CHAPTER 1

Introductory Concepts

At the present time, physicists find it convenient to try to describe the real world in terms of mathematics. Before we also explore the properties of the real world, we must have a firm grasp of the kinds of mathematical concepts that have been useful for physicists. These fundamental building blocks[1–3] are presented in this chapter, together with examples.

I. Basic Building Blocks

1. SET. A **set** is a collection of objects that do not necessarily have any additional structure or properties. For example, a collection of n oranges or bananas constitutes a set. So do n people. So do n points. The archetypical example of a set containing n (possibly infinite) objects is the set of n points.

2. GROUP. A **group** G is:

(a) a set $g_1, g_2, \ldots, g_n \in G$

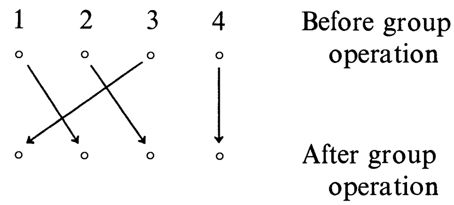
together with

(α) an operation, called group multiplication ($\circ$)

such that

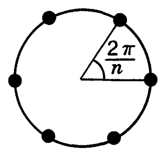
1.	$g_i \in G,\ g_j \in G \Rightarrow g_i \circ g_j \in G$	closure
2.	$g_i \circ (g_j \circ g_k) = (g_i \circ g_j) \circ g_k$	associativity
3.	$g_1 \circ g_i = g_i = g_i \circ g_1$ for all g_i	existence of identity
4.	$g_k \circ g_l = g_l \circ g_k = g_1$	unique inverse $g_l = g_k^{-1}$ (1.1)

Example 1. The collection of all possible permutations of the points 1, 2, 3, 4 constitutes a group with 4! elements, or operations, called P_4

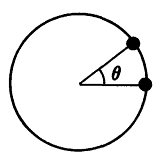


$$g = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix} \qquad (1.2)$$

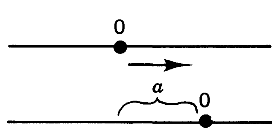
Example 2. The collection of rotations of the circle through multiples of $2\pi/n$ radians constitutes a group with n distinct operations. Such a (finite) group is said to be of **order** n.



Example 3. The collection of rotations of the circle through an angle θ $(0 \leq \theta < 2\pi)$ is an example of a continuous group. The group operations $g(\theta)$ exist in 1-1 correspondence with points on the interval $0 \leq \theta < 2\pi$.



Example 4. The set T_a of rigid translations of the straight line through a distance a is another example of a continuous group. The group operations exist in 1-1 correspondence with the points on the line $-\infty < a < +\infty$.



Example 5. The set of real numbers, excluding 0, forms a group under the operation of multiplication. So do the complex numbers, provided we exclude 0. The identity operation in both groups is 1. But under the operation of addition, both the real and complex numbers form groups with identity element 0.

Example 6. The set of real $n \times n$ nonsingular matrices under matrix multiplication forms a group called $Gl\ (n, r)$. The subset of these matrices with determinant $+1$ forms a (sub)group called $Sl\ (n, r)$. The collection of $n \times n$ unitary matrices $U(n)$ also forms a group under matrix multiplication.

Comment. For the groups discussed in Examples 2 to 5, the order in which the group operations are applied is immaterial. A group that obeys a fifth postulate in addition to the four just listed is called an **abelian** or **commutative** group:

5. $g_i \circ g_j = g_j \circ g_i$ all $g_i, g_j \in G$ commutativity (1.1′)

In an abelian group it is customary to denote the group multiplication operation as $+$ instead of $\circ$. The groups of Examples 1 and 6 are not abelian.

3. FIELD. A **field** F is

(a) a set of elements $f_0, f_1, f_2, \ldots,$

together with two operations:

(α) $+$ called addition
(β) $\circ$ called scalar multiplication

such that Postulates A and B hold.

Postulate A. F is an abelian group under $+$, with f_0 the identity.

Postulate B

1. $f_i \circ f_j \in F$ $\circ$ closure
2. $f_i \circ (f_j \circ f_k) = (f_i \circ f_j) \circ f_k$ $\circ$ associativity
3. $f_i \circ 1 = 1 \circ f_i = f_i$ $\circ$ identity
4. $f_i \circ f_i^{-1} = 1 = f_i^{-1} \circ f_i, f_i \neq f_0$ $\circ$ inverse, except for f_0
5. $f_i \circ (f_j + f_k) = f_i \circ f_j + f_i \circ f_k$
$(f_i + f_j) \circ f_k = f_i \circ f_k + f_j \circ f_k$ distributive law (1.3)

If Postulate B-6 is also obeyed

6. $f_i \circ f_j = f_j \circ f_i$ commutativity (1.3′)

we say the field is commutative.

