PRELIMINARIES

The invention of coordinates by Pierre de Fermat (1601–1665) and René Descartes (1596–1650) united what had been seen as the separate realms of geometry and algebra. Still deeper connections were revealed by the subsequent development of new kinds of geometry and the systematization of algebra. Before we proceed to examine some of those connections, it will be useful to set down a few basic facts about the geometries and algebraic systems themselves.

A. *Euclidean and other geometries.* Euclid's *Elements* deals with relations among points, lines, and planes and properties of geometric figures such as triangles, circles, and spheres. Among the fundamental concepts of Euclidean plane geometry are *collinearity*, *congruence*, *perpendicularity*, and *parallelism*. A rigorous treatment also involves *order* and *continuity*—relations not explicitly dealt with in the *Elements*. By omitting or modifying some of these concepts, a variety of other geometries can be constructed: the real affine and projective planes, the real inversive sphere, and the so-called non-Euclidean geometries. Like Euclidean geometry, all of these alternative systems have extensions to higher-dimensional spaces.

Any two points in the *Euclidean plane* E^2 are joined by a unique line, and lines are of infinite extent. Distances and areas can be measured with the aid of an arbitrarily chosen unit of length. Right angles provide a standard for angular measure. The Euclidean parallel

postulate is equivalent to the assertion that through any point not on a given line there can be drawn *just one* line that does not intersect it. (The other postulates imply the existence of *at least one* such line.) From these assumptions it follows that the sum of the interior angles of every triangle is equal to two right angles, and that (the area of) the square on the hypotenuse of a right triangle is equal to the sum of (the areas of) the squares on the other two sides.

The real *affine plane* A^2 is the Euclidean plane without perpendicularity. There is no way to measure angles, and distances can be compared only for points on a line or on parallel lines. Nevertheless, areas can still be determined. Up to size, all triangles are equivalent, as are all parallelograms; there is no such thing as a right triangle or a square. Conics can be distinguished only as ellipses, parabolas, and hyperbolas—there are no circles.

By adopting the convention that all the affine lines parallel in a given direction meet in a unique "point at infinity" and that all such points lie on a single "line at infinity," we eliminate parallelism. When we admit the new points and the new line into the fold with the same rights and privileges as all the others, we have the real *projective plane* P^2. Incidences now exhibit a "principle of duality": any two points are joined by a unique line, and any two lines meet in a unique point. Angular measure, distance, and area are all undefined. Not only all triangles but all quadrilaterals are alike, and there is only one kind of nondegenerate conic. Congruence, perpendicularity, and parallelism have all disappeared; only the notion of collinearity remains.

Alternatively, the Euclidean plane can be given the topology of a sphere by adjoining a single "exceptional point" common to all lines. A line may then be regarded as a kind of circle. Extended lines and ordinary circles together form a set of "inversive circles" on the real *inversive sphere* I^2. Any three points lie on a unique inversive circle; points lying on the same circle are *concyclic*. Two circles may meet in two, one, or no real points. The distance between two points cannot be measured, but the angle between two intersecting

circles can be. Thus collinearity has been replaced by concyclicity, and perpendicularity is still meaningful, but congruence and parallelism have been eliminated.

Though long suspected of being a theorem in disguise, the parallel postulate was eventually shown to be independent of the other assumptions governing the Euclidean plane. Replacing it with the contrary hypothesis—that through any point not on a given line there is *more than one* line not intersecting it—we obtain the *hyperbolic plane* of Bolyai and Lobachevsky. Moreover, if we do not assume that lines are of infinite length, we can construct a metrical geometry in which there are *no* nonintersecting lines: in the *elliptic plane* (the projective plane with a metric), any two lines meet in a point.

On the *elliptic sphere* (or simply "the sphere"), points come in antipodal pairs, and the role of lines is played by great circles; any two nonantipodal points lie on a unique great circle, and any two great circles meet in a pair of antipodal points. When antipodal points are identified, great circles of the elliptic sphere become lines of the elliptic plane (the two geometries are sometimes distinguished as "double elliptic" and "single elliptic" planes). Another possibility is the *hyperbolic sphere*, comprising two antipodal hemispheres separated by an "equatorial circle" of self-antipodal points. Two great circles either meet in a pair of antipodal points, are tangent at an equatorial point, or do not meet. Identification of antipodal points converts great circles of the hyperbolic sphere into lines of the hyperbolic plane.

