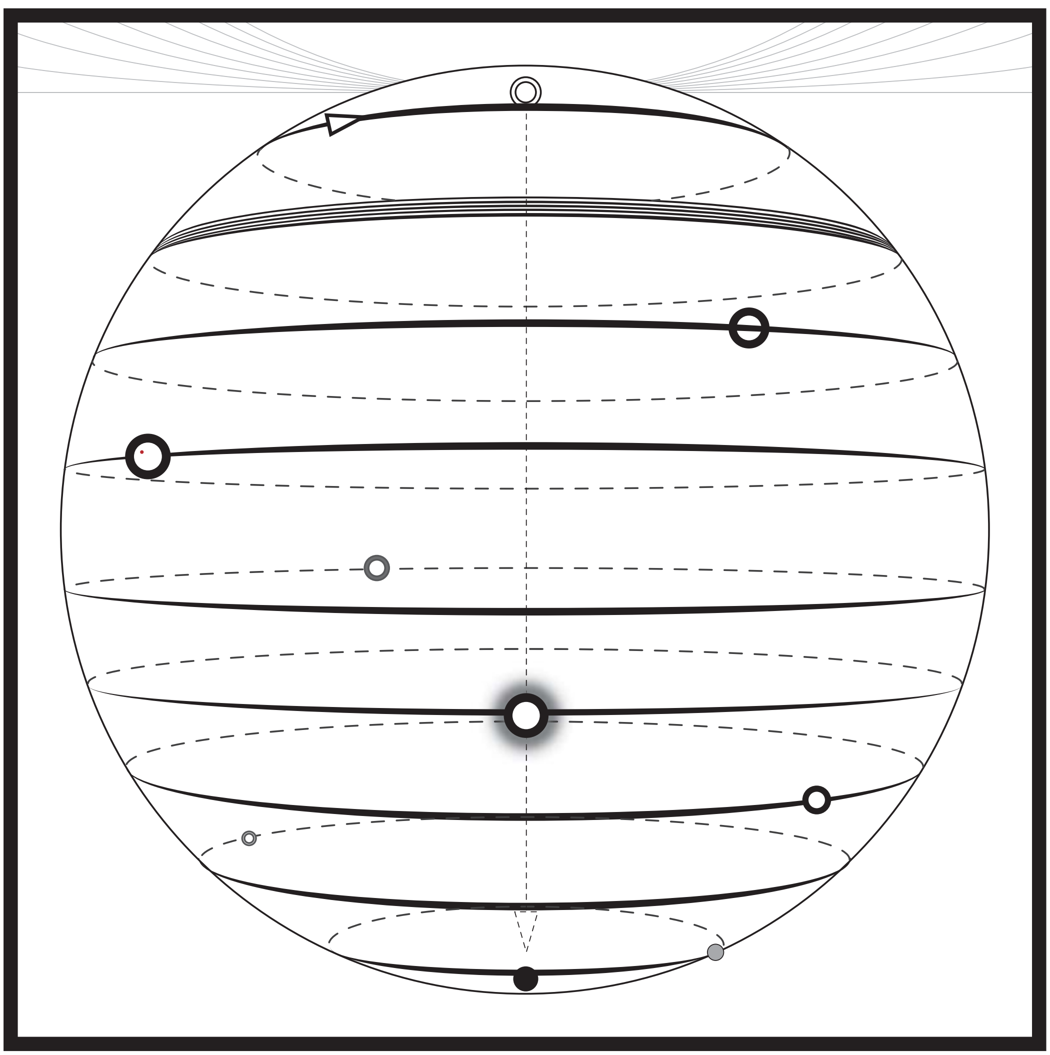
Chapter 1

Spaces: Manifolds

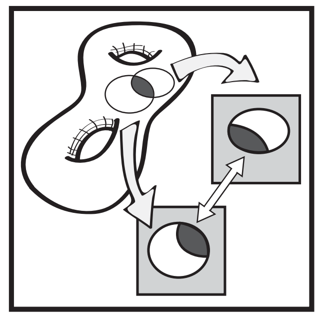


Manifolds are the extension of domains familiar from calculus – curves and surfaces – to higher-dimensional settings. As the most intuitive and initially useful topological spaces, these smooth domains provide a first glimpse of the technicalities implicit in passing from local to global.

1.1 Manifolds

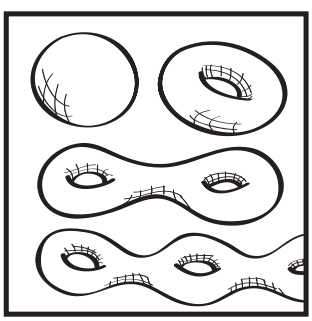
A topological n-**manifold** is a space[1] M locally homeomorphic to $\mathbb{R}^n$. That is, there is a cover $\mathcal{U} = \{U_\alpha\}$ of M by open sets along with maps $\phi_\alpha \colon U_\alpha \to \mathbb{R}^n$ that are continuous bijections onto their images with continuous inverses. In order to do differential calculus, one needs a smoothing of a manifold. This consists of insisting that the maps

$$\phi_\beta \circ \phi_\alpha^{-1} \colon \phi_\alpha(U_\alpha \cap U_\beta) \to \phi_\beta(U_\alpha \cap U_\beta)$$



are smooth (infinitely differentiable, or C^∞) whenever $U_\alpha \cap U_\beta \neq \varnothing$. The pairs (U_α, ϕ_α) are called **charts**; they generate a maximal **atlas** of charts which specifies a smooth structure on M. Charts and atlases are rarely explicitly constructed, and, if so, are often immediately ignored. The standard tools of multivariable calculus – the Inverse and Implicit Function Theorems – lift to manifolds and allow for a simple means of producing interesting examples.

Smooth curves are 1-manifolds, easily classified. Any connected curve is **diffeomorphic** (smoothly homeomorphic with smooth inverse) to either $\mathbb{R}$ or to the circle $\mathbb{S}^1$; thus, compactness suffices to distinguish the two. The story for 2-manifolds – surfaces – introduces two more parameters. Compact surfaces can be orientable or non-orientable, and the existence of holes or handles is captured in the invariant called **genus**.



The **sphere** $\mathbb{S}^2$ is the orientable surface of genus zero; the **torus** $\mathbb{T}^2$ the orientable surface of genus one; their nonorientable counterparts are the **projective plane** $\mathbb{P}^2$ and the **Klein bottle** K^2 respectively. The Classification Theorem for Surfaces states that any compact surface is diffeomorphic to the orientable or non-orientable surface of some fixed genus $g \geq 0$. The spatial universe is, seemingly, a 3-manifold[2]. The classification of 3-manifolds is a delightfully convoluted story [290], with recent, spectacular progress [233] that perches this dimension between the simple (2) and the wholly bizarre (4).