Only three fields are generally used by physicists. These are the real and complex numbers and the quaternions. The properties of the real numbers are assumed to be familiar.

Every complex number can be represented in the form

$$c = a1 + ib$$

where the units 1 and i $(=\sqrt{-1})$ obey

$$\begin{aligned} 1 \cdot 1 &= 1 \\ i \cdot 1 &= 1 \cdot i = i \\ i \cdot i &= -1 \end{aligned} \tag{1.4}$$

and a, b are arbitrary real numbers. Then we have

$$\begin{aligned} c_1 + c_2 &= (a_1 1 + ib_1) + (a_2 1 + ib_2) \\ &= (a_1 + a_2)1 + (b_1 + b_2)i \end{aligned} \tag{1.5}$$

$$\begin{aligned} c_1 c_2 &= (a_1 1 + b_1 i)(a_2 1 + b_2 i) \\ &= (a_1 a_2 - b_1 b_2)1 + (a_1 b_2 + b_1 a_2)i \end{aligned} \tag{1.6}$$

Every quaternion can be represented in the form

$$q = q_0 1 + q_1 \lambda_1 + q_2 \lambda_2 + q_3 \lambda_3 \tag{1.7}$$

where the q_i $(i = 0, 1, 2, 3)$ are real numbers and the λ_i have multiplicative properties defined by

$$\begin{aligned} \lambda_0 \lambda_i &= \lambda_i \lambda_0 = \lambda_i \qquad i = 0, 1, 2, 3 \\ \lambda_i \lambda_i &= -\lambda_0 \\ \lambda_1 \lambda_2 &= -\lambda_2 \lambda_1 = \lambda_3 \\ \lambda_2 \lambda_3 &= -\lambda_3 \lambda_2 = \lambda_1 \\ \lambda_3 \lambda_1 &= -\lambda_1 \lambda_3 = \lambda_2 \end{aligned} \tag{1.8}$$

The sum and the product of two quaternions p and q are

$$\begin{aligned} p + q = (p_0 + q_0)\lambda_0 + (p_1 + q_1)\lambda_1 \\ + (p_2 + q_2)\lambda_2 + (p_3 + q_3)\lambda_3 \end{aligned} \tag{1.9}$$

and

$$\begin{aligned} pq = (p_0 q_0 - p_1 q_1 - p_2 q_2 - p_3 q_3)\lambda_0 \\ + (p_0 q_1 + p_1 q_0 + p_2 q_3 - p_3 q_2)\lambda_1 \\ + (p_0 q_2 + p_2 q_0 + p_3 q_1 - p_1 q_3)\lambda_2 \\ + (p_0 q_3 + p_3 q_0 + p_1 q_2 - p_2 q_1)\lambda_3 \end{aligned} \tag{1.10}$$

The set of eight elements $\pm\lambda_i$ forms a noncommutative group.

Complex conjugation can be defined for quaternions just as for complex numbers. Defining

$$(\lambda_0, \lambda_1, \lambda_2, \lambda_3)^* = (+\lambda_0, -\lambda_1, -\lambda_2, -\lambda_3) \tag{1.11q}$$

in direct analogy to

$$(1, i)^* = (+1, -i) \tag{1.11c}$$

we easily see

$$q^*q = \lambda_0\left(\sum_{i=0}^{3} q_i^{\ 2}\right) \tag{1.12}$$

The product of a quaternion with its conjugate is a real number which is ≥ 0. Also, $q^*q = 0$ implies that q is zero, in exact analogy with complex numbers.

4. LINEAR VECTOR SPACE. A **linear vector space** V consists of

(a) a collection $\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2, \ldots, \in V$, called vectors
(b) a collection $f_1, f_2, \ldots, \in F$, a field

together with two kinds of operations

(α) vector addition, $+$
(β) scalar multiplication, $\circ$

such that Postulates A and B hold.

Postulate A. $(V, +)$ is an abelian group.

