Linear Poisson Geometry
via Matrix Groups and Matrix Algebras

Written by Santiago Cañez

Contents

Introduction

Linear Poisson geometry is the study of Poisson brackets which arise from Lie algebras, and provides a setting where group theory, linear algebra, and geometry blend together nicely. The goal of these notes is to give a crash-course introduction to this area and to the tools it requires. We will assume familiarity only with basic group theory, abstract linear algebra (including abstract vector spaces and linear transformations), and multivariable calculus. We will quite often be somewhat informal and give definitions and results appropriate to the level of the prerequisites listed above, so we will not phrase things in the most general "Lie-theoretic" and "differential-geometric" way possible. We want to simply get to a point where we can do some real computations to get a feel for how everything fits together, and because of this we will focus exclusively on matrix Lie groups.

Here is the basic setup. To a group of matrices G with real entries, we can associate a certain vector space $\mathfrak{g}$ called its *Lie algebra*. This Lie algebra comes equipped with a certain operation known as the *Lie bracket*, which takes two matrices A, B as input and outputs a third matrix $[A, B]$. The Lie algebra of G encodes much "infinitesimal" information about G, from which "global" information about G can often be derived. To the Lie algebra $\mathfrak{g}$ we can further associate a space known as its *dual space* $\mathfrak{g}^*$, on which we can then consider the set $C^\infty(\mathfrak{g}^*)$ of real-valued smooth functions $\mathfrak{g}^* \to \mathbb{R}$. (To be *smooth* simply means that all partial derivatives exist to all orders.) On this set of functions one can construct what's called a *Poisson bracket*, which takes two functions f, g as inputs and outputs a function $\{f, g\}$. This Poisson bracket is constructed from the Lie bracket on $\mathfrak{g}$, and in turn the Lie bracket can be recovered from the Poisson bracket.

We thus have the following correspondence:

$$\underset{\text{group theory}}{G} \rightsquigarrow \underset{\text{Lie theory}}{\mathfrak{g}} \rightsquigarrow \underset{\text{Poisson geometry}}{C^\infty(\mathfrak{g}^*)}$$

The overarching goal in this subject is to understand group-theoretic properties of G from the linear-algebraic properties of $\mathfrak{g}$ and from the Poisson-geometric properties of $C^\infty(\mathfrak{g}^*)$. The main result along these lines is that the orbits of the *coadjoint action* of G are precisely the *symplectic leaves* of the Poisson bracket on $\mathfrak{g}^*$. We aim to build up enough material to understand what this statement means, and to build up enough examples to get a feel for its validity.

Matrix Groups

We denote by $GL_n(\mathbb{R})$ the set of invertible $n \times n$ matrices with real entries, which is a group under matrix multiplication. The "G" stands for "general", so $GL_n(\mathbb{R})$ is called the *general linear* group. We single out the following subgroups of certain $GL_n(\mathbb{R})$ as the main ones of interest:

- $SO_3(\mathbb{R}) := \{A \in GL_3(\mathbb{R}) \mid A \text{ is orthogonal of determinant } 1\}$ is the group of 3-dimensional rotation matrices;
- $SL_2(\mathbb{R}) := \{A \in GL_2(\mathbb{R}) \mid \det A = 1\}$ is the group of 2-dimensional orientation- and area-preserving matrices; and
- $H := \left\{ \left[\begin{smallmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{smallmatrix}\right] \;\middle|\; a, b, c \in \mathbb{R} \right\}$ is what's called the *Heisenberg group* and is a subgroup of $GL_3(\mathbb{R})$.

These are all examples of *Lie groups*, which roughly means they are groups on which we can do calculus. If nothing else, note that matrix multiplication of $n \times n$ matrices, viewed as a function $\mathbb{R}^{n^2} \times \mathbb{R}^{n^2} \to \mathbb{R}^{n^2}$ where we view a matrix as a long column vector, is infinitely differentiable, which is part of what goes into the formal definition of "Lie group". The letter "S" in the notation used for the first two examples stands for "special", and in general refers to the determinant 1 condition; $SO_3(\mathbb{R})$ is thus called the 3-dimensional *special orthogonal* group, and $SL_2(\mathbb{R})$ the 2-dimensional *special linear* group.

