

2. An informal look at Lie groups

2.1. The review is mainly about 19th century developments. The main figures are

Sophus Lie (1842–1899)

Wilhelm Killing (1847–1923)

Friedrich Engel (1861–1941)

Elie Cartan (1869–1951)

The origin of Lie theory is closely related to the view of geometry first adumbrated by *Felix Klein* (1849–1925): *A geometry is completely determined by its motions*. The motions form a group and so there is a direct relationship between groups and their actions on spaces which determine the geometry of that space. Figures that transform into one another by elements of the group are *congruent*.

Euclidean plane geometry: The group is the group of linear affine transformations of \mathbf{R}^2 consisting of translations followed by rotations and reflections. Congruence obviously plays a very central role in this geometry.

Spherical geometry: The space is S^2 , the sphere, and the group is $\mathfrak{o}(3)$, the group of rotations and reflections in 3-space, but restricted to their action on S^2 ,

Non Euclidean geometry: In the Poincaré model the space is the upper half plane \mathcal{H} and the group is $\mathrm{SL}(2, \mathbf{R})$, the group of real 2×2 matrices with determinant 1, actually the quotient of this group by $\{\pm 1\}$, the projective group $\mathrm{PGL}(2, \mathbf{R})$. The action is by Möbius transformations

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} : z \mapsto \frac{az + b}{cz + d}.$$

Sophus Lie, a Norwegian mathematician, had the idea that in order to use analysis to study groups and their actions one should restrict the groups and spaces so that this is possible. The groups were to be such that their points can be described by a finite set of coordinates (t_1, t_2, \dots, t_n) with the property that group multiplication is expressed by differentiable functions of class at least C^2 (this is a technical restriction he needed to

apply his ideas); the spaces also were to be such that their points could be described by a finite set (x_1, x_2, \dots, x_m) of coordinates, and again, the action of the group on the space should be described by differentiable functions of the group and space coordinates. The groups are what we call *Lie groups* nowadays, and the spaces *manifolds*, all of class at least C^2 . In general one cannot use the same set of coordinates to describe the entire group or the manifold, but if one uses different sets of coordinates in different regions, one must make sure that on overlapping regions, the two sets of coordinates are related by differentiable transformations of class at least C^2 . At the time of Lie's work these ideas of groups and manifolds were rather informal but it was clear in any example what was what. All calculations were performed using coordinates and so were *local*. We shall use the notation $g[x]$ for the effect of applying the group element g to the point x .

Lie's problem was very simply stated: *Classify all possible actions of pairs (G, X) where G is a Lie group and X is a manifold, with G acting on X .* Clearly this would give a description of all possible geometries from the Kleinian perspective. Although Lie was a contemporary of Klein, it is not very clear if he had been familiar with the Kleinian view when he formulated his problem and started a deep study of it. In any case very soon after he met Klein and one must assume that he knew where his program fitted into the larger scheme of geometry and group theory.

2.2. In order to explain Lie's basic and beautiful idea, let me consider first the case when the group is \mathbf{R} . We thus have a differentiable manifold X on which \mathbf{R} is acting, what is called an *one parameter group of differentiable maps*. If x is any point of X , the map

$$t \longmapsto t[x]$$

can be differentiated at $t = 0$ to get a tangent vector Z_x at x ; when x varies this gives a vector field Z . Z thus encodes the *infinitesimal action of the group \mathbf{R}* . One can recapture the local action of \mathbf{R} by integrating the vector field. If $x_0 \in X$ and x_1, \dots, x_m are local coordinates near x_0 vanishing (for convenience) at x_0 , then

$$Z = \sum_{1 \leq i \leq m} F_i \frac{\partial}{\partial x_i}$$

in these coordinates; the curves $t \mapsto t[x] = (x_1(t), \dots, x_m(t))$ become the solutions to the initial value problem of ODE's:

$$\frac{dx_i}{dt} = F_i(x_1, \dots, x_m), \quad x_i(0) = 0 \quad (1 \leq i \leq m).$$

One can find a neighborhood U of x_0 and $\varepsilon > 0$ such that the maps $t, x \mapsto t[x]$ are defined and smooth for $x \in U, |t| < \varepsilon$. If X is *compact* we can get a *global action* of \mathbf{R} .