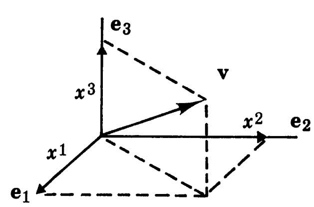
1.	$\mathbf{v}_i, \mathbf{v}_j \in V \Rightarrow \mathbf{v}_i + \mathbf{v}_j \in V$	closure
2.	$\mathbf{v}_i + (\mathbf{v}_j + \mathbf{v}_k) = (\mathbf{v}_i + \mathbf{v}_j) + \mathbf{v}_k$	associativity
3.	$\mathbf{v}_0 + \mathbf{v}_i = \mathbf{v}_i = \mathbf{v}_i + \mathbf{v}_0$	identity
4.	$\mathbf{v}_i + (-\mathbf{v}_i) = \mathbf{v}_0 = (-\mathbf{v}_i) + \mathbf{v}_i$	inverse
5.	$\mathbf{v}_i + \mathbf{v}_j = \mathbf{v}_j + \mathbf{v}_i$	commutativity

Postulate B

1.	$f_i \in F, \mathbf{v}_j \in V \Rightarrow f_i \mathbf{v}_j \in V$	closure′
2.	$f_i \circ (f_j \circ \mathbf{v}_k) = (f_i \circ f_j) \circ \mathbf{v}_k$	associativity′
3.	$1 \circ \mathbf{v}_i = \mathbf{v}_i = \mathbf{v}_i \circ 1$	identity′
4.	$f_i \circ (\mathbf{v}_k + \mathbf{v}_l) = f_i \circ \mathbf{v}_k + f_i \circ \mathbf{v}_l$	bilinearity
	$(f_i + f_j) \circ \mathbf{v}_k = f_i \circ \mathbf{v}_k + f_j \circ \mathbf{v}_k$	(1.13)

Example 1. The most primitive example of a vector is "something that points in some direction."

$$\mathbf{v} = x^1\mathbf{e}_1 + x^2\mathbf{e}_2 + x^3\mathbf{e}_3 \tag{1.14}$$



Example 2. If we associate V with F in the definition of a linear vector space, we see that the real and complex numbers and quaternions are linear vector spaces. The complex numbers form a vector space over the field of real numbers (basis 1, i) or complex numbers (basis 1); the quaternions form a vector space over the field of real numbers (bases $\lambda_0, \lambda_1, \lambda_2, \lambda_3$) or the quaternion field itself (basis 1).

Example 3. Let $\mathscr{L}$ be any linear differential or integral operator:

$$\mathscr{L}(\alpha\phi_1 + \beta\phi_2) = \alpha\mathscr{L}(\phi_1) + \beta\mathscr{L}(\phi_2) \tag{1.15}$$

Then if ϕ_1 and ϕ_2 are solutions to the equation

$$\mathscr{L}(\phi_i) = 0 \tag{1.16}$$

so also is any linear combination. The set of all solutions to the equation

$$\mathscr{L}(\phi) = 0 \tag{1.17}$$

is a linear vector space. Since a large class of the differential and integral operators of mathematical physics has this linearity property, a study of linear vector spaces and their properties is directly relevant for the physicist.

Example 4. The set of functions $f(\phi)$ defined on the circle $(0 \leq \phi < 2\pi)$ forms a linear vector space

$$f(\phi) = \sum_{-\infty}^{+\infty} a_m e^{im\phi} \tag{1.18}$$

where m is an integer.

For that matter, the set of functions defined on any set of points (either finite or infinite) forms a linear vector space.

Example 5. The set of all $N \times M$ matrices forms a vector space under matrix addition. In particular, the sets of $N \times 1$ and $1 \times N$ matrices form vector spaces, V_N and $V_N^\dagger$.

At this point it is convenient to introduce several concepts that are useful for describing the properties of vector spaces.

Definition. The vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are **linearly independent** if

$$\sum \alpha^i \mathbf{v}_i = 0 \Rightarrow \alpha^i = 0 \qquad i = 1, 2, \ldots, n \tag{1.19}$$

In Examples 1 and 4 above, we have

$$x^1\mathbf{e}_1 + x^2\mathbf{e}_2 + x^3\mathbf{e}_3 = 0 \Rightarrow x^1 = x^2 = x^3 = 0 \tag{1.20}$$

$$\sum_{m=-\infty}^{+\infty} a_m e^{im\phi} = 0 \Rightarrow a_m = 0 \qquad m = 0, \pm 1, \pm 2, \ldots, \tag{1.21}$$

Therefore, the vectors $\mathbf{e}_i$ are linearly independent, as are the vectors $\mathbf{e}^{im\phi}$, $0 \leq \phi < 2\pi$.