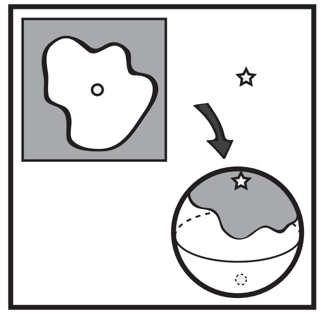
[1] All manifolds in this text should be assumed Hausdorff and paracompact. The reader for whom these terms are unfamiliar is encouraged to ignore them for the time being.

[2] This is a simplification, ignoring whatever complexities black holes, strings, and other exotica produce.

Example 1.1 (Spheres) ⊙

The n-sphere, $\mathbb{S}^n$, is the set of points in Euclidean $\mathbb{R}^{n+1}$ unit distance to the origin. The n-sphere is an n-dimensional manifold.

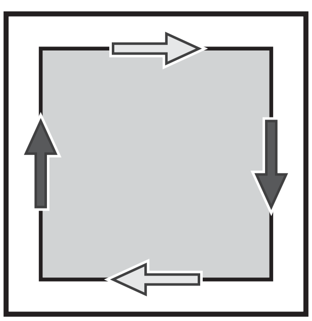
The 0-dimensional sphere $\mathbb{S}^0$ is disconnected – it is the disjoint union of two points. For $n > 0$, $\mathbb{S}^n$ is a connected manifold diffeomorphic to the compactification of $\mathbb{R}^n$ as follows. Consider the quotient space obtained from $\mathbb{R}^n \sqcup \star$, where $\star$ is an abstract point whose neighborhoods consist of $\star$ union the points in $\mathbb{R}^n$ sufficiently far from the origin. This abstract space is diffeomorphic to $\mathbb{S}^n$ via a diffeomorphism that sends the origin and $\star$ to the south and north *poles* of the sphere $\mathbb{S}^n$ respectively. ⊙



Example 1.2 (Projective spaces) ⊙

The **real projective space**, $\mathbb{P}^n$, is defined as the space of all 1-dimensional linear subspaces of $\mathbb{R}^{n+1}$, with the topology that says neighborhoods of a point in $\mathbb{P}^n$ are generated by small open cones about the associated line. That $\mathbb{P}^n$ is an n-manifold for all n is easily shown (but should be contemplated until it appears obvious). Projective 1-space, $\mathbb{P}^1$, is diffeomorphic to $\mathbb{S}^1$. The projective plane, $\mathbb{P}^2$, is a non-orientable surface diffeomorphic to the following quotient spaces:

1. Identify opposite sides of a square with edge orientations reversed;
2. Identify antipodal points on the boundary of the closed unit ball $B \subset \mathbb{R}^2$;
3. Identify antipodal points on the 2-sphere $\mathbb{S}^2$.



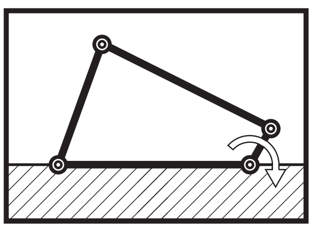
For any n, $\mathbb{P}^n$ is diffeomorphic to the quotient $\mathbb{S}^n/a$, where $a\colon \mathbb{S}^n \to \mathbb{S}^n$ is the **antipodal** map $a(x) = -x$. The space $\mathbb{P}^3$ is diffeomorphic to the space of rotation matrices, SO_3, the group of real 3-by-3 orthogonal matrices with determinant 1. Among the many possible extensions of projective spaces, the **Grassmannian** spaces arise in numerous contexts. The Grassmannian $\mathbb{G}^n_k$ is defined as the space of all k-dimensional subspaces of $\mathbb{R}^n$, with topology induced in like manner to $\mathbb{P}^n$. The Grassmannian is a manifold that specializes to $\mathbb{P}^n = \mathbb{G}^{n+1}_1$. ⊙

1.2 Configuration spaces of linkages

Applications of manifolds and differential topology are ubiquitous in rational mechanics, Hamiltonian dynamics, and mathematical physics and are well-covered in standard texts [1, 15, 217, 231]. A simple application of (topological) manifolds to robotics falls under a different aegis. Consider a planar mechanical linkage consisting of several flat, rigid rods joined at their ends by pins that permit free rotation in the plane. One can use out-of-plane height (or mathematical license) to assert that interior intersections of rods are ignorable. The **configuration space** of the linkage is a topological space that assigns a point to each configuration of the linkage – a relative positioning of the rods up to equivalence generated by rotations and translations in the plane – and which assigns neighborhoods in the obvious manner. A neighborhood of a configuration is all configurations obtainable via a small perturbation of the mechanical linkage. The configuration space of a planar linkage is almost always a manifold, the dimension of which conveys the number of mechanical degrees of freedom of the device.

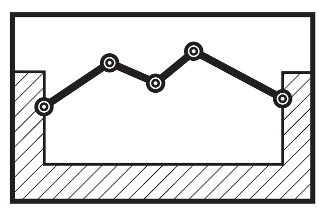
Example 1.3 (Crank-rocker) ⊙

The canonical example of a simple useful linkage is the **Grashof 4-bar**, or **crank-rocker** linkage, used extensively in mechanical components. Four rods of lengths $\{L_i\}_1^4$ are linked in a cyclic chain. When one rod is anchored, the system is seen to have one mechanical degree of freedom. The configuration space is thus one-dimensional and almost always a manifold. If one has a single short rod, then this rod can be rotated completely about its anchor, causing the opposing rod to rock back-and-forth.



This linkage is used to transform spinning motion (from, say, a motor) into rocking motion (as in a windshield wiper). The configuration space of such a linkage is $\mathbb{S}^1 \sqcup \mathbb{S}^1$, the **coproduct** or **disjoint union** of two circles. The second circle comes from taking the mirror image of the linkage along the axis of its fixed rod in the plane and repeating the circular motion there: this forms an entirely separate circle's worth of configuration states. ⊙

Many other familiar manifolds are realized as configuration spaces of planar linkages (with judicious use of the third dimension to mitigate bar crossings). The undergraduate (!) thesis of Walker [299] has many examples of orientable 2-manifolds as configuration spaces of planar linkages. A simple 5-bar linkage has configuration space which can be a closed, orientable surface of genus g ranging between 0 and 4, depending on the lengths of the edges. The reader is encouraged to try building a linkage whose configuration space yields an interesting 3-manifold. The realization question this exercise prompts has a definitive answer (albeit with a convoluted attribution and history):

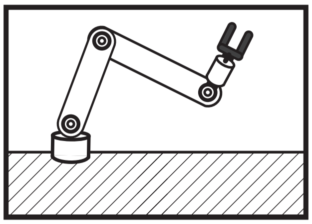


Theorem 1.4 ([189]). *Any smooth compact manifold is diffeomorphic to the configuration space of some planar linkage.*

This remarkable result provides great consolation to students whose ability to conceptualize geometric dimensions greater than three is limited: one can sense all the complexities of manifolds *by hand* via kinematics. The reader is encouraged to build a few configurable linkages and to determine the dimensions of the resulting configuration spaces.