We will want to consider matrices with functions as entries, such as

$$A(t) = \begin{bmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{bmatrix}.$$

We can view this as defining a function $\mathbb{R} \to \{\text{matrices}\}$, which we take to be a *path* in our space of matrices. For example, the 2×2 $A(t)$ above defines a path in the group $SO_2(\mathbb{R})$ of 2-dimensional rotation matrices. The derivative $A'(t)$ is then the usual derivative of this function, and amounts to simply computing the derivative of each entry of $A(t)$; for example, with $A(t)$ defined as above we have

$$A'(t) = \begin{bmatrix} -\sin(t) & -\cos(t) \\ \cos(t) & -\sin(t) \end{bmatrix}.$$

This matrix derivative operation satisfies the usual types of properties you would expect, such as

$$(A(t) + B(t))' = A'(t) + B'(t), \ (cA(t))' = cA'(t),$$

and the fact that the derivative of a "constant" matrix (one which is independent of t) is the zero matrix. It also satisfies the product rule

$$(A(t)B(t))' = A'(t)B(t) + A(t)B'(t).$$

Exercise. Justify the product rule above.

Exercise. Justify the fact that the derivative of the transpose of $A(t)$ is the transpose of $A'(t)$.

Exercise. Justify the fact that if $A(t)$ is invertible, then

$$\frac{d}{dt}(A(t)^{-1}) = -A(t)^{-1}A'(t)A(t)^{-1}.$$

Hint: $A(t)A(t)^{-1} = I$ for all t. (Compare this to $(1/f)' = -(1/f^2)f'$ for scalar-valued functions.)

Lie Algebras

A *Lie algebra* is a vector space $\mathfrak{g}$ equipped with an operation $[,] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ called the *Lie bracket* which satisfies

- $[c_1u_1 + c_2u_2, v] = c_1[u_1, v] + c_2[u_2, v]$ and $[u, c_1v_1 + c_2v_2] = c_1[u, v_1] + c_2[u, v_2]$ for all $c_1, c_2 \in \mathbb{R}$ and $u_1, u_2, u, v_1, v_2, v \in \mathfrak{g}$, which says that the bracket is *linear* in each entry;
- $[u, v] = -[v, u]$ for all $u, v \in \mathbb{R}$, which says that the bracket is *skew-symmetric*; and
- $[u, [v, w]] + [v, [w, u]] + [w, [u, v]] = 0$ for all $u, v, w \in \mathfrak{g}$, which is called the *Jacobi identity.*

The Jacobi identity is the most mysterious of these properties, but is also the most important one, as we'll see. Standard examples of Lie algebras are:

- any vector space V with bracket defined to always be zero (not so interesting),
- the space $\mathfrak{gl}_n(\mathbb{R})$ of $n \times n$ matrices with bracket defined as $[A, B] = AB - BA$,
- $\mathbb{R}^3$ with the Lie bracket defined as the cross product: $[\mathbf{u}, \mathbf{v}] = \mathbf{u} \times \mathbf{v}$.

Exercise. Verify that $\mathfrak{gl}_n(\mathbb{R})$ and $\mathbb{R}^3$ are Lie algebras with the brackets defined above.

Certain Lie algebras arise as subalgebras of other Lie algebras. A *(Lie) subalgebra* of a Lie algebra $\mathfrak{g}$ is a vector subspace $\mathfrak{h}$ of $\mathfrak{g}$ which is closed under the Lie bracket, meaning $[x, y] \in \mathfrak{h}$ for all $x, y \in \mathfrak{h}$. Here are some examples:

- the space $\mathfrak{so}_n(\mathbb{R}) := \{A \in \mathfrak{gl}_n(\mathbb{R}) \mid A^T = -A\}$ of $n \times n$ skew-symmetric matrices with real entries is a subalgebra of $\mathfrak{gl}_n(\mathbb{R})$,
- the space $\mathfrak{sl}_n(\mathbb{R}) := \{A \in \mathfrak{gl}_n(\mathbb{R}) \mid \operatorname{tr} A = 0\}$ of $n \times n$ matrices of trace zero with real entries is a also subalgebra of $\mathfrak{gl}_n(\mathbb{R})$.

Exercise. Verify that each example above is indeed a subalgebra of the given Lie algebra. (You will need to use some properties of transposes and traces.)

