Mathematical notions used in the theory of statistical shape analysis

Enrique del Castillo

Dept. of Industrial Engineering, The Pennsylvania State University, University Park, PA 16802, USA

June 7, 2011

Abstract

We review key mathematical concepts used in the theory of Statistical Shape Analysis (SSA). The treatment is elementary and aims at providing a brief guide to a large number of results and ideas which are dispersed over an also very large literature on Non-Euclidean Geometry, Differential Geometry, and Topology. The goal is to provide an introduction to the ideas SSA touches in these areas to researchers wishing to apply SSA in practice.

0.1 Relations, equivalence relations and equivalence classes

Definition 1. A *relation on a set* A is a subset, R, of $A^2 = A \times A$. Usually, relations are defined by providing a statement that singles out a collection of elements of $A \times A$ for membership in the relation. A relation R on a set A is:

- *reflexive* if for all $x \in A$, xRx.
- *symmetric* if, for all $x, y \in A$, xRy implies yRx.
- *transitive* if, for all $x, y, z \in A$, xRy and yRz imply xRz.
- an *equivalence relation* if R is reflexive, symmetric and transitive.

Example 1. Let $\mathcal{F}$ be the set of fractions of integers. Define $a/b \equiv c/d$ if $ad = bc$. Then $\equiv$ (equality of fractions) is a relation on $\mathcal{F}$. Thus, e.g., the pair $(1/2, 2/4)$ is in the subset of $\mathcal{F}$ defined by $\equiv$. Furthermore, $\equiv$ is an equivalence relation since:

1. it is reflexive: for each $a/b \in \mathcal{F}$, $a/b \equiv a/b$;
2. it is symmetric: for each $a/b, c/d \in \mathcal{F}$, if $a/b \equiv c/d$, then $c/d \equiv a/b$;

3. it is transitive: for each $a/b, c/d, e/f \in \mathcal{F}$, if $a/b \equiv c/d$ and $c/d \equiv e/f$, then $a/b \equiv e/f$.

Equivalence relations are sometimes written with the symbol $\sim$; thus, $x \sim y$ is read "x is equivalent to y". Two elements of a set do not need to be equal to be equivalent, they need only to share a specified property.

Definition 2. Let A be a set and let $\sim$ be an equivalence relation defined on this set. For each $a \in A$, the *equivalence class of* a is a subset, denoted $[a]_\sim$, consisting of all elements of A that are equivalent to a, i.e.,

$$[a]_\sim = \{x \in A : x \sim a\}$$

If there is no ambiguity about the equivalence relation one is talking about, the corresponding equivalence class is written $[a]$. Here the word "class" has been used historically to simply mean a set. Other names for equivalence class is an *orbit*, and, in case the underlying set is a manifold (see 0.3 below) they are also called a *fibre*.

Example 2. In example 1, with $\equiv$ being the equivalence relation, $[1/2] = \{x \in \mathcal{F} : x \equiv 1/2\} = \{a/b \in \mathcal{F} : 2a = b\}$ is the set of all integer fractions equal to $1/2$, which clearly is a subset of $\mathcal{F}$.

Theorem 1. Let $\sim$ be an equivalence relation on A and let $x, y \in A$. Then 1) if $x \sim y$, then $[x] \sim [y]$; 2) if $x \nsim y$, then $[x] \cap [y] = \emptyset$; 3) $A = \bigcup_{x \in A} [x]$.

Proof of 3): each equivalence class is a subset of A by definition. Each $x \in A$ is in the equivalence class $[x]$. Therefore, A is contained in the union of the equivalence classes of all the elements of A. Since from part 2) distinct equivalence classes do not intersect, this union is actually equal to set A.

Part 3) of the result above means that the set of all equivalence classes implied by an equivalence relation $\sim$ forms a *partition* of A. Parts 1) and 2) say that if two equivalence classes have an element in common, then they are identical, or, in other words, that two distinct equivalence classes are always disjoint.

Definition 3. Let $\sim$ be an equivalence relation in A. The set of all equivalence classes is called A *modulo* $\sim$ or the *quotient of A by the equivalence relation* $\sim$, and is denoted $A/\sim$. The *projection map* $\pi : A \to A/\sim$ sends $x \in A$ to its equivalence class $[x]$. If the set $A/\sim$ is closed under arbitrary unions and finite intersections (properties that define a *topology*), this set is called the *quotient space* of A by the equivalence relation $\sim$.

Example 3. Quotient spaces $A/\sim$ (and equivalence classes) are usually created by *identifying a subset of A to a point.* For instance, let $A = [0, 1]$ (unit interval on $\mathbb{R}$) and define the quotient space obtained from A by identifying the two endpoints $\{0, 1\}$ to be equivalent to the same point. Let S^1 be the unit circle on the complex plane $\mathbb{C}$. The function $f : A \to S^1$, $f = \exp(2\pi ix)$ equals the same value (1) at 0 and at 1, and hence it induces a function

$$g : A/\sim \to S^1.$$

Geometrically, points on the unit interval A are being mapped into the points of the unit circle on $\mathbb{C}$, with the two endpoints in A mapping into the same point on $\mathbb{C}$, namely the point (1,0).

When one defines an equivalence relation on a set, one is usually interested in the set $A/\sim$. But to be specific, a given equivalence class has to be described by one of its elements. Thus, if $[x] \in A/\sim$, we need to choose an element $a \in [x]$ to be a *representative* (sometimes called an *icon*) of $[x]$. Representatives are not unique, of course.