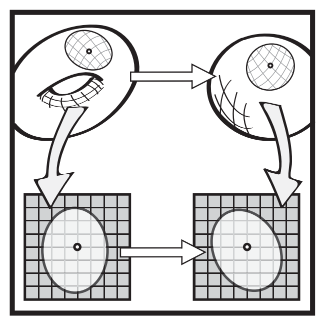
Example 1.5 (Robot arms) ⊙

A robot arm is a special kind of mechanical linkage in which joints are sequentially attached by rigid rods. One end of the arm is fixed (mounted to the floor) and the other is free (usually ending in a manipulator for manufacturing, grasping, pick-and-place, etc.).



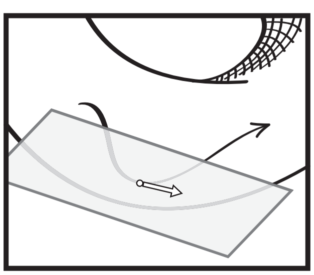
Among the most commonly available joints are pin joints (*cf.* an elbow) and rotor joints (*cf.* rotation of a forearm), each with configuration space $\mathbb{S}^1$. Ignoring the (nontrivial!) potential for collision, the configuration space of such an arm in $\mathbb{R}^3$ has the topology of the n-torus, $\mathbb{T}^n := \prod_1^n \mathbb{S}^1$, the **cartesian product** of n circles, where n is the number of rotational or pin joints. There are natural maps associated with this configuration space, including the map to $\mathbb{R}^3$ which records the location of the end of the arm, or the map to SO_3 that records the orientation (but not the position) of an asymmetric part grasped by the end manipulator. ⊚

1.3 Derivatives



Derivatives, vector fields, gradients, and more are familiar constructs of calculus that extend to arbitrary manifolds by means of localization. Differentiability is a prototypical example. A map between manifolds $f\colon M \to N$ is **differentiable** if pushing it down via charts yields a differentiable map.

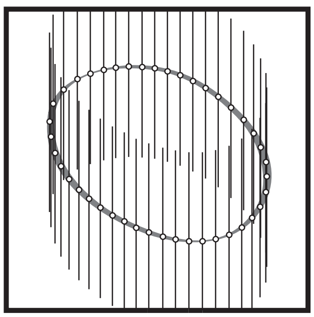
Specifically, whenever f takes $p \in U_\alpha \subset M$ to $f(p) \in V_\beta \subset N$, one has $\psi_\beta \circ f \circ \phi_\alpha^{-1}$ a smooth map from a subset of $\mathbb{R}^m$ to a subset of $\mathbb{R}^n$. The derivative of f at p is therefore defined as the derivative of $\psi_\beta \circ f \circ \phi_\alpha^{-1}$, and one must check that the choice of chart does not affect the result.



It suffices to use charts and coordinates to understand derivatives, but it is not satisfying. A deeper inquiry leads to a significant construct in differential topology. The **tangent space** to a manifold M at a point $p \in M$, T_pM, is a vector space of tangent directions to M at p, where the origin $0 \in T_pM$ is abstractly identified with p itself.

This notion is the first point of departure from the calculus mindset – in elementary calculus classes there is a general confusion between tangent vectors (*e.g.*, from vector fields) and points in the space itself. It is tempting to illustrate the tangent space as a vector space of dimension $\dim M$ that is *tangent* to the manifold, but this pictorial representation is dangerously ill-defined – in what larger space does this tangent space reside? Do different tangent spaces intersect? There are several ways to correctly define a tangent space. The most intuitive uses smooth curves. Define T_pM to be the vector space of equivalence classes of differentiable curves $\gamma\colon \mathbb{R} \to M$

where $\gamma(0) = p$. Two such curves, γ and $\tilde{\gamma}$ are equivalent if and only if $\gamma'(0) = \tilde{\gamma}'(0)$ in some (and hence any) chart. An element of T_pM is of the form $\xi = [\gamma'(0)]$, where $[\cdot]$ denotes the equivalence relation. The vector space structure is inherited from that of the chart in $\mathbb{R}^n$. A tangent vector coincides with the intuition from calculus in the case of $M = \mathbb{R}^n$. Invariance with respect to charts implies that the derivative of $f: M \to N$ at $p \in M$ is realizable as a linear transformation $Df_p: T_pM \to T_{f(p)}N$. In any particular chart, a basis of tangent vectors may be chosen to realize Df_p as the Jacobian matrix of partial derivatives at p.



The next step is crucial: one glues the disjoint union of all tangent spaces T_pM, $p \in M$, into a single space T_*M called the **tangent bundle** of M. An element of T_*M is of the form (p, V), where $V \in T_pM$. The natural topology on T_*M is one for which a neighborhood of (p, V) is a product of a neighborhood of V in T_pM with a neighborhood of p in M. In this topology, T_*M is a smooth manifold of dimension equal to $2\dim M$. For example, the tangent bundle of a circle is diffeomorphic to $\mathbb{S}^1 \times \mathbb{R}^1$. However, it is *not* the case that $T_*\mathbb{S}^2 \cong \mathbb{S}^2 \times \mathbb{R}^2$. That this is so is not so obvious.

1.4 Vector fields

The formalism of tangent spaces and tangent bundles simplifies the transition of calculus-based ideas to arbitrary manifolds; a ready and recurring example is the topology and dynamics of vector fields. A **vector field** on M is a choice of tangent vectors $V(p) \in T_p(M)$ which is continuous in p. Specifically, $V: M \to T_*M$ is a map satisfying $\pi \circ V = \mathrm{Id}_M$, for $\pi: T_*M \to M$ the projection map taking a tangent vector at p to p itself. Such a map V is called a **section** of T_*M.

As with sufficiently smooth differential equations on $\mathbb{R}^n$, vector fields can be integrated to yield a flow. Given V a vector field on M, the **flow** associated to V is the family of diffeomorphisms $\varphi_t: M \to M$ satisfying:

1. $\varphi_0(x) = x$ for all $x \in M$;
2. $\varphi_{s+t}(x) = \varphi_t(\varphi_s(x))$ for all $x \in M$ and $s, t \in \mathbb{R}$;
3. $\frac{d}{dt}\varphi_t(x) = V(x)$.

