1} = 0$. Hence, $L^{-1}_{ij} = 0$ for all $i < j$, making it a lower triangular matrix. In addition, the j^{th} row reads

$$1 = L_{j1}x_{j1} + \cdots + L_{j,j-1}x_{j,j-1} + L_{jj}x_{jj}.$$

All terms but the last one are 0, so the diagonal elements satisfy $L_{jj}L^{-1}_{jj} = 1$. Like before, for a lower triangular matrix U, use $(U^t)^{-1} = (U^{-1})^t$.

Finally, the product of two triangular matrices of the same kind give another triangular matrix of the same kind. Suppose that A and B are two lower triangular matrices. The ij^{th} element of their product AB is

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

Now, $a_{ik} = 0$ for all $i < k$ and $b_{ki} = 0$ for all $k < j$. Thus, when $i < j$, we have $c_{ij} = 0$, hence AB is also lower triangular. Again for upper triangular matrices X, Y, use $(XY)^t = Y^tX^t$.

(a) Suppose that $A = LU = L'U'$ are two LU decompositions of A. Note that $\det A \neq 0$ from its invertibility, hence L, U, L', U' are all invertible. This gives

$$L^{-1}LU = L^{-1}L'U', \qquad U = L^{-1}L'U', \qquad U(U')^{-1} = L^{-1}L'.$$

Now, the left side is upper triangular while the right side is left triangular. Also, the left side has all 1's along its diagonal. This forces

$$U(U')^{-1} = \mathbb{I}_n = L^{-1}L', \qquad U = U', \quad L = L'.$$

(b) The elements of L and U can be obtained by brute force, solving the system $A = LU$ with $n(n+1)/2 + (n-1)n/2 = n^2$ unknowns.

(c) Note that after performing Gaussian elimination on an invertible matrix A, we are left with an upper triangular matrix U with 1's along its diagonal. Also, each elementary operation we performed can be represented by a lower triangular matrix. This is because all scaling matrices are diagonal, and in all cases where we added one row to another we always added higher row to ones lower down. Thus, the product of all these elementary matrices is a lower triangular matrix L, which means $LA = U$. This gives the desired decomposition, $A = L^{-1}U$.

However, we may have to exchange rows while performing the elimination process, which happens when one of the diagonal elements becomes zero. By performing this permutation of rows at the very end, we have actually decomposed $PLA = U$. The inverse of a permutation is another permutation, hence we have the desired decomposition $A = L^{-1}P^{-1}U$.

Exercise 7. Consider a system of n linear equations in n unknowns: $AX = B$, where A and B have *integer* entries. Prove or disprove the following.

(a) The system has a rational solution if $\det A \neq 0$.

(b) If the system has a rational solution, then it also has an integer solution.

Solution.

(a) If $\det A \neq 0$, then A is invertible. Since A has integer entries, its determinant is an integer and its adjoint has integer entries, which means that $A^{-1} = (\operatorname{adj} A)/\det A$ has rational entries. Also, B has integer entries so the solution $X = A^{-1}B$ must also be rational.

(b) This is false. Consider the system

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

This has the unique solution $x = y = \frac{1}{2}$.

Exercise 8. Let A, B be $m \times n$ and $n \times m$ matrices. Prove that $\mathbb{I}_m - AB$ is invertible if and only if $\mathbb{I}_n - BA$ is invertible.

Solution. Note that

$$B(\mathbb{I}_m - AB) = B - BAB = (\mathbb{I}_n - BA)B,$$
$$A(\mathbb{I}_n - BA) = A - ABA = (\mathbb{I}_m - AB)A.$$

Set $X = \mathbb{I}_m - AB$, $Y = \mathbb{I}_n - BA$, whence

$$BX = YB, \qquad AY = XA.$$

First suppose that X is invertible. If A is invertible, then $AY = XA$ gives $Y = A^{-1}XA$, so we can check that $Y^{-1} = A^{-1}X^{-1}A$.

$$(A^{-1}X^{-1}A)Y = A^{-1}X^{-1}A\,A^{-1}XA = \mathbb{I}_n.$$

If A is not invertible but B is invertible, then use $BX = YB$ to write $Y = BXB^{-1}$, so we can check that $Y^{-1} = BX^{-1}B^{-1}$.

$$(BX^{-1}B^{-1})Y = BX^{-1}B^{-1}\,BXB^{-1} = \mathbb{I}_n.$$

Now suppose that neither A nor B is invertible. Consider the products

$$(\mathbb{I}_n + BX^{-1}A)Y = Y + BX^{-1}AY = Y + BX^{-1}XA = Y + BA = \mathbb{I}_n,$$

$$Y(\mathbb{I}_n + BX^{-1}A) = Y + YBX^{-1}A = Y + BXX^{-1}A = Y + BA = \mathbb{I}_n.$$

Thus, $Y^{-1} = \mathbb{I}_n + BX^{-1}A$.

Chapter 2

Groups

2.1 The Definition of a Group

Exercise 1.