The hyperbolic and elliptic planes and the elliptic sphere constitute the classical non-Euclidean geometries. Along with the hyperbolic sphere, they share with the Euclidean plane the notions of collinearity (or concyclicity), congruence, and perpendicularity. One notable difference is that the sum of the interior angles of a non-Euclidean triangle depends on its area, being proportionally greater than two right angles for an elliptic (spherical) triangle or proportionally less for a hyperbolic one.

Other properties of Euclidean and non-Euclidean geometries ("real metric spaces") and how they are related to real affine, projective, or inversive geometry will be described in greater detail beginning in Chapter 1. Although each geometry can be based on a selected set of postulates, a more instructive approach characterizes geometries by their transformation groups.

B. *Algebraic systems.* A *group* is a nonempty set G and a binary operation $G \times G \to G$, with $(a,\ b) \mapsto ab$, satisfying the associative law $(ab)c = a(bc)$, having an identity element, and with each element having a unique inverse. The group operation is commonly taken as multiplication, with the identity element denoted by 1 and the inverse of a by a^{-1}. The commutative law $ab = ba$ may or may not hold; when it does, the group is said to be *abelian*. Additive notation is sometimes used for abelian groups, with the identity denoted by 0 and the inverse of a by $-a$. The number of elements in a group G is its *order* $|G|$.

A subset of a group G that is itself a group with the same operation is a *subgroup*. For each element a of a (multiplicative) group G, the set of all distinct integral powers of a forms an abelian subgroup $\langle a \rangle$; the order of $\langle a \rangle$ is the *period* of a. The *center* $\mathrm{Z}(G)$ is the subgroup of elements that commute with every element of G.

A subgroup H of a group G is said to be *normal* (or "self-conjugate") if for any element $h \in H$ its conjugate ghg^{-1} by any element $g \in G$ is in H, i.e., if $gHg^{-1} \subseteq H$ for all $g \in G$; we write this as $H \lhd G$. If H is a normal subgroup of G, then for every $g \in G$, the left coset $gH = \{gh : h \in H\}$ is the same as the right coset $Hg = \{hg : h \in H\}$, and the set G/H of all such cosets is a group—the *quotient group* (or "factor group") of G by H—with $(g_1H)(g_2H) = (g_1g_2)H$. The number $|G : H|$ of cosets is the *index* of H in G. The center $\mathrm{Z}(G)$ of any group G is always normal, and $G/\mathrm{Z}(G)$ is the *central quotient group*.

If a and b are two elements of a group G, the element $a^{-1}b^{-1}ab$ is their *commutator*; this differs from the identity precisely when

$ab \neq ba$. The set of commutators generates a normal subgroup of G, the *commutator subgroup* (or "derived group") G'. The abelian group G/G' is the *commutator quotient group*.

If H and K are subgroups of a (multiplicative) group G having only the identity element 1 in common, if every element $g \in G$ is the product of some $h \in H$ and some $k \in K$, and if $hk = kh$ for every $h \in H$ and every $k \in K$, then G is the *direct product* of H and K, and we write $G = H \times K$. Necessarily both H and K are normal subgroups of G.

A group G may be presented in terms of a subset S of *generators*, and we write $G = \langle S \rangle$, if every element of G can be expressed as a product of (positive or negative) powers of elements of S. The generators satisfy certain *relations*, and G is the largest group for which the specified relations (but no others independent of them) hold. Thus the *cyclic* group C_p, of order p, is the group generated by a single element a satisfying the relation $a^p = 1$, while the *dihedral* group D_p, of order $2p$, is generated by elements a and b satisfying the relations $a^2 = b^2 = (ab)^p = 1$. For $p \geq 3$, these are, respectively, the rotation group and the full symmetry group (including reflections) of a regular p-gon.