The reason for the notations $\mathfrak{gl}, \mathfrak{so}$, and $\mathfrak{sl}$ in the examples above comes from their relation to the groups described previously. In general, to any matrix (Lie) group G we can associate a Lie algebra in the following way. Consider a path $A(t)$ of matrices in G which begins at the identity, meaning that $A(0) = I$. The *Lie algebra* $\operatorname{Lie}(G)$ of G by definition consists of the derivatives of all such paths at $t = 0$:

$$\operatorname{Lie}(G) := \{A'(0) \mid A(t) \text{ is a path in } G \text{ such that } A(0) = I\}.$$

(In multivariable calculus, derivatives of paths give tangent vectors to paths, so in a certain sense elements of $\operatorname{Lie}(G)$ are "tangent vectors" to G at the identity, and altogether make up what we might call the "tangent space" to G at the identity.)

To turn $\operatorname{Lie}(G)$ into a Lie algebra, we must first turn it into a vector space, meaning we must give meaning to addition and scalar multiplication in $\operatorname{Lie}(G)$. If $A'(0)$ and $B'(0)$ are two elements of $\operatorname{Lie}(G)$, then by definition they are the derivatives at time 0 of two paths $A(t)$ and $B(t)$ in G which begin at the identity. The product $A(t)B(t)$ is then another path in G starting at $A(0)B(0) = I$, and its derivative at time 0 is

$$\frac{d}{dt}\Big|_{t=0} A(t)B(t) = A'(0)B(0) + A(0)B'(0) = A'(0)I + IB'(0) = A'(0) + B'(0)$$

by the product rule, so the sum $A'(0) + B'(0)$ is itself an element of $\operatorname{Lie}(G)$. Thus addition in $\operatorname{Lie}(G)$ makes sense. Moreover, with $A(t)$ as above, if c is any real scalar, then $A(ct)$ is a path in G starting at $A(c0) = A(0) = I$ and

$$\frac{d}{dt}\Big|_{t=0} A(ct) = cA'(c0) = cA'(0),$$

so $cA'(0) \in \text{Lie}(G)$ and scalar multiplication makes sense.

The final ingredient needed to turn $\text{Lie}(G)$ into a Lie algebra is the Lie bracket, which we take to be the one given by $[A, B] = AB - BA$. We will see later how this bracket indeed arises from G as well so that this is not a definition we make by choice, but we postpone this discussion for now. Note in particular that at this point it is not obvious that the space $\text{Lie}(G)$ so defined should be closed under this particular Lie bracket, but we will see why this is true in general soon enough.

The matrices

$$A(t) = \begin{bmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{bmatrix}$$

define a path in $SO_2(\mathbb{R})$ starting at $A(0) = I$, and the derivative of this path at time 0 is

$$\frac{d}{dt}\Big|_{t=0} A(t) = \begin{bmatrix} -\sin(0) & -\cos(0) \\ \cos(0) & -\sin(0) \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},$$

which is an element of $\mathfrak{so}_2(\mathbb{R})$ since it is skew-symmetric. This suggests that elements of $\mathfrak{so}_2(\mathbb{R})$ comes from differentiating paths in $SO_2(\mathbb{R})$, which should mean that $\mathfrak{so}_2(\mathbb{R})$ is the Lie algebra of $SO_2(\mathbb{R})$. Indeed, this is true for general n: $\text{Lie}(SO_n(\mathbb{R})) = \mathfrak{so}_n(\mathbb{R})$, and hence is the reason why we use the notation $\mathfrak{so}$ to refer to skew-symmetric matrices. To see this, take a path $A(t)$ in $SO_n(\mathbb{R})$ starting at $A(0) = I$. Then each $A(t)$ is an orthogonal matrix so

$$A(t)A(t)^T = I \text{ for all } t.$$

Differentiating both sides at $t = 0$—making use of the product rule and relation between derivatives and transposes on the left—gives

$$A'(0)A(0)^T + A(0)A'(0)^T = 0.$$

Since $A(0) = I$, this simplifies to $A'(0) + A'(0)^T = 0$, or $A'(0)^T = -A'(0)$, which means that $A'(0) \in \text{Lie}(SO_n(\mathbb{R}))$ is indeed skew-symmetric.