(a) Verify (1.17) and (1.18) by explicit computation.

(b) Make a multiplication table for S_3.

Solution.

(a) We see that

$$1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad x = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \qquad y = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Calculate

$$\begin{aligned} x^2 &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \\ xy &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \\ x^2y &= \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}. \end{aligned}$$

These six matrices cover all possible permutations of the three rows.

(b)

$\times$	1	x	x^2	y	xy	x^2y
1	1	x	x^2	y	xy	x^2y
x	x	x^2	1	xy	x^2y	y
x^2	x^2	1	x	x^2y	y	xy
y	y	x^2y	xy	1	x^2	x
xy	xy	y	x^2y	x	1	x^2
x^2y	x^2y	xy	y	x^2	x	1

Exercise 2.

(a) Prove that $GL_n(\mathbb{R})$ is a group.

(b) Prove that S_n is a group.

Solution.

(a) We show that $GL_n(\mathbb{R})$ is a group under matrix multiplication. Note that if $\det A \neq 0$ and $\det B \neq 0$, then $\det AB = \det A \det B \neq 0$, so $GL_n(\mathbb{R})$ is closed under multiplication. This composition is also associative, by virtue of the associativity of matrix multiplication. The identity matrix $\mathbb{I}_n$ serves as the group identity, since $\mathbb{I}_n A = \mathbb{I}_n = A\mathbb{I}_n$ for all $A \in GL_n(\mathbb{R})$. Finally, all non-singular matrices are invertible, with the inverse also being non-singular, which means that every element $A \in GL_n(\mathbb{R})$ has an inverse $A^{-1} \in GL_n(\mathbb{R})$. Thus, $GL_n(\mathbb{R})$ forms a group.

(b) We show that S_n is a group under function composition. Note that each element S_n is a bijection $f\colon \{1,\ldots,n\} \to \{1,\ldots,n\}$. Since the composition of two bijections is also a bijection, we see that S_n is closed under composition. Also note that function composition is associative, with $(f \circ g) \circ h = g \circ (g \circ h)$. The identity map $\mathbb{I} \in S_n$ which maps each integer to itself serves as the identity element, since $\mathbb{I} \circ f = \mathbb{I} = f \circ \mathbb{I}$ for all $f \in S_n$. Finally, all bijections have an inverse which is also a bijection, hence every element $f \in S_n$ has n inverse $f^{-1} \in S_n$. Thus, S_n forms a group.

Exercise 3. Let S be a set with an associative law of composition and with an identity element. Prove that the subset of S consisting of invertible elements is a group.

Solution. Let $S' \subset S$ be the set of all invertible elements of S. Note that by construction, composition in S' is associative. The identity element $e \in S$ must also belong to S', since $ee = e$, hence e is invertible with $e^{-1} = e$. This serves as an identity for all elements in S' as well. Also, all elements in S' are invertible, and their inverses must also be in S', because if $a^{-1} = b$, then $ba^{-1} = e = a^{-1}b$, so $b^{-1} = a$ making $b = a^{-1}$ invertible. Finally, S' is closed under composition, because the product of invertible elements is invertible, with $(ab)^{-1} = b^{-1}a^{-1}$. This means that S' is a group.

Exercise 4. Solve for y, given that $xyz^{-1}w = 1$ in a group.

Solution. Write

$$\begin{aligned} xyz^{-1}w &= 1, \\ xyz^{-1}ww^{-1} &= w^{-1}, \\ xyz^{-1} &= w^{-1}, \\ xyz^{-1}z &= w^{-1}z, \\ xy &= w^{-1}z, \\ x^{-1}xy &= x^{-1}w^{-1}z, \\ y &= x^{-1}w^{-1}z. \end{aligned}$$

Exercise 5. Assume that the equation $xyz = 1$ holds in a group G. Does it follow that $yzx = 1$? That $yxz = 1$?

Solution. We have $xyz = 1$, so

$$1 = x^{-1}x = x^{-1}1x = x^{-1}(xyz)x = (x^{-1}x)yzx = yzx.$$

It is not necessarily true that $yxz = 1$. Consider the group S_3 as seen in Exercise 1. We have $(xy)(x)(y) = 1$, but $(x)(xy)(y) = x^2y^2 = x^2 \neq 1$.

Exercise 6. Write out all ways in which one can form a product of four elements a, b, c, d in the given order.

Solution.

$$(ab)(cd) \quad (a(bc))d \quad ((ab)c)d \quad a(b(cd)) \quad a((bc)d)$$

Exercise 7. Let S be any set. Prove that the law of composition defined by $ab = a$ is associative.

Solution. It is sufficient to show that $(ab)c = a(bc)$ for all $a, b, c \in S$. We have

$$(ab)c = ac = a, \qquad a(bc) = a.$$

Exercise 8. Give an example of 2×2 matrices such that $A^{-1}B \neq BA^{-1}$.