A *transformation* is a permutation of the elements of an arbitrary set—e.g., some or all of the points of a space—or a mapping of one set into another. All the permutations of a given set S form the *symmetric* group $\mathrm{Sym}(S)$, any subgroup of which is a *transformation group* acting on S. A permutation of a finite set is *even* or *odd* according as it can be expressed as the product of an even or an odd number of transpositions interchanging two elements. The symmetric group on a set with n elements is denoted by S_n and has order $n!$. For $n \geq 2$, the even permutations constitute a subgroup of index 2 in S_n, the *alternating* group A_n, of order $\frac{1}{2}n!$.

When the points of a geometry are assigned suitable coordinates, each transformation preserving the fundamental properties of the geometry is represented by a particular type of invertible matrix, and groups of transformations correspond to multiplicative groups of matrices. Coordinates and matrix entries may be real numbers, or they

may belong to more general number systems, e.g., *rings*. If we define a *semigroup* as a nonempty set with an associative binary operation but possibly lacking an identity element or inverses, then a ring R is a set whose elements form both an additive abelian group and a multiplicative semigroup, satisfying the distributive laws $a(b+c) = ab + ac$ and $(a+b)c = ac + bc$. The additive identity of R is its *zero*, and the multiplicative identity (if any) is its *unity*. The *trivial* ring 0 has only one element.

A ring in which multiplication is commutative is a *commutative ring*. An *integral domain* is a nontrivial commutative ring with unity without zero divisors, i.e., such that $ab = 0$ implies that either $a = 0$ or $b = 0$. An integral domain in which every element $a \neq 0$ has an inverse a^{-1}, so that the nonzero elements form a multiplicative group, is a *field*. Among the systems of interest are the integral domain $\mathbb{Z}$ of integers and the fields $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ of rational, real, and complex numbers. For each prime or prime power q there is a finite field $\mathbb{F}_q$ with q elements; such systems, also called *Galois fields* and denoted GF(q), were first investigated by Évariste Galois (1811–1832).

A ring R is *ordered* if it has a nonempty subset R^+ of *positive* elements, closed under both ring operations, such that for each element a in R just one of three cases holds: $a \in \mathsf{R}^+$, $a = 0$, or $-a \in \mathsf{R}^+$; in the last case, a is said to be *negative*. Then $a < b$ if and only if $b - a$ is positive. The ring $\mathbb{Z}$ and the fields $\mathbb{Q}$ and $\mathbb{R}$ are ordered, but the field $\mathbb{C}$ is not. (Either i or $-$i would have to be positive, but their squares both equal the negative number -1.) No finite ring can be ordered. The real field $\mathbb{R}$, which has additional properties of continuity, is a *complete* ordered field.

A transformation T mapping the elements of a group, ring, or other algebraic system U to a similar system V, written $T : U \to V$, is a *homomorphism* if it preserves the system operation(s), carrying sums or products in the *domain* U into sums or products in the *codomain* V. The *kernel* Ker T is the set of elements in U that are mapped into the identity element of V (the zero element in the case

of a ring homomorphism). The *image* (or "range") Img T is the set of elements in V to which elements of U are mapped. When the systems are groups, Ker T is a normal subgroup of U and Img T is a subgroup of V.

The mapping $T : U \to V$ is a *monomorphism* or "one-to-one" transformation if Ker T contains only the identity element of U; it is an *epimorphism* or "onto" transformation if Img T is the entirety of V. A homomorphism taking U to V that is both one to one and onto has an inverse taking V to U that is also a homomorphism. Such a mapping is called an *isomorphism* (we write $U \cong V$) or, if U and V are the same system, an *automorphism*. If a is a fixed element of a group G, the mapping $x \mapsto axa^{-1}$ ("conjugation by a") is an *inner automorphism* of G.

When transformations of geometric points are expressed as algebraic homomorphisms, successive operations are normally carried out from left to right, as in the diagram

$$U \xrightarrow{T_1} V \xrightarrow{T_2} W$$

The *product* (T_1 followed by T_2) of the homomorphisms $T_1 : U \to V$ and $T_2 : V \to W$ is then the homomorphism $T_1T_2 : U \to W$, with $x(T_1T_2)$ defined as $(xT_1)T_2$.* Multiplication of homomorphisms is always associative: $(T_1T_2)T_3 = T_1(T_2T_3)$. Any algebraic system has at least the identity automorphism $x \mapsto x$, and every automorphism has an inverse. Thus the set of all automorphisms of an algebraic system forms a group.