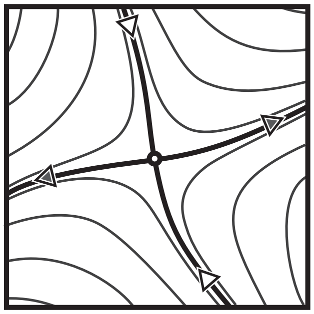
One thinks of $\varphi_t(x)$ as determining the location of a particle starting at x and moving via the velocity field V for t units of time. For M non-compact or V insufficiently smooth, one must worry about existence and uniqueness of solutions: such questions are not considered in this text. Smooth vector fields on compact manifolds yield smooth flows whose dynamics links topology and differential equations.

Example 1.6 (Equilibria) ⊚

The primal objects of inquiry in dynamics are the equilibrium solutions: a vector field is said to have a **fixed point** or **equilibrium** at p if $V(p) = 0$. An isolated fixed point may have several qualitatively distinct features based on stability. The **stable**

manifold of a fixed point p is the set

$$W^s(p) := \{x \in M : \lim_{t\to\infty} \varphi_t(x) = p\}. \quad (1.1)$$

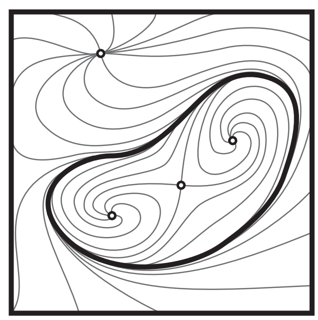


For "typical" fixed points of a "typical" vector field, $W^s(p)$ is in fact a manifold, as is the **unstable manifold**, $W^u(p)$, defined by taking the limit $t \to -\infty$ in (1.1) above [258]. A *sink* is a fixed point p whose stable manifold contains an open neighborhood of p; such equilibria are fundamentally stable solutions. A *source* is a p whose unstable manifold contains an open neighborhood of p; such equilibria are fundamentally unstable. A *saddle* equilibrium satisfies $\dim W^s(p) > 0$ and $\dim W^u(p) > 0$; such solutions are balanced between stable and unstable behavior. ⊙

Example 1.7 (Periodic orbits) ⊙

If a continuous-time dynamical system – a flow – is imagined to be supported by the *skeleton* of its equilibria, then the analogous *circulatory system* would be comprised of the periodic orbits.

A **periodic orbit** of a flow is an orbit $\{\varphi_t(x)\}_{t\in\mathbb{R}}$ satisfying $\varphi_{t+T}(x) = \varphi(x)$ for some fixed $T > 0$ and all $t \in \mathbb{R}$. The minimal such $T > 0$ is the **period** of the orbit. Periodic orbits are submanifolds diffeomorphic to $\mathbb{S}^1$. One may classify periodic orbits as being stable, unstable, saddle-type, or degenerate. The existence of periodic orbits – in contrast to that of equilibria – is a computationally devilish problem. On $\mathbb{S}^3$, it is possible to find smooth, fixed-point-free vector fields whose set of periodic orbits is all of $\mathbb{S}^3$ or empty: see Example 8.11. ⊙



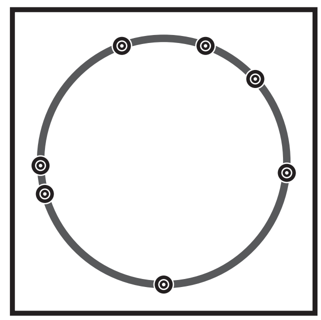
The dynamics of vector fields goes well beyond equilibria and periodic orbits (see, *e.g.*, [167, 258]); however, for typical systems, the skeleton of periodic orbits and equilibria, together with the musculature of their stable and unstable manifolds, give the basic frame for reasoning about the body of behavior.

1.5 Braids and robot motion planning

A different class of configuration spaces is inspired by applications in multi-agent robotics. Consider an automated factory equipped with mobile robots. A common goal is to place several such robots in motion simultaneously, controlled by an algorithm that either guides the robots from initial positions to goal positions (in a warehousing application), or executes a cyclic pattern (in manufacturing applications). These robots are costly and cannot tolerate collisions. As a first step at modeling such a system, assume the location of each robot is a point in a space X (typically a domain in $\mathbb{R}^2$ or $\mathbb{R}^3$). The **configuration space** of n distinct labeled points on X, denoted $\mathcal{C}^n(X)$, is the space

$$\mathcal{C}^n(X) := \left(\prod_1^n X\right) - \Delta \quad ; \quad \Delta := \{(x_k)_1^n \colon x_i = x_j \text{ for some } i \neq j\}. \tag{1.2}$$

The set Δ, the **pairwise diagonal**, represents those configurations of n points in X which experience a collision – this is the set of illegal configurations for the robots. Of course, robots are not point-like, and near-collisions are unacceptable. From the point of view of topology, however, removing a sufficiently small neighborhood of Δ gives a space equivalent to $\mathcal{C}^n(X)$, and the configuration space $\mathcal{C}^n(\mathbb{R}^2)$ forms an acceptable model for robot motion planning on an unobstructed floor.

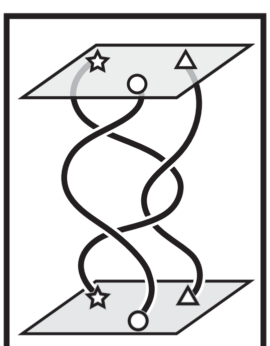


There are applications for which labeling the points is important, with warehousing, in which robots move specific packages, being a prime example. However, in some settings, such as mobile security cameras in a building, anonymity is not detrimental – any camera will do. The **unlabeled configuration space**, denoted $\mathcal{UC}^n(X)$, is defined to be the quotient of $\mathcal{C}^n(X)$ by the natural action of the symmetric group S_n which permutes the ordered points in X:

$$\mathcal{UC}^n(X) := \mathcal{C}^n(X)/S_n.$$

This space is given the **quotient topology**: a subset in $\mathcal{UC}^n(X)$ is open if and only if the union of preimages in $\mathcal{C}(X)$ is open. Configuration spaces of points on a manifold M are all (non-compact) manifolds of dimension $\dim \mathcal{C}^n(M) = n \dim M$. The space $\mathcal{C}^n(\mathbb{S}^1)$ is homeomorphic to $(n-1)!$ disjoint copies of $\mathbb{S}^1 \times \mathbb{R}^{n-1}$, while $\mathcal{UC}^n(\mathbb{S}^1)$ is a connected space.