Solution. Consider

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

Now,

$$A^{-1}B = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}, \qquad BA^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}.$$

Exercise 9. Show that if $ab = a$ in a group, then $b = 1$, and if $ab = 1$, then $b = a^{-1}$.

Solution. If $ab = a$, then

$$b = 1b = (a^{-1}a)b = a^{-1}(ab) = a^{-1}a = 1.$$

If $ab = 1$, then it suffices to show that $ba = 1$ to conclude $b = a^{-1}$.

$$ba = 1(ba) = (a^{-1}a)(ba) = a^{-1}(ab)a = a^{-1}a = 1.$$

Exercise 10. Let a, b be elements of a group G. Show that the equation $ax = b$ has a unique solution in G.

Solution. The existence of a solution is guaranteed by the inverse of a, namely $a^{-1} \in G$ such that $aa^{-1} = 1 = a^{-1}a$. Thus, $(a^{-1})ax = a^{-1}b$, hence $x = a^{-1}b$.

Now suppose that $ax = ay = b$. Again, left multiplying by a^{-1} gives $x = y$, guaranteeing that the solution is unique.

Exercise 11. Let G be a group, with multiplicative notation. We define an *opposite group* G° with law of composition $a \circ b$ as follows: The underlying set is the same as G, but the law of composition is the opposite; that is, we define $a \circ b = ba$. Prove that this defines a group.

Solution. The composition in G° is closed, since $a \circ b = ba \in G$ and G° shares all elements with G. Note that composition is associative in G, so for all $a, b, c \in G$,

$$c(ba) = (cb)a, \qquad (a \circ b) \circ c = a \circ (b \circ c).$$

This shows that composition is associative in G°. Next, the identity in G serves as the identity in G°, since for any $a \in G^\circ$,

$$1 \circ a = a1 = a = 1a = a \circ 1.$$

Each $a \in G$ has an inverse $a^{-1} \in G$, and this guarantees that each $a \in G^\circ$ has the same inverse $a^{-1} \in G^\circ$, since

$$a \circ a^{-1} = a^{-1}a = 1 = aa^{-1} = a^{-1} \circ a.$$

Thus, G° is a group.

2.2 Subgroups

Exercise 1. Determine the elements of the cyclic group generated by the matrix $\begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}$ explicitly.

Solution. Set

$$1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad x = \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}, \qquad x^2 = \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix}, \qquad x^3 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix},$$

$$x^4 = \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix}, \qquad x^5 = \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix}, \qquad x^6 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Therefore, the elements of the cyclic group generated by x are the matrices $\{1, x, x^2, x^3, x^4, x^5\}$. This group has order 6.

Exercise 2. Let a, b be elements of a group G. Assume that a has order 5 and that $a^3b = ba^3$. Prove that $ab = ba$.

Solution. We have $a^5 = 1$, therefore

$$ab = 1ab = a^5ab = a^6b = a^3(a^3b) = a^3(ba^3),$$

$$ba = ba1 = baa^5 = ba^6 = (ba^3)a^3 = (a^3b)a^3.$$

These are equal by associativity.

Exercise 3. Which of the following are subgroups?

(a) $GL_n(\mathbb{R}) \subset GL_n(\mathbb{C})$.

(b) $\{1, -1\} \subset \mathbb{R}^\times$.

(c) The set of positive integers in $\mathbb{Z}^+$.

(d) The set of positive reals in $\mathbb{R}^\times$.

(e) The set of all matrices $\begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix}$, with $a \neq 0$, in $GL_2(\mathbb{R})$.

Solution. It can be shown that (a), (b), (d) meet the axioms required for a subgroup. In (c), the additive identity 0 is missing. In (e), the identity matrix $\mathbb{I}_2$ is missing (and none of the described matrices belong to $GL_2(\mathbb{R})$ in any case).

Exercise 4. Prove that a non-empty subset H of a group G is a subgroup if for all $x, y \in H$ the element xy^{-1} is also in H.

Solution. Since H is non-empty, we can pick $x \in H$. Then, $x \in H$ and $x \in H$, so $xx^{-1} = 1 \in H$. Next, for any $x \in H$, we have $1 \in H$ and $x \in H$, so $1x^{-1} = x^{-1} \in H$. Finally, for any $x, y \in H$, we have $x \in H$ and $y^{-1} \in H$, so $x(y^{-1})^{-1} = xy \in H$. This means that $H \subseteq G$ is a subgroup.

Exercise 5. An nth root of unity is a complex number z such that $z^n = 1$. Prove that the nth roots of unity form a cyclic subgroup of $\mathbb{C}^\times$ of order n.

Solution. Let the set of all nth roots of unity be G. First we have $1^n = 1$ so $1 \in G$. Next, if $x, y \in G$, then $x^n = y^n = 1$, so $(xy)^n = 1$, thus $xy \in G$. Finally, note that $x^{-1} = x^{n-1}$, because $xx^{n-1} = x^{n-1}x = x^n = 1$. To see that this is a cyclic subgroup, set $x = e^{1\pi i/n}$, then $G = \{1, x, \ldots, x^{n-1}\}$. There are no other elements of G, since the polynomial $x^n - 1$ has at most n distinct complex roots.