Suppose now the we have a Lie group of dimension n acting on a manifold of dimension r (we shall come to formal definitions later). So in local coordinates we have a map

$$(t, x) \longrightarrow y$$

where

$$y_i = Y_i(t - 1, \dots, t_n, x_1, \dots, x_m).$$

We introduce the vector field

$$Z_\alpha = \sum_i Z_{i\alpha} \partial_i \quad (\partial_i = \partial/\partial x_i)$$

where

$$Z_{i\alpha} = \left. \frac{\partial Y_i(t, x)}{\partial t_\alpha} \right|_{t=0}.$$

We regard this vector field as the *infinitesimal motion* in the direction of t_α . The individual vector fields depend on the choice of the coordinate (t_β) on G near the identity (we assume that $t = 0$ represents the identity). But, since we do not disturb the x -coordinates, it is an easy consequence of the chain rule that if we had used a different set of coordinates (u_γ) near the identity of G , the Z_α change over to vector fields T_γ which are *constant linear combinations of the Z_α* . Hence the *it finite dimensional space*

$$\mathcal{L} := \sum_\alpha \mathbf{R}Z_\alpha$$

is determined independent of the coordinates we use. It is thus an *invariant of the action of G on X* . Lie discovered the following remarkable fact.

Theorem. \mathcal{L} is closed under the formation of the Lie bracket of vector fields. More precisely, if $A, B \in \mathcal{L}$, then $[A, B] \in \mathcal{L}$ also.

We shall not prove this at this time; we shall do it after we develop some machinery of Lie groups and their actions. Note however that the space \mathcal{L} is well defined for *any* smooth map $G \times X \rightarrow X$ where G need not even be a group. The fact that it is closed under brackets however needs that G is a group and the map represents an action. However to add some intuition to this let us look at a few examples where we can verify this result explicitly.

EXAMPLE 1: $G = X = \mathbf{R}^n$. The action is by translations $t, x \mapsto x + t$. The vector fields are ∂_i .

EXAMPLE 2: $G = \text{SL}(2, \mathbf{R})$, $X = \mathcal{H}$, the Poincaré upper half plane, the action by Möbius transformations

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} : z \mapsto \frac{az + b}{cz + d}.$$

Near the identity, $t = a - 1, b, c$ are coordinates with $d = (1 + bc)(1 + t)^{-1}$, $t \neq -1$. The vector fields corresponding to t, b, c are

$$T = 2z \frac{d}{dz}, \quad B = \frac{d}{dz}, \quad C = -z^2 \frac{d}{dz}.$$

We have

$$[T, B] = 2B, \quad [T, C] = -2C, \quad [B, C] = -T.$$

EXAMPLE 3: $G = \text{SO}(n)$, $X = \mathbf{R}^n$, action by linear transformations corresponding to the matrices. Here we shall look into whether this is a Lie group (according to our informal definition). For $n = 3$ there is a parametrization by *Euler angles*. A better way is to use the method of *Cayley*.

Lemma. *The map*

$$S \mapsto R = (I - S)(I + S)^{-1} = \frac{I - S}{I + S}$$

is a bijection from the vector space of real skew symmetric $n \times n$ matrices S to the open set of all rotations R such that $I + R$ is invertible, with inverse map

$$S = (I - R)(I + R)^{-1} = \frac{I - R}{I + R}.$$

Multiplication is given by rational functions of S .

Problems

1. Write $S = (s_{ij})$ with $s_{ij} = -s_{ji}$ and compute, for the left action of $G = \text{SO}(n)$ on itself, the vector fields X_{ij} corresponding to s_{ij} ($i < j$).
2. For $n = 3$ write multiplication in the Cayley coordinates explicitly.
3. For the $\text{SO}(n)$ action on \mathbf{R}^n show that the vector fields on \mathbf{R}^n corresponding to the Cayley coordinates s_{ij} are (writing $\partial_r = \partial/\partial x_r$)

$$Z_{ij} = x_j \partial_i - x_i \partial_j, \quad [Z_{ij}, Z_{rs}] = \delta_{ir} Z_{js} + \delta_{js} Z_{ir} + \delta_{jr} Z_{is} + \delta_{is} Z_{jr}.$$

4. Find the analogues of the Cayley coordinates for the groups $\text{U}(n)$ and $\text{SU}(n)$.

Given a Lie group we have to decide if it is connected, and if so, what is its fundamental group is, and construct the universal covering group. We shall only look at some examples at this stage.