Example 1.8 (Braids) ⊙



The configuration space of points on $\mathbb{R}^2$ is relevant to mobile robot motion planning (*e.g.*, on a factory floor); it is also among the most topologically interesting configuration spaces [38]. Consider the case of n robots which begin and end at fixed configurations, tracing out non-colliding routes in between. This complex motion corresponds to a path in $\mathcal{C}^n(\mathbb{R}^2)$, or, perhaps, $\mathcal{UC}^n(\mathbb{R}^2)$, if the robots are not labeled. (If the path has the same beginning and ending configuration, it is a loop, an image of $\mathbb{S}^1$ in the configuration space.) How many different ways are there for the robots to wind about one another *en route* from their starting to ending locations? The space-time graph of a path in configuration space is a **braid**, a weaving of strands encapsulating positions. A deformation in the motion plan equals a homotopy of the path (fixing the endpoints), which itself corresponds to moving the braid strands in such a way that they cannot intersect. From this, and a few sketches, the reader reasons correctly that there are infinitely many fundamentally inequivalent motion plans between starting and ending configurations. See §8.3 for a more algebraic description. ⊙

Example 1.9 (Navigation fields) ⊙

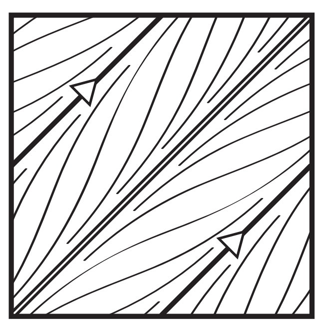
Configuration spaces in robotics are widespread, both as models of complex articulated agents or automated guided vehicles. One of many useful techniques for controlling robotic systems to execute behaviors and avoid collisions is to place a vector field on the configuration space and use the resulting flow to guide systems towards a goal configuration. The task: given a configuration space X and a specified goal point or loop $G \subset X$, construct an explicit vector field on X having G as a global **attractor**, so that (almost) all initial conditions will converge to and remain near the goal. Programming such a vector field is a challenge, usually requiring explicit coordinate systems. Additional features are also desirable: near the collision set (*e.g.*, Δ from Equation (1.2)), the vector field should be pointed away from the collision set, so that Δ is a **repeller**.

For example, consider the problem of mobilizing a pair of robots on a circular track so as to patrol the domain while remaining as far apart as possible. This can be done by means of a vector field on $\mathcal{C}^2(\mathbb{S}^1)$. Coordinatize $\mathcal{C}^2(\mathbb{S}^1) \subset \mathbb{T}^2$ as $\{(\theta_1, \theta_2) ; \theta_1 \neq \theta_2\}$. The reader may check that the following vector field has $\{\theta_1 = \theta_2 \pm \pi\}$ as an attracting periodic solution and $\Delta = \{\theta_1 = \theta_2\}$ as a repelling set:

$$\dot{\theta}_1 = 1 + \sin(\theta_1 - \theta_2) \quad ; \quad \dot{\theta}_2 = 1 + \sin(\theta_2 - \theta_1).$$

All initial conditions evolve to this desired state of circulation on $\mathbb{S}^1$. This is an excellent model not only for motion on a circular track, but also for an alternating gait in legged locomotion, where each leg position/momentum is represented by $\mathbb{S}^1$.

One reason for specifying a vector field on the entire configuration space (instead of simply dictating a path from initial to goal states) is so that if the robot experiences an unexpected failure in executing a motion, the vector field automatically corrects for the failure. Enlarging the problem from *path-planning* to *field-planning* returns stability and robustness [248]. This technique has been successfully applied in visual servoing [75], in robots that juggle [250], in automated guided vehicles [149], and in hopping [195] and insectoid [177] robot locomotion. This approach to robot motion planning makes extensive use of differential equations and dynamics on manifold configuration spaces. ⊙

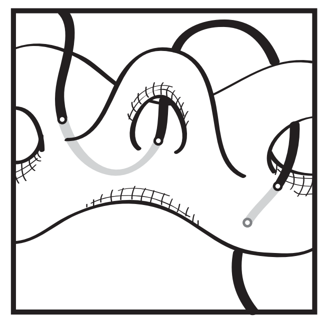


1.6 Transversality

Genericity is often invoked in applications, but seldom explained in detail. Intuition is an acceptable starting point: consider the following examples of generic features of smooth manifolds and mappings:

1. Two intersecting curves in $\mathbb{R}^2$ generically intersect in a discrete set of points.
2. Three curves in $\mathbb{R}^2$ generically do not have a point of mutual intersection.
3. Two curves in $\mathbb{R}^n$ generically do not intersect for $n > 2$.
4. Two intersecting surfaces in $\mathbb{R}^3$ generically intersect along curves.
5. A real square matrix A is generically invertible.
6. Critical points of a $\mathbb{R}$-valued function on a manifold are generically discrete.
7. The roots of a polynomial are generically non-repeating.
8. The fixed points of a vector field are generically discrete.
9. The configuration space of a planar linkage is generically a manifold.
10. A generic map of a surface into $\mathbb{R}^5$ is injective.

Some of these seem obviously true; others less obviously so. All are provably true with precise meaning using the theory of transversality.

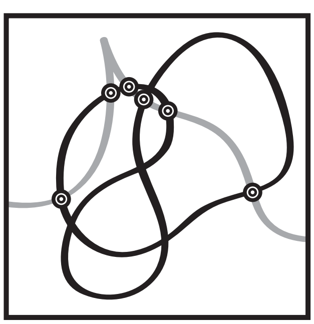


Two submanifolds V, W in M are transverse, written $V \pitchfork W$, if and only if,

$$T_pV + T_pW = T_pM \quad \forall\, p \in V \cap W. \tag{1.3}$$

Otherwise said, at an intersection, the tangent spaces to V and W span that of M. For example, two planes in $\mathbb{R}^3$ are transverse if and only if they are not identical. Note that the absence of intersection is automatically transverse. A central theme of topology is the lifting of concepts from spaces to maps between spaces. The notion of transverse maps is the first of many examples of this principle. Two smooth maps $f\colon V \to M$ and $g\colon W \to M$ are transverse, written $f \pitchfork g$, if and only if:

$$Df_v(T_vV) + Dg_w(T_wW) = T_pM \quad \forall\, f(v) = g(w) = p. \tag{1.4}$$



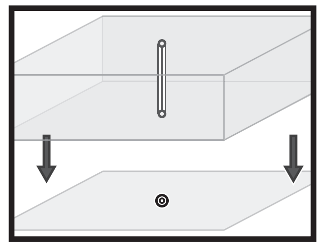
This means that the degrees of freedom in the intersection of images of f and g span the full degrees of freedom in M. Note that submanifolds $V, W \subset M$ are transverse if and only if the inclusion maps $\iota_V\colon V \to M$ and $\iota_W\colon W \to M$ are transverse as maps. Likewise, a map f is transverse to a submanifold $W \subset M$ if and only if $f \pitchfork \iota_W$. This map-centric definition does not constrain the images of maps to be submanifolds. This is one hint that differential tools are efficacious in the management of singular behavior. A point $q \in N$ is a **regular value** of $f\colon M \to N$ if $f \pitchfork \{q\}$. This is equivalent to the statement that, for each $p \in f^{-1}(q)$, the derivative is a surjection – the matrix of Df_p is of rank at least $n = \dim N$.

