SOME NOTES ON THE 0^{th} HOMOTOPY GROUP π_0

K.M. HALPERN

July 22, 2025

Contents

1. Intro

Let X be a topological space. Its 0^{th} homotopy group $\pi_0(X)$ is defined just like the other homotopy groups, but $S^0 = \{-1, +1\}$ consists of two discrete points rather than a path-connected sphere, so $\pi_0(X)$ behaves quite differently from the other homotopy groups.

To start with, unlike $\pi_n(X)$ for $n > 0$, $\pi_0(X)$ need not be a group — so calling it the "0^{th} homotopy group", though common practice, is a misnomer. $\pi_0(X)$ is just a quotient space. It has the same formal definition as the other π_n's, but the latter have a natural group structure via path concatenation while π_0

does not.

As we will see, $\pi_0(X)$ has slightly more algebraic structure than a mere set but falls short of being a group. Like all the $\pi_n(X)$'s, it is a "pointed set", though the notation obscures this fact. However, the other $\pi_n(X)$'s are actual groups. We will discuss this shortly. We will also see that when X is a topological group $\pi_0(X)$ acquires a quotient group structure from X via a completely different mechanism.

We'll begin by reviewing some relevant background math. The quotient topology is pertinent because π_0 is a quotient space, the topological notions of connectedness and path-connectedness are pertinent since π_0 is the set of path-connected components, semidirect products are pertinent because the groups that appear in physics (such as the Poincare group and Lorentz group) are semidirect products of their identity-connected component and π_0, pointed sets are pertinent because they constitute the minimal structure on π_0 that allows it to appear in long exact sequences, and homotopy is pertinent because π_0 — even if quite different in many ways — is constructed in the same fashion as the other homotopy groups, topological groups are pertinent because Lie groups are topological groups and confer a group structure on π_0, and zero-dimensional manifolds are pertinent because π_0 can be regarded as one (and thus as a Lie group).

With the necessary background material out of the way, we'll discuss $\pi_0(X)$ in general, for topological manifolds, for topological groups, and lastly for Lie groups. Next, we'll work through several physically relevant examples in detail: $GL(\mathbb{R}, n)$, $O(3)$, $O(3,1)$, and the Poincare group. For each, we'll construct the set of connected components, the identity component, the center, and the semidirect product structure (or, where possible, a direct product structure).

Finally, an appendix discusses a number of practical aspects of semidirect products which are relevant but would disrupt the flow of our main discussion. This can be thought of as a supplement to the notes [1], which the reader is encouraged to peruse for a more detailed and comprehensive discussion of semidirect products and group extensions. In particular, we discuss the nonassociativity of the semidirect product, restrictions of semidirect products, and whether and when nested semidirect products can be rearranged.

2. Review of Some Relevant Math

2.1. **Quotient Topology.**

Let's briefly review the notion of a quotient space. Recall that a partition is the same as an equivalence relation. Each defines the other. Denote by $\tilde{X}$ a partition of set X, and by $[x]$ a class of $\tilde{X}$.

If the partition arises from an equivalence relation $\sim$, we may write $\tilde{X} = X/\sim$.

The classes $[x]$ can be regarded as either points in $\tilde{X}$ or subsets of X, depending on the context. Where ambiguous, we'll explicitly state our intended meaning.

The **singleton partition** of X assigns each $x \in X$ its own class, and the **trivial partition** has just one class containing all of X. For any partition, the **quotient map** $q : X \to \tilde{X}$ takes each x to the unique $[x]$ containing it.

If X is a topological space, any partition $\tilde{X}$ of X can be endowed with a natural topology called the **quotient topology** and defined as follows. Let T be the topology on X. We say that a topology $T_\sim$ on $\tilde{X}$

is compatible with T iff $q^{-1}(O')$ is open in X for every open $O' \subseteq \tilde{X}$ (i.e. iff $q^{-1}(O') \in T$ for all $O' \in T_\sim$). Put another way, q must be continuous relative to T and $T_\sim$.

Compatibility tells us that we are only looking at those open sets of T which are unions of classes. I.e., they respect $\sim$.

The indiscrete topology on $\tilde{X}$ is manifestly compatible with any topology on X, so there is always at least one compatible topology on $\tilde{X}$. It can be shown that there is a unique maximal compatible topology on $\tilde{X}$, which contains all other compatible topologies as subtopologies. We define our quotient topology $\tilde{T}$ on $\tilde{X}$ to be this maximal compatible topology.

Let's define $T_q \equiv \{q^{-1}(O'); O' \in \tilde{T}\}$ to be the corresponding subtopology of T. Define $\hat{q} : T_q \to \tilde{T}$ via $\hat{q}(O) \equiv q(O)$, with O viewed as an element of T_q on the left and as a subset of X on the right.

The individual classes of $\sim$ may not be open in T, so T itself may not respect $\sim$. We can think of T_q as being generated by the smallest unions of classes that are open in T. I.e., it is the largest subtopology of T that is compatible with $\sim$ in the sense of respecting its classes.

Prop 2.1: Given partition $\tilde{X} = X/\sim$ of topological space X, let $\tilde{T}$, T_q, and $\hat{q}$ be defined as above. Then (i) T_q is a subtopology of T. (ii) Every open set in T_q is a union of classes of $\sim$. (iii) $\hat{q}$ is a bijection. (iv) $\hat{q}$ preserves arbitrary unions and finite intersections in both directions.

Pf: (ii) The inverse image of a point in $\tilde{X}$ is a whole $\sim$-class in X. Therefore, the inverse image of any subset of $\tilde{X}$ is a union of classes.

Pf: (iii) $\hat{q}$ is surjective by construction, since q is surjective and T_q consists of the inverse image of every open set in $\tilde{T}$. Suppose $O_1, O_2 \in T_q$ and $\hat{q}(O_1) = \hat{q}(O_2)$, and call it O'. Then $q^{-1}(O') \supseteq O_1 \cup O_2$. Therefore, some $O \supseteq O_1 \cup O_2$ appears in T_q, and neither O_1 nor O_2 can unless they equal O, because they wouldn't be the inverse of an open set of $\tilde{T}$. If $O_1 \subset O_2$ then O_1 cannot appear in T_q, if $O_2 \subset O_1$ then O_2 cannot appear in T_q, and if neither is a subset of the other then neither can appear in T_q. Only if $O_1 = O_2$ can both be members of T_q, as required by our premise. As mentioned, T_q can't see below the resolution of $\tilde{T}$. Therefore, $\hat{q}$ is injective and thus bijective.

Pf: (i) By the definition of T_q, it is a subset of T. We therefore need only show that it is a topology. Since q is surjective, $q^{-1}(\tilde{X}) = X$. Vacuously, $q^{-1}(\emptyset) = \emptyset$. So $X, \emptyset \in T_q$. Let $\cup O_i$ be an arbitrary union of $O_i \in T_q$. By the definition of T_q, each $O_i = q^{-1}(O_i')$ for some $O_i' \in \tilde{T}$. Therefore, $\cup O_i = \cup q^{-1}(O_i')$. However, $q^{-1}(\cup O_i') = \cup q^{-1}(O_i')$ for any function q, so $\cup O_i = q^{-1}(\cup O_i')$. Since $\tilde{T}$ is a topology, $\cup O_i' \in \tilde{T}$ for an arbitrary union. Therefore $q^{-1}(\cup O_i') \in T_q$. Now, let $\cap O_i$ be a finite intersection of $O_i \in T_q$. Then, by the same token, $\cap O_i = \cap q^{-1}(O_i')$. However, $q^{-1}(\cap O_i') = \cap q^{-1}(O_i')$ for any function q, so $\cap O_i = q^{-1}(\cap O_i')$. Since $\tilde{T}$ is a topology, $\cap O_i' \in \tilde{T}$ for a finite intersection. Therefore $q^{-1}(\cap O_i') \in T_q$.

Pf: (iv) From (iii), $\hat{q}$ is bijective, and we can speak of its inverse as a function. Any function preserves arbitrary unions, and the inverse of any function preserves arbitrary unions and intersections. Applying this to q, we get that $\hat{q}$ preserves arbitrary unions and its inverse preserves arbitrary unions and finite intersections. All that remains is to show that $\hat{q}$ preserves finite intersections in the forward direction. For a general function f between sets, $f(\cap S_i) \subseteq \cap f(S_i)$, with equality if f is injective. However, injectivity is not a necessary condition, merely a sufficient one. Equality holds for *all* collections $\{S_i\}$ iff f is injective, but it can hold for *some* collections even if f is not injective. In our case, q is not injective, but we only need equality for a specific set of collections — the finite subsets of T_q. The obstruction to equality can be seen in the example of a singleton collection. In that case, if $x \neq y$ we can have $f(x) = f(y)$, which then results in $f(\{x\} \cap \{y\}) = f(\emptyset) = \emptyset$ while $f(\{x\}) \cap f(\{y\}) = \{f(x)\} \neq \emptyset$. Returning to our case, consider $q(O_1 \cap O_2)$ for $O_1, O_2 \in T_q$. By construction, $O_1 = q^{-1}(O_1')$ and $O_2 = q^{-1}(O_2')$ for some O_1' and O_2'. Since $\hat{q}$ is bijective, O_1' and O_2' are distinct (i.e. they differ in at least one element). In order to have a proper subset $q(O_1 \cap O_2) \subset q(O_1) \cap q(O_2) = O_1' \cap O_2'$, we need some $x \in O_1$ and $y \in O_2$ s.t. $q(x) = q(y)$ *and* s.t. no other points in O_1 or O_2 map to the relevant image. I.e., we need two sets O_1 and O_2 containing x and y but such that $x' = q(x) = q(y)$ doesn't appear in $q(O_1 \cap O_2)$. If any other points in $O_1 \cap O_2$ map to x', the obstruction is removed. In our case, the very definition of T_q thwarts this obstruction. By construction, the noninjectivity of q is intra-class only. Suppose $q(x) = q(y)$. Then x and y are in the same class. In order to have a proper subset $q(O_1 \cap O_2) \subset O_1' \cap O_2'$, we require that some $x' \in O_1' \cap O_2'$ doesn't have any inverse image in $O_1 \cap O_2$. However, O_1 and O_2 respect classes. Therefore, the inverse image of x' as a whole must be in both. There can be no x and y, with only $x \in O_1$ and only $y \in O_2$ s.t. $x' = q(x) = q(y)$. The whole class $q^{-1}(x')$ must appear in both or in neither or in only one. It can't be split between them. In none of these cases does $q(O_1 \cap O_2)$ lose any points that appear in $O_1' \cap O_2'$. I.e., for our particular collection T_q, $\hat{q}$ preserves finite intersections.

Prop 2.2: $\tilde{X}$ has the discrete topology iff all of its classes are open in X.

Pf: If $\tilde{T}$ is the discrete topology on $\tilde{X}$ then, by definition, $q^{-1}(x') \in T$ for every $x' \in \tilde{X}$, so every class is open in X. Going the other way, suppose that every class is open in X. Then $q^{-1}(x')$ is open in X for every $x' \in \tilde{X}$. This means that the discrete topology on $\tilde{X}$ is compatible with T. The discrete topology is the largest possible topology on $\tilde{X}$, so it must therefore be the quotient topology.

We say that T_q is **quotient-equivalent** to $\tilde{T}$, meaning that there exists a bijection between these topologies that preserves unions and finite intersections and maps X (as an element of T_q) to $\tilde{X}$ (as an element of $\tilde{T}$) and $\emptyset$ to $\emptyset$. As topologies, $\tilde{T}$ and T_q have identical information and structure, but not due to a homeomorphism of the underlying spaces.

I.e. $\tilde{T}$ captures the full information content of T_q. As far as topology goes, there is no information in T_q below the resolution of the $\sim$-classes (and possibly a bit above it, unless those classes are all open).

If $\tilde{X}$ has the discrete topology (i.e. every class of $\sim$ is open in X), then the classes of $\sim$ form a partition basis for T_q. In that case, we say that T_q is quotient-equivalent to the discrete topology. This means that the whole structure of T_q is captured in the discrete topology of some set (in this case $\tilde{X}$) and, as far as topology goes, there is no information in T_q below the resolution of the classes of $\sim$.

If $T_q = T$, then T can't see below the level of certain clusters of classes of $\sim$. If in addition, $\tilde{T}$ has the discrete topology, then T is quotient-equivalent to the discrete topology on $\tilde{X}$. I.e., it can't see below the level of the individual classes of $\sim$, and the latter form a partition basis for it.

Returning to the general case, note that a set O' is open in $\tilde{X}$ iff $q^{-1}(O')$ is open in X. One direction is part of the definition, but the other follows as well — because the quotient topology is the *maximal* compatible topology.

If $O' \in \tilde{T}$ then, by definition, $q^{-1}(O') \in T$. Suppose that $O \equiv q^{-1}(S')$ is open in X (i.e. $O \in T$) for some nonopen $S' \subset \tilde{X}$ (i.e. $S' \notin \tilde{T}$). Let $\tilde{T}'$ be generated by $\tilde{T} \cup \{S'\}$. Then, $\tilde{T}'$ is larger than $\tilde{T}$ but still compatible with T, violating our assumption that $\tilde{T}$ is the maximal compatible topology.

Note that this does *not* imply that q is an open map. We'll see that for topological groups it is, but in general it need not be. Nor does this imply that all of T is the inverse image of $\tilde{T}$ (i.e. that $T_q = T$). While it is true that $q(O)$ is open for $O \in T_q$, it need not be for $O \in T$. Since q is surjective but not injective, $q^{-1}(q(O)) \supseteq O$. If O projects to a set of cosets but isn't their *complete* inverse image (i.e. there's something outside O that projects to them too), then O cannot be in the inverse image of any open set of $\tilde{T}$. I.e., the open sets of T that appear in T_q are precisely those which are unions of not just cosets but the precise clusters of cosets which comprise the open sets of $\tilde{T}$.

2.2. **Connectedness and Path-connectedness.**

2.2.1. *Definitions.* The following are some properties that a topological space X can have:

- **Path-connected**: For every $x_1, x_2 \in X$, there exists a continuous function $f : [0,1] \to X$ s.t. $f(0) = x_1$ and $f(1) = x_2$.
- **Connected**: X cannot be written as the union of two or more disjoint open sets.
  Since unions of open sets are open, this is the same as saying that X can't be written as the union of two disjoint open sets. The "or more" is superfluous.
- **Locally path-connected**: Given any point x, every neighborhood U of x contains (or equals) a path-connected neighborhood of x.
  Recall that a "neighborhood" of a point x is any set containing an open set which contains x. I.e., V is a neighborhood of x iff $\exists O$ that is open s.t. $x \in O \subseteq V$.

The same terminology applies to subspaces. We say that $Y \subset X$ is path-connected or connected or locally path-connected iff Y has the corresponding property in the subspace topology.

We'll refer to "connected sets" and "connected spaces" interchangeably (and ditto for "path-connected"). We always mean in the subspace topology.

We say that two points $x, y \in X$ are path-connected iff there exists a continuous function $f : [0, 1] \to X$ s.t. $f(0) = x$ and $f(1) = y$. Such a continuous function is called a "path" and we say that x and y are "path-connected".

I.e. a space is path-connected iff every pair of points in it is path-connected.

The following general lemma will prove useful in our discussion:

Lemma 2.3: Let X and Y be topological spaces and let $X_1, X_2 \subset X$ be closed subsets s.t. $X_1 \cup X_2 = X$. (i) If $f : X \to Y$ is continuous when restricted to X_1 and X_2, it is continuous on X. (ii) Let $f_1 : X_1 \to Y$ and $f_2 : X_2 \to Y$ agree on $X_1 \cap X_2$ (i.e. $f_1(x) = f_2(x)$ for all $x \in X_1 \cap X_2$). Then we can define $h : X \to Y$ via $h(x) \equiv f_1(x)$ on X_1 and $h(x) \equiv f_2(x)$ on X_2 since there is no ambiguity on the overlap. If f_1 and f_2 are continuous, then so is h.

Pf: (i) Let C' be closed in Y. Since $X = X_1 \cup X_2$, $f^{-1}(C') = (f^{-1}(C') \cap X_1) \cup (f^{-1}(C') \cap X_2)$. However, $f|_{X_1}$ and $f|_{X_2}$ are continuous, so $(f^{-1}(C') \cap X_1)$ and $(f^{-1}(C') \cap X_2)$ are closed in X_1 and X_2 respectively. Since X_1 and X_2 are closed in X, any closed subsets of them are closed in X as well. The union of two closed sets is closed, so $f^{-1}(C')$ is closed in X. Since the inverses of closed sets are closed, f is continuous.

Pf: (ii) Let f_1 and f_2 be continuous, with $f_1 = f_2$ on $X_1 \cap X_2$, and define h as specified. h meets the criterion of (i) since $h|_{X_1} = f_1$ and $h|_{X_2} = f_2$ are both continuous, so h is continuous.

Path-connected implies connected, but the converse need not hold. Path-connectedness is stricter. However, connected *and* locally path-connected implies path-connected. Put another way, in the presence of local path-connectivity, the concepts of path-connected and connected are the same. This and a number of other important properties are codified in the following proposition:

Prop 2.4: The following properties hold:

- (i) A space X is connected iff its only clopen subsets are $\emptyset$ and X.
  I.e. this is an entirely equivalent condition. It is sometimes taken as the definition of "connected".
- (ii) The union of two non-disjoint connected sets is connected.
- (iii) The union of two non-disjoint path-connected sets is path-connected.
- (iv) Every path-connected space is connected.
- (v) Every connected and locally path-connected space is path-connected.
- (vi) If $f : X \to Y$ is continuous, it takes connected sets in X to connected sets in Y.
- (vii) If $f : X \to Y$ is continuous, it takes path-connected sets in X to path-connected sets in Y.
- (viii) A finite product of connected sets is connected.
- (ix) A finite product of path-connected sets is path-connected.
- (x) Every open subset of a locally path-connected set is locally path-connected.
- (xi) $\emptyset$ is vacuously connected, path-connected, and locally path-connected, as is any singlet $\{x\}$.
- (xii) Given an arbitrary (i.e. possibly uncountable) collection of connected sets $\{S_\alpha\}$ whose intersection $\cap_\alpha S_\alpha \neq \emptyset$, the union $\cup_\alpha S_\alpha$ is connected.

- (xiii) Given an arbitrary (i.e. possibly uncountable) collection of path-connected sets $\{S_\alpha\}$ whose intersection $\cap_\alpha S_\alpha \neq \emptyset$, the union $\cup_\alpha S_\alpha$ is path-connected.
- (xiv) The closure of a connected set is connected.

Regarding (ii), (iii), (x), and (xi), note that it is *not* true that the union of connected sets is connected in general or that the union of path-connected sets is path-connected. Nor are intersections of connected sets connected or intersections of path-connected sets path-connected. Ex. think of a solid donut intersecting a transverse plane. The intersection is two disjoint disks, even though the donut and the plane are both connected and path-connected.

Regarding (ii) and (iii), note that they extend to any finite union but not to countable or uncountable unions. Specifically, if $\{S_i\}$ is any finite collection of connected (path-connected) sets s.t. it cannot be divided into two disjoint subunions, we're fine. We do *not* require that every set overlap all the others (i.e. that pairwise intersections are nonempty) or that all the sets share a point in common (i.e. that the mutual intersection is nonempty). We just need to be able to stitch the S_i's into a set that has no disjoint pieces.

Regarding (iv) and (v), note that it is *not* the case that the inverse image of a connected set is connnected or the inverse image of a path-connected set is path-connected, even if f is continuous.

Pf: (i) (forward) Suppose X can't be written as the union of two disjoint open sets. Consider nonempty proper subset $Y \subset X$. If Y is clopen, then it is closed and its complement is disjoint and open, so we have a disjoint union of two open sets. Therefore, X is the only nonempty clopen set. (backward) Suppose X is the only nonempty clopen set. If $X = A \cup B$ for disjoint open sets, then both A and B are clopen, since their complements are open.

Pf: (ii) If Z is clopen in X and $Y \subset X$, then $Z \cap Y$ is clopen in the subspace topology on Y, since open sets restrict to open sets and closed sets restrict to closed sets. Let $C \subset A \cup B$ be a proper nonempty clopen subset. Then $C \cap A$ is clopen in A and $C \cap B$ is clopen in B. The only way to avoid violating our premise is if $C \cap A = \emptyset$ or A and likewise for $C \cap B$. This means that C can only be $\emptyset$, A, B, or $A \cup B$. We posited that it's nonempty and proper, so suppose it equals A. Then $C \cap B \neq \emptyset$ since $A \cap B \neq \emptyset$, so $C \cap B$ is clopen in B. This violates our premise unless $C \cap B = B$, in which case $A = B$. However, we already know that no such C exists in that case, since A is connected.

Pf: (iii) Let A and B be path-connected subsets of X, with $A \cap B \neq \emptyset$. Let $z \in A \cap B$. Pick any $x \in A$ and $y \in B$. Then there exists a path f from x to z and a path g from z to y. Define $h : [0,1] \to A \cup B$ via $h(t) = f(2t)$ for $t \in [0, 1/2]$ and $h(t) = g(2t-1)$ for $t \in [1/2, 1]$. This is unambiguous at $t = 1/2$, since $f(1) = g(0) = z$. The stated reparametrizations of f and g are continuous (ex. $k : [0, 1/2] \to [0, 1]$ given by $k(t) = 2t$ is continuous, so the composition $f \circ k$ is too), so $h|_{[0,1/2]}$ and $h|_{[1/2,1]}$ are continuous from those closed sets to Y. Lemma 2.3 then tells us that h is continuous on $[0, 1]$. I.e., h is a path from x to y. Therefore $A \cup B$ is path-connected.

Pf: (iv) Suppose that X is path-connected but not connected. Then $X = A \cup B$ for some disjoint open sets A and B. Let $x \in A$ and $y \in B$. There exists a continuous $f : [0,1] \to X$ s.t. $f(0) = x$ and $f(1) = y$. $f^{-1}(A)$ is open in $[0,1]$ and $f^{-1}(B)$ is open in $[0,1]$. They are disjoint, since the inverse images of disjoint sets are disjoint for any function. However, $f^{-1}(X) = [0,1]$, so $[0,1]$ is a union of disjoint open sets. $[0,1]$ is connected, so this is impossible. Therefore, X cannot be a union of disjoint open sets and is connected.

Pf: (v) Suppose that X is connected and locally path-connected. Pick some $x \in X$ and let P denote the subset of X that is path-connected to x (i.e. the maximal path-connected subset of X containing x). Since X is locally path-connected, every $x' \in X$ has a path-connected open set containing x'. If this open set overlaps P, then P must contain it (or by (iii) we would have to expand P to contain it, which violates our premise that P is maximal). We therefore may write both P and $X - P$ as unions of such open sets, one for each point in X. They are disjoint and cover X, so X is either a union of two disjoint open sets or $P = X$ (P can't be $\emptyset$, since it contains x). Therefore $P = X$, and X is path-connected.

Pf: (vi) Suppose $U \subset X$ is connected and $f(U)$ is not. Then we can write $f(U) = A \cup B$, where A and B are disjoint. f^{-1} respects disjointness, so $X = f^{-1}(U) = f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B)$, which are disjoint. Since f is continuous, $f^{-1}(A)$ and $f^{-1}(B)$ are open, so X is the union of two disjoint open sets, contradicting our assumption.

Pf: (vii) Suppose $U \subset X$ is path-connected and $f(U)$ is not. Then there exist $x', y' \in f(U)$ with no path between them. Pick any $x \in f^{-1}(x')$ and $y \in f^{-1}(y')$. There exists a path p between these, because U is path-connected. Since $p : [0,1] \to X$ is continuous and f is continuous, $f \circ p$ is continuous and has $(f \circ p)(0) = x'$ and $(f \circ p)(1) = y'$. Therefore $f \circ p$ is a path from x' to y', contradicting our premise.

Pf: (viii) This is adapted from [2]. Let X and Y be connected. Define $U_x \equiv (x, Y) \cup (X, y)$. By (xii) (as proved independently below), U_x is connected (since $(x, Y) \cap (X, y) = \{(x, y)\} \neq \emptyset$). Now, write $X \times Y = \cup_{x \in X} U_x$. Since $\cap_{x \in X} U_x = (X, y) \neq \emptyset$, we again apply (xii) to see that the union (i.e. $X \times Y$) is connected.

Pf: (ix) Let X and Y be path-connected. Suppose $X \times Y$ (in the product topology) is not path-connected. Then there exists (x, y) and (x', y') between which no path exists. However, there exists a path p_1 from x to x' in X and a path p_2 from y to y' in Y. Define $p : [0, 1] \to X \times Y$ via $p(t) = (p_1(t), p_2(t))$. By construction, this is continuous relative to $X \times Y$. [Any open set in $X \times Y$ is a union $\cup(U_i, V_i)$ with each U_i open in X and each V_i open in Y, and the inverse of this is a union $\cup p^{-1}(U_i, V_i) = \cup_i(p_1^{-1}(U_i) \cap p_2^{-1}(V_i))$. Each term is the intersection of two open sets and thus open in $[0, 1]$, and the subsequent union is therefore open as well.] Clearly, $p(0) = (x, y)$ and $p(1) = (x', y')$. We therefore have a path from (x, y) to (x', y'), violating our premise.

Pf: (x) Let X be locally path-connected, and let $U \subset X$ be an open subset of X. A subset $O \subset U$ is open in U iff it is open in X, because the open sets of U (in the subspace topology) are just the intersections of open sets of X with U (which is itself open). Suppose $x \in U$. Pick any neighborhood $W \subset U$ containing x. Then W contains an open set of U containing x. However that open set is also open in X, so W is a neighborhood of x in X as well. Since X is locally path-connected, W contains a neighborhood V of x that is path-connected. V is a subset of U since W is, so it's a neighborhood of x in U too. Therefore, W contains a path-connected neighborhood of x in U, and U is locally path-connected.

Pf: (xi) The only subset of $\emptyset$ that is clopen is $\emptyset$, because it's the only subset. For every pair of points there exists a path, because there are no pairs of points. Every point has a neighborhood with the local path-connectivity condition, because there *are* no points. Any single-point set has only one possible topology, and the same reasoning applies to this.

Pf: (xii) Let $U = \cup_\alpha S_\alpha$, and suppose it can be written $U = A \cup B$ where A and B are open and disjoint. Let $x \in \cap_\alpha S_\alpha$, which we know is nonempty. x is either in A or B, so let's assume it is in A. Since $U = A \cup B$, there is some S_α which intersects B. However, this S_α also intersects A, because $x \in A$ and every S_α contains x. In the subspace topology on this S_α, $A \cap S_\alpha$ and $B \cap S_\alpha$ are, by definition, open. They also must be disjoint, and their union is S_α. I.e., we've partitioned S_α into disjoint open sets, contradicting the premise that it is connected. Therefore U can't be written as a disjoint union of open sets, and it is connected.

Pf: (xiii) Let $U = \cup_\alpha S_\alpha$ and let $x \in \cap_\alpha S_\alpha$, which we know is nonempty. Consider any $y, z \in U$. Each must appear in some member of the union. Let's say that $y \in S_\alpha$ and $z \in S_\beta$. Every S contains x, so S_α is path-connected and contains x and y, and S_β is path-connected and contains x and z. Let f be a path between y and x in S_α and let g be a path between x and z in S_β. We know they agree on $f(1) = g(0) = x$. Define $h : [0, 1] \to U$ via $h(t) = f(2t)$ for $t \in [0, 1/2]$ and $h(t) = g(2t - 1)$ for $t \in [1/2, 1]$. They agree at $t = 1/2$, so this is well-defined. The functions $k_1 : [0, 1/2] \to [0, 1]$ given by $k_1(t) = 2t$ and $k_2 : [1/2, 1] \to [0, 1]$ given by $k_2(t) = 2t - 1$ are continuous, so the compositions $h|_{[0,1/2]} = f \circ k_1$ and $h|_{[1/2,1]} = g \circ k_2$ are continuous. We thus have a function $h : [0, 1] \to U$ which meets the conditions of Lemma 2.3 and is therefore continuous. h defines a path from x to z in U, so U is path-connected.

Pf: (xiv) Suppose everything sits in some space X. Let U be a connected set. Its closure U_C is the intersection of all closed sets containing U. Suppose the closure is not connected. Then $U_C = A \cup B$ for some disjoint open sets A and B. However, $A \cap U$ and $B \cap U$ are disjoint and open in the subspace topology on U, and their union is U. This violates our premise that U is connected unless $A \cap U = \emptyset$ and $B \cap U = U$ or vice versa. Suppose it is the first. Then U is contained in B, and A is disjoint from it. Since A is open and doesn't intersect U, $X - A$ is closed and contains U. Therefore $X - A$ is one of the closed sets contributing to the intersection that defines the closure. I.e., the closure must be a subset of it. However, if $U_C \subseteq X - A$, then $U_C \cap A = \emptyset$, violating our premise that $U_C = A \cup B$.

Every topological manifold is locally path-connected, so for any manifold the notions of "connected" and "path-connected" are the same.

2.2.2. *Components.*

Clearly, any topological space can be written as a union of disjoint path-connected "components" and likewise as a union of disjoint connected "components". This is codified in the following proposition.