$\text{SO}(n)$: We want to prove it is connected for all $n \geq 1$. For $n = 1$ the group consists only of the identity. For $n = 2$ it is the circle group S^1 which is obviously connected. For $n \geq 3$ there are two ways to do this. One is to use the relation $\text{S}(n+1)/\text{SO}(n) = S^n$ and use induction. The other is to observe that given any rotation one can choose an ON basis of \mathbf{R}^n in which it is a direct sum of a number of 2-dimensional rotations

$$\begin{pmatrix} \cos t_k & -\sin t_k \\ \sin t_k & \cos t_k \end{pmatrix} \quad (1 \leq k \leq r),$$

1-dimensional identities, and *an even number* of one dimensional transformations equal to -1 ; if there are $2s$ of these, we can write their direct sum as a direct sum of s rotations

$$\begin{pmatrix} \cos \pi & -\sin \pi \\ \sin \pi & \cos \pi \end{pmatrix}.$$

It is obvious that such a rotation can be connected to the identity by a continuous path by replacing the angle t_k by tt_k and π by $t\pi$ for $0 \leq t \leq 1$.

Double connectivity of $\text{SO}(3)$. One of the important properties for a connected Lie group is its simple connectivity. We shall see later that given

a connected Lie group G , we can find a connected and simply connected Lie group G^\sim with a homomorphism $f(G^\sim \rightarrow G)$ such that f is surjective and has a discrete kernel D . This will make G^\sim the *simply connected covering group* of G . The group D will then be the *fundamental group* $\pi_1(G)$ of G .

One can show that the fundamental groups of $\mathbf{MSO}(n)$ are given by

$$\pi_1(\mathrm{SO}(n)) = \begin{cases} \mathbf{Z}_2 & \text{for } n \geq 3 \\ \mathbf{Z} & \text{for } n = 2 \\ \{1\} & \text{for } n = 1. \end{cases}$$

Since $\mathrm{SO}(n)$ is $\{1\}$ for $n = 1$ and the circle group for $n = 2$, we have only to consider $n \geq 3$. Here we shall look at $n = 3$.

We want to prove that there is a two-fold covering homomorphism

$$\mathrm{SU}(2) \longrightarrow \mathrm{SO}(3)$$

which is surjective and has as its kernel the group $\{\pm 1\}$. This is an example of the *spin group covering of $\mathrm{SO}(n)$* . Let \mathcal{H}_0 be the real vector space of 2×2 hermitian matrices of trace 0 on which $G = \mathrm{SU}(2)$ acts by $g, X \mapsto gXg^{-1}$. If we write

$$X = \begin{pmatrix} x_3 & x_1 + ix_2 \\ x_1 - ix_2 & -x_3 \end{pmatrix}$$

then $\mathcal{H}_0 \simeq \mathbf{R}^3$, and the action preserves $-\det(X) = x_1^2 + x_2^2 + x_3^2$. If R_g is the action on $\mathbf{R}^3 \simeq \mathcal{H}_0$ by $g \in \mathrm{SU}(2)$, then R_g is in $\mathrm{O}(3)$ and we have a homomorphism $G \mapsto R_g$ of $\mathrm{SU}(2)$ into $\mathrm{O}(3)$. For $R_g = 1$ the condition is that g commutes with all of \mathcal{H}_0 ; as any 2×2 matrix is a linear combination of I and elements of \mathcal{H}_0 , the condition is that g commutes with all 2×2 matrices, so that $g \in \{\pm 1\}$. The map is into $\mathrm{SO}(3)$ because the image is a connected subgroup of $\mathrm{O}(3)$, hence $\subset \mathrm{SO}(3)$.