One benefit of transversality is that localized linear-algebraic results can be pulled back to global results. The following local result from linear algebra about dimensions of intersecting subspaces of a vector space is crucial:

Lemma 1.10 (Rank-Nullity Theorem). *For a linear transformation $T : U \to V$ between finite-dimensional vector spaces,*

$$\dim \ker T - \dim \operatorname{coker} T = \dim U - \dim V. \tag{1.5}$$

By applying Lemma 1.10 pointwise along the inverse image of transverse maps, one obtains a fundamental useful theorem:



Theorem 1.11 (Preimage Theorem). *Consider a differentiable map $f : M \to N$ between smooth manifolds. If $f \pitchfork W$ for $W \subset N$ a submanifold, then $f^{-1}(W)$ is a submanifold of M of dimension*

$$\dim f^{-1}(W) = \dim M - \dim N + \dim W. \tag{1.6}$$

This provides an effective means of constructing manifolds without the need for explicit charts: it is often used in the context of a regular value of a map.

1. The sphere $\mathbb{S}^n$ is the inverse image of 1 under $f : \mathbb{R}^{n+1} \to \mathbb{R}$ given by $f(x) = \|x\|$. It is a manifold of dimension $n = (n+1) - 1 + 0$.
2. The torus $\mathbb{T}^n$ is the inverse image $f^{-1}(1, \ldots, 1)$ of the map $f : \mathbb{C}^n \to \mathbb{R}^n$ given by $f(z) = (\|z_1\|, \ldots, \|z_n\|)$. Its dimension is $n = 2n - n + 0$.
3. The matrix group O_n of rigid rotations of $\mathbb{R}^n$ (both orientation preserving and reversing) can be realized as the inverse image $f^{-1}(\mathrm{Id})$ of the identity under the map from n-by-n real matrices to symmetric n-by-n real matrices given by $f(A) = AA^T$. The dimension of O_n is $n^2 - \frac{1}{2}n(n+1) + 0 = \frac{1}{2}n(n-1)$.
4. The determinant map restricted to O_n is in fact a smooth map to the 0-manifold $\mathbb{S}^0 = \{\pm 1\}$. As such, the special orthogonal group SO_n, as the inverse image of $+1$ under this restricted $\det$, is a manifold of the same dimension as O_n; thus, O_n is a disjoint union of two copies of SO_n.

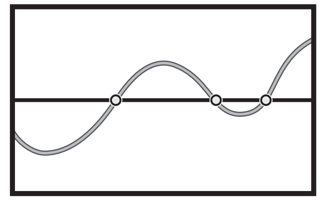
In the above examples, the transversality condition is checked by showing that the mapping f has a derivative of full rank at the appropriate (regular) value. Such regularity seems to fail rarely, for special *singular* values. This intuition is the driving force behind the Transversality Theorem. A subset of a topological space is said to be **residual** if it contains a countable intersection of open, dense subsets. A property dependent upon a parameter $\lambda \in \Lambda$ is said to be a **generic** property if it holds for λ in a residual subset of Λ – even when that subspace is not explicitly given. For reasonable (*e.g.*, **Baire**) spaces, residual sets are dense, and hence form a decent notion of topological typicality.

Theorem 1.12 (Transversality Theorem). *For M and N smooth manifolds and $W \subset N$ a submanifold, the set of smooth maps $f : M \to N$ with $f \pitchfork W$ is residual in $C^\infty(M, N)$, the space of all smooth maps from M to N. If W is closed, then this set of transverse maps is both open and dense.*

The proof of this theorem relies heavily on **Sard's Lemma**: the regular values of a sufficiently smooth map between manifolds are generic in the codomain.

Example 1.13 (Fixed points of a vector field) ⊙

The fixed point set of a differentiable vector field on a compact manifold M is generically finite, thanks to transversality.



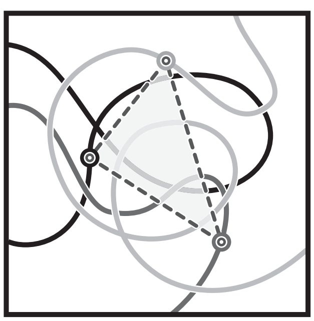
Recall from §1.4 that a smooth vector field is a section, or smooth map $V\colon M \to T_*M$ with $\pi(V(p)) = p$ for all $p \in M$. The **zero-section** $Z \subset T_*M$ is the set $\{(p, 0)\}$ of all zero vectors. The fixed point set of V is therefore the preimage $V^{-1}(Z)$. For a generic perturbation of V, this set is a submanifold of dimension

$$\dim \operatorname{Fix}(V) = \dim M + \dim M - \dim T_*M = 0.$$

A 0-dimensional compact submanifold is a finite point set. With more careful analysis of the meaning of transversality, it can be shown that the type of fixed point is also constrained: on 2-d surface, only source, sinks, and saddles are generic fixed points. ⊙

Example 1.14 (Beacon alignment) ⊙

Consider three people walking along generic smooth paths in the plane $\mathbb{R}^2$. How often are their positions collinear? This is relevant to robot navigation via beacon triangulation.



One considers the map $f\colon \mathcal{C}^3(\mathbb{R}^2) \to \mathbb{R}$ which computes the (signed) area of the triangle spanned by the three locations at an instant of time:

$$f(v_1, v_2, v_3) = \det [v_2 - v_1; v_3 - v_1] .$$

This map has zero as a regular value, and Theorem 1.11 implies that the set of collinear configurations is a submanifold $W = f^{-1}(0)$ of $\mathcal{C}^3(\mathbb{R}^2)$. A set of three paths is a generic map from $\mathbb{R}$ (time) into $\mathcal{C}^3(\mathbb{R}^2)$. It follows from Theorems 1.12 and 1.11 that, generically, one expects collinearity at a discrete set of times, since

$$\dim \mathbb{R}^1 + \dim W - \dim \mathcal{C}^3(\mathbb{R}^2) = 1 + 5 - 6 = 0.$$

This, moreover, implies a stability in the phenomenon of collinearity: at such an alignment, a generic perturbation of the paths perturbs where the alignment occurs, but does not remove it. ⊙

1.7 Signals of opportunity

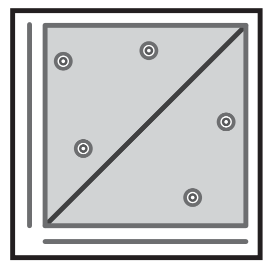
Applications of transversality are alike: (1) set up the correct maps/spaces; (2) invoke transversality; (3) count dimensions. This has some simple consequences, as in computing the generic intersection of curves and surfaces in $\mathbb{R}^3$. Other consequences are not so obvious. The following theorem states that any continuous map of a source manifold into a target manifold has, after generic perturbation, a submanifold as its image when the dimension of the target is large enough. How large? The critical bound comes, as it must, from the proper transversality criterion and a dimension count:

Theorem 1.15 (Whitney Embedding Theorem). *Any continuous function $f : M \to N$ between smooth manifolds is generically perturbed to a smooth injection when* $\dim N > 2\dim M$.