Definition 4. A *homeomorphism* is a mapping in Euclidean space from one object onto another that is continuous and one to one, i.e., it establishes a one to one correspondence between points in each figure. The inverse mapping has the same properties. For example, a sphere in $\mathbb{R}^3$ and a cube are homeomorphic.

The concept of homeomorphism is used to define the properties of objects (figures) that remain unchanged under continuous deformation ("rubber band deformations"). These properties are called the *topological* properties of the objects. Topological properties stand in contradistinction with *metrical* properties, which are associated with distances between points, angles between lines, and edges of a figure, properties that are preserved under rigid body transformations only.

Definition 5. An *n-sphere* S^n is a set of points in $(n + 1)$-Euclidean space such that $S^n = \{x \in \mathbb{R}^{n+1} : ||x|| = r\}$ where the radius r is usually set to one (giving the unit n-sphere). A similar definition exists in case the base space is complex: a complex n-sphere is defined as $S^n = \{z \in \mathbb{C}^{n+1} : ||z|| = r\}$. The notation S^n refers to the dimension of the surface of the sphere. The n-sphere can be described as $S^n = \mathbb{R}^n \cup \{\infty\}$, which is n-dimensional Euclidean space plus a single point representing infinity in all directions (this representation gives origin to the real projective space, see Definition 7). Alternatively, if a single point is removed from an n-sphere, it becomes homeomorphic to $\mathbb{R}^n$.

0.2 Groups and transformations

Definition 6. A *group* is a set G with a binary operation $*$ (sometimes called "multiplication") such that the operation: a) is associative, b) has an identity, and c) has an inverse operation. If in addition, d) $*$ is commutative, then the group is said to be *Abelian*, otherwise it is *non-Abelian*.

Example 4. An instance of an Abelian group is the integers with addition as the $*$ operation.

In Geometry, a *transformation* is a one-to-one correspondence $P \rightarrow P'$ among all the points in the plane (or space), i.e., a rule for associating pairs of points, where each pair has a first point belonging to P and a second point belonging to P'. The most trivial transformation is the *identity* transformation, which leaves each point unchanged. A set of transformations is said to form a group if it contains the inverse of each and the product of any two. For instance, the *symmetry operations*, which leave a figure unchanged while permuting its parts, forms a group, the so-called *symmetry group* (or group of symmetries) of the figure.

Example 5. An instance of a non-Abelian group of transformations is $SO(n)$, the *special orthogonal group* (also called rotation group), which consists of all n-dimensional rotation matrices (orthogonal matrices with determinant equal to one) under the "composition of rotations" operation. Performing a rotation defined by matrix R_1 in a given direction followed by a second one R_2 and a third one R_3 satisfies the associative condition since $(R_1 * R_2) * R_3 = (R_1 * R_2) * R_3$, we clearly have an inverse rotation for every rotation: $R_1^{-1} * R_1 (= I)$ leaves an object in its original position, and we have an identity matrix I which is the zero rotation $R_0 = I$, with $R_0 * R_1 = R_1$. However, rotations do not commute: $R_1 * R_2 \neq R_2 * R_1$ as a 3-dimensional example can demonstrate.

$SO(n)$ can be understood as the group of symmetries of a n-sphere S^n excluding reflections. Note this is a *continuous group*, in contrast to Example 4, where the group is clearly discrete. Continuous groups are called *Lie groups*, after Sofus Lie (1842-1899). Thus, for instance, if $n = 3$, $SO(3)$ is the set of all possible rotations of a 3-dimensional sphere. $SO(n)$ is a *subgroup* of $O(n)$, the orthogonal group. The elements of this set are all $n \times n$ orthogonal matrices, not only those that have determinant one. Hence, $SO(n)$ is a subgroup of $O(n)$. The orthogonal group contains not only the *non-reflective* symmetries $SO(n)$ but also the reflective ones. Both $SO(n)$ and $O(n)$ are in turn subgroups of $GL(n)$, the general linear space of all non-singular $n \times n$ matrices.

Quotient spaces (see Def. 3) can be defined by the *action* of a group on the elements of some manifold $\mathcal{M}$ (see 0.4 below for a definition of manifold). If G is a group, then we define two points x, y in $\mathcal{M}$ to be equivalent if there is a $g \in G$ such that $y = gx$ and this defines the quotient space $\mathcal{M}/G$. In this case, the left action of G on elements of $\mathcal{M}$ define the equivalence relation and hence, the quotient space.

Euclidean geometry is only one of many possible geometries. Felix Klein, in his inaugural address a Erlangen in 1872 proposed the classification of geometries according to the groups of transformations in which the primitive concepts of each geometry remain invariant. In particular, Euclidean geometry is characterized by the group of *similarity transformations*; these are transformations that preserve the *angles* of a figure. Similarity transformations include *isometries*, that is, transformations that preserve the *lengths* of an object, such as *translations* and *rotations* (these are the so-called *rigid-body* transformations, which stand in contrast to the type of continuous or "rubber band" transformations studied in Topology, referred above), and *reflections*. Two objects are *congruent* if and only if they can be transformed into each other by an isometry. Similarity transformations include the isometries but also include *dilatations* (or *dilations*) which transform the scale of the objects. Two objects are *similar* if and only if one can be transformed into the other by similarity transformations. Similarity transformations not only preserve angles, they also preserve *ratios of distances* (for this reason some authors say they preserve the "shape" of an object).

It was Möbius who early in the XIX century showed that sequences of motions on the plane could be understood as "products" that transform the space, and who began the systematic study of congruential (length preserving), similarity (shape preserving) and affine (parallelism preserving) transformations. He showed that the most general continuous transformation that preserves "straightness" are the projective transformations, discussed next. The restatement of Möbius ideas in terms of groups only occurred until 1872 by Klein, once the concept of group was recognized.