$\mathrm{SU}(2)$ is a Lie group. In fact it is easy to see that $\mathrm{SU}(2)$ is precisely the group of all matrices of the form

$$\begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \quad a\bar{a} + b\bar{b} = |a|^2 + |b|^2 = 1.$$

This shows that $\mathrm{SU}(2)$ can be identified with S^3 and hence a Lie group. Moreover this identification shows that it is simply connected. It is not

difficult to verify that it is simply connected. In fact one shows that the images of the 3 one-parameter subgroups

$$\begin{pmatrix} e^{it} & 0 \\ 0 & e^{-it} \end{pmatrix}, \quad \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix}, \quad \begin{pmatrix} \cos t & i \sin t \\ i \sin t & \cos t \end{pmatrix}$$

are the rotations around the x_3, x_2, x_1 -axes in \mathbf{R}^3 and one uses the fact that these rotations generate $\mathrm{SO}(3)$.

For $n > 3$ one has to use *Clifford algebras* to generalize this construction.

2.3. Topology of Lie groups. We shall find that the topology of a Lie group is a critical aspect of understanding their structure. For one thing, the smooth structure is already completely determined by the topology.

Theorem. *A continuous homomorphism between two Lie groups is automatically smooth. In particular the real Lie group structure on a topological group, if it exists, is unique.*

This is one of our main results. The result is not true for the *complex* Lie groups: \mathbf{C} has another complex structure, namely the conjugate one, compatible with its additive group structure and topology.

Theorem (Von Neumann). *Any closed subgroup of $\mathrm{GL}(n, \mathbf{R})$ is a Lie group. More precisely, if G is a closed subgroup of $\mathrm{GL}(n, \mathbf{R})$, then its connected component G^0 is open in G and G^0 is a Lie group.*

We shall also prove this later.

It follows from the first theorem that it makes sense to ask whether a given topological group G is a Lie group. Clearly it is necessary that G is *locally Euclidean*, namely it has an open neighborhood of the identity that is homeomorphic to a ball in a Euclidean space. Hilbert proposed the problem (fifth in his famous list) whether this is sufficient. Von Neumann proved this for compact groups and Chevalley for solvable groups. The affirmative answer was found in the 1950's by the efforts, notably of Gleason, Yamabe, Montgomery–Zippin, and Iwasawa. See the book [MZ] in the historical review for full details. The problem makes sense for p -adic Lie groups also where the question was settled by Lazard [Laz].

Problems

1. Fill in the details for the proof that $SU(2)$ is a two-fold cover for $SO(3)$.
2. Let G be a product of countably many copies of the circle group S^1 . Prove that G is not locally Euclidean and is hence not a Lie group. (*Hint.* show that any neighborhood of the identity contains subsets of arbitrarily large dimension.)
3. Show that \mathbf{Q}_p , the additive group of p -adic numbers, is not a Lie group.
4. Find an example of a topological group of dimension 1 which is not a Lie group.

2.4. Compact Lie groups. The most important Lie groups for physics and many other fields of mathematics are the *compact* Lie groups. It is remarkable that they can be given a very explicit description. Now finite covers of compact Lie groups are compact Lie groups, and so, for any compact Lie group G , any finite cover of $G \times T$ is also a compact Lie group, where T is a *torus*, i.e., a group which is a finite product of the circle groups. Mirroring the Cartan-Killing classification of simple Lie algebras is the theorem that any compact Lie group can be obtained using the above construction from the *compact simply connected groups* which belong to 4 families (the *classical compact Lie groups*) and 5 isolated ones (the *exceptional compact Lie groups*).

The classical compact Lie groups may be viewed as the isometry groups of the Euclidean spaces over \mathbf{R} , \mathbf{C} , and \mathbf{H} , the algebra of quaternions. The classical theorem of *Frobenius* that these three are precisely all the associative division algebras over \mathbf{R} is relevant here.