Exercise 6.

(a) Find generators and relations analogous to (2.13) for the Klein four group.

(b) Find all subgroups of the Klein four group.

Solution.

(a) Write

$$e = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad x = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \qquad y = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad z = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}.$$

It can be verified that $x^2 = y^2 = z^2 = e$. Also, $xy = z = yx$, $xz = y = zx$, $yz = x = zy$. Thus, this group is abelian. Also, any two of x, y, z suffice to generate the third, and hence the entire Klein four group. Any one is not sufficient, since every element has order 2.

(b) The Klein four group has two trivial subgroups, namely $\{e\}$ and the whole group. Also, each of $\{e, x\}$, $\{e, y\}$, $\{e, z\}$ form a subgroup, since all of them contain the identity e, each element is their own inverse, and multiplication is closed. This gives a total of 5 subgroups.

Exercise 7. Let a and b be integers.

(a) Prove that the subset $a\mathbb{Z} + b\mathbb{Z}$ is a subgroup of $\mathbb{Z}^+$.

(b) Prove that a and $b + 7a$ generate the subgroup $a\mathbb{Z} + b\mathbb{Z}$.

Solution.

(a) We have

$$G = a\mathbb{Z} + b\mathbb{Z} = \{ar + bs : r, s \in \mathbb{Z}\}.$$

First note that $0 = a0 + b0 \in G$. Next, for any two elements $x, y \in G$, we can write $x = ar_1 + bs_1$ and $y = ar_2 + bs_2$ for integers r_1, r_2, s_1, s_2, so the sum $x + y = a(r_1 + r_2) + b(s_1 + s_2) \in G$. Finally, for $x = ar + bs \in G$, we have $-x = a(-r) + b(-s) \in G$, with $x + (-x) = 0$. This proves that G is a subgroup of $\mathbb{Z}^+$.

(b) Let H be the group generated by a and $b + 7a$. Note that a, $a + a = 2a$, $2a + a = 3a$, $\ldots$, $6a + a = 7a$, $\ldots$, $(n-1)a + a = na$ are all in H. Similarly, if $na \in H$, then $-na \in H$ for all positive integers n. Thus, $0 = a + (-a) \in H$, and $b = (b + 7a) + (-7a) \in H$. We repeat this process with b to see that $nb \in H$ for all integers n. Thus, $ar + bs \in H$ for all integers r and s, which means that $H \subseteq a\mathbb{Z} + b\mathbb{Z}$. On the other hand, for all $ar + bs \in a\mathbb{Z} + b\mathbb{Z}$, we have $ar + bs = a(1 - 7s) + (b + 7a)s \in H$, so $a\mathbb{Z} + b\mathbb{Z} \subseteq H$. This establishes that $H = a\mathbb{Z} + b\mathbb{Z}$.

Exercise 8. Make a multiplication table for the quaternion group H.

Solution. Use $i^2 = j^2 = k^2 = ijk = -1$, $ij = k$, $ji = i^3 j = -k$.

$\times$	1	i	j	k	-1	$-i$	$-j$	$-k$
1	1	i	j	k	-1	$-i$	$-j$	$-k$
i	i	-1	k	$-j$	$-i$	1	$-k$	j
j	j	$-k$	-1	i	$-j$	k	1	$-i$
k	k	j	$-i$	-1	$-k$	$-j$	i	1
1	-1	$-i$	$-j$	$-k$	1	i	j	k
i	$-i$	1	$-k$	j	i	-1	k	$-j$
j	$-j$	k	1	$-i$	j	$-k$	-1	i
k	$-k$	$-j$	i	1	k	j	$-i$	-1

Exercise 9. Let H be the subgroup generated by two elements a, b of a group G. Prove that if $ab = ba$, then H is an abelian group.

Solution. Since H is generated by a and b, every element in H can be written in the form

$$a^{m_1}b^{n_1}a^{m_2}b^{n_2}\dots a^{m_k}b^{n_k},$$

where $k \in \mathbb{N}$ and $m_i, n_i \in \mathbb{Z}$. First, we claim that every such element can be simplified to the form $a^m b^m$. To see this, note that since a and b commute, we have $ab(b^{-1}a^{-1}) = 1 = ba(a^{-1}b^{-1}) = ab(a^{-1}b^{-1})$, so cancelling ab gives $b^{-1}a^{-1} = a^{-1}b^{-1}$. Now, $ab = ba$ means $a = bab^{-1}$. Thus, for any positive power $a^m = ba^m b^{-1}$, and $a^{-m} = ba^{-m}b^{-1}$. Thus, $a^m b^n = ba^m b^{-1}b^n = ba^m b^{n-1}$. Repeating this another $n-1$ times gives $a^m b^n = b^n a^m$ for positive n. For negative n, simply note that $a^m = b^n(b^{-n}a^m) = b^n(a^m b^{-n})$, hence $a^m b^n = b^n a^m$. This means that the powers a^m and b^n commute. Thus, we can commute all powers b^{n_i} in our general expression to the right, yielding