We'll refer to these as "path-components" and "connected components" or, when they are the same — as in the case of manifolds — just "components".

Prop 2.5: Path-connectedness (connectedness) defines a partition of X via the largest path-connected (connected) set containing each point. For path-connectedness, this is the same partition as that generated by the equivalence relation $x \sim y$ iff x and y are path-connected.

It is easy to see that this is indeed an equivalence relation. If $f : [0, 1] \to X$ is a path from x to y, then $f(1 - t)$ is a path from y to x. The constant function $f(t) = x$ is a path from x to itself. By using a similar path-concatenation method to the proof of proposition 2.4 part (iii), we can prove transitivity as well.

Equivalently, the class in which x resides (and which we'll denote $[x]$) is the union of all path-connected (connected) sets containing x.

Pf: (path-connectedness): Define $[x] \subseteq X$ to be the union of all path-connected sets containing the point x. By construction, such classes must cover X. By proposition 2.4, part (xiii), a union of path-connected sets with nonempty mutual intersection is path-connected. Therefore, each $[x]$ is path-connected. Given $x \neq y$, $[x] = [y]$ or $[x] \cap [y] = \emptyset$, because if the same y appears in two classes, we can use it to form a path between any points in those classes, just as we did in the proofs of proposition 2.4, parts (iii) and (xiii). The classes therefore form a partition of X.

Pf: (connectedness): Via the same argument as for path-connectedness (but using proposition 2.4, part (xii) rather than (xiii)), we have that the classes are connected and cover X. Consider $x \neq y$, and suppose that $[x]$ and $[y]$ partially overlap. Since $[x]$ and $[y]$ are both connected and are not disjoint, by proposition 2.4, part (ii), $[x] \cup [y]$ is connected. However, this contains both x and y and is larger than the maximal set in $[x]$ or $[y]$, which means that it would supplant it. I.e., we have a contradiction. Therefore $[x] = [y]$ or $[x] \cap [y] = \emptyset$. This establishes a partition of X.

Pf: (equivalence to $x \sim y$ for path-connectedness): Let $[x]$ denote the path-component defined as above, and let $x \sim y$ iff there exists a path between x and y. Suppose $y \in [x]$. Since $[x]$ is path-connected, there exists a path between x and y, so $x \sim y$. Therefore, $x, y \in [x] \implies x \sim y$. Now, suppose that $x \sim y$ and $y \notin [x]$. Given any $z \in [x]$, $z \sim x$ and $x \sim y$ so $z \sim y$ by the transitivity of $\sim$. Therefore, z is path-connected to y. As a result, $[x] \cup \{y\}$ is path-connected, so $[x]$ isn't maximal and we have a contradiction.

The sets of path components and connected components need not be the same in general. Nor need the individual path components be homeomorphic (or even homotopy-equivalent) to one another, and ditto for the individual connected components.

Prop 2.6: (i) Connected components are always closed. (ii) Every path component sits inside a single connected component. (iii) The partition of X into path components is a refinement of the partition into connected components.

Connected components need not be open in general. Path components need be neither open nor closed in general.

To help visualize their relationship, note that each connected component is a disjoint union of path-connected components.

Pf: (i) $[x]$ is the union of all connected sets containing x. From proposition 2.4 part (xii), this union is connected, and from part (xiv) of that same proposition the closure of $[x]$ (call it $[x]_C$) is connected. However, $[x] \subseteq [x]_C$. Since $[x]_C$ contains x, it appears in the union that defines $[x]$. So $[x]_C \subseteq [x]$. Therefore $[x] = [x]_C$, which means it is closed.

Pf: (ii) Denote by $[x]_c$ and $[x]_{pc}$ the connected component and path component containing x. $[x]_c = \cup_\alpha A_\alpha$ over all connected sets containing x, and $[x]_{pc} = \cup_\beta B_\beta$ over all path-connected sets containing x. However, proposition 2.4 part (iv) tells us that every path-connected set is connected. Every set in the union for $[x]_{pc}$ also appears in the union for $[x]_c$. Therefore, $[x]_{pc} \subseteq [x]_c$.

Pf: (iii) This follows immediately from (ii). We have two partitions, and every class of one is a subset of some class of the other.

Prop 2.7: If X is locally path-connected, then (i) every path-component is open, (ii) every path-component is a connected component (and vice versa), (iii) every component is clopen.

Since the path and connected components are identical in this case, we'll just refer to them as "components" when it comes to locally path-connected spaces.

Pf: (i) In a locally path-connected space, every point sits in a path-connected open set. The condition nominally is stricter than this, but this is all we'll need. Consider path-component $X_i \subseteq X$ and point $x \in X_i$. There exists a path-connected open set $O \subseteq X$ that contains x. If O is a subset of X_i, we can take its union with X_i (since they share the point x) and obtain a bigger path-connected set — violating our premise that X_i is a path-component. Therefore, $O \subseteq X_i$. We can therefore write X_i as a union of such open sets, one containing each point in it — which makes it open.

Pf: (ii) Because locally path-connected plus connected implies path-connected, the proof is identical to that of proposition 2.6 part (ii), and we get that every connected component is a subset of some path component. This in conjunction with proposition 2.6 part (ii) tells us that any connected component must be a path component and vice versa.

Pf: (iii) Every connected component is closed, and by (i) in a locally path-connected space every path-component is open. Since the two types of component are the same for a locally path-connected space, all the components are both closed and open.

Prop 2.8: Let $f : X \to Y$ be a continuous function, and let X_i and X_i' be a connected component and path component of X. Then (i) $f(X_i) \subseteq Y_j$ for some connected component Y_j of Y, and (ii) $f(X_i') \subset Y_j'$

for some path component Y'_j of Y. Furthermore, if f is a homeomorphism, (iii) $f|_{X_i}$ is a homeomorphism from X_i to Y_j and (iv) $f|_{X'_i}$ is a homeomorphism from X'_i to Y'_j.

Pf: (i) A continuous function takes connected sets to connected sets. Consider $x \in X_i$. $X_i = \cup S_k$ is the union of all connected sets S_k containing x. Therefore $f(\cup S_k) = \cup f(S_k)$. Since each S_k is connected, $f(S_k)$ is connected and $\cup f(S_k)$ is a union of connected sets, all containing $f(x)$. Therefore, it is a piece of the union that forms some connected component of Y.

Pf: (ii) The same exact argument as (i) holds for path components and path-connected sets.

Pf: (iii) If f is a homeomorphism, then $f(X_i) \subseteq Y_j$ and $f^{-1}(Y_j) \subseteq X_i$, which means $f(f^{-1}(Y_j)) \subseteq f(X_i)$. Although $f(f^{-1}(Y_j)) \subseteq Y_j$ in general, it equals Y_j if f is surjective. Since f is a homeomorphism, it is surjective, so $f(f^{-1})(Y_j) = Y_j$, meaning $Y_j \subseteq f(X_i)$. Since we saw that $f(X_i) \subseteq Y_j$, this means $f(X_i) = Y_j$. The restriction of a homeomorphism is a homeomorphism to its image, so $f|_{X_i}$ is a homeomorphism to Y_j.

Pf: (iv) The same exact argument as (iii) holds for path components.

Note that the situations in which path-connectedness differs from connectedness or where the components are not clopen tend to be pathological and contrived.

Every manifold is locally path-connected, so the notions of connected and path-connected are the same for a manifold and we needn't distinguish between path components and connected components. Moreover, every component is clopen. However, the individual components still need not be homeomorphic (or even homotopy-equivalent) to one another.

For example, the disjoint union $S^n \cup D^n$ is an n-manifold but has non-homotopic components.

As we'll see, for a topological group, the path-connected components are all homeomorphic and the connected components are all homeomorphic, but the connected and path-connected components need not be homeomorphic to one another. For a Lie group, the two are the same, and everything will not only be homeomorphic, but diffeomorphic.

2.2.3. *Quotient and Sets of Components.*

As we saw, path-connectivity and connectivity define two partitions of X into components. We can take the quotient of X by either partition by converting the latter into an equivalence relation.

Denote by $\sim_{pc}$ the equivalence relation on X induced by path-connectivity and by $\sim_c$ the equivalence relation on X induced by connectivity. We can write $\tilde{X}_{pc} \equiv X/\sim_{pc}$ and $\tilde{X}_c \equiv X/\sim_c$. What do these look like? Consider $\tilde{X}_{pc}$ (the same discussion applies to $\tilde{X}_c$). A subset of $\tilde{X}_{pc}$ consists of a set of equivalence classes, aka a set of path-components.

Let T be the topology on X and let T_{pc} be the quotient topology on $\tilde{X}_{pc}$. Denote $T' \equiv p^{-1}(T_{pc})$, in the obvious sense of a set of sets. By construction, T' is a subtopology of T. Specifically, it consists of all open sets of T which are unions of path components.

As we saw in our earlier discussion of quotient topologies, the path components need not form a basis for T'. I.e. T_{pc} need not be the discrete topology on $\tilde{X}_{pc}$. The individual path components may not be open in T, so not all unions of them may be represented in T'. If they *are* all open in T, then T' is quotient-equivalent to the discrete topology on $\tilde{X}_{pc}$.

Again, bear in mind that the classes of X under $\sim_{pc}$ (or $\sim_c$) need not be homeomorphic to one another. They can be quite different. This is true even if X is a manifold.

On the subject of quotients, the quotient of a connected space by any $\sim$ is connected and the quotient of a path-connected space by any $\sim$ is path-connected in the quotient topology. This is codified in the following proposition.

Prop 2.9: Let $\sim$ denote any equivalence relation on X, let $Q = X/\sim$ in the quotient topology, and let $q : X \to Q$ be the quotient map. Then (i) if X is connected, Q is connected, and (ii) if X is path-connected, Q is path-connected.

This may seem counterintuitive. What if we take the quotient of $[0, 1]$ by a $\sim$ which partitions it into $[0, 1/2)$ and $[1/2, 1]$? The quotient is two points, one for each piece, which seems pretty disconnected. However, it's actually not. Let $Q = \{a, b\}$, labeling the two half-intervals. The quotient topology is the largest topology on Q s.t. $q^{-1}(O')$ is open in $[0, 1]$ for every open $O' \subset Q$. Well, $q^{-1}(a) = [0, 1/2)$ and $q^{-1}(b) = [1/2, 1]$. Only the former is open in $[0, 1]$. The only open sets in the topology on Q are $\emptyset$, Q, and $\{a\}$. $\{b\}$ is closed. Therefore, neither $\{a\}$ nor $\{b\}$ is clopen. I.e., the quotient topology is not the discrete topology. In the quotient topology, Q is indeed connected.

Pf: By the definition of the quotient topology, $q^{-1}(O')$ is open in X for every open $O' \subseteq \tilde{X}$. Therefore, q is always continuous. Proposition 2.4, parts (vi) and (vii) tell us that q takes connected sets to connected sets and path-connected sets to path-connected sets. Since it is surjective, if X is connected, so is Q, and if X is path-connected, so is Q.

Prop 2.10: (i) $X/\sim_{pc}$ has the discrete topology iff every path-component of X is open in X. (ii) $X/\sim_c$ has the discrete topology iff every connected component of X is open in X.

Since connected components are always closed, we can replace "open" in (ii) with "clopen".

Pf: These both follow immediately from proposition 2.2.

2.2.4. *Products and Components.*

Prop 2.11: Suppose we have two topological spaces X and Y, and let $Z = X \times Y$ with the product topology. (i) $\tilde{Z}_c = \tilde{X}_c \times \tilde{Y}_c$ and the connected components of Z are $X_i \times Y_j$, where X_i and Y_j are connected components of X and Y. (ii) $\tilde{Z}_{pc} = \tilde{X}_{pc} \times \tilde{Y}_{pc}$ and the path components of Z are $X'_i \times Y'_j$, where X'_i and Y'_j are path components of X and Y.

I.e., we get them the obvious way.

Pf: (i) We know from proposition 2.4, (viii) that the product of two connected sets is connected. Therefore, $X_i \times Y_j$ (aka (X_i, Y_j)) is connected in Z. Moreover, $(X_i, Y_j) \cap (X_k, Y_l) = (X_i \cap X_k, Y_j \cap Y_l)$. This is $\emptyset$ if either $X_i \cap X_k = \emptyset$ or $Y_j \cap Y_l = \emptyset$ (since $(A, \emptyset) = (\emptyset, B) = \emptyset$). This is the case unless $i = k$ and $j = l$. We therefore see that the (X_i, Y_j) are disjoint. They trivially cover $X \times Y$ since the X_i's cover X and the Y_j's cover Y. We therefore have a disjoint cover of connected sets, so these must be the components.

Pf: (ii) We know from proposition 2.4, (ix) that the product of two path-connected sets is path-connected, so the exact same proof holds for path components.

2.2.5. *Number of Components.*

It is perfectly possible to have an uncountable number of connected components (and thus an uncountable number of path components) even if X is second countable. A basis need only generate (via unions) all open sets. If our components are open (and thus clopen), then they indeed must be generated by the basis. However, if they are not, then all bets are off. Only open sets need be expressible as unions of basis sets. Put another way, the components (and path-components) need not be in the topology. This is what it means for them not to be open.

A standard example is $X = \mathbb{R} - \mathbb{Q}$ as a subspace of $\mathbb{R}$. Since $\mathbb{R}$ is second-countable, its subspace X is too. However, every point in X is a connected component.

However, if the connected components are all open, then second-countability implies a countable number of components. This is the case, for example, for any locally path-connected space (which includes all manifolds).

Suppose B is our countable basis, and consider some connected component O. Since O is open, it can be written as a union of basis elements $\cup b_i$. However, the connected components are disjoint, so any b_i appearing in this union cannot intersect any other component. I.e., the connected components induce a partition of the basis. A partition cannot have greater cardinality than the underlying set — so the connected components are countable.

2.3. **Semidirect Products and Exact Sequences.**

Let's quickly review exact sequences and semidirect products. For a more extensive treatment, see [1]. Also, see section 5 for an additional discussion of some practical aspects of semidirect products.

2.3.1. *Exact Sequences.*

Recall that an **exact sequence** of groups is of the form $\cdots \xrightarrow{f_{i+1}} G_{i+1} \xrightarrow{f_i} G_i \to \cdots$, where the G_i's are groups and the f_i's are homomorphisms, and for each pair of adjacent arrows, Im $f_{i+1} = \ker f_i$. The exact sequence may be infinite or finite.

Although exact sequences most frequently occur in the context of groups, the notion readily extends to pointed sets. We'll discuss this in more detail later.

A **short exact sequence** (SES) is of the form $1 \to H \xrightarrow{f} G \xrightarrow{g} K \to 1$, which tells us that Im $f = \ker g$, f is injective, and g is surjective.

An SES is just a group extension of K by H.

An SES $1 \to H \xrightarrow{f} G \xrightarrow{g} K \to 1$ **right-splits** if there exists a homomorphism $j : K \to G$ s.t. $g \circ j = Id_K$, and it **left-splits** if there exists a homomorphism $k : G \to H$ s.t. $k \circ f = Id_H$. It turns out that left-splitting is much stricter and implies right-splitting. For our purposes, we will mostly work with right-splitting SES's.

A quick point of notation. In these notes, we'll care about group extensions of topological groups as well as direct products of them. As a result, we'll sometimes need to speak of $H \times K$ as merely a set or topological product space vs $H \times K$ as a direct product (which includes set, topological product space, *and* group multiplication). Unfortunately, $\times$ is used for both, $\oplus$ is typically only used for a direct sum of abelian groups, and $\otimes$ refers to a tensor product, which is something altogether different. For this reason, we'll use $\overline{\times}$ to denote a direct product group and $\times$ to refer to a cartesian set product or topological product space.

2.3.2. *Group Extensions, Semidirect Products, and Direct products.*

A direct product is a type of semidirect product, which is a type of group extension, and all three correspond to short exact sequences with corresponding constraints.

There are two views of group extensions, semidirect products, and direct products. In the "external" (or "outer") view, we are given two groups and some additional gluing information (in the form of some maps), and we then construct a new group. In the "internal" (or "inner" view), we start with a group G and a normal subgroup $N \triangleleft G$, and then construct the quotient.

To avoid confusion, it's important to keep in mind exactly which of the two views is in play. As we'll discuss, they are equivalent, but they differ in the information we are given, the roles of the groups, and what we are trying to accomplish.

In the internal view, a group extension is nothing other than the standard normal-quotient relationship. Given G and $N \triangleleft G$, we have a canonical SES $1 \to N \xrightarrow{i} G \xrightarrow{q} G/N \to 1$, where i is the subset inclusion map and q is the quotient map that takes each $g \in G$ to the coset gN (which equals Ng since N is normal in G). For a given G and N, this SES has no discretionary degrees of freedom. The internal view is best reflected in the canonical SES.

Although the most general SES has the form $1 \to H \xrightarrow{f} G \xrightarrow{g} K \to 1$, it adds no real generality. The use of H and K and f and g constitute syntactic sugar. We're just renaming the normal subgroup $N = f(H)$ to the isomorphic H and the quotient group $G/f(H)$ to the isomorphic K.

In the external view, we are given groups H and K and some gluing maps. We then construct a group G, the extension of K by H, via those particular gluing maps. The resulting G has a copy of H as a normal subgroup, and the resulting quotient group is isomorphic to K. I.e., it looks like the "more general" SES, but is materially equivalent to $1 \to f(H) \xrightarrow{i} G \xrightarrow{q} G/f(H) \to 1$.

At this point, it may seem as if there's an information mismatch. The canonical SES of the internal view has no moving parts once we're given G and N, but the external view allows us the freedom of choosing gluing maps for the given H and K. Nor is this proscribed in any way. A given H and K can, in fact, produce distinct group extensions enacted by distinct gluing maps. I.e., we have freedom in the external view but none in the internal view. How can both be true if the two views are equivalent?

As a trivial example, any H and K can always produce a direct product, apart from whatever other semidirect products or general group extensions may be formed from them.

There is degeneracy the other way too. We can produce isomorphic G's from different H's and K's via group extensions.

To be clear, the difference isn't merely the aforementioned syntactic sugar. There's a key difference. In the internal view, we are given a specific $N \triangleleft G$. This locks everything down (sans any syntactic sugar we choose to add). In the external view, we don't have a fixed G and N. Instead, we have two unrelated groups H and K. Different gluing maps can produce different (and non-isomorphic) G's from the same H and K. Each such G has its own canonical SES. For example, suppose we have two sets of gluing maps for the same H and K, and these produce G and G' as group extensions of K by H. We haven't explained what the gluing maps are or how extensions are constructed using them, but for now we'll simply state that there exist clearcut criteria for gluing maps and a clearcut procedure for producing a group extension from them. The group extension not only produces G, but gives us a normal subgroup of G that is isomorphic to H. Denote the corresponding internal views for G and G' (and the relevant normal subgroups N and N') $1 \to N \xrightarrow{i} G \xrightarrow{q} G/N \to 1$ and $1 \to N' \xrightarrow{i'} G' \xrightarrow{q'} G/N' \to 1$. Both G and G' are setwise just

$H \times K$. However, the SES notation hides the fact that the group multiplications on G and G' may be quite different. We'll see some details of this below.

In either the internal or external view, (i..e H and K or N and G/N), G setwise looks like $H \times K$ (or $G \times G/N$). The meat is in the multiplication. Labeled this way, the multiplication is codified in the rules that define h'', k'' in $(h'', k'') = (h, k)(h', k')$.

This extends to topology and manifolds. If we're dealing with topological groups (ex. Lie groups), then G looks like $H \times K$ topologically as well as setwise. If they are manifolds, then G looks like the product manifold $H \times K$.

We'll discuss topological groups shortly.

Suppose G is an extension of K by H (via some provided gluing maps), and we label the elements (h, k) (either via the attaching maps or, in the internal view, via (n, k), with $K = G/N$). The copy of H in G is (H, e).

In the internal view, if G/N is isomorphic to a subgroup of G s.t. $G = NK$, then we can use that copy as K instead of G/N itself. This is the case for direct and semidirect products, but not general group extensions. The choice of K is tantamount to a right-splitting map in the corresponding SES.

- External view: If H and K are topological groups then, regardless of the choice of gluing maps, every extension G of K by H is setwise $H \times K$, its elements can be labeled (h, k), and it is a topological group in the product topology.
- External view: If H and K are Lie groups, then G is a Lie group.
- Internal view: If G and N are topological groups then, G is setwise $G \times G/N$, and G/N is a topological group in the quotient topology.
- Internal view: If G and N are Lie groups, then G/N is a Lie group.

Put simply, we don't need to worry about preserving "kind" when it comes to topology and manifold structure. If we use topological groups, we end up with topological groups, using the obvious product topology and quotient topologies. Ditto for Lie groups.

If we have a mix of a topological and Lie group, we get a topological group, but typically not a Lie group.

Given a normal subgroup N of G, G/N is always a group. There are three possibilities: (a) G/N is isomorphic to a normal subgroup K of G s.t. $NK = G$ (i.e. every element of G can be written nk for $n \in N$ and $h \in K$), (b) G/N is isomorphic to a non-normal subgroup of G s.t. $NK = G$, or (c) G/N is not isomorphic to such a subgroup of G.

Note that it is not enough for G/N to be isomorphic to a subgroup of G. It has to be isomorphic to a subgroup K s.t. $NK = G$. For example, if G/N happens to be isomorphic to a subgroup of N, that wouldn't work.

It doesn't matter whether we demand that $NK = G$ or $KN = G$. Since N is normal in G, the two constraints are the same.

2.3.2.1. Direct Product.

A direct product (aka direct sum when the number of groups is finite) is the simplest outcome. It corresponds to the case where G can be decomposed into two normal subgroups $G = NK$.

I.e. every $g \in G$ has an expression nk. It can be shown that this expression is unique.

- Internal view: Case (a) above, where G/N is isomorphic to a normal subgroup $K \subset G$ s.t. $G = NK$.
- SES: A left-splitting SES.
- External view: A direct product $G = H \overline{\times} K$. The copy of H in G is (H, e) and the copy of K in G is (e, K), and the two commute. Multiplication is $(h, k)(h', k') = (hh', kk')$, the identity is (e, e), and the inverse is $(h, k)^{-1} = (h^{-1}, k^{-1})$.

2.3.2.2. Semidirect Product.

A semidirect product is the next simplest outcome. In this case, multiplication twists as we move around K. Now, G can be decomposed into a normal subgroup and a possibly non-normal subgroup. A direct product is a type of semidirect product, where the latter subgroup also is normal.

As with a direct product, every $g \in G$ has a unique expression nk.

- Internal view: Cases (a) and (b) above, where G/N is isomorphic to a (possibly non-normal) subgroup $K \subset G$ s.t. $G = NK$.
- SES: A right-splitting SES. The right-split map produces the relevant image of G/N in G. I.e., it gives us K.
- External view: A semidirect product $G = H \rtimes_\phi K$, where ϕ is the glue map we need for this construction. We'll discuss the details shortly.

Because the internal-view canonical SES admits only a single possible glue map ϕ (which we'll explicitly construct shortly), it is common to write $G = N \rtimes K$ or $G = N \rtimes G/N$, where it is understood that K or G/N is the relevant copy of G/N in G and ϕ is the unique choice consistent with this construction.

From the standpoint of an SES, the right-split map $j : G/N \to G$ gives us a copy of G/N as a subgroup of G. It also can be thought of as identifying the counterpart to e in each coset. Specifically, $j(d)$ is the counterpart of $e \in N$ in the coset $dN = Nd$.

2.3.2.3. General Group Extension.

A group extension is the most general form of the normal-quotient relationship.

- Internal view: Cases (a)-(c) are all group extensions. Generally, G/N need not be isomorphic to a suitable (in the sense that $NK = G$) subgroup of G.
- SES: Any SES is a group extension of K by H. However, the notation obscures the specific gluing maps and the multiplication on G.
- External view: We are given H and K and two gluing maps. The construction (the gluing maps, multiplication, and inverse) are a bit involved. See [1] for the gory details.

2.3.3. *Details of the Semidirect Product Construction.*

Our goal at present is to develop some facility with semidirect products, since these appear a lot in physics and can be counterintuitive at times. In keeping with this, we blitzed through a lot of the background material above. For a far more detailed and thorough exposition, See [1]. We'll now provide a little more detail for the case of semidirect products.

2.3.3.1. External View. Although we'll mostly work in the internal view, let's briefly discuss the external view construction of a semidirect product.

Given groups H and K and a group homomorphism $\phi : K \to Aut(H)$, we can define $G = H \rtimes_\phi K$. Setwise, $G = H \times K$, as for any group extension. As mentioned, if H and K are topological groups, G is a topological group in the product topology and if H and K are Lie groups, G is a Lie group as a product manifold. For convenience, we'll write $\phi_k(h)$ for $\phi(k)(h)$. Bear in mind that ϕ is a homomorphism in its subscript *and* any given ϕ_k is an automorphism of H.

With this product set labeling, multiplication on G takes the form $(h,k)(h',k') = (h\phi_k(h'), kk')$ and the inverse is $(h,k)^{-1} = (\phi_{k^{-1}}(h^{-1}), k^{-1})$. As mentioned, (H,e) is the copy of H in G and (e,K) is the copy of K in G. These do *not* commute unless we have a direct product, in which case $\phi_k = Id_H$ for every k.

In the corresponding internal view $N = (H,e)$ and (e,K) is the internal copy of $G/(H,e)$. In this notation, our usual $G = NK$ expression (with K having the internal view meaning) is $G = (H,e)(e,K)$. Each $g \in G$ can be written $(h,k) = (h,e)(e,k)$ for some $h \in H$ and $k \in K$. This often is written $g = hk$.

As expected, $(h,e)(e,k) = (h\phi_e(e), ek) = (h,k)$ and $(h,e)(h',e) = (h\phi_e(h'), ee) = (hh',e)$ and $(e,k)(e,k') = (e\phi_k(e), kk') = (e,kk')$ since ϕ is a homomorphism (so $\phi_e = Id_H$) and each ϕ_k is an automorphism, so $\phi_k(e) = e$. We also have that $(h,e)^{-1} = (\phi_e(h^{-1}), e) = (h^{-1}, e)$ and $(e,k)^{-1} = (\phi_{k^{-1}}(e), k^{-1}) = (e,k^{-1})$. I.e., everything looks as we expect.

Note that (h,k) is *not* equal to $(e,k)(h,e)$, because (H,e) and (e,K) don't commute. $(e,k)(h,e) = (\phi_k(h),k)$.

2.3.3.2. Internal View. In the internal view, N and $K \approx G/N$ are subgroups of G s.t. $G = NK$. It is not hard to show that each element $g = nk$ in a unique way. I.e., it's perfectly consistent to label elements (n,k) as in the external view.

Where it will aid clarity, we'll mean by "mode 1" treating N and K as subgroups of G and by "mode 2" the labeling (n,k). We can translate between the two via $nk \leftrightarrow (n,k)$.

We can deduce the relevant ϕ for the corresponding external view as follows. (n,k) in mode 2 corresponds to element nk in mode 1. The mode 2 multiplication is $(n,k)(n',k') = (n\phi_k(n'), kk')$. Converting everything to mode 1, we have $nkn'k' = n\phi_k(n')kk'$. I.e., $\phi_k(n') = kn'k^{-1}$. Since N is normal in G, this is an element of N, as needed.

We also can verify the inverse. The mode 2 inverse is $(n,k)^{-1} = (\phi_{k^{-1}}(n^{-1}), k^{-1}) = (k^{-1}n^{-1}k, k^{-1})$. In mode 1, the latter is $k^{-1}n^{-1}kk^{-1} = k^{-1}n^{-1} = (nk)^{-1}$, just as expected.

2.4. **Pointed Sets.**

A **pointed set** is just a set with a preferred point, and a **pointed space** is just a topological space with a preferred point. We'll write them (X,x), where x is the preferred point.

For any set with an algebraic structure that involves an identity element, that identity element is a natural candidate for a preferred point. Accordingly, every group G can be viewed as a pointed set (G, e) in a natural way by forgetting the rest of the group structure. Modules are pointed sets via the abelian identity 0. If two algebraic identity elements are present (ex. a unital ring or unital algebra has a 0 and a 1), then either is a candidate for the preferred point, and we can view the underlying set as a pointed set in two natural ways. Which of these, if either, is preferable depends on the application. Most commonly, it is the additive identity 0, but it need not be.

Note that it is irrelevant whether the object is a group under multiplication, or even closed under it. The mere existence of a "special element", such as 0 or 1, provides a natural pointed-set structure.

A homomorphism f from pointed set (X, x) to pointed set (Y, y) is just a set map $f : X \to Y$ s.t. $f(x) = y$. An isomorphism of pointed sets is a bijective homomorphism (or, equivalently, a homomorphism with an inverse that is a homomorphism). Since algebraic homomorphisms take identity elements to the corresponding identity elements, they are homomorphisms of pointed sets too.

Homomorphisms of pointed sets are sometimes called "pointed maps" or "based maps".