C. *Linear algebra.* Of primary importance in our study of geometries and transformations are *vector spaces*, additive abelian groups

* Homomorphisms may be distinguished from ordinary functions, which typically precede their arguments and so are normally composed from right to left. Besides allowing them to be carried out in the order they are written, left-to-right composition of point mappings is compatible with transformations of dual systems (e.g., left and right vector spaces) as well as systems in which multiplication is noncommutative.

whose elements ("vectors") can be multiplied by the elements of a field ("scalars"). The set of vectors is closed under such "scalar multiplication"; i.e., if $\boldsymbol{V}$ is a (left) vector space over a field F, then for all scalars λ in F and all vectors $\mathbf{x}$ in $\boldsymbol{V}$, $\lambda\mathbf{x}$ is also in $\boldsymbol{V}$. Scalar multiplication also has the properties

$$\lambda(\mathbf{x}+\mathbf{y}) = \lambda\mathbf{x}+\lambda\mathbf{y}, \quad (\kappa+\lambda)\mathbf{x} = \kappa\mathbf{x}+\lambda\mathbf{x}, \quad (\kappa\lambda)\mathbf{x} = \kappa(\lambda\mathbf{x}), \quad 1\mathbf{x} = \mathbf{x}.$$

For each positive integer n, the canonical vector space F^n comprises all lists $(x_1,\ldots,x_n)$ of n elements of F (the "entries" of the list), with element-by-element addition and scalar multiplication. The vector space F^1 is the field F itself.

Given an ordered set $[\mathbf{x}_1,\ldots,\mathbf{x}_k]$ of vectors in a vector space $\boldsymbol{V}$, for any ordered set of scalars $(\lambda_1,\ldots,\lambda_k)$ the vector $\lambda_1\mathbf{x}_1+\cdots+\lambda_k\mathbf{x}_k$ is a *linear combination* of the vectors. If S is a subset of $\boldsymbol{V}$ and if every vector $\mathbf{x}$ in $\boldsymbol{V}$ can be expressed as a linear combination of vectors in S, the set S *spans* $\boldsymbol{V}$. If a linear combination of distinct vectors $\mathbf{x}_i$ in a set S is the zero vector only when all the scalar coefficients λ_i are zero, the vectors in S are *linearly independent*. If the vectors in an ordered set S spanning a vector space $\boldsymbol{V}$ are linearly independent, S is a *basis* for $\boldsymbol{V}$, and the expression for each $\mathbf{x}$ in $\boldsymbol{V}$ is unique. Every vector space has a basis, and the number of basis vectors, called the *dimension* of the vector space, is the same for any basis. (The empty set is a basis for the zero-dimensional vector space $\mathbf{0}$.)

A vector-space homomorphism, preserving vector sums and scalar multiples, is a *linear transformation*. (When scalars are written on the left, linear transformations go on the right, and vice versa.) If $\boldsymbol{V}$ is a vector space over F, a linear transformation $\boldsymbol{V}\to\mathsf{F}$ is a *linear form* on $\boldsymbol{V}$. In our treatment, coordinates of points and hyperplanes function as row and column vectors, and geometric operations—expressed algebraically as linear transformations of coordinates—are represented by matrices. Basic geometric properties, such as distances and angles, are defined by means of *bilinear forms*, functions $\boldsymbol{V}\times\boldsymbol{V}\to\mathsf{F}$ that map pairs of vectors into scalars, preserving linear combinations. The

relevant theory of finite-dimensional vector spaces will be developed beginning in Chapter 4.

Many algebraic systems can be dualized. In particular, corresponding to each linear form on a given vector space $\boldsymbol{V}$ is a *covector* of the *dual* vector space $\check{\boldsymbol{V}}$. The *annihilator* of a vector $\mathbf{x} \in \boldsymbol{V}$ is the set of covectors $\check{\mathbf{u}} \in \check{\boldsymbol{V}}$ for which the corresponding linear form maps $\mathbf{x}$ to 0. If $\boldsymbol{V}$ is a left vector space, its dual $\check{\boldsymbol{V}}$ is a right vector space, and vice versa. If the elements of $\boldsymbol{V}$ are rows, the elements of $\check{\boldsymbol{V}}$ are columns. When $\boldsymbol{V}$ is finite-dimensional, the dual of $\check{\boldsymbol{V}}$ is isomorphic to $\boldsymbol{V}$, so that $\boldsymbol{V}$ and $\check{\boldsymbol{V}}$ are mutually dual vector spaces.