The groups $U(n), SU(n)$ as linear isometry groups of \mathbf{C}^n : We regard \mathbf{C}^n as a Euclidean (Hilbert) space with the hermitian scalarproduct

$$a \cdot b := (a, b) = \sum_{1 \leq r \leq n} a_r^* b_r.$$

Here and in what follows we use $*$ to denote complex conjugation and also adjoints. The group of linear isometries is $U(n)$. In quantum mechanics it is also important to consider anti-linear isometries such as complex conjugation in \mathbf{C}^n . Both $U(n)$ and $SU(n)$, the subgroup of $U(n)$ of unitary matrices of determinant 1, are connected, while $SU(n)$ is simply connected.

Quaternions. If k is a field of characteristic $\neq 2$, a *quaternion algebra over k* is an algebra of the form $k[a, b]$ where $a^2 = \alpha, b^2 = \beta, ab = -ba$. Here α, β are two elements of k^\times which are not squares. If $k = \mathbf{R}, \alpha = \beta = -1$ we obtain the standard quaternion algebra over \mathbf{R} first discovered by **W. R. Hamilton** when he solved the problem of modeling rotations in space. It is an important arithmetic and algebraic problem to classify quaternion algebras up to isomorphism over an arbitrary field.

We write \mathbf{H} for the algebra of quaternions of elements

$$q = a_0 + a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}, \quad \mathbf{k} = \mathbf{ij} = -\mathbf{ji}$$

where the a_r are real numbers. \mathbf{H} is an associative non commutative algebra. For any quaternion we write q^* for its *conjugate*:

$$q^* = a_0 - a_1\mathbf{i} - a_2\mathbf{j} - a_3\mathbf{k}$$

and it is easy to verify that

$$qq^* = q^*q = (a_0^2 + a_1^2 + a_2^2 + a_3^2) =: N(q).$$

In particular \mathbf{H} is a *division algebra*, the inverse of a nonzero q being given by

$$q^{-1} = N(q)^{-1}q^*.$$

We also have

$$(q + q')^* = q^* + q'^*, \quad (qq')^* = q'^*q^*, \quad N(qq') = N(q)N(q').$$

There is an identification of \mathbf{H} with \mathbf{C}^2 via

$$q = z_1 + \mathbf{j}z_2, \quad z_1 = a_0 + a_1i, z_2 = a_2 - ia_3.$$

Let us identify \mathbf{C} as a subalgebra of \mathbf{H} with i identified with \mathbf{i} and use 1 and \mathbf{j} as a basis for \mathbf{H} as a right vector space. Then left multiplication ℓ_q by the quaternion q becomes the matrix

$$\begin{pmatrix} a_0 + a_1i & -a_2 - a_3i \\ a_2 - a_3i & a_0 - a_1i \end{pmatrix}$$

so that the group of *unit quaternions*, namely, the group of quaternions q with $N(q) = 1$, which we denote by $\text{Sp}(1)$, becomes isomorphic to $\text{SU}(2)$.

The *imaginary quaternions* form a 3-dimensional space with basis $\mathbf{i}, \mathbf{j}, \mathbf{k}$, and is precisely the space of quaternions of *trace* 0; here

$$\mathrm{tr}(q) = q + q^* = 2a_0, \quad \mathrm{tr}(qq'q^{-1}) = \mathrm{tr}(q').$$

The space of imaginary quaternions is stable under the action $q' \mapsto qq'q^{-1}$ and gives the geometric interpretation of rotations in 3-space that Hamilton was after. The action of $\mathrm{SU}(2)$ on \mathbf{R}^3 that we obtain is nothing but the map $\mathrm{SU}(2) \rightarrow \mathrm{SO}(3)$ described earlier.

$\mathrm{Sp}(n)$ as the linear isometry group of \mathbf{H}^n : We regard \mathbf{H}^n , the space of column vectors with quaternion components, as a *right vector space* over the division algebra \mathbf{H} and its endomorphisms as matrices of quaternions acting on quaternion column vectors by multiplication from the left. We introduce the scalar product

$$\mathbf{a} \cdot \mathbf{b} := \sum_{1 \leq r \leq n} a_r^* b_r$$

where \mathbf{a}, \mathbf{b} are column vectors with components $a_r, b_r \in \mathbf{H}$. We write

$$\|a\|^2 = \sum_{1 \leq r \leq n} a_r^* a_r.$$

The dot product is additive in each variable and satisfies the non commutative hermitian properties:

$$\mathbf{a} \cdot \mathbf{b}q = (\mathbf{a} \cdot \mathbf{b})q, \quad (\mathbf{a}q) \cdot \mathbf{b}q = q^*(\mathbf{a} \cdot \mathbf{b}).$$