Why do we care about such a primitive object? A pointed set is the minimum structure needed to define the kernel of a function. For a group homomorphism, $\ker f = f^{-1}(e)$ and for a module homomorphism, $\ker f = f^{-1}(0)$, etc. The essential feature is that there is a "special element", allowing us to define $\ker f$ as f^{-1} of this special element.

The presence of a kernel allows us to generalize the notion of an exact sequence. For simplicity, we'll denote the trivial pointed set $(\{x\}, x)$ by 1_P. The exact sequence $\cdots \xrightarrow{f_{i+1}} (X_{i+1}, x_{i+1}) \xrightarrow{f_i} (X_i, x_i) \to \cdots$ means that the f_i's are pointed-set homomorphisms and $\text{Im } f_{i+1} = \ker f_i$. The image is a purely set-theoretic concept, but the kernel requires, at bare minimum, a pointed set.

As usual, the exact sequence (or piece of exact sequence) $\xrightarrow{f} (X, x) \to 1_P$ tells us that f is surjective. However, $1_P \to (X, x) \xrightarrow{f}$ only tells us that $\ker f = \{x\}$. Unlike for a group homomorphism, this does *not* imply that f is injective. For that to follow, we need additional algebraic structure that more heavily constrains f.

Ex. let $(X, x) = (\{1, 2, 3, 4\}, 1)$ and $(Y, y) = (\{a, b\}, a)$. The map $f : (1, 2, 3, 4) \to (a, b, b, b)$ is a pointed-set homomorphism and has $\ker f = \{1\}$, but is not injective.

We can define an SES, $1 \to (X, x) \xrightarrow{f} (Y, y) \xrightarrow{g} (Z, z) \to 1$, but this does not have any significant algebraic implications (as it does for groups). There are no pointed-set counterparts to normal subgroups or quotient groups. It is too simple a structure. All that a pointed set allows us to do is meaningfully speak of kernels and exact sequences.

2.5. **Homotopy.**

2.5.1. *Homotopic Functions.*

Recall that two continuous functions f and g between topological spaces X and Y are **homotopic** if one can be continuously transformed into the other. Formally, there must exist a continuous function (relative

to the product topology) $h : [0,1] \times X \to Y$ for which $h(0,x) = f(x)$ and $h(1,x) = g(x)$. Such an h is termed a **homotopy between f and g.**

Prop 2.12: Homotopy defines an equivalence relation on the set of continuous functions from X to Y.

Pf: (reflexive) $h(t,x) = f(x)$ for all $t \in [0,1]$ is a homotopy from f to itself, so $f \sim f$. (symmetric) If $f \sim g$, and h is the relevant homotopy from f to g, then $h(1-t,x)$ is a homotopy from g to f, so $g \sim f$. (transitive) Let $f_1 \sim f_2$ and $f_2 \sim f_3$, with h and h' the relevant homotopies. Define $p_1 : [0,1/2] \to [0,1]$ via $p_1(t) = 2t$ and define $p_2 : [1/2,1] \to [0,1]$ via $p_2(t) = 2t-1$. Both are patently continuous. Define $h''(t,x) \equiv h_1(p_1(t),x)$ for $t \in [0,1/2]$ and $h''(t,x) \equiv h_2(p_2(t),x)$ for $t \in [1/2,1]$ (i.e. $h''(t,x) = h_1(2t,x)$ for $t \in [0,1/2]$ and $h''(t,x) = h_2(2t-1,x)$ for $t \in [1/2,1]$). Since $h_1(1,x) = h_2(0,x) = f_2(x)$, this is well-defined at the overlap point $t = 1/2$. The restriction of h'' to each of the two closed sets $[0,1/2] \times X$ and $[1/2,1] \times X$ is the composition of continuous functions (for each x) and therefore continuous. By lemma 2.3, we therefore see that h'' is continuous on $[0,1] \times X$. Since $h''(0,x) = f_1(x)$ and $h''(1,x) = f_3(x)$, h'' is a homotopy between f_1 and f_3. Therefore, $f_1 \sim f_3$, and we have transitivity.

If we pick a base point $x \in X$ and a base point $y \in Y$, then we can restrict ourselves to continuous functions which take x to y (i.e. continuous functions which are pointed-set homomorphisms from (X,x) to (Y,y)). The proof above still holds, so we get an equivalence relation on this set of functions. Note that the classes of functions under this new equivalence relation can be both fewer and smaller than those under the unrestricted homotopy equivalence relation. Nor does this happen solely due to the restricted class of functions under examination.

Let $C(X,Y)$ denote the set of all continuous functions, let $C_{x,y}(X,Y)$ denotes the set of base-point preserving continuous functions (i.e. those with $f(x) = y$), and let $[f]$ denote a homotopy class in $C(X,Y)$. It may be tempting to think that $[f] \cap C_{x,y}(X,Y)$ (which possibly is empty) is the corresponding homotopy class in $C_{x,y}(X,Y)$. I.e., that we simply weed out all non-basepoint-preserving functions from the existing classes and then discard any classes which are consequently empty. Unfortunately, it's not this simple. The problem is that $\sim$ now requires not only that f and g be basepoint preserving and homotopic, but that a homotopy exists between them that only consists of basepoint-preserving functions (i.e., $h(t,x) = y$ for all $t \in [0,1]$). Under this criterion, there are fewer eligible homotopies. It's quite possible that f and g could both be basepoint-preserving and $f \sim g$ under the original homotopy definition, but $f \not\sim g$ if we require a homotopy involving only basepoint-preserving functions. Stated more concisely, f and g may be free-homotopic but not base-homotopic. See [3] for an example.

The terms "based homotopy" and "based homotopic" are commonly used for the basepoint-preserving cases. The non-basepoint-preserving cases are sometimes referred to as "free homotopy" and "free homotopic".

2.5.2. *Loops and n-loops.*

A loop in 1-dimension is just a path in X s.t. $f(0) = f(1)$. I.e., a loop is a continuous map $f : [0,1] \to X$ s.t. $f(0) = f(1)$. Equivalently, it is a continuous function $f : S^1 \to X$. Note that it may cross itself.

We may generalize the notion of a 1-dimensional path to an n-dimensional path. This is a continuous map $[0,1]^n \to X$. We'll call this an n-path, for lack of a better term.

Note that $[0,1]$ will play two roles in our discussion, and they should not be confused. For homotopies between maps, it always appears as $[0,1]$ in the expression $h : [0,1] \times X \to Y$. However, as the domain for a path, $[0,1]$ only applies to 1-paths and 1-loops, which are used in the definition of the fundamental group π_1. For any higher homotopy group π_n with $n > 1$, we use n-paths which involve $[0,1]^n$ as their domain.

There are a few candidates for the n-dimensional version of a loop. They are easiest to visualize in $n = 2$, so lets do that. Consider $X = \mathbb{R}^3$. $f([0,1]^2)$ defines a surface in $\mathbb{R}^3$. This surface may be bounded or unbounded, open, closed, or neither. The boundary of this surface is the image under f of the four sides of the square $[0,1]^2$. I.e., $A = [0,1] \times \{0\}$, $B = [0,1] \times \{1\}$, $C = \{0\} \times [0,1]$, and $D = \{1\} \times [0,1]$. We now have several obvious candidates for the analogy with a 1-loop. We could require that (i) $f(t,0) = f(t,1)$ for $t \in [0,1]$ or (ii) $f(0,t) = f(1,t)$ for $t \in [0,1]$ or (iii) both or (iv) that $f(A \cup B \cup C \cup D)$ is constant (i.e. f is constant on the boundary of the square).

Choices (i) and (ii) are materially the same and correspond to homotopies of 1-loops. Consider (i). For each $t \in [0,1]$, we have a 1-loop given by $f(t,s)$ with that t fixed, because $f(t,0) = f(t,1)$ and $f(t,s)$ is continuous. Overall, f itself is a homotopy from the loop $f(0,s)$ to the loop $f(1,s)$. Topologically, we can think of this as gluing one edge of $[0,1]^2$ to its opposing edge. We thus have a map from the cylinder $S^1 \times [0,1]$ to X. Consider choice (iii). Topologically, this is like gluing both pairs of opposing edges. This results in a function from the torus $S_1 \times S_1 \to X$. (iv) acts like the one-point compactification of $(0,1)^2$, gluing the entire boundary square of $[0,1]^2$ to a point, which results in S^2. We then have a map from S^2 to X.

It turns out that (iv) is the most useful generalization. Our n-loop is defined as a continuous map $f : [0,1]^n \to X$ s.t. f is constant on the boundary of $[0,1]^n$. Equivalently, f can be viewed as a continuous map from $S^n \to X$, just as a 1-loop is a continuous map from $S^1 \to X$.

This also is consistent with the intuitive interpretation of the homotopy group π_n as capturing the n-dimensional spherical holes in X.

Let's now consider a pointed space (X, x_0) and anchor our loop at x_0. An n-loop requires that the constant value on the boundary equals x_0. In the case of a 1-loop, the boundary is just $\{0,1\}$, and we're requiring that the loop start and end at x_0. We'll call these "based n-loops" or "based loops". Where needed, we'll denote by s_0 the basepoint of S^n corresponding to the collapsed boundary of $[0,1]^n$. I.e., it is the compactification point.

Bear in mind that the n-loop maps need not be injective. n-loops can cross themselves.

Since each n-loop is a basepoint-preserving map from (S^n, s_0) to (X, x_0), we can speak of a based homotopy of n-loops. Every continuous map from S^n to X is an n-loop, so every continuous map that takes s_0 to x_0 is a based n-loop. In coming up with a relevant notion of homotopy between based n-loops, we therefore don't need anything further in terms of restrictions. A based homotopy is automatically a homotopy that is, for every t, a based n-loop.

We thus have our usual based-homotopy equivalence relation on the space of based n-loops on (X, x_0).

2.5.3. *Algebra of Loops.*

We can endow the set of based n-loops on (X, x_0) with a group structure via loop-concatenation. This is easiest to describe if we regard n-loops as functions with domain $[0,1]^n$ rather than S^n (though it can be defined directly for the latter as well, with a bit of extra effort).

Let f and g be two based n-loops. Each is a map $[0,1]^n \to X$ that takes the entire boundary of $[0,1]^n$ to x_0. Denote by $t = (t_1, \dots, t_n)$, a point in $[0,1]^n$. We concatenate two loops by defining $h(t) \equiv f(2t_1, t_2, \dots, t_n)$ on the half of $[0,1]^n$ with $t_1 \leq 1/2$ and $h(t) \equiv g(2t_1 - 1, t_2, \dots, t_n)$ on the half of $[0,1]^n$ with $t_1 \geq 1/2$. This is well-defined since $f(1, \dots) = g(0, \dots) = x_0$. Since each half as defined is a closed set and h restricted to each half is continuous, lemma 2.3 tells us that h is continuous. It is also clearly equal to x_0 on the entire boundary. Therefore, it is a based n-loop. This operation is commonly denoted $f * g$ for $n = 1$ and $f + g$ for $n > 1$. We'll use $*$ in both cases.

The reason for this difference in notation is that $\pi_n(X)$ is always abelian for $n > 1$, so additive notation is appropriate. $\pi_1(X)$ may or may not be abelian, so multiplicative notation is retained.

In terms of spheres, if we think of an n-loop as a balloon in X anchored at x_0, concatenation simply means tying two balloons together at their bases (and reparametrizing them so each gets half the air). A concatenation of multiple 1-loops is commonly called a "bouquet", because it looks like a bunch of flower petals attached at x_0.

We could just as well split the $[0,1]^n$ block in any other way, as long as we can fill the two halves in a continuous manner that allows us to apply lemma 2.3.

The obvious candidate for an "identity" element is the trivial n-loop that takes all of $[0,1]^n$ to x_0 (i.e. $I([0,1]^n) = x_0$). For each n-loop f, we can define a candidate inverse n-loop via $f^{-1}(t_1, \dots t_n) = f(1-t_1, t_2, \dots, t_n)$. Why are these merely candidates? We don't quite have a group yet. The problem is the reparametrization.

For any n-loop f, $I * f$ and $f * I$ result in a reparametrization of f. Setwise the result is the same, and we traverse the new n-loop along t_1 in the same direction as the original loop, but the resulting function isn't identical to f. For example, in $f * I$, we hurry along t_1 to finish f twice as quickly, and then we just sit at x_0 for the remaining half of t_1's time.

Similarly, $f^{-1} * f$ and $f * f^{-1}$ result in traversing f forward and the backward. The resulting loop involves both operations, each squeezed into half of the $[0,1]$ interval for t_1. Concatenation isn't even associative. $f * (g * k)$ is setwise the same as $(f * g) * k$, but each is parametrized differently. $f * (g * k)$ traverses f twice as quickly as normal and each of g and k four times as quickly, while $(f * g) * k$ traverses f and g four times as quickly as normal and k twice as quickly.

The gist is that the raw concept of loop-concatenation is too specific for our purpose. However, it can be adapted to our needs without much effort. Both concatenation and our candidate inverse operation respect the based homotopy classes of n-loops. Class concatenation turns out to be just the right concept for the task. I.e., we need to work in the quotient loop space rather than the loop space itself. Denote by L_n the space of all based n-loops on (X, x_0).

Prop 2.13: Let f, f_1, f_2, f_3, and g be based n-loops on (X, x_0), and let $\sim$ denote that two based n-loops are based-homotopic. Then

- (i) $f_1 \sim f_2$ iff $f_1 * g \sim f_2 * g$
- (ii) $f_1 \sim f_2$ iff $g * f_1 \sim g * f_2$
- (iii) $f \sim g$ iff $f^{-1} \sim g^{-1}$
- (iv) $f^{-1} * f \sim I$ and $f * f^{-1} \sim I$
- (v) $f * I \sim f$ and $I * f \sim f$
- (vi) $f_1 * (f_2 * f_3) \sim (f_1 * f_2) * f_3$

Pf: (i) Given based-homotopy $h : [0,1] \times [0,1]^n \to X$ from f_1 to f_2, define $h' : [0,1] \times [0,1]^n \to X$ via $h'(s,(t_1,t_2,\ldots,t_n)) = h(s,(2t_1,t_2,\ldots,t_n))$ for $t_1 \in [0,1/2]$ and $h'(s,(t_1,t_2,\ldots,t_n)) = g(2t_1-1,t_2,\ldots,t_n)$ independent of s for $t_1 \in [1/2,1]$. This is well-defined because $h(s,(1,t_2,\ldots,t_n)) = x_0$ (since h is a based homotopy) and $g(0,t_2,\ldots,t_n) = x_0$ (because g is an n-loop based at x_0). It is easy to see that h' is continuous (via the same reparametrization arguments and use of lemma 2.3 as before). By construction, $h'(0,t) = (f_1 * g)(t)$ and $h'(1,t) = (f_2 * g)(t)$. Moreover, each $h'(s,-)$ is of the form $h(s,-)*g$, which is a based n-loop. Therefore h' is a based homotopy between $f_1 * g$ and $f_2 * g$. Going the other way, if we start with $f_1 * g \sim f_2 * g$ via based-homotopy h, we define $h'(s,(t_1,\ldots,t_n)) = h(s,(t_1/2,t_2,\ldots,t_n))$. This clearly is continuous and takes the values f_1 and f_2 at $s=0$ and $s=1$. For a given s, $h'(s,-)$ is something for the first half, followed by g. However, we specified that f_1 and f_2 are n-loops and h' is a based homotopy. Therefore, $h(s,-)$ must be an n-loop. The only way to get an n-loop whose latter half is the n-loop g is for the former half to be an n-loop in its own right. I.e., $h(s,-) = f_s * g$ for some based n-loop f_s. This yields $h'(s,-) = f_s$, so h' is a based homotopy from f_1 to f_2.

Pf: (ii) The same exact argument works on the other side by swapping which half of the t_1 interval is assigned to g.

Pf: (iii) Given based-homotopy $h : [0,1] \times [0,1]^n \to X$ from f to g, define $h'(s,(t_1,t_2,\ldots,t_n)) \equiv h(s,(1-t_1,t_2,\ldots,t_n))$. Then $h'(0,t) = f^{-1}(t)$ and $h'(1,t) = g^{-1}(t)$ according to our definition of $-^{-1}$. Since h' is patently continuous and $h'(s,-) = h(s,-)^{-1}$ is a based n-loop, h' is a based homotopy from f^{-1} to g^{-1}. Since $(f^{-1})^{-1} = f$ by our definition, the same argument holds the other way, and we have our iff.

Pf: (iv) This one's a bit less intuitive. The trick is that since we're traversing the same loop one way and then the other, we can pick any point along the first traversal and just start reversing from there. I.e., we can get tired and turn around on our hike at any point. However, our hike must have a fixed duration, so we have to rest for a while before turning around. We implement this by creating a constant piece that grows in the middle. As we expand this outward, it moves toward x_0, because we turn around closer and closer to our starting point. Define $h(s,t) = f(2t_1,\ldots,t_n)$ for $t_1 \in [0,(1-s)/2]$, $h(s,t) = f^{-1}(2t_1-1,\ldots,t_n) = f(2-2t_1,t_2,\ldots,t_n)$ for $t_1 \in [(1+s)/2,1]$ and $h(s,t) = f((1-s),t_2,\ldots,t_n)$ for $t_1 \in [(1-s)/2,(1+s)/2]$. This is well-defined because $f^{-1}(2((1+s)/2)-1,t_2,\ldots,t_n) = f(2-2(1+s)/2,t_2,\ldots,t_n) = f(1-s,t_2,\ldots,t_n)$. All the functions involved are continuous, and $h(s,-)$ is indeed a based n-loop (albeit a simple one, where we move out along a path and then retrace our steps). $h(0,-)$ is just $f * f^{-1}$ since it traverses f for $t_1 \leq 1/2$ and f^{-1} for $t_1 \geq 1/2$. $h(1,0) = I$ since it equals the constant x_0 for all t_1. We thus have a homotopy from $f * f^{-1}$ to I. An identical argument holds for $f^{-1} * f$.

Pf: (v) Define $h(s,(t_1,\ldots,t_n)) \equiv f(2t_1/(1+s),t_2,\ldots,t_n)$ for $t_1 \in [0,(1+s)/2]$ and $h(s,(t_1,\ldots,t_n)) = x_0$ for $t_1 \in [(1+s)/2,1]$. This is well-defined for $t_1 = (1+s)/2$ because $f(1,t_2,\ldots,t_n) = x_0$. It is patently continuous in the range $s \in [0,1]$ and has $h(0,-) = f * I$ and $h(1,-) = f$, and it is easy to see that $h(s,-)$ is a based n-loop followed by I (but not traversed over equal time portions as would be $f * I$). Therefore, $f * I \sim I$. The same argument holds for $I * f$. Note that if we try to play the same game by expanding the I half instead of the f half (i.e. try to prove that $f * I \sim I$, which obviously must be false), we do not get a based homotopy. Let's try to reverse their roles and shrink $I * f$ to I in the same fashion. In that case, we would need something like $h(s,(t_1,\ldots,t_n)) = x_0$ for $t_1 \in [0,(1+s)/2]$ and $h(s,(t_1,\ldots,t_n)) \equiv f((2t_1-(1+s))/(1-s),t_2,\ldots,t_n)$ for $t_1 \in [(1+s)/2,1]$. However, this is not continuous at $s=0$, so we don't have a homotopy. We'd run into the same problem with any other approach. This just reflects the fact that an arbitrary based n-loop need not retract to a point.

Pf: (vi) The approach is the same as in the rest of our proof above, but the details are tedious and unenlightening and we won't explicitly work through them here.

2.5.4. *Homotopy groups* $\pi_n(X,x_0)$.

Define $\pi_n(X,x_0) \equiv L_n/\sim$ to be the set of based homotopy classes. Since the operations in question respect $\sim$, define $[f] * [g] \equiv [f * g]$ and $[f]^{-1} \equiv [f^{-1}]$ and $e \equiv [I]$, where [] denotes a based homotopy class of n-loops (i.e. a subset of L_n or an element of $\pi_n(X,x_0)$). Proposition 2.13 tells us that these definitions impose a group structure on $\pi_n(X,x_0)$.

This is *not* a quotient group. We saw that concatenation fails to endow L_n itself with a group structure. Only by passing to the quotient set $\pi_n(X,x_0)$ can we do so.

The $\pi_n(X,x_0)$'s are called the **homotopy groups** (or sometimes the "based homotopy groups") of X with basepoint x_0. Intuitively, $\pi_n(X,x_0)$ measures the n-dimensional spherical holes in X. $\pi_1(X,x_0)$ is the most important of these, and is termed the **fundamental group**.

How can we speak of the π_n's as measuring holes in X, when they are tied to a particular x_0? The answer is that we can't in general. However, within each path-connected component we can.

Prop 2.14: If x_0 and x_1 are in the same path-connected component of X, then $\pi_n(X, x_0) \approx \pi_n(X, x_1)$.

Pf: We'll just provide a proof sketch here. The intuition is simple. We use the path between x_0 and x_1 to construct an isomorphism. Let $p : [0,1] \to X$ be a path from x_0 to x_1. Suppose that $[f]$ is an element of $\pi_n(X, x_0)$. Pick any $f \in [f]$ and construct the x_1-based loop $f' \equiv p * f * p^{-1}$ (appropriately reparametrized in the obvious way, and treating our path p as an n-dimensional map $p : [0,1]^n \to X$ by simply ignoring the other $n-1$ dimensions). It is easy to see that $g \sim f$ iff $g' \sim f'$. This establishes a bijection. The usual arguments then show that concatenation of classes is preserved, establishing a homomorphism from $\pi_n(X, x_0)$ to $\pi_n(X, x_1)$. A bijective homomorphism is an isomorphism, which gives us what we need. Note that the specific isomorphism does depend on p. For example, if there is a hole, p may pass on either side. It is impossible to homotopically move between such paths. We therefore end up with distinct isomorphisms.

For a path-connected space, we can write $\pi_n(X)$ without worrying about the basepoint, since all such groups are isomorphic. However, for a non-path-connected space, we cannot.

There is *no such thing* as $\pi_n(X)$ for a general non-path-connected space, since the homotopy groups (for $n \geq 1$) only have meaning for a given path-component. In that case, we must write $\pi_n(X, x_0)$. However, there is an important class of cases where we can forego this.

It can be shown that if X and Y are homeomorphic, their homotopy groups are identical. If the path-components are all homeomorphic, then the π_n's must be isomorphic across path-components, not just within them, and we may meaningfully speak of $\pi_n(X)$ without reference to a basepoint.

One important example of this is when X is a topological group. As we'll discuss shortly, for a topological group all the path-components are homeomorphic to one another. In that case (which includes any Lie group), $\pi_n(X)$ may be written instead of $\pi_n(X, x_0)$.

Note that we don't actually care for this purpose *what* the group structure is. If the topology on X *admits* any topological group structure, then the path components must be homeomorphic. Admitting a topological group structure is highly constraining.

The homotopy groups $\pi_n(X, x_0)$ (for $n > 0$) are groups under $*$, so they are also pointed sets. The special point is $e = [I]$ (with I the aforementioned identity based n-loop that takes all of S^n to x_0). Conceptually, this is the set of all n-spheres in X that are anchored at x_0 and can be continuously shrunk to the point x_0 (also known as "deformation retracting to a point") If there is an n-dimensional hole, we can't shrink any balloon containing it to a point. The balloon gets stuck on it and can't collapse further.

We're not working with manifolds, so we must be careful what we mean by an "n-dimensional hole". Simplicial or cell complexes offer a clear meaning, but we won't delve into them here. See [4] for an excellent discussion of this (and of all of homotopy theory).

2.5.5. *Extending this to* $\pi_0(X, x_0)$.

So far, we've only considered n-loops for $n > 0$. With care, we can extend our machinery to 0-loops. To do so, we generalize $S^n \to X$ to include the case of S^0. Topologically, S^0 is two points in the discrete topology, usually written $\{0, 1\}$. Every map from a discrete topology is continuous relative to it, so we don't need to worry about continuity. We need to anchor at least one point to x_0 for it to be a 0-loop based at x_0. What about the other? It need not be x_0. Just as only one point in S^n is anchored to x_0, only one point in S^0 is. I.e., $f : \{0, 1\} \to X$ is a based 0-loop iff $f(0) = x_0$. This means that a based 0-loop is just a choice of point $f(1) \in X$.

Consider two based 0-loops: f and g. A based homotopy h from f to g is a continuous function $h : [0,1] \times \{0,1\} \to X$ s.t. $h(t, 0) = x_0$, $h(0, 1) = f(1)$ and $h(1, 1) = g(1)$. I.e., it's a continuous function

$h(t,1) : [0,1] \to X$ with $f(1)$ and $g(1)$ as the endpoints. But that's just the definition of a path from $f(1)$ to $g(1)$. I.e., a based homotopy of based 0-loops is just a path between $f(1)$ and $g(1)$. A based homotopy exists iff a path exists.

By definition, a path exists iff $f(1)$ and $g(1)$ are in the same path-component. Therefore, the based homotopy classes of based 0-loops are just the path-components of X. Unlike for $n > 0$, this is true regardless of x_0 or its path-component. We don't care if x_0 is path-connected to $f(1)$ or $g(1)$ — just whether $f(1)$ and $g(1)$ are path-connected to one another.

We therefore can write $\pi_0(X)$ without any qualms, though we sometimes benefit from keeping the basepoint x_0 — as we'll see momentarily. $\pi_0(X)$ is naturally bijective with the set of path-components of X, and we'll just treat it *as* that set. However, unlike the other π_n's, π_0 is not a group. There is no notion of concatenation. We have one degree of freedom in a 0-loop and we cannot compress and reparametrize it.

The best we can do is treat $\pi_0(X)$ as a pointed set. If we have a basepoint x_0 (perhaps one already in use for the other π_n's), we have a special path-component: that containing x_0. This lends π_0 the structure of a pointed set, thus endowing it with notions of "kernel" and "homomorphism". These, in turn, allow π_0 to appear in long exact sequences. This turns out to be quite handy, because the generalization of π_n to π_0 plays well with the standard long exact sequences which arise in topology. This saves us from having to separate out the 0-dimensional edge case and give it special treatment. It also conceptually unifies π_0 with the other π_n's.

As discussed earlier, if π_0 appears in a long exact sequence, it is almost always as a pointed set. On the other hand, if it appears in a short exact sequence as $1 \to G_0 \xrightarrow{i} G \xrightarrow{q} \pi_0(G) \to 1$, π_0 has a distinct quotient group structure inherited from G and not related to the homomotopy group construction of the higher π_n's. We'll discuss this shortly. For now, we simply observe that it is a distinct concept.

More generally, if we have an exact sequence of $\pi_n(X)$'s, there are three typical cases:

- If it is a long exact sequence and either X is path-connected or all the path-components are homeomorphic, there is no ambiguity. Everything is being viewed as a pointed set, including π_0 (which is a single-point). In this case, $\pi_n(X)$ (including $\pi_0(X)$) may appear without an explicit choice of basepoint.
- If we have a long exact sequence of π_n's and X is not path-connected and the path-components are not homeomorphic, then we need based homotopy groups. It is $\pi_n(X, x_0)$ (and $\pi_0(X, x_0)$) which must appear in the sequence. Everything is a pointed set.
- If we have a topological group G, we'll see that we have a short exact sequence of the form $1 \to G_0 \xrightarrow{i} G \xrightarrow{q} \pi_0(G) \to 1$. In this case, G_0 is the identity path component of G and $\pi_0(G)$ is the quotient group G/G_0. As far as pointed sets go, $\pi_0(G)$ is the same as the topological pointed set $\pi_0(X, x_0)$ if we choose the group identity element of G (i.e. $x_0 = e$) as the basepoint. However, now $\pi_0(G)$ is actually a group.

A common example of a long exact sequence is the "FEB" sequence for a fiber bundle. Let E be a fiber bundle with base B and projection map $p : E \to B$. Let $b_0 \in B$ be our choice of base-point for B. Let $e_0 \in p^{-1}(b_0)$ be our choice of preferred point in E. We'll designate the canonical fiber as $F = p^{-1}(b_0)$ and its base-point as $f_0 = e_0 \in F$. We then have a long exact sequence of homotopy groups $\cdots \to \pi_n(F, f_0) \to \pi_n(E, e_0) \to \pi_n(B, b_0) \to \pi_{n-1}(F, f_0) \to \cdots \to \pi_0(F, f_0) \to \pi_0(E, e_0) \to \pi_0(B, b_0) \to 0$. Up until the π_0's, we have group homomorphisms (which also are pointed-set homomorphisms). The map from $\pi_n(E, e_0)$ to $\pi_n(B, b_0)$ is just the homomorphism induced by the projection map $p : E \to B$. The map from $\pi_n(F)$ to $\pi_n(E)$ is the homomorphism induced by the subset inclusion $i : F \to E$. The last homomorphism from $\pi_n(B, b_0)$ to $\pi_{n-1}(F, f_0)$ is induced by something called the "boundary map", which we won't go into here (see [5] for a discussion). Now, consider the π_0 part. The map from $\pi_0(F, f_0)$ to $\pi_0(E, e_0)$ is easy enough. Each path component of the canonical fiber over b_0 sits in a unique path component of E. Obviously, the path-component containing f_0 sits in the path-component of E containing $e_0 = f_0$. The map therefore is a pointed-set homomorphism. Of course, things that look disconnected within F may be path-connected away from F, so the map from $\pi_0(F, f_0)$ to $\pi_0(E, e_0)$ need not be injective. Next, consider the map from $\pi_0(E, e_0)$ to $\pi_0(B, b_0)$. p^{-1} of the path components of B are disjoint — though each can contain more than one path-component of E (since F can introduce further path-disconnectedness) so $\pi_0(B, b_0)$ is mapped to by a set of path components of E. The map is induced by p in the obvious way, since a path in E projects to a path in B. Since e_0 projects to b_0, the path-component containing e_0 projects to the path-component containing b_0, and we have a pointed-set homomorphism. The last part ($\pi_0(B, b_0) \to 0$) simply tells us that this homomorphism is surjective — which is obvious because every path-component of B lifts to a least one path-component of E. The only remaining term is the $\pi_1(B, b_0) \to \pi_0(F, f_0)$. Such a map assigns to each homotopy class of loops in B based at b_0 a path-component of the fiber over b_0.