0.3 Projective geometry and Complex projective space

Contrary to the transformations in Euclidean space, Projective geometry deals with transformations that do *not* preserve angles and lengths, namely, projections. In addition, and as it was first known during the Renaissance with perspective painting, there exist points at infinity ("vanishing points") where parallels met. Thus, projective geometry allows infinity to be put on the same footing as the finite points of the plane [7]. A natural question, first raised by Alberti during the Renaissance in his study on perspective, is this: if projections do not preserve angles and lengths, what is preserved? What is preserved under projections is the *cross ratio* of four points A,B,C,D on a line, defined by $\frac{\frac{CA}{CB}}{\frac{DA}{DB}}$.

Homogeneous coordinates (invented by Möbius) give a natural extension of the Cartesian plane $\mathbb{R}^2$ by assigning new coordinates to the points already present and creating new points including points at infinity. They are the coordinates used in projective geometry.

Definition 7. The homogeneous coordinates of a point $(X, Y) \in \mathbb{R}^2$ are all the real triplets (Xz, Yz, z) with $z \neq 0$, i.e., all real triplets (x, y, z) with $x/z = X$, and $y/z = Y$. ([7], p. 134).

If we take X, Y to be the x, y coordinates in the plane $z = 1$, then the coordinates (Xz, Yz, z) are just the coordinates of points on the line in $\mathbb{R}^3$ from the origin to (X, Y). Thus, homogeneous coordinates give a one-to-one correspondence between points $(X, Y) \in \mathbb{R}^2$ and nonhorizontal lines through the origin in $\mathbb{R}^3$. The horizontal lines, those with coordinates $(x, y, 0)$, correspond to the points at infinity. In geometrical terms, we have enlarged the $\mathbb{R}^2$ Euclidean space to the *Real Projective Space* $\mathbb{RP}^2$ by "adding a point" to $\mathbb{R}^2$ to represent infinity. From the construction of homogeneous coordinates above, $\mathbb{RP}^2$ is the set of straight lines of $\mathbb{R}^3$ which pass through the origin $(0, 0, 0) \in \mathbb{R}^3$, that is, $\mathbb{RP}^2$ is the space of all possible "directions" of $\mathbb{R}^3$ (a sphere). It is the quotient space of $\mathbb{R}^3 - \{0\}$ by the equivalence relation $(x_1, x_2, x_3) \equiv (\lambda x_1, \lambda x_2, \lambda x_3), \quad \lambda \in \mathbb{R}, \quad \lambda \neq 0$.

One can consider either real projective spaces $\mathbb{P}^n = \mathbb{RP}^n$ or complex ones ($\mathbb{P}^n = \mathbb{CP}^n$.).

Example 6. One example of a complex projective space is the so-called *Riemann sphere* (also called sometimes the Gauss sphere), which is $\mathbb{CP}^1$. The Riemann sphere arises as the space of ratios of complex numbers (w, z), not both zero, which is the space of complex lines through the origin in $\mathbb{C}^2$. The Riemann sphere can be thought as a one-to-one correspondence established between the points on a sphere sitting on $\mathbb{C}$ and the points in $\mathbb{C}$, obtained by *stereographic projection* of the plane into the sphere. This is achieved by drawing lines from the "north pole" N of the sphere into the plane $\mathbb{C}$ below. Any such nonhorizontal line pierces the sphere and touches it in one point, which is then projected into the complex plane into a single point. As the line becomes more horizontal, the point on the sphere is closer to N and the point on the plane is farther away in $\mathbb{C}$, with a horizontal line at N not touching $\mathbb{C}$ and corresponding to infinity on the plane. We thus have enlarged $\mathbb{C}$ to $\mathbb{CP}^1 = \mathbb{C} \cup \{\infty\}$. The projective completion of $\mathbb{C}$, $\mathbb{CP}^1$, is therefore topologically (i.e., quantitatively) equivalent to a sphere.

More generally, any projective space can be assigned homogeneous coordinates like those illustrated in Definition 7 for the case $\mathbb{P}^n$. These are the n independent *ratios* of the coordinates $z^0, z^1, ..., z^n$ for the $n+1$-dimensional space from which $\mathbb{P}^n$ arises:

$$z^0 : z^1 : z^2 : ... : z^n$$

(where the z's are not all zero) rather than the values of the individual z's themselves. If the z's are all real, then these coordinates describe $\mathbb{RP}^n$; if they are all complex then they describe $\mathbb{CP}^n$.

0.4 Manifolds, tangent space, submersions and immersions, parallel transport

Informally, a *manifold* is a space that can be thought as "curved" in various ways, but where, locally, (i.e., in the vicinity of each of its points) it can be approximated by ordinary Euclidean space. Manifolds can be thought of as a set of "points" tied together continuously and differentially, so that the points in any sufficient small region can be put into a one-to-one correspondence with an open set of $\mathbb{R}^n$. This correspondence furnishes a coordinate system for the neighborhood. The ideas of manifolds, their *charts* and *atlases*, were developed by Gauss when working in geodesy and cartography. In the same way that the curvilinear surface of the Earth is approximately represented by planar maps that describe small regions of the globe, which are then "glued" together to form a consistent *Atlas*, similar concepts explain the structure of a general manifold. A formal definition refers to the standard type of manifold, the *Hausdorff space*. A Hausdorff space has the defining property that, for two distinct points on the space, there are open sets containing each which do not intersect.