Definition. A vector space is N-**dimensional** if it is possible to find a set of N nonzero linearly independent vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N$, but every set of $N + 1$ nonzero vectors is linearly dependent.

Definition. Any such maximal set of vectors is called a **basis**, or **coordinate system**.

Then any vector $\mathbf{v}$ can be expanded in terms of a basis. For if

$$\beta\mathbf{v} + \sum_{i=1}^{N} \alpha^i \mathbf{v}_i = 0 \tag{1.22}$$

there is a nontrivial solution.

1. If β is zero,

$$\sum \alpha^i \mathbf{v}_i = 0 \Rightarrow \alpha^i = 0$$

and this is the trivial solution.

2. Therefore $\beta \neq 0$, and

$$\mathbf{v} = \sum_{i=1}^{N} \left(-\frac{\alpha^i}{\beta}\right)\mathbf{v}_i \tag{1.23}$$

is the unique expansion of $\mathbf{v}$ in terms of the basis $\mathbf{v}_i$.

By a fundamental[4] theorem of algebra, all N-dimensional vector spaces over the same field are isomorphic to each other. In particular, they are isomorphic to the canonical* N-dimensional vector space of $N \times 1$ matrices, with bases†

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \\ \vdots \\ 0 \end{pmatrix}; \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \\ \vdots \\ 0 \end{pmatrix}; \text{—}; \mathbf{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \tag{1.24}$$

Therefore, we can learn all the properties of any N-dimensional vector space merely by studying its faithful canonical representation V_N. The foregoing vector spaces with bases $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_N$, over the field of real numbers, complex numbers, and quaternions, are denoted R_N, C_N, and Q_N, respectively.

* Canonical means standard. A canonical form is one that has been standardized through use or convenience.

† The term "bases" is used in place of "basis vectors" whenever no confusion will result.

5. ALGEBRA. A **linear algebra** A consists of

(a) a collection $\mathbf{v}_1, \mathbf{v}_2, \ldots, \in V$, called vectors
(b) a collection $f_1, f_2, \ldots, \in F$, a field,

together with three kinds of operations

(α) vector addition, $+$
(β) scalar multiplication, $\circ$
(γ) vector multiplication, $\square$

such that we can state Postulates A to C.

Postulate A. Postulates $A1$ to $A5$ for a vector space hold.

Postulate B. Postulates $B1$ to $B4$ for a vector space hold.

Postulate C.

1. $\mathbf{v}_1, \mathbf{v}_2 \in V \Rightarrow \mathbf{v}_1 \square \mathbf{v}_2 \in V$ closure″ (1.25)
2. $(\mathbf{v}_1 + \mathbf{v}_2) \square \mathbf{v}_3 = \mathbf{v}_1 \square \mathbf{v}_3 + \mathbf{v}_2 \square \mathbf{v}_3$
   $\mathbf{v}_1 \square (\mathbf{v}_2 + \mathbf{v}_3) = \mathbf{v}_1 \square \mathbf{v}_2 + \mathbf{v}_1 \square \mathbf{v}_3$ bilinearity′ (1.26)

Different varieties of algebras may be obtained, depending on which additional postulates are also satisfied.

3. $(\mathbf{v}_1 \square \mathbf{v}_2) \square \mathbf{v}_3 = \mathbf{v}_1 \square (\mathbf{v}_2 \square \mathbf{v}_3)$ associativity″
4. $\mathbf{v}_1 \square \mathbf{1} = \mathbf{v}_1$ existence of identity″; in general, this identity is not equal to the identity under $+$ or $\circ$
5. $\mathbf{v}_1 \square \mathbf{v}_2 = \pm \mathbf{v}_2 \square \mathbf{v}_1$ $\left\{\begin{matrix}\text{symmetric}\\ \text{antisymmetric}\end{matrix}\right\}$ under interchange
6. $\mathbf{v}_1 \square (\mathbf{v}_2 \square \mathbf{v}_3) = (\mathbf{v}_1 \square \mathbf{v}_2) \square \mathbf{v}_3 + \mathbf{v}_2 \square (\mathbf{v}_1 \square \mathbf{v}_3)$ derivative property

Example 1. The set of real $n \times n$ matrices forms a real n^2-dimensional vector space under matrix addition and scalar multiplication by real numbers. If we adjoin to this vector space the additional operation defined simply by matrix multiplication

$$(A \square B)_{ik} = \sum_{j=1}^{n} A_{ij} B_{jk} \tag{1.27}$$

this space becomes an associative algebra. The identity vector under $+$ is 0, the identity vector under $\square$ is the unit matrix I

$$(I)_{ik} = \delta_{ik}$$

The identity under $\circ$ is 1.