As mentioned, it can be shown that $\pi_n(X)$ is abelian for $n > 1$. However, $\pi_1(X)$ need not be abelian and generally is not. It turns out that for topological groups, π_1 is both abelian and finitely generated.

Note that $\pi_0(X)$ need not have the discrete topology, and it may be countable or uncountable.

Prop 2.15: $\pi_0(X)$ has the discrete topology iff every path-component of X is open in X.

Pf: $\pi_0(X) = X/\sim_{pc}$, so this just restates proposition 2.10, part (i), which itself is just a case of proposition 2.2.

2.6. **Topological Groups.**

Recall that a topological group is a topological space X with a group structure such that multiplication $\times : X \times X \to X$ and the inverse $-^{-1} : X \to X$ are continuous functions.

As a group, a topological group has a notion of an identity element. Unlike in a general topological space, we have a preferred path-component and a preferred connected component: those containing the identity element. These are called the **identity path-component** and the **identity connected-component**.

Being a topological group is quite restrictive because of the continuity conditions on $\times$ and $-^{-1}$. In fact, it forces the path-components to all look the same and the connected components to all look the same.

Prop 2.16: For a topological group G and any $h \in G$, the maps $L_h, R_h : G \to G$ defined by $L_h(g) = h \cdot g$ and $R_h(g) = g \cdot h$ are homeomorphisms.

These are known as left and right translations.

Pf: We'll prove it for L_h, and the proof is materially identical for R_h. (i) Surjective: Given $g \in G$, $L_h(h^{-1}g) = g$. (ii) Injective: Let $hg = hg'$. Multiply on the left by h^{-1} and we have $g = g'$. (iii) L_h continuous: Consider an open set $O \subseteq G$. $L_h^{-1}(O)$ is the set of $g \in G$ s.t. $hg \in O$. We know that $\times : G \times G \to G$ is continuous, so $U \equiv \times^{-1}(O)$ is open in $G \times G$. Consider $(h, G) \subset G \times G$ in the subspace topology, which consists of all open sets of $G \times G$ intersected with (h, G). Obviously, (h, G) is open in its own subspace topology (as $(G \times G) \cap (h, G)$). Since U is open in $G \times G$, $U \cap (h, G)$ is open in (h, G). Denote it $(h, U') \equiv (h, G) \cap U$. Then U' is precisely $L_h^{-1}(O)$, the set of g's such that $hg \in O$. There is a trivial homeomorphism $(h, g) \to g$ between (h, G) and G. Therefore, $U' \subset G$ is open in G iff (h, U') is open in (h, G). Since (h, U') is open in (h, G), $U' \subset G$ is open in G. We thus see that $L_h^{-1}(O)$ is open in G. (iv) $(L_h)^{-1}$ continuous: $L_{h^{-1}}(hg) = g$, so $(L_h)^{-1}$ (as an inverse map) is just $L_{h^{-1}}$. We've already shown that L_h is continuous for all h, which includes the inverse of any given h. (v) We thus have a bijective continuous map with continuous inverse, which makes it a homeomorphism.

Prop 2.17: For a topological group, $-^{-1}$ is a homeomorphism from G to itself.

Pf: For notational convenience, define $f : G \to G$ to be the inverse. I.e., $f(g) \equiv g^{-1}$. Since every g has a unique multiplicative inverse, f is a bijection. We already know that f is continuous for a topological group. We therefore need only show that f^{-1} is continuous — i.e. that f is an open map. However, $f^2 = Id_G$, since $f(f(g)) = g$. Therefore, $f = f^{-1}$. Since f is continuous, $f^{-1} = f$ is.

Prop 2.18: For a topological group: (i) all the path-connected components are homeomorphic and (ii) all the connected components are homeomorphic.

However, the path-connected and connected components still need not be homeomorphic to one another.

Pf: (i) Let G'_0 denote the identity path-component and let G'_1 be some other path-component. Pick any $g \in G'_0$ and $g' \in G'_1$. Then $(g'g^{-1})g = g'$, so the element $h = g'g^{-1}$ takes us from g to g'. Consider $L_h : G \to G$ given by $L_h(g) = hg$. We saw in proposition 2.16 that any L_h is a homeomorphism from G to itself. The restriction of a homeomorphism is a homeomorphism to the image (relative to the subspace topologies on both ends), so $L_h|_{G'_0}$ is a homeomorphism between G'_0 and $L_h(G'_0)$. Consider any point $k' \in G'_1$. Define $l \equiv g'^{-1}k'$, so $k' = g'l = hgl$. If $gl \in G'_0$, then $k' = L_h(gl)$ is in the image. However, $gl \in G'_0$ iff there exists a path between g and gl. Since G'_0 is a subgroup of G, it is closed under inverses, so $g^{-1} \in G'_0$. Since G'_0 is also closed under multiplication, multiply on the left by g^{-1} to get that $gl \in G'_0$ iff $l \in G'_0$. I.e., we must show that $l \in G'_0$. Since g' and k' are in the same path-component, there exists a path p between g' and k'. This is a continuous function from $[0,1]$ to G'_1 s.t. $p(0) = g'$ and $p(1) = k'$. Continuous functions compose, so take $L_{g'^{-1}} \circ p : [0,1] \to G$. This creates a path from e to l in G, proving that l is in the same path-component as e. I.e. $l \in G'_0$. We therefore have shown that $L_h(G'_0)$ contains G'_1. Consider any point $k \in G'_0$. There exists a path p from e to k, so there exists a path $L_h \circ p$ from h to hk. Therefore, $L_h(G'_0)$ consists of a single path-connected component: the one containing h. Since $g \in G'_0$ and $L_h(g) = g' \in G'_1$, this component must be G'_1. Therefore, $L_{g'g^{-1}}$ furnishes a homeomorphism from G'_0 to G'_1. Since every path-connected component is homeomorphic to G'_0, they all are homeomorphic to one another.

Pf: (ii) Let G_0 denote the identity connected-component and let G_1 be some other connected component. The proof mirrors that of part (i) up to the point where L_h restricts to a homeomorphism from G_0 to $L_h(G_0)$. All that remains is to show that $L_h(G_0) = G_1$. A continuous function takes connected sets to connected sets, and a homeomorphism does so in both directions, thus restricting to a homeomorphism between connected sets. L_h is a homeomorphism from G to G, so $L_h(S)$ is connected iff S is connected. G_0 is the union of all connected sets containing g and G_1 is the union of all connected sets containing g', and L_h takes g to g'. Therefore, S is a connected set containing g iff $L_h(S)$ is a connected set containing g'. Let S_i be the set of connected sets containing g. Then $L_h(\cup S_i) = \cup L_h(S_i)$ is a union of connected sets containing g'. Moreover, if S'_i is a connected set containing g', then $L_h^{-1}(S'_i)$ is a connected set containing g (since L_h^{-1} is continuous in its own rite), and thus appears amongst the S_i sets. Therefore, L_h takes a maximum union to a maximum union, and $L_h(G_0) = G_1$. Once again, we have a homeomorphism.

Prop 2.19: Any subgroup of a topological group is a topological group in the subspace topology.

Pf: Let H be a subgroup of topological group G. We need only show that $\times$ and $^{-1}$ are continuous in the subspace topology. Since H is a subgroup, $\times$ restricted to $H \times H$ has image H and $-^{-1}$ restricted to H has image H. The restriction of a continuous function is continuous in the subspace topology, so these restrictions are automatically continuous relative to the subspace topologies.

Prop 2.20: Let G be a topological group and let H be a normal subgroup of G, and let $q : G \to G/H$ be the quotient map. Then: (i) H is a topological group, (ii) q is continuous, (iii) q is an open map, and (iv) G/H is a topological group.

Note that (i)-(iii) holds for non-normal subgroups too, though G/H isn't a group in that case (i.e. we don't have an SES). However, (iii) does not hold for general quotient spaces. The group structure of G is a sufficient, but not necessary condition for q to be open.

We know what $X/\sim$ means topologically, and we know what G/H means from group theory. In the case of topological groups, the two are consistent setwise and have the same quotient map q if we define the obvious equivalence relation $g \sim h$ iff they are in the same coset. This yields a $G/\sim$ which is indeed the set of cosets and a q which takes each g to its coset.

Pf: (i) From proposition 2.19, any subgroup of a topological group is a topological group.

Pf: (ii) By the definition of the quotient topology, $q^{-1}(O')$ is open in G for any open O' in G/H.

Pf: (iii) First, we'll observe that if $O \subset G$ is open then gO is open for any $g \in G$. This follows because L_g is a homeomorphism from G to G and thus takes open sets to open sets. Therefore, $L_g(O) = gO$ is open. The same goes for Og using $R_g(O)$. $O' \equiv q(O)$ is some subset of G/H. $q^{-1}(O') = q^{-1}(q(O)) \supseteq O$ since q is surjective but not injective. However, the noninjectivity is only intra-coset. In fact, clearly $q^{-1}(q(O)) = OH$. This can be written $OH = \cup_{h\in H} Oh$. Let O be open. We saw that each Oh is then open. We thus have a union of open sets, so OH is open in G. Since $OH = q^{-1}(O')$, and $q^{-1}(O')$ is open iff O' is open (see our earlier discussion of the quotient topology for clarification of this point), $O' = q(O)$ is open in the quotient topology. Therefore, q is an open map.

Pf: (iv) general: To avoid confusion, in what follows we'll use $[gH]$ to denote an element of G/H and gH for the corresponding set in G. For a quotient group, multiplication is defined via $[gH][g'H] = [(gg')H]$ and the inverse is defined via $[gH]^{-1} = [g^{-1}H]$ and the identity is $[eH] = [H]$. We know from basic group theory that this defines a group structure on G/H that is consistent with G being a group extension of G/H by H. The present proposition requires us to prove that the multiplication and inverse thus defined are continuous relative to the quotient topology.

Pf: (iv) inverse: Denote the inverse on G by $f : G \to G$ (with $f(g) = g^{-1}$) and the inverse on G/H by $f' : G/H \to G/H$ (with $f'([gH]) = [gH]^{-1} = [g^{-1}H]$). Since G is a topological group, f is continuous. Because f is continuous and bijective, with continuous inverse, it is a homeomorphism from G to G. Similarly, f' is a bijection from G/H to itself. Consider an open set $O' \in G/H$. $f'^{-1}(O') = \{[gH] \in G/H; [g^{-1}H] \in O'\}$. Since O' is open and q is continuous, $q^{-1}(O')$ is open in G. This is just $q^{-1}(O') = \cup_{[gH]\in O'} gH$. Define $U \equiv f^{-1}(q^{-1}(O')) = \cup_{[gH]\in O'} f^{-1}(gH)$. Since f is continuous and $q^{-1}(O')$ is open, U is open too. As sets, $f^{-1}(gH) = H^{-1}g^{-1} = Hg^{-1} = g^{-1}H$ (with $H^{-1} = H$ because H is a subgroup of G and $Hg^{-1} = g^{-1}H$ because H is normal in G). We therefore have $U = \cup_{[gH]\in O'} g^{-1}H$. Bear in mind that it doesn't matter which representative g's we use from the cosets; the resulting cosets are the same, as is the union. Now, take $q(U)$, which is trivially obtained by dropping each coset in the union to its corresponding element of G/H. I.e. $q(U) = \{[g^{-1}H]; [gH] \in O'\}$. Since (from (iii)) q is an open map, this is open in G/H. The role of g as a variable in $q(U)$ is irrelevant, so we can just observe that $\{[g^{-1}H]; [gH] \in O'\} = \{[gH]; [g^{-1}H] \in O'\} = \{[\eta H]; [\eta^{-1}H] \in O'\}$ or any other symbol we choose. The key relationship is that we end up with the set of inverses to whatever was in O'. I.e., $q(U) = f'^{-1}(O')$. Since we just showed that $q(U)$ is open, f' is continuous.

Pf: (v) multiplication: Denote multiplication on G by $f : G\times G \to G$ (with $f(g_1, g_2) = g_1 g_2$) and multiplication on G/H by $f' : G/H\times G/H \to G/H$ (with $f'([g_1H], [g_2H]) = [g_1H][g_2H] = [(g_1g_2)H]$). Since G is a topological group, f is continuous (relative to the product topology). Consider an open set $O' \in G/H$. $f'^{-1}(O') = \{([g_1H], [g_2H]) \in G/H \times G/H; [g_1H][g_2H] \in O'\}$. By the multiplication rule, the condition is the same as $[(g_1g_2)H] \in O'$. Since O' is open and q is continuous, $q^{-1}(O')$ is open in G. This is just $q^{-1}(O') = \cup_{[gH]\in O'} gH$. Define $U \equiv f^{-1}(q^{-1}(O')) = \{(k, l) \in G \times G; kl \in \cup_{[gH]\in O'} gH\}$. I.e., it consists of all pairs k and l that multiply to an element of one of the cosets in O'. Since f is continuous relative to the product topology, U is open in $G \times G$. Now, consider $(q, q)(U)$ (meaning $\{(q(k), q(l)) \in G/H \times G/H; (k, l) \in U\}$). This respects classes, because $kl \in gH$ iff $(kH)l \in gH$ (and actually equals it). I.e., if a given k appears, so does all of kH, and ditto for a given l. We thus can read off $(q, q)(U)$ as $\{([kH], [lH]); [kH][lH] \in O'\}$, and the latter condition is just $[(kl)H] \in O'$. I.e., naming of our irrelevant variables k and l aside, we just have $f'^{-1}(O')$. Now that we know they are the same, let's show that $(q, q)(U)$ is open. To do so, we'll need to observe that a product of open maps is open. [Consider open maps $f, g : X \to Y$, and consider $(f, g) : X \times X \to Y \times Y$. An open set in $X \times X$ is a union of pairs (O_1, O_2), with O_1 and O_2 open. Any map takes unions to unions, so we need only show that a particular pair is mapped to an open set, and the product topology will take care of the rest. However, we already know it is. $(f(O_1), g(O_2))$ is a pair of open sets — and thus open in the product topology on $Y \times Y$ — since f and g are open maps.] We therefore have that $(q, q)(U)$ (and thus $f'^{-1}(O')$) is open, so we've shown that f' is continuous.

A **discrete group** is just a topological group that has the discrete topology. Due to the group structure, this turns out to be equivalent to the requirement that the identity e is topologically isolated (i.e. there exists an open set containing e and no other points). A discrete group need not be abelian or have a simple group structure, nor need it be countable.

If e is topologically isolated, then $L_g(\{e\}) = \{g\}$, which is homeomorphic to it, must be topologically isolated as well. So every point is topologically isolated, and we have the discrete topology. On the other hand, if G has the discrete topology, then every point (including e) is, by definition, topologically isolated.

Any group is a discrete group in the discrete topology.

A product of discrete spaces is discrete in the product topology, and every function from a discrete space is continuous, so $\times : G \times G \to G$ and $-^{-1} : G \to G$ are automatically continuous.

2.6.1. *Note on Lie groups.*

Recall that a real Lie group is a smooth manifold with a group structure for which $\times$ and $-^{-1}$ are smooth functions, and a complex Lie group is a complex-analytic manifold with a group structure for which $\times$ and $-^{-1}$ are complex-analytic.

Why "smooth" and not differentiable or real-analytic? It turns out that there is no difference, and the nomenclature is somewhat historical.

Early in the development of the theory of manifolds, it was shown that any C^1 (i.e. differentiable) real manifold contains a nested doll of unique (modulo diffeomorphism) subatlases of each higher-level of differentiability up to smoothness C^∞.

I.e., we can choose a maximal subatlas of C^n that is C^{n+1}.

This is one of Whitney's theorems. See [6], theorem 2.9 of section 2.2 (page 51).

Due to this, there is little reason to distinguish between C^n and C^∞ for $n > 0$, and many treatments of differential geometry just assume a smooth structure. If there is a differential structure, then we can always work with a smooth subatlas.

With far greater effort, it was later established that this extends to C^ω (i.e. real-analyticity). However, by then the definitions and language in terms of smoothness were well-established. Moreover, for many purposes real-analyticity is an unnecessary nicety, and there's no need to work with a real-analytic atlas.

This is a theorem of Morrey and Grauert. See [7] for a detailed discussion and lots of references.

A Lie group structure is highly constraining. There are very few manifolds that are Lie groups. In fact, aside from π_0 and π_1 (i.e. the set of path-components and the fundamental group), a Lie group is determined exclusively by its Lie algebra — which is a purely algebraic and highly-constrained object. I.e. the rest of the topology is automatic once we have π_0, π_1, and the Lie algebra.

We're being a little glib here. We need π_0 and π_1, along with the relevant group extension maps that glue the components of G together and extract the correct set of deck transformations (i.e. copy of $\pi_1(G)$) from the universal cover of G_0.

The same actually holds, though much less obviously, for topological groups that are topological manifolds. These too are highly constrained, and it turns out that this is related.

Hilbert's fifth problem (in one of its forms) asks whether any topological group that is a topological manifold is a real Lie group. In the 1950's the answer was shown to be yes. This amounts to the following statement: for topological groups that are topological manifolds, there is a unique real-analytic structure relative to which the group operations are real-analytic.

This was proved by Montgomery-Zippin and Gleason. See [8] for a detailed discussion.

This, of course, descends to any C^n as well.

This gives us a couple of special features. First, it is by no means obvious that if a smooth (or C^n or real-analytic) manifold admits a topological group the group operations will be correspondingly smooth (or C^n or real-analytic). A topological group merely requires them to be continuous, and the fact that we choose to work with a differentiable subatlas doesn't change that. There's no clear reason to expect these

maps to be differentiable relative to such a subatlas. Now we know that they must be, at least for some such subatlas.

Second, for a real Lie group, our indifference to the specific level of differentiability (from C^1 to C^ω) now extends to C^0. As far as real Lie groups are concerned, there's no need to distinguish between topological manifolds and real-analytic manifolds. If a manifold has a topological group structure, it is a real Lie group.

Part of what this implies is that any topological manifold which admits a topological group structure must contain a real differential (and thus C^n, C^∞, and C^ω) structure.

When it comes to complex-analytic manifolds, we must be a little careful interpreting this. A topological manifold has no notion of "complex" vs "real". It is locally homeomorphic to some $\mathbb{R}^n$ or $\mathbb{C}^n$, but $\mathbb{C}^n$ is homeomorphic to $\mathbb{R}^{2n}$, so specifying it in terms of $\mathbb{C}^n$ tells us nothing other than that we have an even "real" dimension. Until we start working with differentiable subatlases, we have no meaningful constraint. A topological space either has a unique topological manifold structure or none at all.

The statement that every topological group on a topological manifold is a real Lie group is unambiguous. It simply means that there always exists some smooth subatlas relative to which the group operations are smooth. However, the complex case is a bit more tricky. Recall that a complex-analytic n-manifold is a real-analytic $2n$-manifold with a "complex structure" on it (which implements the Cauchy-Riemann equations). Not every real-analytic $2n$-manifold admits a complex structure.

A topological group on a topological manifold is a complex Lie group iff there exists a complex-analytic subatlas of that topological atlas under which the group operations are complex analytic. Put another way, there must exist a (necessarily even-dimensional) real-analytic subatlas that admits a complex structure *and* relative to which the group operations are complex analytic. This may or may not be the case. There certainly exist complex Lie groups (ex. $SL(n, \mathbb{C})$), but not every topological group on an even-dimensional topological manifold is one.

In summary, when it comes to topological groups on topological manifolds:

- Every topological group on a topological manifold is a real Lie group.
- We don't care about the type of manifold (i.e. C^0, C^n, C^∞, or C^ω), when it comes to real Lie groups.
- A given topological group on a $2n$-dimensional topological manifold may or may not be a complex-analytic n-dimensional Lie Group.

2.6.2. *Aside: Group G with different topologies.*

Bear in mind that a given group G (i.e. a set with a particular group structure on it) may be a topological group relative to one natural topology but not another. The identity connected-component, identity path-component, the numbers of components, etc, all may vary with the choice of topology.

"Algebraic Groups" furnish a common example of this. An "algebraic group" is defined by a set of polynomial equations. For example, the matrix groups are algebraic groups.

Even $GL(K, n)$ turns out to be an algebraic group, despite the usual defining condition being a polynomial inequality. There is a way to define it in terms of polynomial equations in $n^2 + 1$ rather than n dimensions. See [9].

As the solution set of a collection of polynomials, such a group is an algebraic variety and therefore has a natural Zariski topology.

Recall that the Zarsiki topology on an affine space A^n (often just $\mathbb{R}^n$ or $\mathbb{C}^n$ as an affine space over itself) has as its closed sets the varieties. I.e., for each set of polynomial equations in n A-valued variables, the solution set is closed in the Zariski topology. Note that the open sets are very large.

If the relevant polynomials are over (i.e. have coefficients in) a topological field K (which, just like "topological group", simply means that the relevant operations are continuous relative to the topology), such as $K = \mathbb{R}$ or $K = \mathbb{C}$, then we also have a natural topology inherited from the product topology on K^n, where n is the number of variables appearing in the polynomials. I.e., we view the solution set of the polynomials as a subset of K^n.

A given algebraic group G may be a topological group relative to the K^n topology but not the Zariski topology. This is commonly the case, because the product of Zariski topologies is not a Zariski topology in general, whereas the product of K^n topologies is a $K^{n'}$ topology. The continuity of multiplication is relative to the product topology on $G \times G$, so this matters.

A "linear algebraic group" is an algebraic group that is a subgroup of $GL(K, n)$ for some field K. Such groups are always topological groups relative to the K^{n^2} topology (aka the topology on $gl(K, n)$), but they generally are not topological groups relative to the Zariski topology. Because they inherit a manifold structure from K^{n^2} as well, they are Lie groups relative to the K^{n^2} topology.

When one speaks of a group G as an "algebraic group", it means that it is being considered as an algebraic variety, and *usually* in the context of the Zariski topology. However, when one speaks of the same group G as a "Lie group", it means as a topological group (and manifold) relative to the topology inherited from K^{n^2}.

The "usually" when it comes to Zariski topology is because it depends on what we are doing. "Algebraic group" can be the way we are looking at it (i.e. variety, Zariski topology, etc) — or it can be a property that a group does or does not have. I.e., sometimes, people will speak of G as an "algebraic group" in the mere sense of a classification — as a group which can be defined that way — even though we are working with it as a Lie group. For example, we earlier said that the matrix groups are algebraic groups. In this context, we meant that they have that property, not that we will work with them in the Zariski topology (even though we happened to go on and mention it).

Ex. $O(n)$ is an algebraic group in the sense that it can be defined via the polynomial equations $M^T M - I = 0$ in its n^2 elements as a subgroup of $GL(\mathbb{R}, n)$. However, we almost never work with the Zariski topology, and almost always mean that it is a Lie group. I.e., the "algebraic group" moniker merely specifies that it has the property of being one, not that we will work in the Zariski topology.

Again, the properties of a given group G may differ between the standard and Zariski topologies.

Ex. $GL(\mathbb{R}, n)$ is a Lie group and has two path-components in the standard topology, but it is not a topological group and is path-connected in the Zariski topology.

2.7. **Aside: Zero-dimensional Manifolds.**

Any set may be considered a 0-dimensional topological manifold. In fact, any set can be considered a smooth manifold or complex-analytic manifold as well.

If we include second-countability in the definition of "manifold", the above statements only hold for countable sets.

More precisely, any discrete topological space can be viewed as a 0-dimensional manifold of any type. Since any set can admit the discrete topology, this applies. Note that if we don't require second-countability in the definition of a manifold, then this holds for things like $\mathbb{R}$ in the discrete topology too (i.e. viewed not as a vector space, but as a set of points). This can be a bit counterintuitive — but topological spaces with uncountable bases usually are.

Let's understand why the edge-case of zero-dimensionality works this way. We'll show why a discrete topological space is a smooth manifold, but the idea is the same for a complex-analytic manifold.

A n-dimensional real "FOO"-manifold M consists of a topological space and a maximal atlas of charts (U_α, ψ_α), where the open sets U_α cover M and $\psi_\alpha : U_\alpha \to \mathbb{R}^n$ is a homeomorphism and the map $\psi_\alpha \circ \psi_\beta^{-1}$ between subsets of $\mathbb{R}^n$ induced from each nonempty overlap $U_\alpha \cap U_\beta$ satisfies the condition "FOO".

Note that not every open set need appear in some chart, and the same open set can appear in more than one chart.

$\mathbb{R}^0$ is a 0-dimensional vector space, meaning it is a vector space with real coefficients and no basis vectors. There are only two possible meanings to this: $\emptyset$ or $\{0\}$. The latter is the correct definition from many standpoints. For example, if we successively reduce dimension by removing basis vectors, we move from $\mathbb{R}^n$ to $\mathbb{R}^{n-1}$ but never lose the origin. Following this procedure, we end up with $\{0\}$ rather than $\emptyset$. Similarly, the trivial abelian group is $\{0\}$, not $\emptyset$, because any abelian group must have an identity element. As a module (and thus an abelian group under addition), every vector space must have an additive identity. The natural choice is therefore to define $\mathbb{R}^0$ to be the single-point set $\{0\}$.

We may be tempted to think that any discrete set could count as $\mathbb{R}^0$, but this is not the case. A $\mathbb{R}^0$ defined with more than one point would cease to be a vector space, and wouldn't serve as a suitable edge-case. Moreover, we would have no reason to favor any multi-point discrete set over any other in that case. Only the 1-point set would be special. Either way, it is the clear choice.

Since $\mathbb{R}^0$ consists of a single point, there is only one possible topology on it: the discrete topology (which equals the indiscrete topology in this case).

In the discrete topology, every set is open and every function to another topological space is continuous. A 0-dimensional real manifold has $\psi_\alpha : U_\alpha \to \{0\}$, so there is only one possible such ψ_α per U_α.

For U_α to be homeomorphic to $\{0\}$, we need a bijective function that is continuous both ways. Since we're dealing with two sets both in the discrete topology, any bijective function between them is a homeomorphism. A function to $\{0\}$ is bijective iff its domain has a single point. Therefore, each U_α consists of a single point.

For a discrete space M, the maximal atlas is therefore the only possible atlas on M. It has one chart per point $x \in M$ and the chart map is $\psi_x : \{x\} \to \{0\}$ given by $\psi_x(x) = 0$.

Since there are no overlaps, the overlap condition "FOO" is vacuously satisfied for any "FOO". Therefore, any set in the discrete topology is a "FOO"-manifold for any "FOO" (again, subject to the aforementioned second-countability caveat).

3. $\pi_0(X)$

3.1. General Considerations.

We've seen that $\pi_0(X) = X/\sim_{pc}$ is the set of path-connected components of X and does not have a homotopy-induced group structure like the other π_n's. If we pick a basepoint $x_0 \in X$ (or if X is path-connected), $\pi_0(X)$ has the structure of a pointed-set, with the special point being the path-component containing x_0.

As mentioned, when we encounter π_0 in a long exact sequence involving other homotopy groups, it appears as a pointed set, thus engendering suitable notions of homomorphism and kernel. When we encounter it in a short exact sequence involving topological groups, it is typically endowed with a quotient group structure by the topological group in question. This group structure is of a completely different origin than the group structure of the other π_n's.

We also saw (in proposition 2.15) that π_0 has the discrete topology iff all the path components are open.

As discussed, having the discrete topology does *not* mean that a set is countable. Any set can be endowed with the discrete topology. In general, π_0 can be finite, denumerably infinite, or uncountable.

3.2. $\pi_0(X)$ for Topological Manifolds.

Every topological manifold is locally path-connected, so certain simplifications can be made:

- Path-connected and connected are the same.
- The set of path-connected components is the same as the set of connected components, so we can just refer to them as "components".
- Every component is clopen.
- $\pi_0(X)$ has the discrete topology.

  This follows from proposition 2.15 since every component is open. Once again, this does *not* mean that $\pi_0(X)$ need be countable.
- $\pi_0(X)$ is a "FOO" real or complex manifold for any choice of overlap condition "FOO".

  We saw that this is true of any set in the discrete topology. As mentioned, if we include second-countability in the definition of "manifold" then we must also stipulate that $\pi_0(X)$ is countable.