Example 7. The simplest example of a manifold is an open region in Euclidean space, for instance, that described by sets of solutions of systems of equations in $\mathbb{R}^n$. A more interesting example is the space of all $n \times n$ real matrices, $GL(n)$, defined as

$$GL(n) = \{\boldsymbol{X} \in \mathbb{R}^{n \times n} : det(X) \neq 0\} = det^{-1}(\mathbb{R} - \{0\}).$$

Since the determinant function,

$$det : \mathbb{R}^{n \times n} \to \mathbb{R}$$

is continuous, $GL(n)$ is an open subset of $\mathbb{R}^{n \times n}$, and is therefore an n-dimensional manifold. Likewise, a subgroup of $GL(n)$ such as the rotation group $SO(n)$, whose "points" (elements) are the $n \times n$ matrices $\{\boldsymbol{X} : \boldsymbol{X}'\boldsymbol{X} = I, \det(\boldsymbol{X}) = 1\}$ also constitutes an n-dimensional manifold.

Definition 8. (formal definition of a manifold) A Hausdorff space $\mathcal{M}$ is called a n-*dimensional manifold* if it is represented by the union of its open subsets U_i and if at each of these subsets there is a homeomorphism $\phi_i : U_i \to D^n$ which maps U_i onto an open disc in Euclidean space $\mathbb{R}^n$. The homeomorphism ϕ_i is called the *coordinate map*, the sets U_i are called a *coordinate neighborhood*, the pair (U_i, ϕ_i) is called a *chart*, and the union of charts $\bigcup\{(U_i, \phi_i)\}$ is called the *atlas* on $\mathcal{M}$. The number n is the *dimension* of the manifold. Each coordinate map determines coordinates on the set U_i since the map determines on the chart the family of continuos functions $x^1(P), ..., x^n(P)$ at point $P \in \mathcal{M}$ which can be regarded as coordinates of a variable point P.

With each atlas on a manifold there is associated the concept of *transition functions*. Consider two charts U_i and U_j and their intersection $U_i \cap U_j$. On this intersection the two coordinate maps $\phi_i(U_i \cap U_j) \in D^n$ and $\phi_j(U_i \cap U_j) \in D^n$ are defined (recall D^n is a n-dimensional open disk in $\mathbb{R}^n$). The composition of functions $\phi_{ij} = \phi_j \phi_i^{-1}$, mapping the set $\phi_i(U_i \cap U_j)$ onto the set $\phi_j(U_i \cap U_j)$ is also defined in D^n. The composite functions ϕ_{ij} are called the *glueing functions* or the *transition functions* of the atlas. If ϕ_{ij} is a smooth function(i.e., it is differentiable any number of times), the charts U_i and U_j are said to be *compatible*. Determination of the glueing functions allows to restore the whole manifold if individual charts and coordinate maps are only available.

With each point x on a smooth manifold $\mathcal{M}$ there is associated a linear n-dimensional space call the tangent space.

Definition 9. If in the neighborhood of a given point a coordinate system $x^1, ..., x^n$ is fixed (note the convention of indexing coordinates with superscripts), then at this point there naturally arise n linearly independent *tangent vectors* $e_i = \partial/\partial x^i$ that correspond to differentiations along the coordinate lines passing through the point x. The set of all tangent vectors to a point x in an n-dimensional manifold $\mathcal{M}$ forms a linear space of dimension n. This space is called the *tangent space* to the manifold at x, and is denoted $T_x M$. A *tangent bundle* $T_* M$ is the set of all pairs (x, a) where $x \in \mathcal{M}$ and a is a vector tangent at x.

Definition 10. Let f be a function whose domain is a set A. The function f is *injective* if for all a and b in A, if $f(a) = f(b)$, then $a = b$; that is, $f(a) = f(b)$ implies $a = b$. Equivalently, if $a \neq b$, then $f(a) \neq f(b)$. Thus, an injective function preserves distinctness; it never maps distinct elements of its domain to the same element of its codomain. A canonical injective function is the *inclusion function* $i : A \to B$ defined, for every $x \in A \subset B$, as $i(x) = x \in B$. That is, A is a subset of B and all elements of A are treated as elements of B as well.

Definition 11. A *surjective function* (or *onto* function) is a function whose image is equal to its codomain. Equivalently, a function f with domain X and codomain Y is surjective if for every $y \in Y$ there exists at least one $x \in X$ with $f(x) = y$. A surjective function is called a *surjection*. In a surjective function every point in the codomain is the value of $f(x)$ for at least one point x in the domain.