Proof. The configuration space $\mathcal{C}^2 M = M \times M - \Delta_M$ of two distinct points on M is a manifold of dimension $2\dim M$. The graph of f induces a map

$$\mathcal{C}^2 f : \mathcal{C}^2 M \to \mathcal{C}^2 M \times N \times N \quad : \quad (x, y) \mapsto (x, y, f(x), f(y)).$$

The key step is this: observe that the set of points on which f is non-injective is $(\mathcal{C}^2 f)^{-1}(\mathcal{C}^2 M \times \Delta_N)$, where $\Delta_N \subset N \times N$ is the diagonal. According to the appropriate transversality theorem (specifically, the **multi-jet** transversality theorem [161, Thm. 4.13]), generic perturbations of f induce generic perturbations of $\mathcal{C}^2 f$. From Theorems 1.11, 1.12, and the hypothesis that $\dim N > 2\dim M$, the generic dimension of the non-injective set of f is:



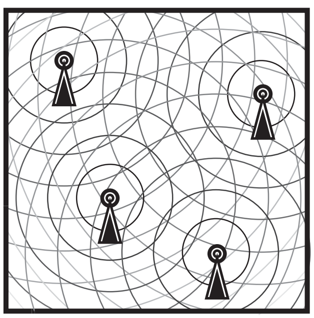
$$\dim \mathcal{C}^2 M - \dim(\mathcal{C}^2 M \times N \times N) + \dim(\mathcal{C}^2 M \times \Delta_N) = 2\dim M - \dim N < 0.$$

Thus, there are no self-intersections, and the result is a smooth submanifold. ⊙

A simple application of Whitney's theorem [257] informs a problem of localization via signals. Consider a scenario in which one wants to determine location in an unknown environment. Certainly, a GPS device would suffice. Such a device works by receiving signals from multiple satellite transmitters and utilizing known timing data about the transmitters to determine location, within the tolerances of the signal reception. This sophisticated system requires many independent components to operate, including geosynchronous satellites, synchronized clocks, and more, all with nontrivial power constraints. Though wonderfully useful, GPS devices are not universally available: they do not operate underwater or indoors; they are unreliable in *urban canyons*; the need for satellites and synchronized clocks can limit availability.

There is no reason, however, why other signals could not be used. In fact, passive signals – arising naturally from TV transmissions, radio, even ionospheric waves induced by lightning strikes [99] – provide easily measured pulses with which to attempt to reconstruct location. There is a small but fascinating literature on the use of such **signals of opportunity** to solve localization and mapping problems.

The following is a simple mathematical model for localization via signals of opportunity. Consider a connected open domain $\mathcal{D} \subset \mathbb{R}^k$ which is a k-manifold. Assume there exist N transmitters of fixed location which asynchronously emit pulses whose times of arrival can be measured by a receiver at any location in $\mathcal{D}$. Given a receiver located at an unknown point of $\mathcal{D}$, do the received TOA (time of arrival) signals uniquely localize the receiver? Consider the signal profile mapping $T : \mathcal{D} \to \mathbb{R}^N$ which records the TOA of the (received, identified, and ordered) transmitter pulses. If one assumes that the generic placement of transmitters associated with this system provides a generic perturbation to T within $C^\infty(\mathcal{D}, \mathbb{R}^N)$, then the resulting perturbation embeds the domain $\mathcal{D}$ smoothly for $N > 2k$. Thus, the mapping T is generically injective, implying unique *channel response* and the feasibility of localization in $\mathcal{D}$ via TOA. For example, this implies that a receiver can be localized to a unique position in a planar domain $\mathcal{D}$ using only a sequence of *five* or more transmission signals, globally readable from generically-placed transmitters. Note: it is not assumed that the signals move along round waves, nor does the receiver compute any distances-to-transmitters.



This result is greatly generalizable to a robust *signals embedding theorem* [257]. First, one may modify the codomain to record different signal inputs. For example, using TDOA (time difference of arrival) merely reduces the dimension of the signal codomain by one and preserves injectivity for sufficiently many pulses. Second, one may quotient out the signal codomain by the action of the symmetric group S_N to model inability to identify target sources. This does not change the dimension of the signals codomain: the system has the same number of degrees of freedom. The only change is that the codomain has certain well-mannered singularities inherited from the action of S_N. It follows from transversality that *any* reasonable signal space of sufficient dimension preserves the ability to localize based on knowledge of the image of $\mathcal{D}$ under T. Transversality and dimension-counting provide a critical bound on the number of signals needed to disambiguate position, independent of the types of signals used.

1.8 Stratified spaces

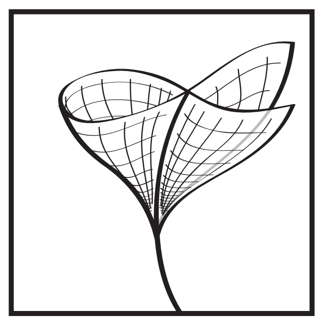
The application of Whitney's Theorem to signal localization in the previous subsection is questionable in practice. Signals do not propagate unendingly, and the physical realities of signal reflection/echo, multi-bounce, and diffraction conspire to make manifold theory suboptimal in this setting. The addition of signal noise further frustrates a differential-topological approach. Finally, the assumption that $\mathcal{D}$ is a manifold is a poor one. In realistic settings, the domain has a boundary: signals are bouncing off of walls, building exteriors, and other structures that, at best, are piecewise-manifolds.