3.3. $\pi_0(G)$ for Topological Groups.

Although $\pi_0(G)$ has no homotopy-induced group structure, for a topological group it inherits a natural quotient group structure.

Bear in mind that this group structure is entirely unrelated to the usual homotopy group construction for $\pi_n(X)$ with $n > 0$. Nor do the higher homotopy groups inherit a second group structure in this manner. It is specific to π_0.

Prop 3.1: Let G be a topological group, and denote by G_0 the identity connected component and by G_0' the identity path component. Then:

- (i) G_0 is a topological closed normal subgroup of G.

- (ii) G_0' is a topological normal subgroup of G.
- (iii) G_0' is a topological normal subgroup of G_0.
- (iv) G/G_0 is a topological group consisting of the connected components as cosets.
- (v) G/G_0' is a topological group consisting of the path components as cosets.
- (vi) G_0/G_0' is a topological group.
- (vii) G/G_0' is naturally bijective with $\pi_0(G)$ (i.e. with $G/\sim_{pc}$).
- (viii) G/G_0 is naturally bijective with $G/\sim_c$.
- (ix) $\pi_0(G)$ has a natural topological group structure inherited from G.
- (x) G/G_0 is a discrete group iff G_0 is open (and thus clopen).
- (xi) $\pi_0(G)$ is a discrete group iff G_0' is open.
- (xii) G/G_0 is a quotient group of G/G_0'.

For (x) and (xi), note that for a topological group, G/G_0 and $\pi_0(G) = G/G_0'$ are *always* totally disconnected spaces. However, this is *not* the same as being discrete. Every discrete space is totally disconnected but not vice versa. Ex. the irrationals $\mathbb{R} - \mathbb{Q}$ in the subspace topology.

Pf: (i) (closed): Every connected-component is topologically closed, so G_0 is as well.

Pf: (i) (subgroup): By definition, $e \in G_0$. From proposition 2.17, $-^{-1}$ is a homeomorphism from G to itself. By proposition 2.8 it thus restricts to a homeomorphism between connected components. Since it takes e to e, it must restrict to a homeomorphism from G_0 to G_0. Therefore, G_0 is closed under the group inverse. Pick some $g \in G_0$. From proposition 2.16, L_g is a homeomorphism from G to G. Again by proposition 2.8, it restricts to a homeomorphism from G_0 to some connected component. Since $g \in G_0$, L_g takes e to $g \in G_0$, so this component must be G_0. Therefore L_g is a homeomorphism from G_0 to itself. Since this holds for every $g \in G_0$, G_0 is closed under multiplication.

Pf: (i) (normal): Pick any $g \in G$. By 2.16, L_g and $R_{g^{-1}}$ are homeomorphisms, so by proposition 2.8, each restricts to a homeomorphism between components. The composition of homeomorphism is a homeomorphism, so $L_g R_{g^{-1}}$ (which takes the form $h \to ghg^{-1}$) is a homeomorphism too, and thus restricts to a homeomorphism between components. However, $L_g R_{g^{-1}}(e) = e$, so it must take G_0 to G_0. Therefore $gG_0g^{-1} = G_0$ for every $g \in G$, and G_0 is normal in G.

Pf: (ii) The exact same argument holds for path components, since $e \in G_0'$ and proposition 2.8 holds for path components as well as connected components.

Pf: (iii) We know from topology that $G_0' \subseteq G_0$ (since each path-component is a subset of a connected component, and both identity components contain the identity). We therefore know that $G_0' \subseteq G_0 \subset G$. Since G_0' and G_0 are subgroups of G, this makes G_0' a subgroup of G_0. However, we also know (from (ii)) that $G_0' \triangleleft G$. From group theory, if G_0' is normal in G, then it is normal in any subgroup of G that contains it (this is obvious since if $gG_0'g^{-1} = G_0'$ holds for all $g \in G$, it certainly holds for all $g \in G_0 \subset G$). Therefore, $G_0' \triangleleft G_0$.

Pf: (iv) Since G_0 is normal in G, G/G_0 is a group. Consider some $h \in G_0$. Since L_h is a homeomorphism from G to itself, it restricts to one between components. $he = h$, so L_h restricts to a homeomorphism between G_0 and the connected component containing h. However, $L_h(G_0) = hG_0$ is just the left coset (which is the same as the right coset, since G_0 is normal). I.e., the connected components *are* the cosets, so G/G_0 labels the connected components. Moreover, it does so in a natural way via the quotient map. From proposition 2.20, we know that G/G_0 is a topological group.

Pf: (v) The proof of (iv) holds equally well for G_0'.

Pf: (vi) From proposition 2.20, G_0/G_0' is a topological group.

Pf: (vii) $\pi_0(G) = G/\sim_{pc}$ is the set of path-connected components. From (v), we saw that the path-connected components are nothing other than the cosets. Therefore, $\pi_0(G)$ *is* G/G_0'. The natural bijection is a purely pedantic mapping association of G_i in its capacity as coset with G_i in its capacity as path-component.

Pf: (viii) The same exact argument holds as for (vii), though now the relevant set of components is denoted $G/\sim_c$ rather than $\pi_0(G)$.

Pf: (ix) We showed this in (vii). It has the topological group structure of G/G_0'.

Pf: (x) Any group is a discrete group if we endow it with the discrete topology. We saw from (viii) and proposition 2.20 that G/G_0 is the same topological space as $G/\sim_c$. Our statement is therefore equivalent to "$G/\sim_c$ has the discrete topology iff G_0 is open". Since all the components are homeomorphic for a topological group, G_0 is open iff all the components are open. We saw in proposition 2.15 that all the components are open iff $G/\sim_c$ has the discrete topology.

Pf: (xi) The same exact argument as for (x) holds for path-components, so $\pi_0(G) = G/G_0'$ is a discrete group iff G_0' is open.

Pf: (xii) This follows from one of the standard isomorphism theorems of group theory. Since $G_0' \triangleleft G_0 \triangleleft G$, G/G_0 is isomorphic to $(G/G_0')/(G_0/G_0')$, so G/G_0 is a quotient group of G/G_0' (by G_0/G_0').

For an example of a topological group that is connected but not path-connected (i.e. the path-components are proper subgroups of the single connected component), see [10].

It turns out that the identity connected component and identity path-component of a topological group are not just normal subgroups. G_0 and G_0' are **fully characteristic** subgroups of G, which implies they are **characteristic subgroups**. Why do we care? Normality is not ordinarily transitive, so $K \triangleleft H \triangleleft G$ does not imply that $K \triangleleft G$. However, if K is a *characteristic subgroup* of H and $H \triangleleft G$, then $K \triangleleft G$ does follow.

A "characteristic" subgroup H of G is a subgroup of G s.t. every group automorphism of G takes H to itself. I.e., $\phi(H) = H$ for every $\phi \in Aut(G)$. This is very restrictive. It implies normality, because $gHg^{-1} = H$ for the inner automorphism $g' \to gg'g^{-1}$. H is "fully characteristic" if it not only preserves H under automorphisms but under homomorphisms from G to G. I.e., $\phi(H) \subseteq H$ for any continuous homomorphism $\phi : G \to G$. This is far more restrictive than "characteristic" since it demands that H be preserved under a far wider range of functions. When speaking of topological groups, we modify these definitions to concern only homeomorphic automorphisms and continuous homomorphisms. Note that this is not universal practice. The definition of "characteristic" is a group-theoretic one. Although it makes sense to weaken the definition for topological groups by restricting the requirement to automorphisms which are homeomorphisms, not all authors do so. For our purposes, the distinction is irrelevant. The inner automorphisms are always homeomorphisms, so either definition implies normality. We'll stick with the weaker definition which tests only homeomorphic automorphisms.

We have two SES's: (i) $1 \to G_0 \xrightarrow{i} G \xrightarrow{q} G/G_0 \to 1$, with i subset inclusion $G_0 \subset G$, and q the coset map $g \to gG_0$ (which equals $G_0 g$ since G_0 is normal in G). This does not right-split in general, even for a Lie group. (ii) $1 \to G_0' \xrightarrow{i'} G \xrightarrow{q'} G/G_0' \to 1$, with i' subset inclusion for $G_0' \subset G$ and q' the relevant coset map $g \to gG_0'$ (which equals $G_0' g$ since G_0' is normal in G). This too need not right-split.

In light of our discussion so far, the following holds:

Prop 3.2: Let G be a topological group. (i) If $H \triangleleft G$, then $H_0 \triangleleft G$. (ii) If $H \triangleleft G$, then $H_0' \triangleleft G$.

Pf: We have $H_0 \triangleleft H \triangleleft G$. However, since H_0 is characteristic in H, we have transitivity, and $H_0 \triangleleft G$. The same holds for H_0', since it too is characteristic in H.

Note that $\pi_0(G)$ need not be abelian.

We can construct counterexamples easily enough. Start with some finite group D. We can view D as a topological group (and, in fact, a Lie group) in the discrete topology. Given any path-connected topological group H on which D has an action (it need not be full or faithful), we can construct the direct product $G = H \overline{\times} D$. As a product space of a discrete space and a path-connected space, the path components are just copies of H. In fact, we've constructed a group G whose identity path component is a copy $H' \approx H$ that is a normal subgroup and whose $\pi_0(G) = D$ is isomorphic to a normal subgroup $D' \subseteq G$ s.t. $G = H'D'$. I.e. $H' = (H, e)$ and $D' = (e, D)$. If we start with a nonabelian D, then, by construction, $\pi_0(G) \approx D$ is nonabelian.

Ex. Let D be the permutation group of the set (x, y, z), consisting of 6 elements (this is commonly known as D_3 or S_3). Denote by s the right shift operator $(x, y, z) \to (z, x, y)$ and by e the identity operator and by a the swap operator $(x, y, z) \to (y, x, z)$ and by b the swap operator $(x, y, z) \to (x, z, y)$. We then have an order-3 subgroup generated by s (i.e. $s^3 = e$). The remaining three elements are a, b, and $ab = ba$. This group is nonabelian. For example, as (viewed as acting from the right on (x, y, z)) gives us (x, z, y) and sa gives us (z, y, x). Now, let H denote the group of quaternions as an additive group. It turns out that the automorphism group of the quaternions as a ring is isomorphic to $SO(3)$ (see [11] and [12]), and the automorphism group of H as a ring is a subgroup of the automorphism group of H as an abelian group. Therefore, $Aut(H)$ contains a copy of $SO(3)$. However, every D_k is a finite subgroup of $SO(3)$, and this includes D_3. We thus can construct a homomorphism $D_3 \to Aut(H)$ via the isomorphism from D_3 to the copy of $D_3 \subset SO(3) \subset Aut(H)$. Therefore, there exists a semidirect product (in the external view) $G = H \rtimes_\phi D_3$. Since H is path-connected (its just $\mathbb{R}^4$ topologically, much as $\mathbb{C}$ is $\mathbb{R}^2$ topologically) and D_3 has the discrete topology, $G_0 = H$ and $\pi_0(G) = D_3$. I.e., we have a nonabelian π_0.

3.4. $\pi_0(G)$ for Lie Groups.

In section 2.6.1, we saw that any topological group on a topological manifold is a real Lie group. As topological groups that are also topological manifolds, Lie groups have the nice properties of both.

There are two types of subgroups of a Lie group, and different authors refer to them differently. We'll follow Kirillov's convention. A **closed Lie subgroup** is a subgroup that is an embedded submanifold (or, equivalently, a subgroup that is topologically closed). A **Lie subgroup** is a subgroup that is an immersed submanifold.

See [13] for a detailed discussion of the distinction.

Prop 3.3: For a Lie group G, every L_g, every R_g, and $-^{-1}$ are diffeomorphisms from G to G and restrict to diffeomorphisms between the components.

Pf: (i) inverse: The same exact argument as in proposition 2.17 holds. Since $f = f^{-1}$ and f is a smooth homeomorphism, $f^{-1} = f$ is smooth too, so f is a diffeomorphism.

Pf: (ii) L_g and R_g: Consider $\times|_{(g,G)}$. The restriction of a diffeomorphism to a submanifold is a diffeomorphism to its image. In this case, the restriction is to (g, G) and has image G (since $gG = G$). Define $\alpha : (g, G) \to G$ to be the trivial diffeomorphism $\alpha(g, h) = h$. Then $L_g = \times|_{g,G} \circ \alpha^{-1}$. As the composition of diffeomorphisms it is a diffeomorphism from G to G. The same argument holds for R_g.

Pf: (iii) Let G_i be a component, and let f be any of L_g, R_g, or $-^{-1}$. The restriction of a diffeomorphism is a diffeomorphism to its image, so all we need do is show that $f(G_i)$ is a component. However, we already know this since (from our earlier discussion of topological groups) each of L_g, R_g, and $-^{-1}$ restricts to a homeomorphism between components. Therefore, that homeomorphism must be a diffeomorphism.

Prop 3.4: For a (second-countable) Lie group G:

- (i) $G_0 = G_0'$ and $G/G_0 = G/G_0' = \pi_0(G)$ and we may speak of "components" without worrying about connected vs path-connected.
- (ii) The components are the cosets, are clopen submanifolds of the same dimension as G, and are diffeomorphic to one another.
  Obviously, only the identity component is a group.
- (iii) $G_0 = G_0'$ is a closed Lie subgroup of G of the same dimension as G, and the inclusion map is smooth.
- (iv) $\pi_0(G) = G/G_0' = G/G_0$ is a discrete (aka 0-dimensional), countable Lie group. The quotient map is smooth.
  If we drop the "second-countability" requirement from the definition of manifold, then $\pi_0(G)$ need not be countable, but it still is discrete (aka 0-dimensional) and a Lie group.

Pf: (i) G is a topological manifold, so these follow immediately.

Pf: (ii) As a topological manifold, all the components are clopen. The n-dimensional submanifolds of an n-manifold are precisely its nonempty open sets. Since the components all are open, they are n-dimensional submanifolds. G is a topological group, so all the components are homeomorphic to one another. We know from proposition 3.3 that L_g is a diffeomorphism from G to itself. We therefore can use the same argument as in proposition 2.18, to show that components are diffeomorphic. Given any $g \in G_0$ and $g' \in G_1$, $L_{g'g^{-1}}$ is a diffeomorphism of G that restricts to a diffeomorphism between G_0 and G_1. This holds for every G_1 using some $g' \in G_1$, so the components are diffeomorphic.

Pf: (iii) Since G_0 is both a subgroup of G and a closed submanifold of G, it is a closed Lie subgroup (and thus is an embedded submanifold). The inclusion map is the restriction of the smooth map Id_G to G_0, and therefore is smooth.

Pf: (iv) We saw that for a topological group, $\pi_0(G)$ is discrete iff G_0' is open, and we saw that for a topological manifold, G_0' is clopen. Therefore, $\pi_0(G)$ is discrete for a Lie group. As discussed in section 2.7, every discrete group can be considered a 0-dimensional real or complex Lie group. If we include second-countability in the definition of "manifold", then we must also require that the group be countable. As for the smoothness of the quotient map, we won't prove that here. It follows from the "Quotient Manifold Theorem". See [14], page 544 (theorem 21.10).

We saw in section 2.2.5 that if all the components are open then second-countability implies that there are a countable number of them. If we include second-countability in their definition, then all Lie groups have a countable $\pi_0(G)$.

However, the number need not be finite. Ex. $SO(3) \times Z$. See [15]. However, for a compact Lie group, the number of connected components must be finite.

Since path components and connected components are the same for a Lie group, we have a single SES, which can be written $1 \to G_0 \xrightarrow{i} G \xrightarrow{q} \pi_0(G) \to 1$, with i the usual subset inclusion, and q the quotient coset map $q(g) = gG_0$ (which equals $G_0 g$ since G_0 is normal in G).

Even for Lie groups, this SES need not right-split. However, it turns out that it does right-split for the classical Lie groups. Therefore, these groups (which include the Lie groups we typically encounter in physics) are all semidirect products $G = G_0 \rtimes_\phi \pi_0(G_0)$, and sometimes direct products.

In fact, one must go to some length to concoct an interesting (i.e. non-discrete) Lie group G for which the SES fails to right-split. See [16] for an example of how to construct a Lie group which is not a semidirect product.

3.5. **The Matrix Lie Groups.**

All of the familiar classical matrix Lie groups are connected or are disconnected only in very simple ways. This is not a coincidence. All of these groups are subgroups of $GL(K, n)$ for $K = \mathbb{R}$ or $K = \mathbb{C}$, which is itself a subset of the unital associative algebra $gl(K, n)$ (and inherits its multiplicative structure). The latter topologically is the same as K^{n^2} in the product topology.

There is no surprise here. The natural topology on the space of $n \times n$ K-valued matrices is a product of n^2 copies of K.

The relevant topology for all of the matrix Lie Groups is just the subspace topology inherited from $gl(K, n)$. I.e., they are topological subspaces of K^{n^2}. The defining condition for $GL(K, n)$ is that $\det M \neq 0$, and for all the others it is of the form $M^T \eta M = \eta$ or $M^\dagger \eta M = \eta$ for some fixed matrix η (possibly supplemented by the condition $\det M = 1$).

Specifically:

- $GL(\mathbb{R}, n)$ is defined by $\det M \neq 0$. It has 2 components.
- $SL(\mathbb{R}, n)$ is defined by $\det M = 1$. It has 1 component.
  This is *not* the identity component of $GL(\mathbb{R}, n)$, which requires $\det M > 0$ rather than $\det M = 1$. However, it is homotopic to it.
- $O(n)$ is defined by $M^T M = I$ (i.e. $\eta = I$). It has 2 components.

- $SO(n)$ is defined by $M^T M = I$ and $\det M = 1$. It has 1 component.

  This is the identity component of $O(n)$.

- $O(n,m)$ is defined by $M^T \eta M = \eta$, where η is diagonal with n ones and m minus-ones. It has 4 components.
- $SO(n,m)$ is defined like $O(n,m)$ but with the additional constraint that $\det M = 1$. It has 2 components.

  This consists of two components of $O(n,m)$, including the identity component. Basically, it is half of $O(n,m)$.

- $U(n)$ is defined by $M^\dagger M = I$ (i.e. $\eta = I$). It has 1 component.

  This is a real Lie group (i.e. a smooth manifold) despite being defined in terms of complex matrices. It has a real parametrization but no complex parametrization.

- $SU(n)$ is defined the same way as $U(n)$ but with the additional constraint that $\det M = 1$. It has 1 component.

  This is *not* the identity component of $U(n)$, as is clear from the fact that both have 1 component and they differ.

- $SP(2n,\mathbb{R})$ is defined by $M^T \eta M = \eta$ with $\eta = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}$ (where I_n is the $n \times n$ identity matrix). It has 1 component.

What we see (stated without proof) here is that only $GL(\mathbb{R},n)$, $O(n)$, $O(n,m)$, and $SO(n,m)$ have nontrivial π_0's. We won't treat $SO(n,m)$ separately, because it's just halfway between $O(n,m)$ and the identity component of $O(n,m)$. This leaves us with $GL(\mathbb{R},n)$, $O(n)$, and $O(n,m)$ as the cases we'll examine in detail. We'll also look at the Poincare group, which, though not typically counted as a "classical" Lie group, is the fundamental space-time symmetry group of physics as well as the ultimate subject of these notes.

The source of two and a half of the nontrivial π_0's is obvious: $\det M$. Let's look at this in a more general context.

We're trying to identify the elements and components of some $G \subseteq gl(K,n)$ (where $\subseteq$ denotes subspace, not subgroup). This G is defined by some physically-motivated constraints (ex. $M^T M = I$).

For notational simplicity, we'll write $X \equiv gl(K,n)$. Suppose that, either as a defining constraint or as a consequence of one, we have a condition of the form $f(X) \subseteq U'$, where $f : X \to K$ is continuous and $U' = \cup S'_i$ is either a disjoint union of open sets in K or a *finite* disjoint union of closed sets in K.

Note that any polynomial in the elements of the matrix is continuous — including $\det M$, all the polynomials constituting $M^T M - I$, and all the polynomials constituting $M^T \eta M - \eta$ for any fixed matrix η. In the complex case, complex conjugation is also continuous — which means that any polynomial in the elements of the matrix and their conjugates is also continuous. This includes $\overline{(\det M)}(\det M) - 1$, all the polynomials constituting $M^\dagger M - I$, and all the polynomials constituting $M^\dagger \eta M - \eta$ for any fixed matrix η. I.e., every constraint appearing in our classical matrix list above involves continuous functions.

We're not requiring that f be surjective to U'. As long as its image is contained in a union of disjoint open sets or a finite union of disjoint closed sets, we're fine. There may be many ways to choose U'. For example, we'll see an example later where we require that $|M_{00}| \geq 1$. Choosing $f : X \to \mathbb{R}$ via $f(M) = M_{00}$, we could pick $U' = \{(-\infty,-1],[1,\infty)\}$ but we equally well could pick $U' = \{(-\infty,-0.5),(0.5,\infty)\}$ or $U' = \{(-\infty,0],[0.5,\infty)\}$. Intuitively, we just need the existence of gaps in the image.

A special case of this is when the image is a finite set of points. For example $\det M = \pm 1$ has image $\{-1,+1\}$. These are closed sets in $\mathbb{R}$, but we equally well could just consider det as a function from $gl(\mathbb{R},n)$ to $\{-1,1\}$, with the latter in the discrete topology. det is continuous in that case too, and the inverse images of the points in the image are clopen sets in the domain.

Let $S_i = f^{-1}(S'_i)$ and $U = \cup S_i$. Note that any given S_i may be empty if none of the image of f falls in S'_i. Inverse images preserve intersections, so the S_i's are disjoint. They also cover $G \subset X$, since every element of G must satisfy $f(x) \in U'$.

However, this need not be the *only* constraint. It is quite possible that the other constraints confine G to a tiny subset of U. U simply is the subset of X that satisfies *this* constraint. G must be a subset of it.

Since f is continuous, the inverse image of an open set is open and the inverse image of a closed set is closed, so the S_i's are either all open or all closed depending on which of the two types the S'_i's are.

This holds not just in $gl(\mathbb{R}, n)$ but in G itself. If S is clopen in X, then $S \cap Y$ is clopen in $Y \subset X$ (in the subspace topology). Therefore, the $S_i \cap G$ sets are either all open or all closed in G (again, according to the type of the S'_i's).

Note that continuous functions preserve connectedness, but the inverse image of a connected set need not be connected. All we can say is that if $S'_i \cap \text{Im } f$ is disconnected then $f^{-1}(S'_i)$ cannot be connected. Note that we need the $\cap$Im f, because if the image of f doesn't fall in the nasty bits of S'_i, then those bits can't obstruct the connectedness of S_i.

If the S'_i's are open, then their inverses (the S_i's) are open in X and thus in the subspace topology on U. Since an arbitrary union of open sets is open, U is open in X too. For each S_i, $U - S_i = \cup_{j \neq i} S_j$ is open too. This means that its complement, S_i, is open in U, making S_i closed and thus clopen in U (note that it need not be closed in X, however). We thus see that every S_i must be clopen in U, yielding a partition of U into clopen sets.

On the other hand, if the S'_i's are closed, then their inverses (the S_i's) are closed in X and thus in the subspace topology on U. Since a *finite* union of closed sets is closed, U is closed in X too. For each S_i, $U - S_i = \cup_{j \neq i} S_j$ is a finite union and thus closed too. This means that its complement, S_i, is closed in U, making S_i open and thus clopen in U (note that it need not be open in X, however). We thus see that every S_i must be clopen in U, once again yielding a partition of U into clopen sets.

This is why we had the caveat that U could be only a finite union of disjoint closed sets. Only finite unions of closed sets need be closed.

Since the S_i's are clopen in $gl(\mathbb{R}, n)$, it immediately follows that the $S_i \cap G$ sets are clopen (though possibly empty) in G. This doesn't tell us that each $S_i \cap G$ is a component. A given S_i may not intersect G at all or $S_i \cap G$ may contain more than one connected component of G. However, it does tell us that any S_i which G intersects must contain at least one connected component. I.e., each $S_i \cap G$ is a (possibly empty) union of connected components of G.

Note that we needn't examine every $S_i \cap G$. Because we know that the components are diffeomorphic for a Lie group, we can see that: (i) if our constraint results in non-diffeomorphic $S_i \cap G$'s, then some must contain more than one component, and (ii) if the $S_i \cap G$'s are diffeomorphic, we need only prove that one is connected to show that they all are.

One way to use this is to extract a set of constraints from the defining equations for G that have the desired form (i.e. each involves a continuous function that must produce values in a disjoint union of open sets of a finite disjoint union of closed sets). Each such constraint then produces a partition of G into corresponding $S_i \cap G$ classes. If the nonempty classes aren't all diffeomorphic, then we certainly have more work to do (because the components must all be diffeomorphic). Each intersection of classes from the different constraints must also be a union of components. By doing this repeatedly, we refine our classes, whittling our way down to single-component classes. At any point when all the classes are diffeomorphic (and thus may possibly be individual components), all we need do is prove that the identity class is path-connected. If we accomplish this, then all the other diffeomorphic classes must be the other components, and we're done. I.e., any additional constraints would either replicate the partitions from existing ones or be coarser. We'll see an example of this procedure when we look at $O(n, m)$.

In the case of $GL(\mathbb{R}, n)$, $\det M \neq 0$ is the only constraint, so the S_i's are the same as the $S_i \cap G$'s. The constraint takes the form discussed, with $f(M) = \det M$, and $U' = (-\infty, 0) \cup (0, \infty)$ being our breakdown. It is easy to see that G intersects each of the two S_i's (I sits in one, and I with any single diagonal element flipped to -1 sits in the other). Since this is the only constraint, we know that we have at least two connected components.

The two S_i's in this case are clearly diffeomorphic. Pick any matrix γ with $\det \gamma = -1$ (i.e. any in the non-identity S_i) and the map $M \to \gamma M$ is a diffeomorphism of the form L_γ. If n is odd, $\gamma = -I$ will work. We therefore need only look at one of our S_i's, call it $S \equiv f^{-1}(0, \infty)$. To show that this is a single component, we simply construct a path from any element M to I. I.e., we need a continuous way of transforming M to I s.t. all the intermediate matrices have positive determinant.

There are many standard linear algebra ways to accomplish this, and we won't go into the details here.

Note that, $GL(\mathbb{C}, n)$ is connected. There, $\det M \neq 0$ tells us that $U = \mathbb{C} - \{0\}$, which is connected. Our constraint therefore doesn't give us more than one S_i. This doesn't *prove* connectivity, of course. However, by the same linear algebra methods which we alluded to (but didn't provide), a path can be constructed from any nonsingular matrix to I.

Before turning to our detailed examples, let's briefly consider another aspect of the classical Lie groups. As mentioned earlier, all of the non-connected matrix groups listed take the form of semidirect products rather than general group extensions of π_0 by G_0. Why do we get $G = G_0 \rtimes \pi_0(G)$ in all these cases? I.e., why does the SES right-split? Is there something special going on?

First, it's important to note that our statement that "all the classical matrix groups" behave this way, while technically accurate, is a bit misleading. As we saw, most of those groups have a single component — so there is no question of a semidirect product or not. It is more accurate to ask: for $GL(\mathbb{R}, n)$, $O(n)$, and $O(n, m)$ (and ultimately for the Poincare group), why do we only get semidirect products?

There is a glib answer and a slightly less-glib answer. The glib answer is: we have three cases, two of which are closely related and involve the tiny $\pi_0(G) = Z_2$. Why are we surprised that they happen to behave nicely? We have a tiny number of cases and lucked out. They're probably not consequences of some deep unifying principle.

And this is essentially true. Even confined to one of the most nicely behaved classes of groups (reduced affine algebraic groups), one can come up with counterexamples.

See [17].

However, we can bring a little more color to the situation by thinking of things from a physical point of view. It's not accidental that we're interested in $O(n)$ and $O(n, m)$. They arise from real-world symmetries. I.e., there is a selection bias in our choice of Lie groups.

In order for the SES to right-split, we need a copy of $\pi_0(G)$ implemented in G and s.t. it moves us between the components. I.e., we need a copy $D \approx \pi_0(G)$ s.t. $G = G_0 D$.

If we think about the physics of where $O(3)$ and $O(3,1)$ come from, it's space-time symmetries. In the case of $O(3)$, we have a notion of rotation and spatial inversion. The latter can be implemented in many ways, but the most common is $-I$ (which *only* generalizes to odd dimensions n). There is no other angle and norm preserving operation that can be performed. I.e., we have only one parity operation. Similarly, in $O(3,1)$, we have only two discrete physical operations, and both can be implemented via standard matrices. I.e., we know that $O(3,1)$ right-splits because we already have the internal copy of π_0 in hand.

There are many ways to pick those discrete matrices, all equivalent. We'll discuss this in detail in our examples.

This dispels some of the mystery. In a very real sense, we constructed $O(3)$ and $O(3,1)$ with specific right-split maps already in hand.

4. Examples

In our discussion of section 3.5, we mentioned that, of the classical Lie groups, only $GL(n,m)$, $O(n)$, and $O(n,m)$ (along with its subgroup $SO(n,m)$) are not connected. We already touched on some aspects of these groups, but let's now look at them in more detail — along with the Poincare group, which is the ultimate subject of interest.