Technically this only shows that $\text{Lie}(SO_n(\mathbb{R}))$ is contained in $\mathfrak{so}_n(\mathbb{R})$, and to finish we would need to know that *all* skew-symmetric matrices arise from differentiating paths of orthogonal matrices in this way. This is true, but we leave the details to an exercise. (One way to show this is to use the notion of a *matrix exponential*, where for a given square matrix A, the exponential e^A is defined to be the infinite sum

$$e^A := I + A + \frac{1}{2}A^2 + \frac{1}{3!}A^3 + \cdots = \sum_{n=0}^{\infty} \frac{1}{n!}A^n.$$

Exponentials in general are used to go from $\text{Lie}(G)$ back to G, and hence go from the "infinitesimal" to the "global".) Note also that in this construction we only used the fact that matrices in $SO_n(\mathbb{R})$ are orthogonal, and not that they have determinant 1; in fact, the group $O_n(\mathbb{R})$ of orthogonal $n \times n$ matrices, allowing for the determinant to be -1 as well, also has $\mathfrak{so}_n(\mathbb{R})$ as its Lie algebra. Thus, different matrix groups can have the same Lie algebra, but this won't be a big concern for our purposes. (The difference in these examples is that $SO_n(\mathbb{R})$ is *connected* and $O_n(\mathbb{R})$ is *disconnected*, but this distinction won't play a role for us.)

For $SL_2(\mathbb{R})$, a path $A(t)$ starting at the identity has entries $A(t) = \begin{bmatrix} a(t) & b(t) \\ c(t) & d(t) \end{bmatrix}$ that satisfy

$$a(t)d(t) - b(t)c(t) = 1 \text{ and } a(0) = 1 = d(0), b(0) = 0 = c(0).$$

Differentiating the first equation with respect to t and evaluating at $t = 0$, while making use of the other equations, gives

$$a'(0)d(0) + a(0)d'(0) - b'(0)c(0) - b(0)c'(0) = 0 \rightsquigarrow a'(0) + d'(0) = 0.$$

Thus $A'(0) = \begin{bmatrix} a'(0) & b'(0) \\ c'(0) & d'(0) \end{bmatrix}$ has trace zero and is hence in $\mathfrak{sl}_2(\mathbb{R})$.

Exercise. Show that the Lie algebra of $SO_n(\mathbb{R})$ is all of $\mathfrak{so}_n(\mathbb{R})$. (Hint: Show that if A is skew-symmetric, then e^A is orthogonal, and consider the derivative of e^{tA}.)

Exercise. Show that the Lie algebra of $SL_n(\mathbb{R})$ is all of $\mathfrak{sl}_n(\mathbb{R})$. You will need a formula for the determinant of a general $n \times n$ matrix.

Exercise. Show that the Lie algebra of the Heisenberg group

$$H := \left\{ \begin{bmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix} \right\}$$

of upper-triangular 3×3 matrices with diagonal entries 1 is the space

$$\text{Lie}(H) = \left\{ \begin{bmatrix} 0 & x & y \\ 0 & 0 & z \\ 0 & 0 & 0 \end{bmatrix} \right\}$$

of 3×3 upper-triangular matrices with diagonal entries 0. Verify also that this $\text{Lie}(H)$ is closed under the Lie bracket $[A, B] = AB - BA$.

It turns out that for the group $GL_n(\mathbb{R})$ of all invertible $n \times n$ matrices, there end up being no constraints on what the entries of the derivative at time 0 of a path starting at the identity can be, so the Lie algebra $\text{Lie}(GL_n(\mathbb{R})) = \mathfrak{gl}_n(\mathbb{R})$ consists of all $n \times n$ matrices.

Certain Lie algebras appear to be different at first, but turn out to be the "same" when viewed in the right way. This is captured by the notion of a *Lie algebra isomorphism*: an isomorphism between two Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$ is a linear isomorphism $T : \mathfrak{g} \to \mathfrak{h}$ (i.e. an invertible linear map) which preserves the bracket in the sense that

$$T([x, y]) = [T(x), T(y)] \text{ for all } x, y \in \mathfrak{g}.$$

(To be clear, the bracket $[x, y]$ on the left is the one on the Lie algebra $\mathfrak{g}$, and the bracket $[T(x), T(y)]$ on the right is the one on $\mathfrak{h}$.) The main example for us is the isomorphism between $\mathfrak{so}_3(\mathbb{R})$ and $\mathbb{R}^3$, where we consider $\mathbb{R}^3$ as a Lie algebra with the cross product as the Lie bracket. A 3×3 skew-symmetric matrix looks like

$$\begin{bmatrix} 0 & x & y \\ -x & 0 & z \\ -y & -z & 0 \end{bmatrix},$$

and hence by associating to this the vector $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$, for example, we do get a linear isomorphism between $\mathfrak{so}_3(\mathbb{R})$ and $\mathbb{R}^3$ as vector spaces. However, this does not quite preserve the Lie bracket in the sense given above, but it *almost* does.