One approach to this last difficulty is to enlarge the class of manifolds. An n-**manifold with boundary** is a space locally homeomorphic to *either* $\mathbb{R}^n$ or $\mathbb{R}^{n-1} \times [0, \infty)$, with the usual compatibility conditions required for a smoothing. The **boundary** of $\mathcal{D}$, $\partial\mathcal{D}$, is therefore a manifold of dimension $n - 1$. Many of the tools and theorems of this chapter (*e.g.*, tangent spaces and transversality) apply with minor modifications to manifolds with boundary. Yet this is not enough in practice: further generalization is needed. An n-**manifold with corners** is a space, each point of which has a neighborhood locally homeomorphic to

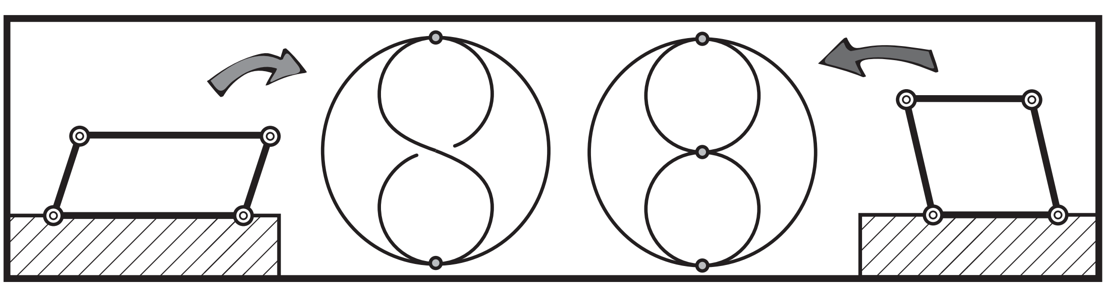
$$\{x \in \mathbb{R}^n : \; x_i \geq 0, \; i = 1, \ldots, m\},$$

for some $0 \leq m \leq n$, where m may vary from point to point. A true manifold has $m = 0$ everywhere; a manifold with boundary has $m \leq 1$. The analogues of smoothings, derivatives, tangent spaces, and other constructs are not difficult to generate. The boundary of a manifold with corners no longer has the structure of a smooth manifold, as, *e.g.*, is clear in the case of a cube or other platonic solid. Note however, that such a boundary is assembled from manifolds of various dimensions, suitably glued together. Such piecewise-manifolds are common in applications. Consider the solution to a polynomial equation $p(x) = 0$, for $x \in \mathbb{R}^n$. An application of transversality theory shows that the solution set is, for generic choices of coefficients of p, a manifold of dimension $n - 1$; however, nature does not always deal out such conveniences. Innumerable applications call for the solution to a specific polynomial equation. The null set of a polynomial, even when not a true manifold, can nevertheless be decomposed into manifolds of various dimensions, glued together in a particular manner.

There is a hierarchy of such **stratified spaces** which deviate from the smooth regularity of a manifold. An intuitive definition of a stratified space is a space X, along with a finite partition $X = \cup_i X_i$, such that each X_i is a manifold. Precise control over how these manifolds are pieced together is needed but is too intricate for this introduction: let the reader think of a stratified space as a piecewise-manifold.



Typical examples of stratified spaces include singular solutions to polynomial or real-analytic equations. More physical examples are readily generated. Recall the setting of planar linkage configuration spaces. The 4-bar mechanism gives a 1-d manifold, *except* when the lengths satisfy $L_1 = L_3$ and $L_2 = L_4$. In this case, the configuration space is a pair of circles with intersections (or, better, *singularities*) – either two or three depending on whether or not all the lengths are the same. These singularities have physical significance: they correspond to configurations which are collinear. Upon building such a linkage, one can *feel* the difference as it passes through a singular point.



Notes

1. Homeomorphic manifolds are not always diffeomorphic. In dimensions three and below, they are. Each $\mathbb{R}^n$ has a unique smooth structure except $n = 4$, which has an uncountable number of *exotic* smoothings [162]. Spheres $\mathbb{S}^7$ of dimension seven have exactly 28 distinct smooth structures; see [225] for a survey. It is unknown at the time of publication if $\mathbb{S}^4$ has non-standard smoothings [130].
2. A manifold is **orientable** if it has an atlas such that all transition maps $\phi_\beta \circ \phi_\alpha^{-1}$ between charts preserve orientation (have derivatives with positive determinant). The projective plane $\mathbb{P}^2$ and Klein bottle K^2 are well-known non-orientable surfaces. Verifying that all transition maps preserve orientation seems difficult: more efficient algebraic means of orienting manifolds will arise in Chapters 4 and 5.
3. Configuration spaces are for many the entrance to applied topology: it is a subject worthy of its own text. It is instructive and highly recommended to build a complex mechanical linkage and investigate its topology by hand. For planar linkages, flat cardboard, wood, metal or plastic with pin joints works well. For 3-d linkages, the author uses wooden dowels with latex tubing for rotational joints. Sadly, no higher spatial dimensions are available for resolving intersections between edges in spatial linkages. With practice, the user can tell the dimension of the configuration space by feel, without explicit computation (past dimension 10, the author's discernment fails).
4. Other tools of advanced calculus follow in patterns similar to those of derivatives, including differential forms, integration of forms, Stokes' Theorem, partial differential equations, and more: see Chapter 6. The reader interested in calculus on manifolds can find excellent introductions [169], some tuned to applications in mechanics [1, 15].
5. Transversality is a topological approach to genericity. Probability theory offers complementary and, often, incommensurate approaches.
6. Lemma 1.10 is the hidden jewel of this chapter, as it enables so much of the machinery in future chapters. This is the first appearance of an Euler characteristic in this text. It is not the last – there are at least half-a-dozen manifestations of this index in this text.
7. One must be careful in proving genericity results, as the specification of a topology on function spaces is required. Tweaking the topology of the function spaces allows for relative versions, which allow perturbations to one domain while holding the function fixed elsewhere.
8. Higher derivative data associated to a map $f: M \to N$ is encoded in the r-jet, $j^r f$, taking values in a **jet bundle** $J^r(M, N)$. This $j^r f$ records, for each $p \in M$, the Taylor polynomial of f at p up to order r. As always, the computations are done at the chart level and shown to be independent of coordinates. The Jet Transversality Theorem says that, for any submanifold $W \subset J^r(M, N)$ of the jet bundle, the set of $f: M \to N$ whose r-jets are transverse to W is a residual subset of $C^\infty(M, N)$ [179].
9. The study (and even the definition) of stratified spaces is *much* more involved than here indicated [163]. Various forms of the Transversality Theorem apply to stratified spaces [161, 216]; they are sufficient to derive a form of Whitney's Theorem on embeddings and allow for unique channel response in a transmitter-receiver system outfitted with corners, walls, and reflections [257].
10. The configuration space of **hard spheres** in a domain gives a wonderful class of stratified spaces whose topology can be quite intricate [23]. The topology of these configuration spaces undergoes large-scale qualitative changes as the number of spheres is increased – changes that mimic phase transitions in matter [58].