There are three key questions associated with each such group:

- (i) What is the set of connected components? I.e., what is the set $\pi_0(G)$?
- (ii) What is the group structure of the quotient group $\pi_0(G) \approx G/G_0$?
- (iii) Is G a semidirect product (or perhaps direct product) of G_0 and $\pi_0(G)$? I.e., is there an internal copy D of $\pi_0(G)$ in G s.t. $G = G_0 D$, and is D normal? Equivalently, does the canonical SES right or left split?

The answers for (i) and (ii) are "small and simply-behaved" and the answer to (iii) is "at worst a semidirect product" in all the cases of interest. We motivated this in section 3.5, but ultimately, there is no fundamental mathematical reason for this. If anything, it comes from the source of the particular groups in question as symmetries (i.e. actions on space or spacetime).

We'll find the following result useful.

Prop 4.1: If the decomposition of Lie group G into G_0 and $\pi_0(G)$ is a semidirect product $G = G_0 \rtimes \pi_0(G)$ and $\pi_0(G)$ is abelian, then G is a direct product $G = G_0 \overline{\times} \pi_0(G)$ iff we can choose a suitable internal copy D of $\pi_0(G)$ in the center $Z(G)$ of G.

Recall that the center $Z(G) \subseteq G$ is the subgroup consisting of those elements which commute with all of G. Obviously, G is abelian iff $Z(G) = G$.

Bear in mind that choosing a different D gives a different semidirect product breakdown of the same G. It is quite possible to have G be a direct product of one copy of D and G_0 but only a semidirect product of another. See [1] for an extensive discussion of what distinct right-splitting maps entail.

Pf: Let $D \approx \pi_0(G)$ and $G = G_0 D$. To get a direct product, we need D to be normal in G. I.e. (as sets) $gD = Dg$ for all $g \in G$. Obviously, if $D \subseteq Z(G)$, $gd = dg$ for every d and g, which certainly implies that $gD = Dg$. Put another way, any subgroup of the center is automatically normal in G. Going the other way, suppose we have a direct product. Then D is normal in G, so $gD = Dg$ as sets. First, we'll show that D and G_0 commute with one another. Consider dh for some $h \in G_0$ and $d \in D$. Since G_0 is normal, this equals $h'd$ for some $h' \in G_0$, and since D is normal, it equals hd' for some $d' \in D$. However, each G has a unique decomposition as $g = h_g d_g$, for some $h_g \in G_0$ and $d_g \in D$. Since $hd' = h'd$, we must have $h' = h$ and $d' = d$. Since $dh = hd' = hd$, we see that D commutes with G_0. Now that we know that G_0 and D commute with one another, consider some general $g = hd$ and multiply it by some $d' \in D$. Then $gd' = hdd' = hd'd$ since D is abelian. This equals $d'hd = d'g$ since G_0 and D commute with one another. Therefore, D commutes with all of G and is a subgroup of $Z(G)$.

In all the examples we'll now examine, the conditions of proposition 4.1 are met: $\pi_0(G)$ is abelian and G is a semidirect product of G_0 and $\pi_0(G)$. Therefore, $G = G_0 \overline{\times} \pi_0(G)$ iff there exists such a D in its center. For $GL(\mathbb{R}, n)$, $O(n)$, and $O(n, m)$ the center is nontrivial and this is a possibility. We'll see that when n (and m, where present) is odd, such a D does, in fact, exist. However, for the Poincare group (or its generalization to (n, m)), the center is trivial and we never have a direct product.

Our programme for each example will be the same:

- (i) From the defining constraints for the group, we'll identify a partition of G into nonempty clopen sets.
- (ii) We'll show that these clopen sets are diffeomorphic to one another. This makes them candidates for components (since all the components of a Lie group must be disjoint, clopen, and diffeomorphic), and sets a lower bound on the number of components. However, each could still be a union of more than one component.

  We'll show they are diffeomorphic by exhibiting one or more g's s.t. L_g takes us between the various clopen sets.
- (iii) We'll show that the identity candidate-component G_0 is path-connected. Since the other candidate components are homeomorphic to it, they too are path-connected. This establishes that the candidate components are the actual components and gives us $\pi_0(G)$ as a set. We also know that $\pi_0(G)$ has the discrete topology.

  There are several approaches to proving path-connectivity. (a) Since components are path-components for a Lie group, we could find some homeomorphic space that we know the number of connected components of. If this equals the number of candidate components, then those candidate components must be connected. (b) We could look at the Lie algebra and show that it generates the entirety of G_0. (c) We could use the techniques of linear algebra (ex. Gram-Schmidt orthogonalization, Polar decomposition, etc) specific to each of the groups in question to construct a path from every element of G_0 to the identity element. We'll allude to this, but we won't delve into the details of it. Ultimately, we'll assert or allude to (iii) rather than proving it, since doing so would take us too far afield.
- (iv) We'll show that $\pi_0(G)$ is abelian and either Z_2 or $Z_2 \oplus Z_2$.

  There is only one 2-element group, so if $\pi_0(G)$ has 2 elements it is Z_2, which is abelian. In the case of $Z_2 \oplus Z_2$, we'll use L_g relations of (ii) to show that the 4-element $\pi_0(G)$ has two generators of degree 2 rather than a single generator of degree 4. This forces it to be $Z_2 \oplus Z_2$, which is abelian. For the record, Z_4 is abelian too.
- (v) We'll explicitly determine the set of all possible internal copies D of $\pi_0(G)$ in G s.t. $G = G_0 D$. The existence of such D's tells us that G is a semidirect product of G_0 and $\pi_0(G)$. Given any such D, $G = G_0 \rtimes D$.
- (vi) We'll determine the center $Z(G)$, the set of elements of G that commutes with all of G.

- (vii) We'll ascertain whether any of the D's from (v) sit in the center. Since our $\pi_0(G)$'s will be abelian, by proposition 4.1, G is a direct product $G = G_0 \overline{\times} D$ iff D is in the center of G. If we can find such a D, then G is a direct product. If not, it is only a semidirect product.

Let's now look at our examples.

4.1. $GL(\mathbb{R}, n)$.

We've already discussed $GL(\mathbb{R}, n)$ to some extent, so we'll keep this brief. $GL(\mathbb{R}, n) \subset gl(\mathbb{R}, n)$ (subset, not subalgebra) is defined by the condition $\det M \neq 0$. We saw that $\det M$ is a continuous map to $(-\infty, 0) \cup (0, \infty)$.

Let's define $G_0 = \{M \in gl(\mathbb{R}, n); \det M > 0\}$ and $G_1 = \{M \in gl(\mathbb{R}, n); \det M < 0\}$. Define the matrix J to be I with the first diagonal entry flipped to -1. Then $I \in G_0$ and $J \in G_1$, so $GL(\mathbb{R}, n)$ intersects both G_0 and G_1. Since $\det M \neq 0$ is the *only* condition, $G = G_0 \cup G_1$. As discussed earlier, the disjoint sets G_0 and G_1 must be clopen and each is a union of components of G.

Note that G_0 is *not* $SL(\mathbb{R}, n)$. The latter is defined by $\det M = 1$ and the former is defined by $\det M > 0$. They are, however, homotopic.

So far, we've just reiterated our previous discussion. Given any matrix $M \in G_1$ (i.e. $\det M < 0$), we have a diffeomorphism $L_M : G_0 \to G_1$. G_0 and G_1 are therefore diffeomorphic, which means they *can* (but need not be) components. The methods of linear algebra can be used to construct a continuous path from any M with $\det M > 0$ to I, thus showing that G_0 is path-connected. Since G_1 is homeomorphic to G_0, it too is path-connected. I.e., G_0 and G_1 are the two components of G.

Details can be found in any of the classical treatments of the subject. See, for example, [18].

There is only a single two-element group: Z_2, so $\pi_0(G) \approx Z_2$ and is abelian. As discussed earlier, $\pi_0(G)$ has the discrete topology.

Let's now turn to the question of whether we have a semidirect product. To have one, we need a copy of Z_2 in G, which we'll call $D \subset G$, s.t. $G = G_0 D$. As a subgroup of G, D contains the identity matrix. Denote its other element P. Since $I \in G_0$, we need $P \in G_1$ (or we'd have no way to move between the components).

We thus need a matrix P with $\det P < 0$. In order to implement the group structure of Z_2, we need $P^2 = I$. We already know that, given such a P, L_P is a diffeomorphism between G_0 and G_1. This means that every element of G_1 has a unique expression as gP for some $g \in G_0$. I.e., given any $g \in G$, it either can be written gI (if $g \in G_0$) or $g'P$ for some $g' \in G_0$ (if $g \in G_1$). We thus have that D is a subgroup of G and $G = G_0 D$.

There exists at least one such P. For example, we could define P to be J. G can be written as a semidirect product of G_0 and $\pi_0(G)$. Each D (i.e. choice of P) gives us a particular semidirect product.

In terms of SES's, each such semidirect product has the same quotient map (taking $G_0 \subset G$ to $[G_0] \in \pi_0(G)$ and $G_1 \subset G$ to $[G_1] \in \pi_1(G)$) but a different right-splitting map $k : \pi_0(G) \to G$ that takes $[G_0]$ to I and

$[G_1]$ to our choice of P. There are many suitable copies of Z_2 in G, each corresponding to a suitable choice of P.

To determine whether G can be written as a direct product, let's first compute its center.

Prop 4.2: $Z(GL(\mathbb{R},n)) = \{cI; c \neq 0\}$.

I.e., the center is all nonzero multiples of I.

Pf: Suppose $N \in Z(G)$. Recall that $\det(I+\epsilon M) = 1+\epsilon(\text{tr } M)+O(\epsilon^2)$. Therefore, for *any* matrix M, $\det(I+\epsilon M)$ is nonzero for sufficiently small ϵ, and $(I+\epsilon M) \in GL(\mathbb{R},n)$. If N is in the center of G, it must commute with every element of $GL(\mathbb{R},n)$, including $I+\epsilon M$. We therefore need that $N(I+\epsilon M) = (I+\epsilon M)N$. It follows that we need $NM = MN$ for *every* matrix $M \in gl(\mathbb{R},n)$. Define $M^{(l,m)}$ to have all 0's except a single 1 in the $(l,m)^{th}$ slot. Obviously, $M^{(l,m)} \notin GL(\mathbb{R},n)$, but — as we just saw — we nonetheless need N to commute with it. Pick some $l \neq m$. Then $(M^{(l,m)}N)_{ij} = \sum_k M^{(l,m)}_{ik} N_{kj} = \sum_k \delta_{li}\delta_{mk}N_{kj} = \delta_{li}N_{mj}$. Similarly, $(NM^{(l,m)})_{ij} = \sum_k N_{ik}M^{(l,m)}_{kj} = \sum_k N_{ik}\delta_{lk}\delta mj = N_{il}\delta_{mj}$. We thus require that, for all $l \neq m$ and i,j, $\delta_{il}N_{mj} = \delta_{mj}N_{il}$. If $i \neq l$ and $m = j$, then $N_{il} = 0$. Pick any l and any $i \neq l$. Since this holds for all $m \neq l$ and j, we just pick any $m \neq l$ and $j = m$. I.e., $N_{il} = 0$ for all $i \neq l$, so N is diagonal. On the other hand, if $i = l$ and $m = j$, then $N_{mj} = N_{il}$, which means $N_{jj} = N_{ii}$ for all i,j. Therefore, N is diagonal with identical entries and thus a multiple of I. It is a nonzero multiple since it is in $GL(\mathbb{R},n)$.

This is from [19].

Since $\pi_0(G)$ is abelian, in order for G to be a direct product of G_0 and $\pi_0(G)$, proposition 4.1 tells us we need D to sit in $Z(G)$. I.e., $P = cI$ for some c. Since $P^2 = I$ and $P \neq I$, $P = -I$ is the only candidate. However, this only has $\det P = -1$ for odd dimensions n.

We can also see this in the external view by noting that $GL(\mathbb{R},n)$ is a direct product iff the relevant $\phi : Z_2 \to Aut(G_0)$ has $\phi_P = Id_{G_0}$ (we already know that $\phi_I = Id_{G_0}$, of course, since ϕ is a homomorphism). From our earlier discussion of semidirect products, we know that $\phi_P(M) = PMP^{-1}$ is the translation between the internal and external views. I.e., we need $PMP^{-1} = M$, or $PM = MP$. P therefore must commute with G_0, and the rest follows as above.

Note that the fact that $P = -I$ is not linearly independent from I is immaterial. We're talking about I and P as elements of a group whose multiplication is matrix multiplication. Considerations of "linearity" or "linear independence" don't come into play.

4.2. $O(n)$.

Consider $G = O(n)$, the orthogonal group of real $n \times n$ matrices. It is defined by the condition $M^TM = I$ and arises as the group of isometries that preserve an inner product. Despite superficial differences, we can tackle $O(n)$ in much the same way as we did $GL(\mathbb{R},n)$. The reason is that the disconnectedness of $O(n)$ arises exclusively from $\det M = \pm 1$, which follows from the defining constraint.

$M^TM = I$ implies $(\det M)^2 = 1$, so $\det M = \pm 1$. Since $\{-1\} \cup \{+1\}$ is a union of a finite number of disjoint closed sets, our discussion from section 3.5 applies. We thus have two clopen, disjoint sets $S_0 \equiv \{M \in gl(\mathbb{R},n); \det M = 1\}$ and $S_1 \equiv \{M \in gl(\mathbb{R},n); \det M = -1\}$. I.e., the sets of matrices with determinant -1 or $+1$. It follows that $G_0 \equiv G \cap S_0$ and $G_1 \equiv G \cap S_1$ are clopen in the subspace topology on G. They thus constitute a disjoint clopen cover of $O(n)$.

Define J to be I but with the first diagonal entry changed to -1. Then, $J^TJ = I$ and $\det J = -1$. Since $J \in G_1$ and $I \in G_0$, both G_0 and G_1 are nonempty.

Given any $M \in G_0$, $M^TM = I$ and $\det M = 1$. Therefore, $(JM)^T(JM) = M^TJ^TJM = M^TM = I$ and $\det(JM) = \det J \det M = -1$. Since L_J is a diffeomorphism from $gl(\mathbb{R},n)$ to itself, it restricts to one from

G_0 to $L_J G_0 = G_1$. We thus see that G_0 and G_1 are diffeomorphic. This means that they are candidates for components, each must be a union of components, and we have at least 2 components.

As with $GL(\mathbb{R}, n)$, it is possible to use methods from linear algebra to prove that G_0 is indeed path-connected, but we won't go into the details here.

See [18] for a detailed discussion of this.

Since G_0 is path-connected, and G_0 and G_1 are homeomorphic, G_1 is path-connected too. We therefore have that $\pi_0(G) = \{G_0, G_1\}$. As discussed earlier, $\pi_0(G)$ has the discrete topology.

The common name for G_0 is $SO(n)$, but we'll (mostly) stick with G_0.

Once again, there is only one two-element group: Z_2, which is abelian.

In order to have a semidirect product, we need a copy of Z_2 in G, call it $D \subset G$, s.t. $G = G_0 D$. The identity of D must be that of G. Let's call the other element P. Since $I \in G_0$, we need $P \in G_1$. To implement Z_2, we need $P^2 = I$. To be an element of G, we need $P^T P = I$. To be an element of G_1 (the non-identity component), we need $\det P = -1$. To have $G = G_0 D$, we need $G_1 = G_0 P$. I.e., the diffeomorphism L_P must swap G_0 and G_1. However, this is automatic since we're already requiring that $P \in G_1$.

The constraints that $P^T P = I$ and $P^2 = I$ tell us that $P = P^T = P^{-1}$. Every symmetric orthogonal matrix has the form AQA^T, where A is orthogonal and Q is diagonal with all entries ± 1. To have $\det P = -1$, we must have $\det Q = -1$.

I.e., we gain no generality by looking beyond diagonal matrices with ± 1 entries. Nonetheless, we will continue to do so. [More precisely, using P as our copy of Z_2 is the same as using the corresponding Q as our copy of Z_2 on an L_A-transformed version of G.].

See [20].

Any choice of P as described yields a suitable copy D of Z_2 in G, and we have $G = G_0 D$. In terms of the internal view SES, different D's give rise to different right-splitting maps $k : \pi_0(G) \to G$. We always have $k(G_0) = I$, and all that differs is our choice of $k(G_1) = P$. As long as P satisfies $\det P = -1$ and $P = P^T = P^{-1}$, we're fine.

Since $\pi_0(G)$ is abelian, proposition 4.1 tells us that G is a direct product iff we can choose a D in the center of G. Let's therefore determine that center.

Prop 4.3: $Z(O(n)) = \{\pm I\}$.

Pf: This is adapted from [21]. Suppose N commutes with $O(n)$. Let $M^{(i)}$ denote a matrix that is I except for a single -1 in the (i,i) slot. Clearly, $M^{(i)} \in O(n)$ for all $i = 1 \ldots n$. $NM^{(i)}$ is just N but with the sign of the i^{th} column flipped and $M^{(i)}N$ is just N with the sign of the i^{th} row flipped. These are equal iff all elements in the i^{th} column and i^{th} row except the (i,i) slot are zero. Therefore, N is diagonal. Next, define a rotation in the (i,j)-plane (aka a "Givens rotation") for $i \neq j$ via $R_{kl} = 0$ except for all 1's along the diagonal other than (i,i) and (j,j) and the four values $R_{ii} = R_{jj} = \cos\theta$ and $R_{ij} = \sin\theta$ and $R_{ji} = -\sin\theta$ for some $\theta \neq 0$. This is easily seen to be an orthogonal matrix. Let v_i denote the i^{th} diagonal element of N. Then RN is the same as N for all columns except i and j. $(RN)_{ii} = v_i \cos\theta$ and $(RN)_{ji} = -v_i \sin\theta$ and the rest of the column is 0. $(RN)_{jj} = v_j \cos\theta$ and $(RN)_{ij} = v_j \sin\theta$ and the rest of the column is 0. On the other hand, (NR) looks the same as N except for rows i and j. $(NR)_{ii} = v_i \cos\theta$ and $(NR)_{jj} = v_j \cos\theta$ as before, so we're fine there. However, $(NR)_{ij} = v_i \sin\theta$ and $(NR)_{ji} = -v_j \sin\theta$. Setting $(NR)_{ij} = (RN)_{ij}$ requires $v_i = v_j$. I.e., $N = cI$ for some $c \neq 0$. However, $N^T N = I$ tells us that $c = \pm 1$. Therefore, the center is $\{\pm I\}$.

The only possible way for D to be in the center of G is if it equals that center. $(-I)^T(-I) = I$ and $(-I)(-I) = I$, but $\det(-I) = -1$ iff n is odd. We therefore have a direct product only in odd dimensions.

The left-splitting homomorphism is $j : G \to G_0$, given by $j(M) = j(-M) = M$ for all $M \in G_0$. Why does this fail if the dimension is even? To be a left-splitting homomorphism, $j \circ i = Id_{G_0}$. I.e., $j|_{G_0} = Id_{G_0}$. If the dimension is even, then both M and $-M$ are in the same component, since $\det(-M) = \det M$. This means that if $M \in G_0$, so is $-M$. However, $j(M) = j(-M) = M$, so $j \circ i(-M) \neq -M$, and j isn't a left-splitting homomorphism.

Once again, the fact that $P = -I$ is not linearly independent of I is immaterial.

4.3. $O(n, m)$.

Now, let's look at a slightly more interesting case: $G = O(n, m)$. G is the group of invertible $(n + m) \times (n + m)$ matrices which preserve a pseudometric. I.e., it is the matrices M which satisfy $M^T \eta M = \eta$ for some diagonal matrix η with n ones and m minus-ones.

This is sometimes called the "indefinite orthogonal group". The cases $O(n, 1)$ and $O(1, n)$ are called the "generalized Lorentz group" and the cases $O(3, 1)$ and $O(1, 3)$ are called the "Lorentz group".

A note on conventions: there are two choices that must be made. $O(n, m)$ is defined using an η with n plus-ones and m minus-ones. However, some people employ $O(m, n)$ in its stead, swapping pluses with minuses. The resulting groups are identical. This is obvious from the defining condition: $M^T \eta M = \eta$ holds for M iff $M^T(-\eta)M = (-\eta)$. For example, consider the usual Lorentz group from physics. Some people prefer $O(3, 1)$ and some people prefer $O(1, 3)$. We'll use $O(3, 1)$. Independent of this is our choice of the ordering of the ± 1 entries in η. For $O(3, 1)$ most people place -1 as the first diagonal entry, but some prefer the last diagonal entry. Regardless of whether we use $O(3, 1)$ or $O(1, 3)$ and how we order the entries, the odd-man out is always the time axis. I.e., for $O(3, 1)$, if the i^{th} diagonal entry is -1, time is the i^{th} component of each 4-vector. We'll use $O(n, 1)$, with the -1 in the first slot — meaning that x_0 is the time component of a $(n + 1)$-vector.

Although vector and matrix indices usually start with 1, in physics it is common to start with 0 when placing the temporal (i.e. -1 in $O(3, 1)$) element in the first diagonal slot of η. The same is true for $O(n, 1)$. We'll stray into this convention occasionally. Our meaning will be clear.

For example, for $O(3, 1)$, we'll use

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

In general, $O(n, m)$ is an $(n+m)(n+m-1)/2$ dimensional real Lie Group. It has 4 connected components. In keeping with physics usage, we'll refer to the negative (i.e. m) dimensions as "temporal" and the positive (i.e. n) dimensions as "spatial". We'll assume in this discussion that $n > 0$ and $m > 0$ (otherwise, we just get $O(n)$ or $O(m)$).

Again, it is important not to conflate the number of -1's with their placement (together referred to as the "signature"). We'll have m minus-ones, and always place them first, but this is just a choice of convention.

As with $O(n)$, the defining condition $M^T \eta M = \eta$ implies that $\det M = \pm 1$. The same exact argument as for $O(n)$ provides us with two diffeomorphic disjoint clopen sets G_0 and G_1, corresponding to $\det M = +1$ and $\det M = -1$. By the usual reasoning, we have that each *can* (but need not) be a component, and is a union of components. However, unlike for $O(n)$, G_0 and G_1 turn out *not* to be components.

G_0 is commonly referred to as the "proper Lorentz group" and denoted $SO(3, 1)$.

There are a number of ways to proceed at this point. One can construct a topological argument for why $O(n,m)$ has 4 components and then search for a suitable constraint to implement the remaining division. However, we can also intuitively see why there should be 4. $O(n,m)$ looks like $O(n)$ and $O(m)$ glommed together in some fashion — and we'll firm up this intuition shortly.

We're not claiming that $O(n,m)$ consists of block matrices or anything so simple.

To get a feel for things, let's first consider the simpler case $O(n,1)$. The constraint $M^T\eta M = \eta$ tells us that $\sum_{j,k} M^T_{ij}\eta_{jk}M_{kl} = \eta_{il}$. Denote $\eta_{ij} = \delta_{ij}\eta_j$, where η_j is the j^{th} diagonal element and there is no implied sum. Then we have $\sum_j M_{ji}M_{jl}\eta_j = \delta_{il}\eta_i$. Consider the diagonal element $l = i$. $\sum_j M_{ji}M_{ji}\eta_j = \eta_i$. This tells us that $-M^2_{0i} + \sum_{k=1}^n M^2_{ki} = \eta_i$. For $i = 0$, we get $-M^2_{00} + M^2_{10} + M^2_{20} + \cdots + M^2_{n0} = -1$. This is of the form $M^2_{00} = 1 + x$, where $x = \sum_{k=1}^n M^2_{k0}$ is nonnegative. As a result, we must have $M^2_{00} \geq 1$.

We're following Arthur Jaffe's treatment of the subject in [22].

We could just as easily have used $i = 1$ or any of myriad other constraints derived from $M^T\eta M = \eta$. There are many ways to identify the two components comprising $SO(n,1)$. This just happens to be an intuitively convenient one. To go from one component to the other requires negation of the temporal element. This can be thought of as "time reversal".

M_{00} is a continuous function, and the constraint tells us that it is from $gl(\mathbb{R},n)$ to $(-\infty,-1]\cup[1,\infty) \subset \mathbb{R}$. Since M_{00} maps to the union of two disjoint closed sets, we once again know that the inverse images of these closed sets are clopen and disjoint in $gl(\mathbb{R},n)$. Call the $M_{00} \geq 1$ set S_0 and the $M_{00} \leq -1$ set S_1. Then, $H_0 = G_0 \cap S_0$, $H_1 = G_1 \cap S_0$, $H_2 = G_0 \cap S_1$, and $H_3 = G_1 \cap S_1$ are all clopen. The fact that each pair (G_0 and G_1, and S_0 and S_1) is disjoint means that their four intersections are disjoint. We thus have a partition of G into 4 clopen sets.

It is easy to see that all four of these sets are nonempty. Define J to be I with the first diagonal element swapped to -1, define J' to be I with the second element swapped to -1, and define J'' to be I with both the first and second diagonal elements swapped to -1. Then $\det I = 1$ and $I_{00} = 1$, $\det J = -1$ and $J_{00} = -1$, $\det J' = -1$ and $J'_{00} = 1$, and $\det J'' = 1$ and $J''_{00} = -1$. I.e., we have an element of $O(n,1)$ in each of the four sets. Therefore, $H_0, \ldots, H_3$ are all nonempty.

Specifically, $I \in H_0$, $J \in H_3$, $J' \in H_1$, and $J'' \in H_2$.

Since J flips the sign of both $\det M$ and M_{00}, L_J restricts to a diffeomorphism between H_0 and H_3 and between H_1 and H_2. Since J' flips only the sign of $\det M$ and leaves M_{00} alone, $L_{J'}$ restricts to a diffeomorphism between H_0 and H_1 and between H_2 and H_3. We thus see that all four subspaces are diffeomorphic. As per our usual reasoning, this qualifies them as candidate components.

Of course, $M^T\eta M = \eta$ is far more constraining than just $\det M = \pm 1$ and $M^2_{00} \geq 1$. Those are simply two pieces of information, and $M^T\eta M = \eta$ gives rise to a large variety of other constraints. However, it turns out that $O(n,1)$ has only four components. From the standpoint of connectedness, all the other constraints are irrelevant or split G into unions of the H_i's.

To see that there are four components, we need only show that H_0 is path-connected. As with $GL(\mathbb{R},n)$ and $O(n)$, it is possible to use methods from linear algebra to construct an explicit path from every $M \in H_0$

to I. We won't do so here.

The problem actually reduces to that of $O(n)$, because one can show that any element $g \in O(n,1)$ has a polar decomposition into either $\begin{pmatrix} 1 & 0 \\ 0 & M \end{pmatrix}$ or $\begin{pmatrix} -1 & 0 \\ 0 & M \end{pmatrix}$, where $M \in O(n)$. Since the $(0,0)$ element partitions us into $H_0 \cup H_1$ and $H_2 \cup H_3$, the remaining $O(n)$ breaks each down into the usual two components by $\det M$ (which is *not* the same as $\det g$ since we also must multiply by the $(0,0)$ element). See [23] for details.

Let's now return to the general case of $O(n,m)$. It is not immediately obvious how to generalize the $|M_{00}| \geq 1$ constraint. Intuitively, we want something that smacks of "spatial" and "temporal" inversions. A general matrix in $O(n,m)$ need not be block diagonal, so we can't cavalierly just use the signs of the determinants of the two pieces. However, it turns out that we can noncavalierly do so.

Any $M \in gl(n,m)$ can be written as a block matrix $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$, where A is $m \times m$, B is $m \times n$, C is $n \times m$, and D is $n \times n$.

Prop 4.4: For any $M \in O(n,m)$ written in this block form, $|\det A| = |\det D| \geq 1$ and $\det M = \det A / \det D = \det D / \det A$.