If $\boldsymbol{V}$ is an n-dimensional vector space over a field F, we may express the fact that a covector $\check{\mathbf{u}} \in \check{\boldsymbol{V}}$ belongs to the annihilator of a vector $\mathbf{x} \in \boldsymbol{V}$ (and vice versa) by writing $\mathbf{x} \lozenge \check{\mathbf{u}}$. A one-to-one linear cotransformation $\boldsymbol{V} \to \check{\boldsymbol{V}}$ mapping each vector $\mathbf{x}$ to a covector $\check{\mathbf{x}}$ is a *polarity* provided that $\mathbf{x} \lozenge \check{\mathbf{y}}$ whenever $\mathbf{y} \lozenge \check{\mathbf{x}}$, and vectors $\mathbf{x}$ and $\mathbf{y}$ are said to be *conjugate* in the polarity. These concepts can be extended to the $(n-1)$-dimensional projective space $\mathrm{P}\boldsymbol{V}$ whose "points" are one-dimensional subspaces $\langle \mathbf{x} \rangle$ spanned by nonzero vectors $\mathbf{x} \in \boldsymbol{V}$.

A *module* has the structure of a vector space except that scalars are only required to belong to a ring. An *algebra* $\mathbf{A}$ is a module over a ring R in which there is also defined a multiplication of module elements, distributive over addition and such that

$$\lambda(\mathbf{xy}) = (\lambda\mathbf{x})\mathbf{y} = \mathbf{x}(\lambda\mathbf{y})$$

for all λ in R and all $\mathbf{x}$ and $\mathbf{y}$ in $\mathbf{A}$. If each nonzero element has a multiplicative inverse, $\mathbf{A}$ is a *division algebra*.

D. *Analysis.* The assignment of coordinates establishes a correspondence between points of a geometric line and elements of some number system. When this system is an ordered field (e.g., $\mathbb{Q}$ or $\mathbb{R}$), sets of collinear points have a definite linear or cyclic *order*, which can be described, following Moritz Pasch (1843–1930), in terms of one

point lying between two others or, following Giovanni Vailati (1863–1909), one pair of points separating another pair. The order relation can be used to define line segments, rays, and the like.

The property that sets real and complex geometries apart from others is *continuity*, which essentially means that no points are "missing" from a line. The notion of continuity is implicit in the theory of proportion developed by Eudoxus (fourth century BC) and presented in Book V of Euclid's *Elements*. As we shall see in Chapter 2, a formal definition can be based on the theory of rational "cuts" invented by Richard Dedekind (1831–1916), so that each point but one of a "chain" corresponds to a unique real number. (If the definition of continuity were modified to allow infinitesimal quantities, one could even identify the points of a line with the "hyperreal" numbers of nonstandard analysis.)

E. *Arithmetic.* A nonzero integer b is a *divisor* of an integer a if there is an integer c such that $a = bc$. A positive integer p greater than 1 whose only positive divisors are 1 and p itself is a *prime*. If a and b are integers with $b > 0$, then there exist unique integers q (the "quotient") and r (the "remainder"), such that $a = bq + r$ with $0 \leq r < b$. The process of determining q and r is called the *division algorithm*.

The *greatest common divisor* of two nonzero integers a and b is the largest integer that is a divisor of both; we denote it by $\gcd(a,\ b)$ or simply (a, b). When $(a,\ b) = 1$, a and b are said to be *relatively prime*. Given two nonzero integers a and b, we can repeatedly apply the division algorithm to obtain a decreasing sequence of positive integers $r_1 > r_2 > \cdots > r_k$, where

$$a = bq_1 + r_1,\ b = r_1q_2 + r_2,\ r_1 = r_2q_3 + r_3,\ \ldots,\ r_{k-1} = r_kq_{k+1} + 0.$$

Then r_k, the last nonzero remainder, is the greatest common divisor (a, b). This process, described in Book VII of the *Elements*, is called the *Euclidean algorithm*.