$$a^{m_1+\dots+m_k}b^{n_1+\dots n_k} = a^m b^n.$$

Now pick arbitrary $x, y \in H$, and write $x = a^m b^n$, $y = a^r b^s$. Then,

$$xy = a^m b^n a^r b^s = a^m(b^n a^r)b^s = a^m a^r b^n b^s = a^{m+r}b^{n+s},$$

$$yx = a^r b^s a^m b^n = a^r(b^s a^m)b^n = a^r a^m b^s b^n = a^{m+r}b^{n+s}.$$

This gives $xy = yx$ for all $x, y \in H$, which means that H is abelian.

Exercise 10.

(a) Assume that an element x of a group has order rs. Find the order of x^r.

(b) Assuming that x has arbitrary order n, what is the order of x^r?

Solution. We first show that if x has order n and $x^m = 1$, then n divides m. Note that $n \le m$, since n is chosen to be the least natural number satisfying $x^n = 1$. Thus, use Euclid's Division Lemma to write $m = nq + r$ for integers $q > 0$, $0 \le r < n$. We now have

$$1 = x^m = x^{nq+r} = (x^n)^q x^r = x^r.$$

Since $r < n$, the only possibility is $r = 0$, hence $m = nq$, proving that n divides m.

(a) Note that $(x^r)^s = x^{rs} = 1$, so the order of x^r divides s. Also, if the order of x^r were less then s, say $t < s$, then we would have $(x^r)^t = 1$, so $x^{rt} = 1$, with $rt < rs$. This would contradict the fact that rs is the order of x. Thus, the order of x^r must be s.

(b) We claim that the order of x^r is $m = n/\gcd(n, r)$. Write $\gcd(n, r) = d$, so $rm = nr/d = nk$ for some integer k since d divides r. Thus,

$$(x^r)^m = x^{rm} = x^{nk} = (x^n)^k = 1.$$

Suppose instead that the order of x^r is some $m' < m$. Then m' divides m, so $m = m'q$ for some integer $q > 1$. We have $1 = (x^r)^{m'} = x^{rm'}$, therefore n must divide rm', say $rm' = nk'$ for some integer k'. Now, $nk = rm = rm'q = nk'q$, so $k = k'q$. Recall that $n = md$, $r = kd$. Now we have found $n = m'(qd)$, $r = k'(qd)$, thus qd divides both n and r. However, $qd > d$, which contradicts the fact that any common factor of n and r must divide d. Hence, the order of x^r must be $m = n/d$.

Exercise 11. Prove that in any group the orders of ab and ba are equal.

Solution. Note that $ab = (b^{-1}b)ab = b^{-1}(ba)b$. It is easily seen by induction that $(ab)^n = b^{-1}(ba)^n b$ for all positive integers n. Therefore, if the order of ab is some integer n, $1 = (ab)^n = b^{-1}(ba)^n b$, hence $(ba)^n = bb^{-1} = 1$. Thus, the order of ba divides n. A similar argument using $(ba)^n = a^{-1}(ab)^n a$ shows that if the order of ba is n', then the order of ab divides n'. This forces $n = n'$ when either ab or ba has finite order.

Note that we shown that if the order of ab is finite, then the order of ba must be finite, and vice versa. This immediately implies that if the order of any one of ab or ba is infinite, then the order of the other must also be infinite. Hence, the order of ab and ba are always the same in any group.

Exercise 12. Describe all groups G which contain no proper subgroup.

Solution. Suppose that G has no proper subgroups, i.e. the only subgroups of G are $\{1\}$ and G. The trivial group of one element $\{1\}$ satisfies this. Otherwise, let G be a non-trivial group, with $x \in G$ such that $x \neq 1$. Then, the group generated by x, i.e. the group of elements $\{\dots x^{-2}, x^{-1}, 1, x, x^2, \dots\}$ is a subgroup of G. Since G has no proper subgroups, this forces this to be equal to G itself. Thus, G is a cyclic group.

Suppose that G has finite order, and furthermore suppose that this order is a composite number ab, where $a \geq b > 1$. Then, it can be shown that the group generated by x^a is a proper subgroup of G, with x not in this subgroup. Therefore, the order of G must be prime.

Suppose that G has infinite order. Now note that the group generated by x^2 is a proper subgroup of G, with x not in this subgroup. This is a contradiction, thus the order of G cannot be infinite.

Exercise 13. Prove that every subgroup of a cyclic group is cyclic.