Pf: Since η is diagonal with m minus ones followed by n plus ones, $M^T \eta M = \eta$ translates into $B^T A = D^T C$ (an $n \times m$ matrix equation) and $A^T A - C^T C = I_m$ (the $m \times m$ identity) and $D^T D - B^T B = I_n$ (the $n \times n$ identity). If D is nonsingular, we can write $\det M = (A - BD^{-1}C) \det D$, and if A is nonsingular, we can write $\det M = \det(D - CA^{-1}B) \det A$ (these are standard matrix results). Both A and D are, in fact, nonsingular. To see this, first note that for square matrix Q, $Qx = \lambda x$ iff $(I+Q)x = (1+\lambda)x$. I.e., the eigenvectors of $I+Q$ are just those of Q and the eigenvalues are $1+\lambda_i$. Since the determinant is the product of the eigenvalues, $\det(I+M) = \prod(1+\lambda_i)$. In our case, we have the equations $D^T D = B^T B + I_n$ and $A^T A = C^T C + I_m$ (bearing in mind that, although $B^T B$ and $C^T C$ are square and symmetric, B and C generally are not). A symmetric matrix need not be nonnegative-definite in general, but a square matrix of the form $C^T C$ must be. Consider an m-vector x. $x^T C^T C x = (Cx)^T (Cx)$, which is just $\sum_{i=1}^n (Cx)_i^2$. This is nonnegative for all x, so $C^T C$ is nonnegative-definite. This means that all its eigenvalues are ≥ 0. Therefore $\prod(1+\lambda_i) \geq 1$. I.e. $\det A^T A = \det(I + C^T C) \geq 1$, so $(\det A)^2 \geq 1$. Using $I + B^T B$, we similarly get $(\det D)^2 \geq 1$. This tells us that both $\det A$ and $\det D$ are nonzero, so the earlier formulae hold. Now, consider $\det(D - CA^{-1}B) = \det(D^T D - D^T C A^{-1} B) / \det D$ (since $\det D^T = \det D$). However $D^T D = I + B^T B$, so the numerator is $\det(I + B^T B - D^T C A^{-1} B)$. Since $D^T C = B^T A$, we have $\det(I + B^T B - B^T A A^{-1} B) = \det(I + B^T B - B^T B) = \det I = 1$. I.e., $\det(D - CA^{-1}B) = 1/\det D$. We therefore have $\det M = \det A \det(D - CA^{-1}B) = \det A / \det D$. The same exact machinations with $\det M = (A - BD^{-1}C) \det D$ (now using $A^T A = I + C^T C$) yield $\det M = \det D / \det A$. This makes sense, since swapping rows and columns doesn't affect the determinant — so it shouldn't matter what we call A or D. Since $\det M = \pm 1$ and $\det M = \det A / \det D$, $\det D = \pm \det A$ (with the $\pm$ determined by whether $\det M = \pm 1$). This also guarantees that $\det M = \det D / \det A$. We therefore have that $|\det D| = |\det A|$, and we already showed that they are ≥ 1.

This is fully consistent with our discussion of $O(n,1)$, where $\det A = M_{00}$. We showed that $|M_{00}| \geq 1$ and that the sign of M_{00} (along with that of $\det M$) labels the connected components.

We now have a plausible approach. Even though M need not be block diagonal, the signs of the determinants of the A and D blocks may give us what we need. Since $I \in G_0$ and $\det I = 1$, any other $N = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ in G_0 must have $\det N = 1$ as well, meaning that $\text{sgn } \det A = \text{sgn } \det D$ (which, in light of the fact that $|\det A| = |\det D|$, actually tells us that $\det A = \det D$). A reasonable hypothesis therefore would be that the components are actually indexed by the signs of $\det A$ and $\det D$. This is, in fact, the case.

Equivalently, they could be indexed by $(\det M, \text{sgn } \det A)$ or by $(\det M, \text{sgn } \det D)$, since any two determine the third. Note that we write $\det M$ rather than $\text{sgn } \det M$ since $\det M = \pm 1$, so the two are the same. Our discussion of the $O(n,1)$ case indexed the components by $(\det M, \text{sgn } \det A)$, but now it will prove more convenient to use $(\text{sgn } \det A, \text{sgn } \det D)$.

$\det A$ and $\det D$ furnish two continuous functions from $gl(\mathbb{R}, n)$ to $(-\infty, -1] \cup [1, \infty) \subset \mathbb{R}$. By the usual argument, each of them partitions $gl(\mathbb{R}, n)$ into two disjoint clopen subsets, and thus also partitions G

into two disjoint clopen subsets in the subspace topology. Denoting by A_0 and A_1 the sgn $\det A = 1$ and sgn $\det A = -1$ sets and by D_0 and D_1 the sgn $\det D = 1$ and sgn $\det D = -1$ sets, we'll define $H_0 = G \cap A_0 \cap D_0$, $H_1 = G \cap A_0 \cap D_1$, $H_2 = G \cap A_1 \cap D_1$, and $H_3 = G \cap A_1 \cap D_0$.

It is easy to see that this is in keeping with our definitions of $H_0, \ldots, H_3$ for $O(n,1)$, since $\det A$ (the generalization of M_{00}) is ≥ 1 for H_0 and H_1, while $\det M = \det A / \det D$ is 1 for H_0 and H_2.

Define J to be I with the first diagonal element flipped to -1 and define J' to be I with the m^{th} diagonal element flipped to -1 and define J'' to be I with both flipped. Then $I \in H_0$, $J' \in H_1$, $J'' \in H_2$, and $J \in H_3$. This shows that the four sets are nonempty. We already know that they are disjoint and clopen and cover G.

L_J restricts to a diffeomorphism between H_0 and H_3 and between H_1 and H_2, and $L_{J'}$ restricts to a diffeomorphism between H_0 and H_1 and between H_2 and H_3. This shows that they are all diffeomorphic and thus are candidate components.

At this point, we can either observe from topological considerations (ex. by finding a space homotopic to $O(n,m)$) that G has only 4 components — in which case these must be them — or we can once again turn to the techniques of linear algebra. We won't perform either here and simply will state that H_0 is indeed path-connected.

$\pi_0(G) = \{H_0, H_1, H_2, H_3\}$ is discrete by the usual argument (i.e. because each component is clopen in G). Modulo isomorphism (i.e. relabeling of elements), there are two groups with four elements: Z_4 and $Z_2 \oplus Z_2$, depending on whether we have a single generator of order 4 or two generators of order 2 each.

We already know the answer (since we have two independent degrees of freedom in sgn $\det A$ and sgn $\det D$ or whichever pair we pick). This follows from the fact (which we didn't prove) that H_0 is path-connected and thus our labeling is meaningful. It is confirmed by the following proposition, which we won't prove.

Prop 4.5: Let $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ and $M' = \begin{pmatrix} A' & B' \\ C' & D' \end{pmatrix}$ be matrices in $O(n,m)$ in block form, and let $\begin{pmatrix} A'' & B'' \\ C'' & D'' \end{pmatrix}$ be the block form of MM'. Then sgn $\det A'' = (\text{sgn } \det A)(\text{sgn } \det A')$ and sgn $\det D'' = (\text{sgn } \det D)(\text{sgn } \det D')$.

I.e., sgn $\det A$ and sgn $\det D$ behave as hoped for under matrix multiplication, which will allow us to move between components using elements of G (and thus establish a semidirect product relationship between G and H_0).

Since $\det M = \det A / \det D$ and $\det M' = \det A' / \det D'$ and $\det MM' = \det M \det M'$, we have $\det A'' / \det D'' = (\det A \det A')/(\det D \det D')$. This is fully consistent with the result but isn't enough to prove it (because it's also consistent with both the numerator and denominator flipping sign).

Multiplication between elements of two cosets H_i and H_j yields a matrix whose block form satisfies sgn $\det A'' = (\text{sgn } \det A)(\text{sgn } \det A')$ and sgn $\det D'' = (\text{sgn } \det D)(\text{sgn } \det D')$. This means that if we label the elements of $\pi_0(G)$ by $(a,d) \equiv (\text{sgn } \det A, \text{sgn } \det D)$, their product is given by $(a,d)(a',d') = (aa', dd')$. This is patently $Z_2 \oplus Z_2$. We therefore have $\pi_0(G) \approx Z_2 \oplus Z_2$ rather than Z_4. It is obviously abelian.

Under this labeling, $H_0 = (1,1)$, $H_1 = (1,-1)$, $H_2 = (-1,1)$, and $H_3 = (-1,1)$.

This $\pi_0(G)$ is sometimes dignified with the fancy title "The Klein 4-group", but it's really nothing other than the direct sum $Z_2 \oplus Z_2$.

Note that the choice of labeling (i.e. $(\det M, \text{sgn } \det A)$ or $(\det M, \text{sgn } \det D)$ or $(\text{sgn } \det A, \text{sgn } \det D)$) doesn't affect the group π_0 or the components. Each choice corresponds to a different copy of $\pi_0(G)$ in G. It is a nice property of $Z_2 \oplus Z_2$ that any pair of non-identity elements can serve as the generators. Note that neither $\pi_0(G)$ nor whatever image of it in G we pick is affected by this choice. All that changes is which elements we designate as generators. We can think of this as similar to a choice of basis for $\mathbb{R}^2$.

We already have evidence that there will exist a suitable copy D of $\pi_0(G)$ in G s.t. $G = H_0 D$, so let's figure out what that copy can look like. Since we're labeling the H_i's by $(\text{sgn } \det A, \text{sgn } \det D)$, we'll call the elements $D = \{I, T, P, TP\}$, where T and P are the generators that flip the signs of $\det A$ and $\det D$ respectively. Since $D \approx Z_2 \oplus Z_2$, $PT = TP$ and $P^2 = T^2 = I$. Since T and P are elements of $O(n,m)$, we need $T^T \eta T = \eta$ and $P^T \eta P = \eta$. Denote by P_A and P_D and T_A and T_D their A and D blocks. To place $T \in H_3$ and $P \in H_1$ as needed, we need $\text{sgn } \det P_A = \text{sgn } \det T_D = 1$ and $\text{sgn } \det P_D = \text{sgn } \det T_A = -1$. Because L_P and L_T restrict to diffeomorphisms between the relevant components, we then have $G = H_0 D$ automatically. *Any* choice (of which there are many) of $P \in H_1$ and $T \in H_3$ which satisfy these constraints defines a suitable D.

As with our previous examples, different choices of D correspond to different right-splitting maps of the short exact sequence.

Let's now consider whether and when $O(n,m)$ is a direct product. Since $\pi_0(G)$ is abelian, proposition 4.1 tells that we have a direct product iff there exists a suitable copy of D that resides in the center of $O(n,m)$. Let's therefore construct that center.

Prop 4.6: $Z(O(n,m)) = \{\pm I\}$.

Pf: We start with the same argument as in proposition 4.3. Suppose N commutes with $O(n,m)$. Let $M^{(i)}$ denote a matrix that is I except for a single -1 in the (i,i) slot. Clearly, $M^{(i)} \in O(n,m)$ for all $i = 1 \ldots (n+m)$. $NM^{(i)}$ is just N but with the sign of the i^{th} column flipped and $M^{(i)}N$ is just N with the sign of the i^{th} row flipped. These are equal iff all elements in the i^{th} column and i^{th} row except the (i,i) slot are zero. Therefore, N is diagonal. This is the same as for $O(n)$. All we need is that η is diagonal, so that ${M^{(i)}}^T \eta M^{(i)} = \eta$ for all $M^{(i)}$. We next replicate the Givens rotation part of the proof, but here we must be a little careful. We can apply that technique to $i, j < m$ (with $i \neq j$, of course) to show that the first m diagonal elements of N must be equal. We also can apply it to $m \leq i, j < (n+m)$ (again with $i \neq j$) to show that the latter n diagonal elements of N must be equal. However, if we try to mix i, j between $< m$ and $\geq m$, the resulting R matrix is not in $O(n,m)$ and cannot be used to probe commutativity. [In that case, consider $R^T \eta R$ and suppose $i < j$ (i.e. $i < m$ and $j \geq m$). Focusing on the relevant 2×2 piece (i.e. the $\begin{pmatrix} (i,i) & (i,j) \\ (j,i) & (j,j) \end{pmatrix}$ slots), we see that $\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & cos\theta \end{pmatrix}^T \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & cos\theta \end{pmatrix} = \begin{pmatrix} \sin^2\theta - \cos^2\theta & -2\sin\theta\cos\theta \\ -2\sin\theta\cos\theta & \cos^2\theta - \sin^2\theta \end{pmatrix} \neq \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$.] We thus get that N must be of the form $\begin{pmatrix} aI_m & 0 \\ 0 & bI_n \end{pmatrix}$ for some a and b. To have $N^T \eta N = \eta$, we need $a^2 = 1$ and $b^2 = 1$. We therefore have $a = \pm 1$ and $b = \pm 1$. Let's see which of these work. I and $-I$ obviously commute with all of $O(n,m)$. Let $Y \equiv \begin{pmatrix} I_m & 0 \\ 0 & -I_n \end{pmatrix}$. Obviously, $-Y$ commutes with $O(n,m)$ iff $-Y$ does, so we need only test Y. $\begin{pmatrix} A & B \\ C & D \end{pmatrix} Y = \begin{pmatrix} A & -B \\ C & -D \end{pmatrix}$ and $Y \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A & B \\ -C & -D \end{pmatrix}$. So we need that $-B = B$ and $-C = C$, forcing $B = 0$ and $C = 0$. There are many members of $O(n,m)$ without $B = 0$ and $C = 0$, and Y doesn't commute with those. Therefore, Y and $-Y$ are not in the center.

There is a brief mention in a paper by Shlomo Sternberg that the center has 4 elements ([24], page 659) and this was mentioned in [25] as well. That is wrong. Either he was talking about something else or misspoke. See [26] for clarification.

Since the center has only two elements and $\pi_0(G)$ has four, it is impossible to pick a D that is in the center. Therefore, $O(n,m)$ (for $n > 0$ and $m > 0$) is never a direct product.

4.4. **Poincare Group.**

The Poincare group is the group of all spacetime symmetries in special relativity. As such, we define it through its action on space-time $\mathbb{R}^4$. However, there is no reason to stick with 4 dimensions. As

with $O(n,m)$, we'll allow n spatial and m temporal dimensions. For lack of standard notation, we'll use $PCR(n,m)$ for the Poincare group.

The occasional author uses $Iso(n,m)$ (for "isometry group") but this seems easy to confuse with various groups of isomorphisms, so we'll stick with our homegrown PCR.

The case of special relativity is $PCR(3,1)$ or $PCR(1,3)$.

It is reasonable to think of $PCR(n,m)$ as $O(n,m)$ along with space-time translations. We'll denote by $\mathbb{R}^{(n,m)}$ the group of translations of $\mathbb{R}^{n+m}$. It is just a direct sum of the translation groups for each dimension separately, but we distinguish the spatial and temporal dimensions notationally because of how $\mathbb{R}^{(n,m)}$ will interact with $O(n,m)$ within $PCR(n,m)$. We'll denote the elements of $\mathbb{R}^{(n,m)}$ via t_v, meaning translation by vector $v \in \mathbb{R}^{n+m}$.

The translation group of $\mathbb{R}$ is just $\mathbb{R}$ itself as an abelian group under addition.

Formally, $PCR(n,m)$ turns out to be a semidirect product: $PCR(n,m) = \mathbb{R}^{(n,m)} \rtimes O(n,m)$. I.e., it is a type of group extension of $O(n,m)$ by $\mathbb{R}^{(n,m)}$. Since it is a semidirect product, there is a normal copy of $\mathbb{R}^{(n,m)}$ in $PCR(n,m)$ and there is a (non-normal) copy of $O(n,m)$ within $PCR(n,m)$. Unless otherwise specified, we'll just refer to those internal copies as $\mathbb{R}^{(n,m)}$ and $O(n,m)$, with the understanding that we mean the relevant subgroups. In the internal view, every element $g \in PCR(n,m)$ can be decomposed in a unique way as $t_v r$ for some $t_v \in \mathbb{R}^{(n,m)}$ and $r \in O(n,m)$.

Since $\mathbb{R}^{(n,m)}$ is normal in $PCR(n,m)$, we can also write $g = t_{v'} r$, but for a different $t_{v'} \in \mathbb{R}^{(n,m)}$.

$O(n,m)$ is defined as a matrix group, so we already know how it acts on $\mathbb{R}^{n+m}$. That is its defining representation. However, there is no full and faithful matrix representation of $\mathbb{R}^{(n,m)}$ on $\mathbb{R}^{n+m}$. There are two common ways to deal with this issue:

- (i) We can describe the action of $\mathbb{R}^{(n,m)}$ directly via a vector + operation. In this case, we must keep careful track of what is being multiplied or added to what. For example, we would write a rotation by r followed by a translation by v as $x \to rx + v$.
- (ii) We can describe all of $PCR(n,m)$ as a matrix group if we work in $\mathbb{R}^{n+m+1}$ rather than $\mathbb{R}^{n+m}$.

Let's consider (ii). This is a standard trick for dealing with affine transformations. We express an $(n+m)$-vector x as $\begin{pmatrix} x \\ 1 \end{pmatrix}$ and write an element $r \in O(n,m)$ as $\begin{pmatrix} r & 0 \\ 0 & 1 \end{pmatrix}$ and a translation t_v as $\begin{pmatrix} I_{n+m} & v \\ 0 & 1 \end{pmatrix}$ (where I_{n+m} is the $(n+m)$-dimensional identity matrix, and v appears as a column vector).

Note that t_v and r do not commute. $\begin{pmatrix} r & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} I_{n+m} & v \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} r & rv \\ 0 & 1 \end{pmatrix}$, while $\begin{pmatrix} I_{n+m} & v \\ 0 & 1 \end{pmatrix}\begin{pmatrix} r & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} r & v \\ 0 & 1 \end{pmatrix}$. I.e., a translation followed by a rotation takes $x \to r(v+x)$ and a rotation followed by a translation takes $x \to rv + x$. This is exactly as expected.

We may equally well write a general element of $PCR(n,m)$ as $g = t_v r$, which gives $\begin{pmatrix} r & v \\ 0 & 1 \end{pmatrix}$, or $g = r t_{v'}$, which gives $\begin{pmatrix} r & rv' \\ 0 & 1 \end{pmatrix}$. In this case, we just have $rv' = v$. Neither form is more general than the other. We'll typically stick with the simpler (i.e. first) one.

Using this approach, matrix multiplication effects the overall group multiplication, even between translations and rotations. Bear in mind that only multiplication has meaning. We cannot add or subtract such affine matrices because they do not form a group under addition, and we likewise cannot multiply them by scalars. We'll primarily adopt this matrix approach.

For our present purposes, the normality of $\mathbb{R}^{(n,m)}$ in $PCR(n,m)$ — and the resulting semidirect product — is a bit of a red herring. $\mathbb{R}^{(n,m)}$ has trivial homotopy and is connected (and, as a manifold, this is the same as being path-connected). Recall that a direct product, semidirect product, and general group extension of K by H all look like $H \times K$ setwise, topologically, and (if applicable) as manifolds. The meat is in the group multiplication. Since $PCR(n,m) = \mathbb{R}^{(n,m)} \rtimes O(n,m)$, it is topologically just $\mathbb{R}^{(n,m)} \times O(n,m)$. However, from the standpoint of connectivity, the components of the product are just the products of the components. Since $\mathbb{R}^{(n,m)}$, has a single component, the components of $PCR(n,m)$ are bijective (in a natural way) with those of $O(n,m)$.

Denoting by $SO^+(n,m)$ the identity component of $O(n,m)$ (what we called H_0 in our general discussion of $O(n,m)$), the corresponding identity component of $PCR(n,m)$ is just $\mathbb{R}^{(n,m)} \rtimes SO^+(n,m)$. The details of this restriction of a semidirect product are discussed in section 5.2, where we learn that the restriction is exactly what we expect it to be. In the internal view, $\mathbb{R}^{(n,m)} \rtimes SO^+(n,m) \subset \mathbb{R}^{(n,m)} \rtimes O(n,m)$. It consists of all elements $t_v r$ s.t. $r \in SO^+(n,m)$ rather than all of $O(n,m)$, and is a subgroup of $PCR(n,m)$.

We can similarly construct each of the components of $PCR(n,m)$ this way, as $H'_i \equiv \mathbb{R}^{(n,m)} \rtimes H_i$, where H_i $(i = 0 \ldots 3)$ is the relevant component of $O(n,m)$.

Such operations are of the form $t_v r$ (i.e. $x \to rx + v$), with $r \in H_i$ and $v \in \mathbb{R}^{n+m}$.

We now can view $PCR(n,m)$ in two equivalent but distinct ways:

- (i) $PCR(n,m) = \mathbb{R}^{(n,m)} \rtimes (SO^+(n,m) \rtimes (Z_2 \oplus Z_2))$, where we first extract $O(n,m)$, and then look at its components.
- (ii) $PCR(n,m) = (\mathbb{R}^{(n,m)} \rtimes SO^+(n,m)) \rtimes (Z_2 \oplus Z_2)$, the usual component SES.

Although semidirect products are not associative (as discussed in section 5.1), these two seem pretty close to it. However, the relevant $Z_2 \oplus Z_2$ actually looks quite different between the two. Let's see what's going on.

Choice (i) corresponds to our description of a general transformation as being $g = t_v r = \begin{pmatrix} r & v \\ 0 & 1 \end{pmatrix}$, with $r \in O(n,m)$. In this case, r can be P or T or $PT = TP$ as discussed in our treatment of $O(n,m)$. If we think of an affine transformation as a rotation followed by a translation, the parity/time-reversal takes place at the rotation stage. Let P' be the matrix which takes us from H'_0 to H'_1. Since t_v is glommed on afterward (via the semidirect product), we expect P to take us from an $r \in H_0$ to $rP \in H_1$, and then glom the relevant v on. I.e., the function $x \to rx + v$ should become $x \to rPx + v$ for $r \in H_0$. Let $P' = \begin{pmatrix} P_1 & P_2 \\ 0 & 1 \end{pmatrix}$, the most general form an element of $PCR(n,m)$ can take. We want $\begin{pmatrix} r & v \\ 0 & 1 \end{pmatrix} P' = \begin{pmatrix} rP & v \\ 0 & 1 \end{pmatrix}$. The left

side is $\begin{pmatrix} rP_1 & rP_2+v \\ 0 & 1 \end{pmatrix}$. This means we want $P_1 = P$ and $P_2 = 0$. There's no surprise here. In order not to affect t_v, we have $P' = \begin{pmatrix} P & 0 \\ 0 & 1 \end{pmatrix}$.

Next, let's consider choice (ii). Our $Z_2 \oplus Z_2$ now operates on the entire copy of $\mathbb{R}^{(n,m)} \rtimes SO^+(n,m)$, so its generators, which we'll call P' and T', aren't incorporated in r. Instead, we confine $r \in SO^+(n,m)$, and construct P' and T' analogously to our development of $O(n,m)$. Consider P'. It is an element of $PCR(n,m)$, and thus has the form $\begin{pmatrix} P_1 & P_2 \\ 0 & 1 \end{pmatrix}$, where $P_1 \in O(n,m)$, not $SO^+(n,m)$, is the relevant constraint, and P_2 is some $(n+m)$-vector. The elements of H_0' are $\begin{pmatrix} r & v \\ 0 & 1 \end{pmatrix}$, now with $r \in SO^+(n,m)$.

To generate Z_2, we need $P'^2 = I$, and to take H_0' to H_1', we need $\det P_A = 1$ and $\det P_D = -1$, where $P_1 = \begin{pmatrix} P_A & P_B \\ P_C & P_D \end{pmatrix}$ is the block decomposition of P_1 from our discussion of $O(n,m)$.

The constraint $P^2 = I$ gives us $\begin{pmatrix} P_1^2 & P_1P_2+P_2 \\ 0 & 1 \end{pmatrix}$, so we need $P_1^2 = I_{n+m}$ and $P_1P_2 + P_2 = 0$. The former is just our $O(n,m)$ constraint for P. I.e., any P that would work for $O(n,m)$ will serve as our P_1. The second constraint is just an eigenvalue equation: $P_1P_2 = -P_2$. Either $P_2 = 0$ or P_2 is an eigenvector of P_1 with eigenvalue -1.

We thus have two classes of permissible choices of P'. The first is of the same form as for (i): $P' = \begin{pmatrix} P & 0 \\ 0 & 1 \end{pmatrix}$, with P any choice that would work for $O(n,m)$. The second is of the form $P' = \begin{pmatrix} P & V \\ 0 & 1 \end{pmatrix}$, where P is any solution that would work for $O(n,m)$ *and* has a -1 eigenvalue, and V is any of its eigenvectors that has eigenvalue -1.

Ex. If n and m are odd and $P = \begin{pmatrix} I_m & 0 \\ 0 & -I_n \end{pmatrix}$, then $P' = \begin{pmatrix} P & V \\ 0 & 1 \end{pmatrix}$ works for any V that has $V_i = 0$ for all $i < m$ (including $V = 0$).

The same considerations hold for T'. We need $T' = \begin{pmatrix} T & V \\ 0 & 1 \end{pmatrix}$ where T is any solution for $O(n,m)$ and V is any eigenvector of T with eigenvalue -1.

The final $Z_2 \oplus Z_2$ constraint is that $P'T' = T'P'$. Let $P' = \begin{pmatrix} P & V_P \\ 0 & 1 \end{pmatrix}$, where V_P is either 0 or an eigenvector of P with eigenvalue -1, and ditto for $T' = \begin{pmatrix} T & V_T \\ 0 & 1 \end{pmatrix}$. Bear in mind that P and T must both be suitable generators for the copy of $\pi_0(O(n,m))$ in $O(n,m)$.

Our $P'T' = T'P'$ constraints takes the form $\begin{pmatrix} PT & PV_T+V_P \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} TP & TV_P+V_T \\ 0 & 1 \end{pmatrix}$. We already know that $PT = TP$, since we stipulated that P and T must be appropriate to $O(n,m)$. The remaining condition is that $TV_P + V_T = PV_T + V_P$. I.e, $(T-I)V_P = (P-I)V_T$. We know that $\det T = \det P = -1$,

so we cannot have $T = I$ or $P = I$. Suppose $V_P \neq 0$. As discussed, it must be an eigenvector of P with -1 eigenvalue. Therefore $-IV_P = PV_P$. This means that $(T + P)V_P = (P - I)V_T$. This holds if $V_P = 0$ too, since $(T + P)V_P = 0$ then as well. By the same considerations on the right, we have $(P + T)V_T$. I.e., our equation can be written $(T + P)(V_P - V_T) = 0$. Note that it is perfectly possible for $T + P$ to be singular and have zero eigenvalues, so $V_P - V_T$ need not equal zero.

Ex. if $T = \begin{pmatrix} -I_m & 0 \\ 0 & I_n \end{pmatrix}$ and $P = -T$, then $P + T = 0$ and we have no additional constraint on V_P and V_T. However, if nonzero, they still must be eigenvectors of P and T with eigenvalue -1. In this case, V_T must be of the form $(a_1, \ldots, a_m, 0, \ldots, 0)$ and V_P must be of the form $(0, \ldots, 0, b_1, \ldots, b_n)$.

How do we interpret something like $P' = \begin{pmatrix} P & V_P \\ 0 & 1 \end{pmatrix}$ with nonzero V_P? $\begin{pmatrix} r & v \\ 0 & 1 \end{pmatrix} P' = \begin{pmatrix} rP & rV_P + v \\ 0 & 1 \end{pmatrix}$. I.e., we take the transform $x \to rx + v$ to the transform $x \to rPx + rV_P + v$. We already know that rPx just flips the parity of the spatial part of r. The only weird term is rV_P. Since V_P is an eigenvector of P with eigenvalue -1, we can replace rV_P with $-rPV_P$. I.e., our transform now is $x \to rP(x - V_P) + v$.

Basically, we can think of V_P as specifying a spatial origin relative to which we perform our rotations and inversions.

As usual, let's ask whether and when the semidirect product $PCR(n, m) = H'_0 \rtimes \pi_0(PCR)$ is a direct product. Since $\pi_0(PCR(n, m)) = Z_2 \oplus Z_2$ is abelian, proposition 4.1 tells us that we have a direct product iff there exists a suitable copy of D that resides in the center of $PCR(n, m)$. Let's therefore construct that center.

Prop 4.7: The center of the Poincare group is trivial.

This may come as a surprise, since $O(n, m)$ has 2 elements in its center, and $\mathbb{R}^{(n,m)}$ is abelian and characteristic in G. Unfortunately, these do not suffice. We already saw that rotations and translations do not commute.

Pf: Although $\mathbb{R}^{(n,m)}$ is abelian and characteristic in $PCR(n, m)$, it does not commute with $O(n, m)$. Consider a general $N = \begin{pmatrix} r & v \\ 0 & 1 \end{pmatrix}$. To commute with $M = \begin{pmatrix} r' & v' \\ 0 & 1 \end{pmatrix}$, we need $NM = \begin{pmatrix} rr' & rv' + v \\ 0 & 1 \end{pmatrix}$ to equal $MN = \begin{pmatrix} r'r & r'v + v' \\ 0 & 1 \end{pmatrix}$. This means that r must commute with every $r' \in O(n, m)$, so r is in the center of $O(n, m)$. We saw that the latter is $\pm I$, so we have $r = \pm I$. We also need $rv' + v = r'v + v'$ for every r' and v'. This requires $(r - I)v' = (r' - I)v$. Suppose $r \neq I$. Pick $r' = I$. We then must have $(r' - I)v' = 0$ for every v', which is impossible unless $r' = I$, violating our premise. Therefore, the only possible choice is $r = I$. This leaves $v' + v = r'v + v'$, so $r'v = v$. However, this must hold for every $r' \in O(n, m)$, which it cannot unless $v = 0$. I.e., any given v cannot be an eigenvector of every possible r'. Therefore, the only element in the center is $N = \begin{pmatrix} I & 0 \\ 0 & 1 \end{pmatrix}$.

Therefore, there is no way a copy of $Z_2 \oplus Z_2$ can fit in the center of $PCR(n, m)$, and $PCR(n, m)$ can never be a direct product of H'_0 and $\pi_0(PCR(n, m))$.

5. Appendix: Some Practical Aspects of Semidirect Products

We've briefly introduced semidirect products and some of their properties. A more detailed discussion of their theoretical underpinnings can be found in [1]. Here, we'll consider some practical aspects that tangentially relate to our earlier discussion of π_0.