Exercise. Show that with the linear isomorphism $T : \mathfrak{so}_3(\mathbb{R}) \to \mathbb{R}^3$ given above, we have

$$T([A, B]) = -(T(A) \times T(B)),$$

where $T(A) \times T(B)$ is the cross product (i.e., Lie bracket) of $T(A), T(B) \in \mathbb{R}^3$. So, we do not quite get the requirement $T([A, B]) = [T(A), T(B)]$ of preserving the Lie bracket. But then find a slight modification of the linear isomorphism $T : \mathfrak{so}_3(\mathbb{R}) \to \mathbb{R}^3$ that *does* preserve Lie brackets, thereby showing that $\mathfrak{so}_3(\mathbb{R})$ and $\mathbb{R}^3$ are isomorphic Lie algebras.

Poisson Brackets

A *Poisson bracket* is a Lie bracket that in addition satisfies a "product rule" property. This product rule—or what is more precisely called the *Leibniz rule*—requires that we be able to multiply elements in our set, so this is not a definition we can give on a general Lie algebra. Rather, we restrict ourselves to considering sets of infinitely-differentiable, i.e. *smooth*, functions on $\mathbb{R}^n$. (For our purposes, "infinitely-differentiable" or "smooth" just means that all partial derivatives of the functions in question exist to all orders.) We denote the set of all such smooth functions $\mathbb{R}^n \to \mathbb{R}$ by $C^\infty(\mathbb{R}^n)$, where the ∞ refers to the fact we are considering infinitely-differentiable functions. On this set $C^\infty(\mathbb{R}^n)$, we can make sense of adding functions, multiplying functions by a scalar, and multiplying two functions together.

A Poisson bracket on $C^\infty(\mathbb{R}^n)$ is a Lie bracket we'll denote using braces $\{,\}$ that in addition to satisfying all the properties in the definition of Lie bracket also satisfies the *Leibniz rule*:

$$\{fg, h\} = \{f, h\}g + f\{g, h\}.$$

(Skew-symmetry of the Lie bracket implies the Leibniz rule is also satisfied for $\{f, gh\}$ where we multiply in the second entry instead.) If we think of the operation of Poisson-bracketing with h as a mapping $f \mapsto \{f, h\}$, the Leibniz rule says that this mapping behaves like a "derivative" since it satisfies the product rule: "differentiating" fg, by which we mean Poisson-bracketing with h, is the same as "differentiating" f and leaving g alone plus leaving f alone and "differentiating" g. The standard example of a Poisson bracket is

$$\{f, g\} = \frac{\partial f}{\partial x}\frac{\partial g}{\partial y} - \frac{\partial f}{\partial y}\frac{\partial g}{\partial x}$$

for $f, g \in C^\infty(\mathbb{R}^2)$, or more generally

$$\{f, g\} = \sum_i \left(\frac{\partial f}{\partial x_i}\frac{\partial g}{\partial y_i} - \frac{\partial f}{\partial y_i}\frac{\partial g}{\partial x_i} \right)$$

for functions on $\mathbb{R}^{2n}$ where we take $(x_1, \ldots, x_n, y_1, \ldots, y_n)$ as the variables on $\mathbb{R}^{2n}$. The fact that Poisson brackets satisfy the Leibniz rule implies that they are *always* given by expressions involving partial derivatives, so that any Poisson bracket on $\mathbb{R}^n$ looks like

$$\{f, g\} = \sum_{i,j} a_{ij} \frac{\partial f}{\partial x_i}\frac{\partial g}{\partial x_j}$$

for some $a_{ij} \in C^\infty(\mathbb{R}^n)$. The motivation for Poisson brackets comes from physics, which we will briefly touch upon in the final section just to provide some context behind all that we will develop.