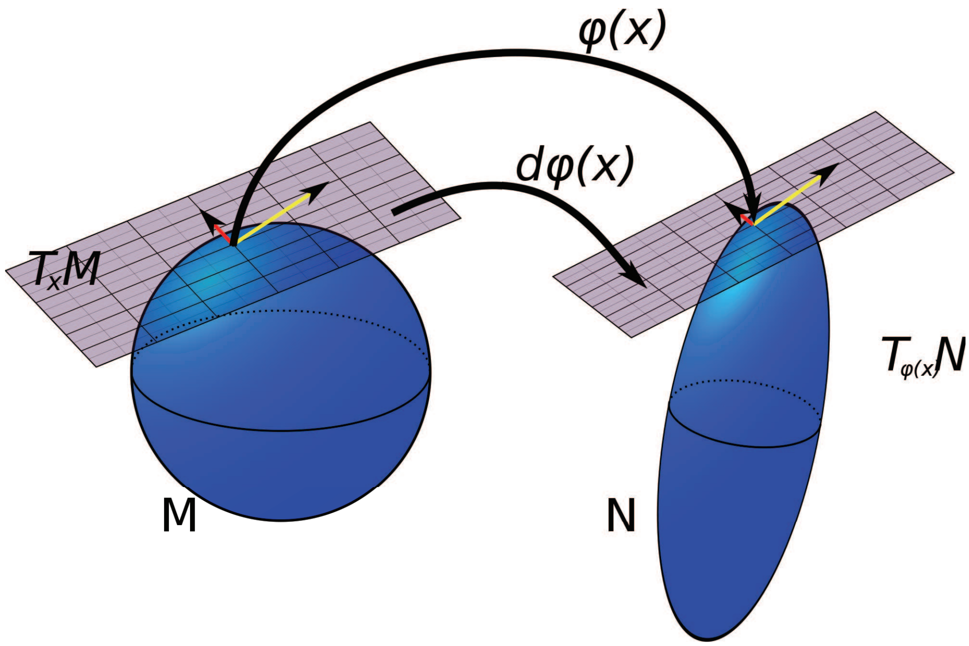

Figure 1: A smooth differential map ϕ between two manifolds. (Source: http://en.wikipedia.org/wiki/File:Pushforward.svg)

Recall that in vector calculus the *Jacobian matrix* is a matrix representation of the differential (or total derivative) of a smooth map ϕ at a point $x \in U \subset \mathbb{R}^m$ between subsets $U \subset \mathbb{R}^m$ and $V \subset \mathbb{R}^n$, i.e.,

$$d\phi_x : \mathbb{R}^m \to \mathbb{R}^n.$$

This idea can be generalized to the case ϕ is a smooth function between two manifolds $\mathcal{M}$ and $\mathcal{N}$.

Definition 12. Let $\phi : \mathcal{M} \to \mathcal{N}$. For some $x \in \mathcal{M}$, the *differential* of ϕ at x is the map

$$\phi : T_x\mathcal{M} \to T_{\phi(x)}\mathcal{N}$$

from the tangent space of $\mathcal{M}$ at x to the tangent space of $\mathcal{N}$ at $\phi(x)$. See Figure 1.

Definition 13. A smooth map between manifolds $f : \mathcal{M} \to \mathcal{N}$ is called an *immersion* if the differential $df : T_x\mathcal{M} \to T_{f(p)}\mathcal{N}$ is injective for every $p \in \mathcal{M}$. If an immersion is homeomorphic to its image it is said to be an *embedding*. The map f is called a *submersion* if df is surjective for every $p \in \mathcal{M}$. Authors speak of the smooth map f as being an immersion or a submersion at a point $x \in \mathcal{M}$, but this means that their differential df at x is injective or surjective, respectively.

Example 8. The prototype of an immersion is the inclusion of $\mathbb{R}^m$ in a *higher* dimensional $\mathbb{R}^n$:

$$i(x^1, ..., x^m) = (x^1, ..., x^m, 0, 0, ..., 0).$$

The prototype of a submersion is the projection of $\mathbb{R}^m$ onto a *lower* dimensional $\mathbb{R}^n$:

$$\pi(x^1, ..., x^n, x^{n+1}, ..., x^m) \to (x^1, ..., x^n).$$

There exist a series of theorems in Topology that indicate how finite-dimensional manifolds can always be embedded in $\mathbb{R}^m$ for sufficiently large m (e.g., the Whitney theorem).

Example 9. An important submersion in shape analysis is the *Hopf submersion* $S^3 \to S^2$ where each distinct point of a 2-sphere comes from a distinct circle (a fibre) on the 3-sphere. This can be explained in two different ways:

- Identify $\mathbb{R}^4$ with $\mathbb{C}^2$ and $\mathbb{R}^3$ with $\mathbb{C} \times \mathbb{R}$ by writing

  $$(x_1, x_2, x_3, x_4) \text{ as } z_0 = x_1 + ix_2, \text{ and } z_1 = x_3 + ix_4$$

  and

  $$(x_1, x_2, x_3) \text{ as } z = x_1 + ix_2 \text{ and } x = x_3.$$

  Thus,

  S^3 is identified with the subset $(z_0, z_1) \in \mathbb{C}^2$ such that $|z_1|^2 + |z_2|^2 = 1$ and

  S^2 is identified with the subset $(z, x) \in \mathbb{C} \times \mathbb{R}$ such that $|z|^2 + x^2 = 1$ (note: $|z|^2 = zz^*$). The *Hopf submersion* $p : S^3 \to S^2$ is then defined as

  $$p(z_0, z_1) = (2z_0 z_1^*, |z_0|^2 - |z_1|^2)$$

  where the first entry on the right hand side is a complex number and the second one is real. Thus, $p(z_0, z_1) \in \mathbb{C} \times \mathbb{R}$, and since $p(z_0, z_1) = 1$[1], it actually lies on $S^2(1)$. Furthermore, since

  $$p(z_0, z_1) = p(\lambda z_0, \lambda z_1)$$

  for some $\lambda \in \mathbb{C}$ such that $|\lambda|^2 = 1$, then different points in S^3 map to the same point on the 2-sphere. Since $|\lambda|^2 = 1$ forms a circle on $\mathbb{C}$, it follows that for each point $w \in S^2$, $p^{-1}(w) = S^1$ (a circle) on S^3. Thus, the 3-sphere is a disjoint union of circular fibres (for this reason this is also called the *Hopf fibration*).

[1]Proof: $2z_0 z_1^* \cdot 2z_0 z_1^* + (|z_0|^2 - |z_1|^2)^2 = 4|z_0|^2|z_1|^2 + |z_0|^4 - 2|z_0|^2|z_1|^2 + |z_1|^4 = 2|z_0|^2|z_1|^2 + |z_0|^4 + |z_1|^4 = (|z_0|^2 + |z_1|^2)^2 = (1)^2 = 1$.