5.1. **Non-Associativity of Semidirect Products.**

Semidirect products are *not* associative. $(A \rtimes_\phi B) \rtimes_\psi C$ is not the same as $A \rtimes_\phi (B \rtimes_\psi C)$. In fact, the two don't even have the same meaning. In the first, we need homomorphisms $\phi : B \to Aut(A)$ and $\psi : C \to Aut(A \rtimes_\phi B)$, while in the second, we need homomorphisms $\phi : (B \rtimes_\psi C) \to Aut(A)$ and $\psi : C \to Aut(B)$. These aren't even of the same type.

We can also see this in terms of the roles the groups play. In both cases, the resulting group is setwise and topologically $G = A \times B \times C$. However, this is where the similarities end. In the external view, we typically end up with two quite different groups. In the internal view, if both expressions hold for a given G, then the placement, interactions, and group multiplication of the internal copies of A and B and C within G will differ.

This should come as no surprise. Given any H and K, all group extensions of K by H are setwise and topologically, just $H \times K$ — and, in the internal view, have $G = HK$. However, they can differ significantly in the group multiplication and the placement of H and K within G. Nor need they be isomorphic. Formally, the group extensions of K by H are classified by the $2nd$ and $3rd$ group cohomology groups (these are similar to, but distinct from, the topogical cohomology groups). Central extensions (i.e. if the internal copy of H is central to the resulting G, for which H must be abelian as a prerequisite) are classified by the $2nd$ group cohomology group. We won't go into this here.

The following are the group relations that describe the placements of the internal copies of A, B, and C in the resulting group. We denote by i the usual inclusion and by j the right-splitting map. Bear in mind that normality is *not* transitive.

We'll use i and j cavalierly, with the understanding that each applies to the relevant SES.

In $(A \rtimes_\phi B) \rtimes_\psi C$:

- $i(A) \triangleleft (A \rtimes_\phi B)$
- $B \approx (A \rtimes_\phi B)/i(A)$
- $A \rtimes_\phi B = i(A)j(B)$ (i.e. all products).
- $i(A \rtimes_\phi B) \triangleleft ((A \rtimes_\phi B) \rtimes_\psi C)$
- $C \approx ((A \rtimes_\phi B) \rtimes_\psi C)/i(A \rtimes_\phi B)$
- $(A \rtimes_\phi B) \rtimes_\psi C = i(A \rtimes_\phi B)j(C)$

While in $A \rtimes_\phi (B \rtimes_\psi C)$:

- $i(A) \triangleleft (A \rtimes_\phi (B \rtimes_\psi C))$
- $(B \rtimes_\psi C) \approx (A \rtimes_\phi (B \rtimes_\psi C))/i(A)$
- $A \rtimes_\phi (B \rtimes_\psi C) = i(A)j(B \rtimes_\psi C)$
- $i(B) \triangleleft (B \rtimes_\psi C)$
- $C \approx (B \rtimes_\psi C)/i(B)$
- $B \rtimes_\psi C = i(B)j(C)$

We can work through the gory details. First, consider $(A \rtimes_\phi B) \rtimes_\psi C$. Setwise $A \rtimes_\phi B$ is just $A \times B$, and we label its elements (a, b). Multiplication is $(a, b)(a', b') = (a\phi_b(a'), bb')$. $\psi : C \to Aut(A \rtimes_\phi B)$, so ψ_c maps (a, b) to some (a', b'). Let's denote this $\psi_c(a, b) = (\psi_c^1(a, b), \psi_c^2(a, b))$. Overall multiplication takes the form $(a, b, c)(a', b', c') = ((a, b), c)((a', b'), c') = ((a, b)\psi_c((a', b')), cc')$. This is $((a, b)(\psi_c^1(a', b'), \psi_c^2(a', b')), cc')$, which gives us $(a, b, c)(a', b', c') = (a\phi_b(\psi_c^1(a', b')), b\psi_c^2(a', b'), cc')$. Now, consider $A \rtimes_\phi (B \rtimes_\psi C)$. Setwise, $(B \rtimes_\psi C)$ is just $B \times C$, so we can label its elements (b, c). Multiplication is $(b, c)(b', c') = (b\psi_c(b'), cc')$. $\phi : B \rtimes_\psi C \to Aut(A)$ consists of automorphisms labeled $\phi_{(b,c)}$. Overall multiplication takes the form $(a, b, c)(a', b', c') = (a, (b, c))(a', (b', c')) = (a\phi_{(b,c)}(a'), (b, c)(b', c'))$, which gives us $(a, b, c)(a', b', c') = (a\phi_{(b,c)}(a'), b\psi_c(b'), cc')$. Unsurprisingly, the first and second components look a lot different. We didn't expect the semidirect product to be associative, because it doesn't even make sense to ask the question (since the types differ, as mentioned).

Let's now ask a more interesting question that *does* make sense: If we have $(A \rtimes_\phi B) \rtimes_\psi C)$, does there exist a ϕ' and ψ' s.t. $A \rtimes_{\phi'} (B \rtimes_{\psi'} C)$ equals it (or vice versa)? The answer, unfortunately, is no. We don't have enough freedom to always pick such a ϕ' and ψ'. We can only do so if ϕ and ψ cooperate in a certain way.

Bear in mind that ψ' and ϕ' are just labels for homomorphisms and are, as mentioned earlier, of different types than ϕ and ψ.

The same is true going the other way too. I.e. if we start with ψ' and ϕ', we cannot derive a suitable ψ and ϕ unless ψ' and ϕ' cooperate a certain way. The calculation is much the same and we won't replicate it here.

Here are the gory details. Comparing the two multiplications, we see that we need: (i) $a\phi_b(\psi^1_c(a',b')) = a\phi'_{(b,c)}(a')$, (ii) $b\psi^2_c(a',b') = b\psi'_c(b')$, and (iii) $cc' = cc'$. Obviously, the third is satisfied. (i) and (ii) must hold for all values, so we need (i) $\phi_b(\psi^1_c(a',b')) = \phi'_{(b,c)}(a')$ and (ii) $\psi^2_c(a',b') = \psi'_c(b')$. This tells us that we need (i) $\psi'_c(b') = \psi^2_c(a',b')$ and (ii) $\phi'_{(b,c)}(a') = \phi_b(\psi^1_c(a',b'))$. In (i), a' doesn't appear on the left, so we can only define such a ψ' if $\psi^2_c(a',b')$ is independent of a'. Similarly, in (ii) the left side is independent of b', so we can only find such a ϕ' if either $\psi^1_c(a',b')$ is independent of b' or ϕ_b somehow removes any dependence. Suppose ψ^1_c is independent of b' and ψ^2_c is independent of a'. Then $\psi_c(a',b') = (\psi^1_c(a'),\psi^2_c(b'))$. In order for this to be an automorphism of $A \rtimes_\phi B$, we need $\psi_c((a,b)(a',b')) = \psi_c(a,b)\psi_c(a',b')$. The left side is $\psi_c(a\phi_b(a'),bb') = (\psi^1_c(a\phi_b(a')),\psi^2_c(bb'))$ and the right side is $(\psi^1_c(a),\psi^2_c(b))(\psi^1_c(a'),\psi^2_c(b')) = (\psi^1_c(a)\phi_{\psi^2_c(b)}(\psi^1_c(a')),\psi^2_c(b)\psi^2_c(b'))$. For the 2nd components to match, we need ψ^2_c to be a homomorphism from B to itself. Since ψ_c is an automorphism, it is easy to see that ψ^2_c must be an automorphism of B. For the 1st components to match, we need $\psi^1_c(a\phi_b(a')) = \psi^1_c(a)\phi_{\psi^2_c(b)}(\psi^1_c(a'))$. This certainly is possible, but we have no reason to expect it to hold in general. We need ϕ and ψ to cooperate in undoing one anothers' twisting. We needn't go so far as to demand that ψ^1_c be an automorphism of A — but if we do, then we need $\psi^1_c(a)\psi^1_c(\phi_b(a')) = \psi^1_c(a)\psi^1_c(\phi_{\psi^2_c(b)}(\psi^1_c(a')))$, meaning that $\psi^1_c(\phi_b(a')) = \psi^1_c(\phi_{\psi^2_c(b)}(\psi^1_c(a')))$. Since ψ^1_c is an automorphism of A (by assumption now), $\psi^1_c(x) = \psi^1_c(y)$ iff $x = y$, so we need $\phi_b(a') = \phi_{\psi^2_c(b)}(\psi^1_c(a'))$. Since ϕ is a homomorphism, $\phi^{-1}_x(\phi_y(-)) = \phi_{x^{-1}y}(-)$. In our case, $a' = \phi_{b^{-1}\psi^2_c(b)}(\psi^1_c(a'))$. In summary, we can't find a suitable ψ' and ϕ' unless ψ and ϕ cooperate in a certain way.

Now, consider the special case where we're dealing with Lie groups and $G = H \rtimes K$ (in the internal view) and we H is a semidirect product of its H_0 with $\pi_0(H)$ via the canonical $H = H_0 \rtimes D$ with D the internal copy of $\pi_0(H)$. Note that these are all assumption; we needn't have semidirect products in general. I.e., we have $G = (H_0 \rtimes D) \rtimes K$. Since H_0 is characteristic in H and H is normal in G, H_0 is normal in G. Nonetheless, we still need not have any breakdown of the form $G = H_0 \rtimes (D \rtimes K)$.

There is no reason that H_0 should equal G_0, and H_0 (though normal in G) need not be characteristic in G. There is no reason to expect G/H_0 to bear any resemblance to a group extension of K by $\pi_0(H)$, let alone a semidirect product. $G = H_0DK$ in the usual sense, but it does not follow that DK is a subgroup of G. The product of subgroups need not be a subgroup. [Ex. let G be the permutation group of $S = (a,b,c)$ Then G has 6 elements consisting of 3 swaps, 2 rotations, and e. Let $D \approx Z_2$ consist of $(e, a \leftrightarrow b)$ and let $K \approx Z_2$ consist of $(e, b \leftrightarrow c)$. Then DK consists of e, $a \leftrightarrow b$, $b \leftrightarrow c$, and $(a,b,c) \rightarrow (b,c,a)$. However, the latter has no inverse in DK, so DK isn't a group.] Because H_0 is normal in G, we *may* have $G = H_0Q$ for Q an internal copy of G/H_0, but only if the relevant SES right-splits. Even if it does, Q need bear no resemblance to DK. We are relabeling $G = H_0 \times Q$ setwise, but this doesn't mean that our labeling is $Q = DK$. I.e., we may not be able to write $Q = DK$. And this is only *if* G is a semidirect product of of H_0 and G/H_0. It very well may be a general group extension, in which case there is no suitable copy of Q inside it, let alone one that looks like DK. Even in this special case, we needn't have any semblance of associativity.

On the other hand, we know that the direct product *is* associative. $(A\overline{\times}B)\overline{\times}C$ is isomorphic to $A\overline{\times}(B\overline{\times}C)$. Let's consider two special cases that potentially generalize this:

- (i) Is $(A \rtimes_\phi B)\overline{\times}C = A \rtimes_{\phi'} (B\overline{\times}C)$ for some suitable derived ϕ' or vice versa?
- (ii) Is $(A\overline{\times}B) \rtimes_\phi C = A\overline{\times}(B \rtimes_{\phi'} C)$ for some suitable derived ϕ' or vice versa?

As the following proposition tells us, the answer to (i) is yes in both directions (though we don't have a roundtrip since many ϕ' choices result in the same ϕ) but the answer to (ii) is yes in only one direction (going from ϕ' to ϕ).

Prop 5.1:

- (i) Given $G = (A \rtimes_\phi B)\overline{\times}C$, there exists ϕ' derived from ϕ s.t. $G = A \rtimes_{\phi'} (B\overline{\times}C)$.
- (ii) Given $G = A \rtimes_{\phi'} (B\overline{\times}C)$, there exists ϕ derived from ϕ' s.t. $G = (A \rtimes_\phi B)\overline{\times}C$.
- (iii) Given $G = A \rtimes_{\phi'} (B\overline{\times}C)$, there exists ϕ derived from ϕ' s.t. $G = (A \rtimes_\phi C)\overline{\times}B$.
- (iv) Given $(A\overline{\times}B) \rtimes_\phi C$, there *need not* exist ϕ' derived from ϕ s.t. $A\overline{\times}(B \rtimes_{\phi'} C)$.
- (v) Given $(A\overline{\times}B) \rtimes_\phi C$, there *need not* exist ϕ' derived from ϕ s.t. $B\overline{\times}(A \rtimes_{\phi'} C)$.
- (vi) Given $A\overline{\times}(B \rtimes_{\phi'} C)$, there exists ϕ derived from ϕ' s.t. $(A\overline{\times}B) \rtimes_\phi C$.

Warning: look very carefully at the subscripts in the discussion and proofs below. The subscripts 'c' and 'e' appear identical from afar, but the distinction between them is critical to our arguments.

Note that the mappings (i) and (ii) (or (i) and (iii)) aren't inverses of one another. In (i), we go $\phi'_{(b,c)} \equiv \phi_b$ and in (ii), we go $\phi_b \equiv \phi'_{(b,e)}$. A roundtrip in one direction does take us $\phi_b \to \phi'_{(b,c)} \to \phi_b$, but a roundtrip in the other takes us $\phi'_{(b,c)} \to \phi_b \to \phi''_{(b,c)}$, where ϕ'' loses all information from ϕ' except for the $\phi'_{(b,e)}$ piece. I.e., $\phi''_{(b,c)} = \phi''_{(b,e)}$ for all c. Put another way, the map from ϕ' to ϕ is highly noninjective, taking may ϕ' choices to the same ϕ. The same considerations hold for (i) and (iii).

Pf: (i) Let $G = (A \rtimes_\phi B)\overline{\times}C$. Our multiplication is $(a,b,c)(a',b',c') = (a\phi_b(a'), bb', cc')$. If $G = A \rtimes_{\phi'} (B\overline{\times}C)$ for some ϕ', then the multiplication must look like $(a,b,c)(a',b',c') = (a\phi'_{(b,c)}(a'), bb', cc')$ Therefore, we need $a\phi'_{(b,c)}(a') = a\phi_b(a')$, so $\phi'_{(b,c)}(a') = \phi_b(a')$. This is easy enough. $\phi' : B\overline{\times}C \to Aut(A)$ can be obtained from $\phi : B \to Aut(A)$ via $\phi'_{(b,c)}(a') \equiv \phi_b(a')$, independent of c. Clearly, $\phi'_{(e,e)} = \phi_e = Id_A$. $\phi'_{(b,c)(b',c')} = \phi'_{(bb',cc')} = \phi_{bb'} = \phi_b \circ \phi_{b'} = \phi'_{(b,c)}\phi'_{(b',c')}$, where in the last step, we could use any (b,c_1) and (b,c_2), including $c_1 = c$ and $c_2 = c'$, so we're fine. Finally, $\phi'_{(b,c)^{-1}} = \phi'_{(b^{-1},c^{-1})} = \phi_{b^{-1}} = (\phi_b)^{-1} = (\phi'_{(b,c)})^{-1}$, again because we can substitute any c (including our original one) in the last step.

Pf: (ii) Let $G = A \rtimes_{\phi'} (B\overline{\times}C)$. We want a suitable ϕ. The equations are the same as in (i), so we need $\phi'_{(b,c)}(a') = \phi_b(a')$. This is easy enough. $\phi : B\overline{\times}B \to Aut(A)$ can be obtained from $\phi' : B\overline{\times}C \to Aut(A)$ via $\phi_b(a') \equiv \phi'_{(b,e)}(a')$. Put another way, the restriction of the homomorphism ϕ' from $B\overline{\times}C$ to the subgroup $B\overline{\times}\{e\}$ keeps it a homomorphism. We therefore just have a restriction of the type we'll discuss in section 5.2.

Pf: (iii) This immediately follows from (ii). Since $B\overline{\times}C = C\overline{\times}B$, the roles of B and C are symmetric on the left. Specifically, we have $\phi_c \equiv \phi'_{(e,c)}$, and the proof carries through for the exact same reasons.

Pf: (iv,v) Let $G = (A\overline{\times}B) \rtimes_\phi C$. Then $(a,b,c)(a',b',c') = ((a,b)\phi_c((a',b')), cc')$. Since $\phi : C \to Aut(A\overline{\times}B) = Aut(A)\overline{\times}Aut(B)$, write $\phi_c(a',b') \equiv (\phi^1_c(a'), \phi^2_c(b'))$, where $\phi^1_c \in Aut(A)$ and $\phi^2_c \in Aut(B)$. It is easy to see that since ϕ is a homomorphism to a product space, ϕ^1 and ϕ^2 are homomorphisms as well. We thus have $(a,b,c)(a',b',c') = (a\phi^1_c(a'), b\phi^2_c(b'), cc')$. In order to have $G = A\overline{\times}(B \rtimes_{\phi'} C)$, we need $\phi' : C \to Aut(B)$. We then have $(a,b,c)(a',b',c') = (aa', b\phi'_c(b'), cc')$. The obvious candidate is $\phi'_c \equiv \phi^2_c$, which we already know is a homomorphism. However, the two forms are only equal if $\phi^1_c = Id_A$. This need not be the case in general. If $\phi^1_c \neq Id_A$ for some c, then we cannot replicate the multiplication using ϕ'. By symmetry, the same holds for (v).

Pf: (vi) Let $G = A\overline{\times}(B \rtimes_{\phi'} C)$. The equations are the same as for (iv) and (v), but now we start with $\phi' : C \to Aut(B)$ and want $\phi : C \to Aut(A\overline{\times}B) = Aut(A)\overline{\times}Aut(B)$. In this direction, we *do* have enough control to guarantee a solution. Just pick $\phi'_c(a,b) = (a, \phi_c(b))$. I.e., pick (Id_A, ϕ_c) as our ϕ'_c. As before, it is easy to see that this is a homomorphism.

5.2. **Restriction of Semidirect Product.**

Suppose $G = H \rtimes K$. There are two possible ways we may wish to restrict this: to a subgroup of H or a subgroup of K. The latter is always possible, but the former requires a bit of care.

5.2.1. *Restriction of $H \rtimes K$ to subgroup of K.*

Let $K' \subset K$ be a subgroup of K. Restricting $H \rtimes K$ to $H \rtimes K'$ is unambiguous and simple.

In the internal view, $G = HK$, so let $G' = HK'$. It is easy to see that G' is a subgroup of G, and that H is normal in it.

In general, a product of subgroups need not be a subgroup. We know that $G = HK$ is a group (by definition, since we started with a semidirect product), but this doesn't guarantee that HK' is a subgroup of it. I.e., we must show that it is. Suppose that $G' \equiv HK'$ is indeed a subgroup of G. Since $e \in K'$ (since K' is a subgroup of K and thus of G), $H = He \subset HK'$ is a subgroup of HK'. We thus have $H \subset G' \subset G$, with $H \triangleleft G$. Although normality isn't transitive, if $A \subset B \subset C$ and $A \triangleleft C$ then $A \triangleleft B$ (since if $cAc^{-1} \in A$ holds for all of $c \in C$, it also holds for all $b \in B \subset C$). Therefore, we have that H is normal in G'. This is *if* G' is a subgroup of G. Let's now show that it is. Any $g'_1, g'_2 \in G'$ can be written $h_1k'_1$ and $h_2k'_2$ for some h_1, h_2, k'_1, and k'_2. $g'_1g'_2 = h_1k'_1h_2k'_2$. Since H is normal in G, and all these elements are in G, $k'_1h_2 = h'_2k'_1$ for some $h'_2 \in H$. We thus have $h_1h'_2k'_1k'_2$. Since H and K' are both subgroups of G, $h_1h'_2 \in H$ and $k'_1k'_2 \in K'$, so the result is of the form hk' and is in G'. Thus G' is closed under multiplication. It contains e, since $e \in H$ and $e \in K'$. Consider the inverse: $(hk')^{-1} = k'^{-1}h^{-1}$. Since H is normal in G and $k'^{-1} \in G$, $k'^{-1}h^{-1} = h'k'^{-1}$ for some $h' \in H$, and this has the necessary form to be in G'. I.e., G' is closed under the inverse. It therefore is a subgroup of G and, as we saw, this means that H is normal in it. Since $G' = HK'$, we have a semidirect product. Therefore, $H \rtimes K$ restricts to $H \rtimes K'$ in the obvious way, resulting in a subgroup of G.

In the external view, the semidirect product is embodied in the homomorphism $\phi : K \to Aut(H)$. The restriction of a homomorphism to a subgroup is a homomorphism, so $\phi|_{K'} : K' \to Aut(H)$ defines a new semidirect product which (as we just saw) is isomorphic to a subgroup of the original one.

5.2.2. *Restriction of $H \rtimes K$ to subgroup of H.*

Now, suppose we wish to restrict $G = H \rtimes K$ to a subgroup $H' \subset H$.

Let's start with the external view. In this case, $\phi : K \to Aut(H)$. However, this doesn't restrict to a homomorphism $\phi' : K \to Aut(H')$ unless each $\phi(k)$ happens to be an automorphism which preserves H'.

If H' is characteristic in H, then (by definition) every element of $Aut(H)$ restricts to an automorphism of H'. This is a sufficient but not necessary condition. We don't actually need all of $Aut(H)$ to restrict to automorphisms of H'. We only require it of the elements of Im ϕ, which is (in general) a small subset of $Aut(H)$.

An example would be if H' is merely normal in H but Im ϕ consists solely of inner automorphisms.

The internal view features an analogous obstruction. In this case, the relevant subset is $G' = H'K$. To be a semidirect product, we need H' to be a normal subgroup of G' and we need G' to be a subgroup of G. Normality is not transitive, so the fact that $H' \triangleleft H \triangleleft G$, doesn't mean that H' is normal in G (which would then imply, since $H' \subset G' \subset G$, that if G' is a subgroup of G then H' is normal in G'). If H' is characteristic in H, this would follow — just as we saw in the external view case — but that is not a necessary condition, just a sufficient one. It is easy to see that if H' is normal in G', then G' is a subgroup of G.

This may sound a bit weird, since normality only has meaning in the context of a subgroup. What we mean is that if $H'g' = g'H'$ for every $g' \in G'$, then G' is a subgroup of G and, by definition, H' is normal in G'. We see this as follows. $e \in H'$ and $e \in K$ since both are subgroups of G, so $e \in G'$. Let $g_1 = h'_1k_1$ and $g_2 = h'_2k_2$. Then $g_1g_2 = h'_1k_1h'_2k_2$. If $H'g' = g'H'$ for every $g' \in G'$, then, since $K = eK \subset G'$, we have $k_1h'_2 = h''_2k_1$ for some $h''_2 \in H'$, giving us the desired form $(h'_1h''_2)(k_1k_2)$, so G' is closed under multiplication. $(h'k)^{-1} = k'^{-1}h'$. Again, if $H'g' = g'H'$ then $k'^{-1}h' = h''k^{-1}$ for some $h'' \in H'$, and we have the desired form, so G' is closed under the inverse.

I.e., if $H'g' = g'H'$ for all $g' \in G'$ then G' is a subgroup of G and (by definition) $H' \triangleleft G'$, but the converse need not hold. It is possible for G' to be a subgroup of G, but for H' not to be normal in G'. However, if H' is characteristic in H, then G' is a subgroup of G *and* H' is normal in G', and we have a semidirect

product. Again, this is a sufficient but not necessary condition. All we need is that G' be a subgroup of G and H' be normal in it.

The gist is that restricting K to K' always gives us a restricted semidirect product that behaves exactly as expected, but restricting H to H' only does so under certain circumstances. These include, but are not restricted to, H' being characteristic in H.

5.3. **Extension of Discrete Group by Connected Lie Group.**

So far, our primary use case has revolved around decomposing an existing Lie group G into an identity component G_0 and the discrete quotient group $\pi_0(G) = G/G_0$. However, we can also try to go the other way.

Suppose we have a connected Lie Group, which we'll guilelessly call G_0, and a discrete group, which we'll call D. For simplicity, we'll assume D is countable.

A few questions immediately present themselves:

- (i) Does there always exist a group extension of D by G_0 into a Lie group G (which then, by construction, would have $\pi_0(G) \approx D$)?
- (ii) If there does exist such an extension, is it unique? I.e., can there be more than one way to twist G_0 while going around D that is consistent with both group structures?
- (iii) Are such group extensions always semi-direct products?
- (iv) If not, is there always at least one extension which *is* a semidirect product?
- (v) If the answer to (iv) is yes, can there be more than one group extension which is a semidirect product? I.e., does being a semidirect product limit us to only one group extension?

The answer to (i) is yes. It is a basic result of algebra that a group extension always exists. For example, we can always form a direct product. The answer to (ii) is no. There can be more than one extension. Otherwise, we would *only* be able to have direct products. The answer to (iii) is no, as we discussed earlier. We need not have a semidirect product, although we mentioned a large class of cases (including any Lie group we're likely to care about) for which we do have one. The answer to (iv) is yes. As just mentioned, we can always form a direct product. The answer to (v) is yes. As long as there are any nontrivial homomorphisms from D to $Aut(G_0)$, we can form nontrivial semidirect products in addition to the direct product. In fact, we saw some nontrivial semidirect products earlier (ex. $O(n)$ for even n, $O(n, m)$, and the Poincare group). The direct product always exists — and this is a type of semidirect product — so the semidirect product only is unique if there are no nontrivial homomorphisms from D to $Aut(G_0)$). Typically, $Aut(G_0)$ is quite large, so this is not likely to happen unless G_0 is very small and D is not.

Bear in mind that different extensions yield different groups G.

Note that the existence of a group extension that is a semidirect product does *not* guarantee that the resulting G will be useful to us. For example, suppose we are working with the defining representation of a matrix group — such as $O(3)$ acting on $\mathbb{R}^3$. Even if G_0 and D have representations on V, it does not follow that G does. For example, if we start with $SO(3)$ and some discrete group D, we can construct

a direct product, perhaps various semidirect products, and various group extensions. Each produces a G in which $SO(3)$ is normal and for which D is the quotient group. However, this G may have no faithful homomorphism to $GL(\mathbb{R}, 3)$. I.e., it may have no faithful representation on $\mathbb{R}^3$. The same holds for actions on manifolds. There may be no faithful homomorphism to a given $Diff(M)$.

References

[1] https://kmhalpern.com/2023/07/08/semidirect-products-split-exact-sequences-and-all-that/.
[2] https://math.stackexchange.com/questions/338027/product-of-connected-spaces.
[3] https://math.stackexchange.com/questions/2600880/confusion-about-free-homotopy-based-homotopy-and-homotopy-groups.
[4] Allen Hatcher. *Algebraic Topology*. Cambridge University Press, 2002.
[5] https://math.stackexchange.com/questions/2248751/what-is-the-boundary-map-in-the-long-exact-sequence-in-homotopy-groups-especia.
[6] Morris W. Hirsch. *Differential Topology*. Springer-Verlag, 1976.
[7] https://mathoverflow.net/questions/8789/can-every-manifold-be-given-an-analytic-structure.
[8] https://terrytao.wordpress.com/2011/06/17/hilberts-fifth-problem-and-gleason-metrics/.
[9] https://math.stackexchange.com/questions/682290/is-gln-mathbbc-algebraic-or-not.
[10] https://en.wikipedia.org/wiki/Solenoid_(mathematics).
[11] https://en.wikipedia.org/wiki/Automorphism.
[12] https://en.wikipedia.org/wiki/Skolem_Noether_theorem.
[13] Jr. Alexander Kirillov. *An Introduction to Lie Groups nad Lie Algebras*. Cambridge University Press, 2008.
[14] John M. Lee. *Introduction to Smooth Manifolds*. Springer, 2nd edition, 2013.
[15] https://math.stackexchange.com/questions/2175501/real-semisimple-lie-group-with-infinitely-many-connected-components.
[16] https://math.stackexchange.com/questions/941066/lie-group-decomposition-as-semidirect-product-of-connected-and-discrete-groups.
[17] https://math.stackexchange.com/questions/692951/does-the-quotient-of-an-algebraic-group-by-its-neutral-component-always-split.
[18] https://virtualmath1.stanford.edu/~conrad/diffgeomPage/handouts/gramconnd.pdf.
[19] https://math.stackexchange.com/questions/299626/the-center-of-operatornamegln-k.
[20] https://math.stackexchange.com/questions/835829/what-can-be-said-about-a-matrix-which-is-both-symmetric-and-orthogonal.
[21] https://math.stackexchange.com/questions/554957/center-of-the-orthogonal-group-and-special-orthogonal-group.
[22] Arthur Jaffe. *Lorentz Transformations, Rotations, and Boosts*. 2013.
[23] https://www.cis.upenn.edu/~cis6100/cis610-15-sl4.pdffordetails.
[24] Shlomo Sternberg. On charge conjugation. *Communications of Mathematical Physics*, 109:649–679, 1987.
[25] https://math.stackexchange.com/questions/1217712/what-is-the-center-of-sop-q.
[26] https://mathoverflow.net/questions/32164/how-big-is-the-center-of-an-orthogonal-group.