We denote by $\mathrm{Sp}(n)$ the group of endomorphisms of \mathbf{H}^n which preserve $\|\cdot\|$:

$$\sigma \in \mathrm{Sp}(n) \iff (\sigma \mathbf{a}, \sigma \mathbf{b}) = (\mathbf{a}, \mathbf{b}).$$

This is the *compact symplectic group*. The reason for the name and the notation will become clear presently.

We identify \mathbf{H}^n with \mathbf{C}^{2n} via the map

$$bfa = (a_r) \iff \mathbf{a}' = (z_1, \dots, z_n, z_{n+1}, \dots, z_{2n})$$

where we write

$$a_r = z_r + \mathbf{j}z_{n+r}, \quad (1 \leq r \leq n).$$

Then any bijection σ corresponds to a bijection σ' of \mathbf{C}^{2n} . If σ is \mathbf{H} -linear, then σ' is \mathbf{C} -linear. So the map $\sigma \mapsto \sigma'$ gives an imbedding of $\mathrm{Sp}(n)$ into $\mathrm{GL}(2n, \mathbf{C})$:

$$\mathrm{Sp}(n) \hookrightarrow \mathrm{GL}(2n, \mathbf{C}).$$

It is easy to determine the image of $\mathrm{Sp}(n)$ under this imbedding. First we have, with $b_r = t_r + \mathbf{j}t_{n+r}$,

$$\mathbf{a} \cdot \mathbf{b} = \sum_{1 \leq r \leq n} z_r^* t_r + \mathbf{j} \sum_{1 \leq r \leq n} (z_r t_{n+r} - z_{n+r} t_r) = (\mathbf{a}, \mathbf{b}') + \mathbf{j}J(\mathbf{a}', \mathbf{b}')$$

where J is the *symplectic bilinear form* on \mathbf{C}^{2n} defined by the skew-symmetric non-singular matrix

$$J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}.$$

Thus the image of $\mathrm{Sp}(n)$ is contained in

$$\mathrm{USp}(2n) := \mathrm{U}(2n) \cap \mathrm{Sp}(2n, \mathbf{C})$$

where as usual we write $\mathrm{Sp}(2n, \mathbf{C})$ for the subgroup of $\mathrm{GL}(2n, \mathbf{C})$ preserving the bilinear form J or, equivalently, satisfying

$$g^T J g = J.$$

The image is exactly this group. For, if σ corresponds to a $\sigma' \in \mathrm{USp}(2n)$, then σ is \mathbf{H} -linear and so belongs to $\mathrm{Sp}(n)$, as an easy argument shows.

The classical simple compact Lie groups are

$$\mathrm{SU}(n), \quad \mathrm{SO}(n), \quad \mathrm{Sp}(n).$$

Using the operations of products with tori and finite covers and passing to quotients by finite central subgroups and using these as well as 5 isolated ones, one can obtain *all* compact connected Lie groups.

For connectivity and simple connectivity one uses two results formulated in the following.

Lemma. *If G is a topological group and $H \subset G$ a closed subgroup, then G is connected if H and the homogeneous space G/H are connected. Moreover, if further G/H is simply connected, then there is a natural map $\pi_1(H) \rightarrow \pi_1(G)$ which is surjective.*

Problems

1. Find a Cayley parametrization of $\mathrm{Sp}(n)$ in the quaternionic context as well as in the context of \mathbf{C}^{2n} .
2. Prove the lemma (see Chevalley[Ch1]).
3. Use problem 2 to show that $\mathrm{SU}(n)$ and $\mathrm{Sp}(n)$ are connected and simply connected.
4. Complete the proof that the image of $\mathrm{Sp}(n)$ under the map described in the text is all of $\mathrm{USp}(2n)$. (*Hint*: Show that $(\sigma\mathbf{c})\cdot(\sigma(\mathbf{a}q) - \sigma(\mathbf{a})q) = 0$.)