- We can consider the complex projective space $\mathbb{CP}^1$ as equal to the quotient space of $\mathbb{C}/\{0\}$ by the equivalence relation that identifies (z_0, z_1) with $(\lambda z_0, \lambda z_1)$ for z_0, z_1 and $\lambda(\neq 0) \in \mathbb{C}$ (set of equivalence classes under multiplication by a non-zero complex number). Then, on any complex line in $\mathbb{C}^2$ (a one dimensional complex subspace that replicates the entire complex space $\mathbb{C}$) there is a unit circle, so the quotient maps circles to points. Alternatively, (z_0, z_1) can be mapped to the point z_0/z_1 (using homogeneous coordinates) on the Riemann sphere $\mathbb{C} \cup \{\infty\}$ (see example 6).

0.5 Intrinsic and extrinsic geometry and geodesics

The concept of intrinsic geometrical properties of an object originated from the work by Gauss, who, in 1827, conceived the idea of defining the curvature of a surface by measurements that take place entirely on the surface and not based on measurements on the ambient space where the surface is embedded, that is, he found a way to detect the curvature of a surface *intrinsically*. For instance, in the time of Gauss, the curvature of the earth was known on the basis of surveyors and explorers, not by viewing it from space [7].

In the case of a curved line on the plane, there is no way to define the curvature of a line by measurements confined to (intrinsic to) the line itself. One needs, for instance, an angle θ of the tangent vector with respect to some fixed direction as a function of the distance s measured along the curve, thus $\theta = \theta(s)$. Then the curvature κ, and its reciprocal, the radius of curvature ρ, are defined as $\kappa = 1/\rho = d\theta(s)/ds$ ([5], p. 335), which is an *extrinsic measure of curvature*. Before Gauss, Euler showed in 1760 how to extend the extrinsic idea of curvature to the case of a surface in three dimensional space by expressing the curvature at a point P on the surface S in terms of "plane" curves by considering sections of S by planes through the normal at P. Among the many possible such curves, there is one of maximum curvature and one of minimum curvature. Euler showed how these two curvatures, κ_1 and κ_2, called the *principal curvatures*, occur in perpendicular sections and that together define the curvature κ along a section at any angle α to one of the principal sections by the expression:

$$\kappa = \kappa_1 \cos^2 \alpha + \kappa_2 \sin^2 \alpha.$$

The sum $\kappa_1 + \kappa_2$ is called the mean curvature, and it is an extrinsic measure of curvature of a surface. This is as far as one can go in terms of extrinsic measures of curvature of surfaces in $\mathbb{R}^3$ [7]. Gauss then showed how the *product* $\kappa_1\kappa_2$ (called the *Gaussian curvature*) can be defined intrinsically, and serve as an intrinsic measure of curvature. For the plane, $\kappa_1 = \kappa_2 = 0$, whereas for the 2-sphere $\kappa_1 = \kappa_2 = 1/\rho = 1/r$, so $\kappa = 1/r^2$ (note this is a positive constant on all points on the 2-sphere).

The Gaussian curvature has the property, proved by Gauss in his "remarkable theorem"

(*Theorema Egregium*), that it is unaffected by bending[2]). Since a cylinder can be obtained from a (rectangular portion of a) plane by bending, we have that a cylinder has also Gaussian curvature equal to zero (since either κ_1 or κ_2 equal to zero). Gauss' theorem then says that if S_1 and S_2 are locally isometric, then S_1 and S_2 have the same Gaussian curvature at corresponding points. This actually provides the following definition of an intrinsic property in $\mathbb{R}^3$.

Definition 14. A property of surfaces in $\mathbb{R}^3$ is called *intrinsic* if it is preserved by *local isometries.* Two surfaces S_1 and S_2 are locally isometric if any sufficiently small portion of S_1 can be mapped isometrically (i.e., preserving arc lengths) into any part of S_2 (thus the map takes any curve on S_1 into a curve in S_2 of equal length). Local isometries between S_1 and S_2 are obtained by a bending transformation that does not include stretching, compressing, or tearing.

Example 10. A plane can be bent into a cylinder, hence they are locally isometric, and hence they have the same Gaussian curvature (zero). A sphere and a plane are not locally isometric (and therefore have different Gaussian curvatures), a fact of great importance in cartography: any planar map of the Earth induces necessarily some distortion: there is no sphere to plane transformation that preserves both angles and areas.

Note that *local* isometry is not the same as isometry (or global isometry): the plane and the cylinder are clearly not isometric (since they are not congruent, see 0.2 above), but they are locally isometric according to Definition 14.

We thus have that Gaussian curvature $\kappa_1\kappa_2$ is an intrinsic property of a surface embedded in $\mathbb{R}^3$. Separately, the principal curvatures κ_1 and κ_2 are extrinsic properties, in contrast, since they do not remain constant after an isometric transformation.

The geodesic lines, or geodesics, of a surface are a generalization of the straight lines of the plane and are fundamental in determining the intrinsic properties of a surface. There are several definitions of geodesics.

Definition 15. A *geodesic* is a curve on a surface such that every sufficiently small portion of it is the shortest path on the surface connecting the end-points of the portion. It follows that the geodesic lines of a surface continue to be geodesic if the surface is subject to bending.