Solution. Let G be a cyclic group generated by the single element x. If $x = 1$, then $G = \{1\}$, which has no proper subgroups. Otherwise, note that we can enumerate

$$G = \{\dots x^{-2}, x^{-1}, 1, x, x^2, \dots\}.$$

Let H be a proper subgroup of G, and let $y \in H$ be a non-trivial element. This means that $y \in G$, so we can write $y = x^k$ for some positive integer k (note that if k were negative, then $y^{-1} = x^{-k} \in H$ too, so use $-k$ instead). Suppose that we have chosen $w = x^m \in H$ such that m is the smallest possible, positive choice. This means that $m \leq k$, so using Euclid's Division Lemma, write $k = mq + r$ for $0 \leq r < m$. Thus, $x^k = (x^m)^q x^r = w^q x^r$, or $x^r = w^{-q} x^k$. Now, $w \in H$ means that all powers of w are also in H, hence $w^{-q} \in H$. This means that $w^{-q} x^k = x^r \in H$. However, recall that m was the smallest positive integer such that $x^m \in H$, which forces $r = 0$. Thus, for any $y \in H$, we see that $y = w^q$. This means that H is generated by the element $w = x^m$, which makes it a cyclic subgroup.

Exercise 14. Let G be a cyclic group of order n, and let r be an integer dividing n. Prove that G contains exactly one subgroup of order r.

Solution. Let G be generated by the element x, so $G = \{1, x, \dots, x^{n-1}\}$, and let H be a subgroup of order r, where $n = mr$ for some positive integer m. We claim that the subgroup generated by x^m is the only subgroup of order r. Note from the previous exercise that H must be cyclic, and thus is generated by some element $x^k \in G$. Thus, if we pick $y \in H$, we must have $y = x^{kq}$ for some non-negative integer q. Since H has order r, we require $y^r = 1$, i.e. $x^{kqr} = 1$. Now, x has order n, hence n must divide kqr, say $np = kqr$ for some positive integer p. Substitute $n = mr$, so $mp = kq$. Thus, $y = x^{kq} = x^{mp}$, so every element of H is a power of x^m. In other words, $H \subseteq \{1, x^m, \dots, x^{m(r-1)}\}$. Also note that H contains exactly r elements, and the right hand side also contains r elements since all x^{mp} are distinct. Thus, we must have an equality, which means that the only subgroup of G with order r is the one generated by x^m.

Exercise 15.

(a) In the definition of subgroup, the identity element in H is required to be the identity of G. One might require only that H have an identity element, not that it is the same as the identity in G. Show that if H has an identity at all, then it is the identity in G, so this definition would be equivalent to the one given.

(b) Show the analogous thing for inverses.

Solution.

(a) Let 1 be the identity element in G, and suppose that $1'$ is the identity element in H. Now, both $1' \in G$ and $1 \in G$, and we demand

$$1x = x \text{ for all } x \in G, \qquad 1'x = x \text{ for all } x \in H.$$

Combining these, we want

$$1 = 1'1 = 1',$$

so the identity in H must be the same as the identity in G.

(b) Let $x \in H$, let y be its inverse in G, and let w be its inverse in H. We want

$$xy = 1 = yx, \qquad xw = 1 = wx.$$

Thus,

$$y = y1 = y(xw) = (yx)w = 1w = w,$$

so the inverse of x in G must be the same in H.

Exercise 16.

(a) Let G be a cyclic group of order 6. How many of its elements generate G?

(b) Answer the same question for cyclic groups of order 5, 8, and 10.

(c) How many elements of a cyclic group of order n are generators for that group?

Solution.

(a) Let $G = \{1, x, \ldots, x^5\}$. Then, the elements x and x^5 (2 elements) generate G. Note that x^2 and x^4 only generate $\{1, x^2, x^4\}$, and x^3 only generates $\{1, x^3\}$.

(b) Using the same notation as before, a cyclic group of order 5 is generated by x, x^2, x^3, x^4 (4 elements). A cyclic group of order 8 is generated by x, x^3, x^5, x^7 (4 elements). A cyclic group of order 10 is generated by x, x^3, x^7, x^9 (4 elements).

(c) A cyclic group of order n is generated by $\phi(n)$ elements, where the Euler totient function ϕ counts the number of positive integers less than n which are co-prime with n. This is because a cyclic group $G = \{1, \ldots, x^{n-1}\}$ of order n is generated by x^r precisely when $\gcd(n, r) = 1$. Recall from Exercise 14 that if the order of a cyclic group G is $n = ab$, then x^a generates a subgroup of order b. Thus, when $b > 1$, the generated subgroup is a proper subgroup. This means that whenever r divides n, x^r cannot generate the entire group G. Similarly, if r and n share a common factor $m > 1$, then m divides n so x^m cannot generate G, and neither can $x^r = x^{mk}$ for some integer k. This only leaves those x^r such that r and n have no common factors. If so, then we can choose integers p and q such that $\gcd(n, r) = 1 = np + rq$, so $x^1 = x^{np+rq} = (x^r)^q$. Thus, any element $x^m \in G$ can be expressed as $(x^r)^{mq}$, so x^r does indeed generate G.

Exercise 17. Prove that a group in which every element except the identity has order 2 is abelian.