In addition to the postulates for an algebra, Example 1 satisfies Postulates C-3 and C-4; it is called a linear associative algebra with identity.

Example 2. The set of $n \times n$ real symmetric matrices, which obey

$$(S_{ij})^t = S_{ji} = +S_{ij} \qquad (S^t = S) \tag{1.28}$$

is a linear subspace of the vector space discussed in Example 1. However, if we adjoin the multiplication operation of Example 1, we do not satisfy Postulate C-1 for an algebra. That is, the product of two symmetric matrices is not in general a symmetric matrix:

$$\begin{aligned}(ST)^t &= (T)^t(S)^t = TS \neq ST \\ (S_{ij}T_{jk})^t &= T_{kj}S_{ji} \neq S_{kj}T_{ji}\end{aligned} \tag{1.29}$$

However, if we define the operation $\square$ by

$$\begin{aligned}S \square T = [S, T]_+ &= ST + TS \\ [S, \alpha T_1 + \beta T_2]_+ &= \alpha[S, T_1]_+ + \beta[S, T_2]_+\end{aligned} \tag{1.30}$$

then both postulates C-1 and C-2 are satisfied. The real symmetric $n \times n$ matrices form an algebra under symmetrization, or **anticommutation**.

Example 3. The set of $n \times n$ real antisymmetric matrices

$$\begin{aligned}A^t &= -A \\ A_{ij} &= -A_{ji}\end{aligned} \tag{1.31}$$

is not closed under matrix multiplication either. But if we define the combinatorial operation $\square$ by antisymmetrization,

$$\begin{aligned}A \square B = [A, B] &= AB - BA \\ [A, \beta B + \gamma C] &= \beta[A, B] + \gamma[A, C]\end{aligned} \tag{1.32}$$

postulates C-1 and C-2 are satisfied and this system forms an algebra.

It is easily verified that this algebra in general has no identity, nor is it associative:

$$\begin{aligned}A \square (B \square C) &= ABC - ACB - BCA + CBA \\ &\;\;\not\parallel \qquad\qquad \not\parallel \\ (A \square B) \square C &= ABC - BAC - CAB + CBA\end{aligned} \tag{1.33}$$

An algebra with the antisymmetric multiplication defined by the commutation relations (1.32) is called a **Lie algebra**, provided this combinatorial operation also obeys Postulate C-6:

$$A \square (B \square C) = (A \square B) \square C + B \square (A \square C)$$

This property, called a **derivation**, may be written more familiarly as

$$[A, [B, C]] = [[A, B], C] + [B, [A, C]]$$

or

$$[A, [B, C]] + [C, [A, B]] + [B, [C, A]] = 0 \tag{1.34}$$

The latter form is called Jacobi's identity.

The process of accreting additional structure and complexity in going from a set to an algebra is shown schematically in Table 1.1. In general, the more highly structured a system is, the more we can prove about it. On the other hand, results that are true for a less structured system are also true, whenever applicable, in more highly structured systems.

TABLE 1.1

THE INCREASING COMPLEXITY OF THE VARIOUS MATHEMATICAL SYSTEMS OF USE TO A PHYSICIST

Number and Kinds of Elements	Number and Kinds of Operations: 0	1 Group Multiplication	2 Group Multiplication for Abelian Groups + (Vector Addition); Scalar Multiplication ∘	3 Abelian Addition +, Scalar Multiplication ∘, Algebraic Multiplication □
1	Set	Group Section I.2 Postulates 1-4	Field Section I.3 Postulates $A1$–$A5$ Postulates $B1$–$B5$	
2			Vector Space Section I.4 Postulates $A1$–$A5$ Postulates $B1$–$B4$	Algebra Section I.5 Postulates $A1$–$A5$ Postulates $B1$–$B4$ Postulates $C1$, $C2$

II. Bases

Bases have been introduced in conjunction with linear vector spaces. This is a matter of convenience, since it is much easier to keep track of a small number of basis vectors than it is to account for every possible vector within a vector space.