[2]Bending (without stretching, compressing or tearing) is a type of isometric transformation, i.e., it preserves arc lengths, but it is *not* a "rigid" transformation, which is also an isometric transformation. Two surfaces that can be transformed into each other by bending are said to be "applicable" or that they can be "applied" to each other. Two applicable surfaces, such as a plane and a cylinder, have the same Gaussian curvature (zero in this case) [4]

Hence geodesics are fundamental in the intrinsic properties of a surface. In fact, all intrinsic properties of a surface (e.g., its Gaussian curvature) can be determined by drawing geodesics and measuring its arc lengths.

The geodesics of the sphere are its great circles. A geodesic is a curve such that its principal normal lying on the surface coincides with the normal to the surface.

0.6 Kendall's Preshape and Shape Spaces

Let $\boldsymbol{X}$ be a $k \times m$ matrix containing the k landmarks (coordinate pairs or triples) of an object in m (2 or 3) dimensions. $\boldsymbol{X}$ is sometimes called a *configuration matrix* (since it is an element of the configuration space, the space of all possible arrangements of k landmarks in m dimensions). With this notation, the shape of a configuration $\boldsymbol{X}$ is obtained, first, by removing location and scale effects by computing the so-called *pre-shape* $\boldsymbol{Z}$:

$$\boldsymbol{Z} = \frac{\boldsymbol{H}\boldsymbol{X}}{||\boldsymbol{H}\boldsymbol{X}||} \tag{1}$$

where $\boldsymbol{H}$ is a $(k-1) \times k$ Helmert submatrix (Dryden and Mardia, 1998) and $||\cdot||$ denotes the Frobenius norm of a matrix. If we define $h_j = -[j(j+1)]^{-1/2}$, then $\boldsymbol{H}$ is a matrix whose jth row is: $(\underbrace{h_j, h_j, ..., h_j}_{j \text{ times}}, -jh_j, \underbrace{0, ..., 0}_{k-j-1 \text{ times}})$ for $j = 1, ...k-1$. Note that $\boldsymbol{H}\boldsymbol{H}' = \boldsymbol{I}_{k-1}$ and that the rows of $\boldsymbol{H}$ are contrasts. Alternatively, one could start with the *centered preshapes*, defined by $\boldsymbol{Z}_c = \boldsymbol{H}'\boldsymbol{Z}$ (these are $k \times m$ matrices), although the development below assumes Helmertized preshapes where one of the k coordinates is eliminated.

Transformation (1) removes location effects via the numerator, and re-scales the configurations to unit length via the denominator. Since we have not removed rotations from $\boldsymbol{Z}$ it is not yet the shape of $\boldsymbol{X}$, hence the name preshape. The centered preshapes are equivalent to centering each coordinate of each configuration by its centroid and dividing each by its norm.

The *shape* of configuration $\boldsymbol{X}$, denoted $[\boldsymbol{X}]$, is defined as the geometrical information that is invariant to similarity transformations except reflections. In the work by Kendall (1984), reflections are *not* considered, thus, two objects, one the mirror image of the other are considered to have different shapes. Therefore, once location and scale effects are filtered as above, the shape is then defined as:

$$[\boldsymbol{X}] = \{\boldsymbol{Z}\boldsymbol{\Gamma} : \boldsymbol{\Gamma} \in SO(m)\} \tag{2}$$

where $\boldsymbol{Z}$ is the preshape of $\boldsymbol{X}$, $\boldsymbol{\Gamma}$ is a rotation matrix (i.e., a matrix such that $\boldsymbol{\Gamma}'\boldsymbol{\Gamma} = \boldsymbol{\Gamma}\boldsymbol{\Gamma}' = \boldsymbol{I}_m$ with $\det(\boldsymbol{\Gamma}) = +1$) and $SO(m)$ is the space of all $m \times m$ rotation matrices that exclude reflections, the special (or non-reflective) orthogonal group. Multiplication by a suitable matrix $\boldsymbol{\Gamma}$ reorients (rotates) the object. Note that a shape is therefore defined as a set.

The following geometrical interpretation of these transformations is due to Kendall (1984 and 1989). Given that preshapes are scaled and centered objects, they can be represented by vectors from the center to the surface of a unit sphere of dimension $(k-1)m$, because the numerator in (1) removes m degrees of freedom for location parameters and the denominator removes one additional degree of freedom for the change of scale. The preshapes, having unit length, form a space (denoted S_m^k), which has $(k-1)m-1$ dimensions by virtue of being on the surface. As one rotates a pre-shape $\boldsymbol{Z}$ via (2), the vectors $\boldsymbol{Z\Gamma}$ describe an *orbit* on S_m^k. All the vectors on an orbit correspond to the same shape, since by definition the shape of an object is invariant to rotations. Thus, the orbits (also called *fibers*) of the preshape space are mapped one to one into single points in the *shape space* (denoted Σ_m^k), the space of all possible shapes of k landmarks in m dimensions. This space in general will be a non-Euclidean M-dimensional manifold. Two objects have the same shape if and only if their preshapes lie on the same fiber. The shape space has dimension $M = (k-1)m - 1 - m(m-1)/2$ since in addition to losing location and dilation degrees of freedom we also lose $m(m-1)/2$ degrees of freedom in the specification of the (symmetric) $m \times m$ rotation matrix $\boldsymbol{\Gamma}$.