Solution. Let G be such a group, and let $x, y \in G$. Note that since every element has order 2, we have $x^2 = y^2 = (xy)^2 = (yx)^2 = 1$. Also,

$$(xy)(yx) = xy^2x = x^2 = 1, \qquad (xy)(xy) = (xy)^2 = 1.$$

Equating and cancelling xy from the left gives $yx = xy$, showing that G is abelian.

Exercise 18.

(a) Prove that the elementary matrices of the first and third types suffice to generate $GL_n(\mathbb{R})$.

(b) The *special linear group* $SL_n(\mathbb{R})$ is the set of $n \times n$ matrices whose determinant is 1. Show that $SL_n(\mathbb{R})$ is a subgroup of $GL_n(\mathbb{R})$.

(c) Use row reduction to prove that the elementary matrices of the first type generate $SL_n(\mathbb{R})$. Do the 2×2 case first.

Solution.

(a) First, we show that elementary matrices of the second type, i.e. permutation matrices, can be generated by elementary matrices of the other two types, i.e. row adding and scaling. In order to transpose rows i and j, consider the following row operations: $R_i \to R_i + R_j$, $R_j \to R_i - R_j$, $R_i \to R_i - R_j$. This can be achieved using only row addition and scaling, hence any transposition matrix can be obtained by applying this to the identity matrix. Furthermore, any permutation matrix can be expressed as the product of transposition matrices. This proves that elementary matrices of the first and third types generate those of the second type. Now, if A is invertible, then the process of Gauss-Jordan elimination can be applied to reduce it to an identity matrix, hence $E_1 \cdots E_k A = \mathbb{I}$. This means that $A = E_k^{-1} \cdots E_1^{-1}$, hence any matrix in $GL_n(\mathbb{R})$ can be expressed as a product of elementary matrices. In addition, any product of elementary matrices is invertible, which means that the elementary matrices generate $GL_n(\mathbb{R})$.

(b) First note that the identity matrix $\mathbb{I}_n \in SL_n(\mathbb{R})$, since it has determinant 1. Using $\det AB = \det A \det B$, conclude that if $A, B \in SL_n(\mathbb{R})$, then $AB \in SL_n(\mathbb{R})$. Finally, use $\det A^{-1} = 1/\det A$ to conclude that if $A \in SL_n(\mathbb{R})$, then $A^{-1} \in SL_n(\mathbb{R})$. This proves that $SL_n(\mathbb{R})$ is a subgroup of $GL_n(\mathbb{R})$.

(c) Consider an arbitrary matrix

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in SL_2(\mathbb{R}),$$

where $ad - bc = 1$. Perform $R_1 \to R_1 - (b/d)R_2$ to get

$$\begin{bmatrix} a - bc/d & 0 \\ c & d \end{bmatrix}.$$

Note that $a - bc/d = 1/d$. Next, perform $R_2 \to R_2 - cdR_1$.

$$\begin{bmatrix} 1/d & 0 \\ 0 & d \end{bmatrix}.$$

Note that if $d = 0$, we could have performed analogous operations using b instead (both $b = d = 0$ is not possible, since we require $ad - bc = 1$). Now, perform $R_2 \to R_2 + dR_1$.

$$\begin{bmatrix} 1/d & 0 \\ 1 & d \end{bmatrix}.$$

Next, perform $R_1 \to R_1 - R_2/d$.

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

Next, perform $R_2 \to R_2 + dR_1$.

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

Finally, perform $R_1 \to R_1 + R_2$, $R_2 \to R_2 - R_1$, $R_1 \to R_1 + R_2$.

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

Thus, we have shown that any matrix in $SL_2(\mathbb{R})$ can be generated using only elementary matrices of the first type.

Analogously, consider some $A = [a_{ij}] \in SL_n(\mathbb{R})$. We have already seen that $R_i \to R_i + R_j$, $R_j \to R_j - R_i$, $R_i \to R_i + R_j$ transposes rows with a change of sign. Thus, transpose rows such that let $a_{11} \neq 0$ (if all $a_{1j} = 0$, then $\det A = 0$). Performing $R_i \to R_i - (a_{i1}/a_{11})R_1$ for all $i = 2, \ldots, n$ makes sure that all elements except the first one are zero.

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} - a_{12}a_{21}/a_{11} & \cdots & a_{2n} - a_{1n}a_{21}/a_{11} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2} - a_{12}a_{n1}/a_{11} & \cdots & a_{nn} - a_{1n}a_{n1}/a_{11} \end{bmatrix}.$$

Now, note that the determinant of A is $a_{11} \det M_{11} \neq 0$, which means that the first column of M_{11} contains a non-zero element. We thus repeat the above process, transposing rows if necessary n times, finally giving us an upper triangular matrix. Relabel the entries as b_{ij}.

$$\begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ 0 & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & b_{nn} \end{bmatrix}.$$

Note that $b_{11}b_{22} \ldots b_{nn} = 1$. Thus, for each $j = n, n-1, \ldots, 2$, perform $R_i \to R_i - (b_{ij}/b_{jj})R_j$ for all $i = 1, \ldots, j-1$. This yields a diagonal matrix,

$$\begin{bmatrix} b_{11} & 0 & \cdots & 0 \\ 0 & b_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & b_{nn} \end{bmatrix}.$$

Now, look at the upper left 2×2 block. Performing $R_2 \to R_2 + R_1/b_{11}$, $R_1 \to R_1 - b_{11}R_2$, $R_2 \to R_2 + R_1/b_{11}$ yields

$$\begin{bmatrix} 0 & -b_{11}b_{22} & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & b_{nn} \end{bmatrix}.$$

Performing our transposition with sign change yields

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & b_{11}b_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & b_{nn} \end{bmatrix}.$$

Repeating this down the line, we see that $b_{11}b_{22}\dots b_{nn}$ bunches up at the lower right, but this is just 1, so we get the identity matrix. Thus, $SL_n(\mathbb{R})$ is generated by elementary matrices of the first type.

Exercise 19. Determine the number of elements of order 2 in the symmetric group S_4.

Solution. Use the notation $(a\ b\ c\ d)$ to denote a cycle $a \rightsquigarrow b$, $b \rightsquigarrow c$, $c \rightsquigarrow d$, $d \rightsquigarrow a$. Note that any transposition of two elements has order 2, hence all two-cycles

$$(1\ 2) \quad (1\ 3) \quad (1\ 4) \quad (2\ 3) \quad (2\ 4) \quad (3\ 4)$$

have order 2. If any two of them commute, then their product will also be of order 2 (Exercise 11), therefore the products of disjoint two-cycles

$$(1\ 2)(3\ 4) \quad (1\ 3)(2\ 4) \quad (1\ 4)(2\ 3)$$

also have order 2. This gives a total of 9 elements.

Exercise 20.

(a) Let a, b be elements of an abelian group of orders m, n respectively. What can you say about the order of their product ab?

(b) Show by example that the product of elements of finite order in a non-abelian group need not have finite order.

Solution.

(a) Set $d = \gcd(m, n)$. The order of ab will always divide $k = mn/d$. Note that in an abelian group,
$$(ab)^k = a^k b^k = a^{mn/d} b^{mn/d} = (a^m)^{n/d}(b^n)^{m/d} = 1.$$
Note that the order of ab is not always k. Consider $a = b = i \in \mathbb{C}^\times$, and note that the orders of a and b are 4. However, the order of their product $i^2 = -1$ is 2.

(b) Select the following elements from $GL_2(\mathbb{R})$.
$$A = \begin{bmatrix} -1 & 1 \\ 0 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} -1 & -1 \\ 0 & 1 \end{bmatrix}.$$
Note that $A^2 = B^2 = \mathbb{I}_2$. However,
$$AB = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}, \qquad (AB)^2 = \begin{bmatrix} 1 & 4 \\ 0 & 1 \end{bmatrix}, \qquad (AB)^3 = \begin{bmatrix} 1 & 6 \\ 0 & 1 \end{bmatrix}, \qquad \dots$$
In general for $n > 0$, we have
$$(AB)^n = \begin{bmatrix} 1 & 2n \\ 0 & 1 \end{bmatrix}.$$
Thus, AB has infinite order.

Exercise 21. Prove that the set of elements of finite order in an abelian group is a subgroup.

Solution. Let G be abelian, and let H be the set of all elements with finite order. Note that $1 \in H$. Also, if a and b have finite order, so does ab by the previous exercise. Finally, if a has finite order n, then $a^n = 1$ forces $1 = a^{-n} = (a^{-1})^n$, hence a^{-1} has finite order n. Thus, H is a subgroup of G.

Exercise 22. Prove that the greatest common divisor of a and b, as defined in the text, can be obtained by factoring a and b into primes and collecting the common factors.

Solution. Suppose that

$$a = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_n^{\alpha_n} \cdots , \qquad b = p_1^{\beta_1} p_2^{\beta_2} \cdots p_n^{\beta_n} \cdots ,$$

where p_i are the prime numbers, and α_i, β_i are non-negative integers. We have $\alpha_{i \geq M} = \beta_{i \geq N} = 0$ eventually, for a and b to be finite. Setting $\gamma_i = \min\{\alpha_i\, \beta_i\}$, we claim that the greatest common factor of a and b is

$$d = p_1^{\gamma_1} p_2^{\gamma_2} \cdots p_n^{\gamma_n} \cdots .$$

First note that d divides both a and b, therefore d must divide $d' = \gcd(a, b)$. Write

$$d'/d = p_1^{\gamma_1'} p_2^{\gamma_2'} \cdots p_n^{\gamma_n'} \cdots .$$

Suppose that $\gamma_k' > 0$ for some k. This means that the power of p_k in d' is $\gamma_k + \gamma_k' > \min\{\alpha_k, \beta_k\}$. This means that d' cannot divide one or more of a and b, so $\gamma_i' = 0$ for all i. Thus, $d = d'$.