Example 11. Preshape space and shape space for lines. In order to explain these ideas, consider one of the simplest possible cases, where we have 2 lines in $\mathbb{R}^2$ (see Figure 2). Thus, we have that $m = 2$ and $k = 2$, where the obvious landmarks are the endpoints of the lines. After centering and scaling the two lines using (1), one obtains the preshapes with matrices $\boldsymbol{Z}_1$ and $\boldsymbol{Z}_2$. Since the original objects evidently have the same shape (that of a line in Euclidean space) these two preshapes lie on the same fiber or orbit, generated as the preshapes are rotated using (2). The preshape space S_2^2 is of dimension $(k-1)m-1 = 1$, namely, the circumference of a unit circle. As the preshapes rotate (they can rotate clockwise or counterclockwise) they will eventually coincide, which corresponds to the centered and scaled lines coinciding. Finally, since there is a single shape, the shape space Σ_2^2 is the simplest possible, namely, a single point (dimension is $M = (k-1)m - 1 - m(m-1)/2 = 0$, i.e., a 0-manifold).

In general, the shape space Σ_m^k will be a nonlinear space, the Riemannian M-manifold formed by the landmarks modulo similarity transformations, of reduced dimension than the always spherical preshape space. That is, the shape space is defined as a quotient space, i.e., $\Sigma_m^k = \mathbb{R}^{km}/\mathcal{G} = S_m^k/SO(m)$, where $\mathcal{G}$ is the group of similarity transformations that exclude reflections. While the step of going from configuration space (the km-manifold of all possible arrangements of the landmarks) to preshape space is easy to understand, going from preshape space to shape space is a non-trivial step. For instance, for planar shapes Kendall (1984) showed that $\Sigma_2^k = \mathbb{C}P^{k-2}(4)$, the complex projective space of sectional curvature 4 (thus in the previous example, $\Sigma_2^2 = \mathbb{C}P^0(4)$, a one-point space). See Kendall et al. [1] for a detailed discussion of the

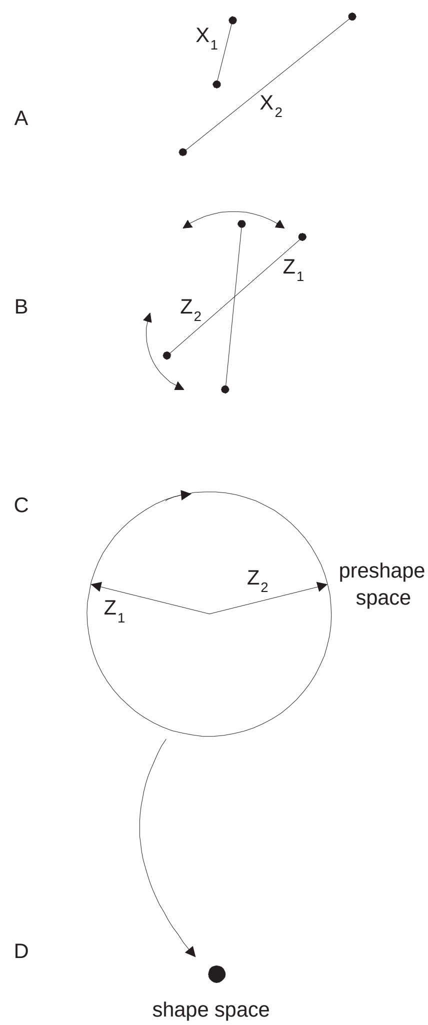

Figure 2: One of the simplest illustrations of preshape and shape space. A) two lines in the original 2-dimensional space; B), preshapes on 2-dimensional Euclidean space, after centering and scaling; C) the corresponding pre-shape space is the (one-dimensional) circumference of a unit circle. The two pre-shapes lie on the single fiber or orbit generated as the preshapes are rotated, hence there is a single shape; D) the shape space for the two lines (Σ_2^2) is zero dimensional (a single point) and corresponds to the only shape that exists in this example.

geometry of shape spaces. For most applications in manufacturing, the shapes will typically be very close in shape space, and therefore the nonlinearity of the space can be neglected. There might be, however, applications of SSA in micro-manufacturing where the assumption of low between shape variability is false.

Example 12. Preshape and shape space for planar triangles. The map $S_2^3 \to \Sigma_2^3$ is the *Hopf submersion* of Example 9, which is a map from each non-overlapping circular fibre (the preshapes) to the points in shape space. Each fibre $[x]$ in S_2^3 (point in Σ_2^3) corresponds to a particular triangular shape, the equivalence class generated by the quotient space $S_2^3/SO(2)$.

0.7 Other representations of shape

The Kendall shape space is appropriate when the landmarks of a configuration are only weakly related. D.G. Kendall originally developed the ideas behind SSA theory for the case of triangles, with application in Archeology. In other applications, specifically, manufacturing applications, landmarks that delineate outlines which do not self-intersect have a stronger relationship and this hints at the need of a different approach at representing the shapes [2].

References

[1] Kendall, D.G., Barden, D., Carne, T.K., and Le, H., (1999), *Shape and shape theory.* New York, NY: John Wiley & sons.

[2] Kendall, W.S., and Molchanov, I., (eds.), (2010). *New Perspectives in Stochastic Geometry*, Oxford University Press, NY (chapter 12).

[3] Fomenko, A., *Visual Geometry and Topology*, Springer Verlag, 1994.

[4] Hilbert, D. and Cohn-Vossen, S. (1952). *Geometry and the Imagination*, Reprinted (1999) by AMS Chelsea Publishing, Providence, Rhode Island.

[5] Misner, C.W., Thorne, K.S., and Wheeler, J.A., (1973). *Gravitation*, W.H.Freeman and Co., NY.

[6] Penrose, R., *The Road to Reality. A complete guide to the laws of the universe.* Vintage books, 2004.

[7] Stillwell, J., *Mathematics and its History*, Springer, 2000.

[8] Tu, L.W., *An introduction to manifolds*, Springer, 2007.