Exercise. Show that if $D : C^\infty(\mathbb{R}^n) \to C^\infty(\mathbb{R}^n)$ is a linear map that satisfies the Leibniz rule $D(fg) = D(f)g + fD(g)$, then $D(1) = 0$ (where 1 denotes the constant function 1) and D is of the form

$$D(f) = \sum_i a_i \frac{\partial f}{\partial x_i}$$

for some $a_i \in C^\infty(\mathbb{R}^n)$. (Such a map D is called a *derivation*, and the point is that derivations are just given by ordinary partial derivatives.) For this you will need a multivariable version of Taylor's theorem from calculus.

Exercise. Use the exercise above to show that Poisson brackets are always of the form

$$\{f, g\} = \sum_{i,j} a_{ij} \frac{\partial f}{\partial x_i} \frac{\partial g}{\partial x_j}$$

for some $a_{ij} \in C^\infty(\mathbb{R}^n)$. (In fact, these functions a_{ij} are themselves Poisson brackets of coordinate the coordinate functions x_i, x_j.) Moreover, use the skew-symmetry of the Poisson bracket to show that $a_{ij} = -a_{ji}$ and hence that the sum above can be written as

$$\{f, g\} = \sum_{i<j} a_{ij} \left(\frac{\partial f}{\partial x_i} \frac{\partial g}{\partial x_j} - \frac{\partial f}{\partial x_j} \frac{\partial g}{\partial x_i} \right)$$

where now the sum is taken only over indices $i < j$.

It is true that the Poisson bracket of any function with a constant function is always zero, which we can see from the derivative interpretation for example. We say that a Poisson bracket is *non-degenerate*, or *symplectic*, if it is only constant functions that have this property: $\{,\}$ is non-degenerate if $\{f, g\} = 0$ for all g implies that f is constant.

Exercise. Show that the standard Poisson brackets on $\mathbb{R}^2$ and $\mathbb{R}^{2n}$ described above are non-degenerate. Hint: What does the condition $\{f, g\} = 0$ for all g say about f when you take g to be coordinate functions?

A Lie bracket on $\mathbb{R}^n$ induces a Poisson bracket in the following way. (The description we give here is fairly ad-hoc, but will start to give us the insight we need. Later we will return to this and see the *true* way of constructing a Poisson bracket from a Lie bracket in the context of *dual spaces*.) First, take the standard basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ of $\mathbb{R}^n$ and compute the Lie bracket between any two of these:

$$[\mathbf{e}_i, \mathbf{e}_j] = \sum_k c_{ij}^k \mathbf{e}_k$$

for some constants c_{ij}^k. We define the Poisson bracket of the corresponding coordinate functions $x_1, \ldots, x_n \in C^\infty(\mathbb{R}^n)$ via

$$\{x_i, x_j\} = \sum_k c_{ij}^k x_k.$$

In effect, we are just literally just viewing the given Lie bracket as a Poisson bracket, where instead of working with vectors $\mathbf{e}_i$ we work with the corresponding coordinate functions x_i. The resulting Poisson bracket is said to be *linear* since the bracket of any two linear functions, such as the coordinate functions x_i and x_j, is still linear.

Next we can extend the Poisson bracket to quadratic functions like x_1^2 and $x_2 x_3$ using the Leibniz rule: in order for the Leibniz rule to hold, it must be true that $\{x_1^2, x_2 x_3\}$, for example, is given by

$$\begin{aligned}\{x_1^2, x_2 x_3\} &= \{x_1^2, x_2\} x_3 + x_2 \{x_1^2, x_3\} \\ &= [\{x_1, x_2\} x_1 + x_1 \{x_1, x_2\}] x_3 + x_2 [\{x_1, x_3\} x_1 + x_1 \{x_1, x_3\}]\end{aligned}$$

where we think of x_1^2 as $x_1 x_2$. All remaining Poisson brackets deal with only linear terms, and these we already know how to compute from above. Similarly, we can then extend to cubic functions using the Leibniz rule, and so on: we get a unique Poisson bracket on all polynomial functions which extends the one we defined above on coordinate functions. To jump to a Poisson bracket defined for all smooth functions on $\mathbb{R}^n$, such as $e^{x_1x_2}$ for example, we then use the fact that polynomials are dense in the space of all smooth functions, or more simply that all smooth functions can be approximated to whatever accuracy we want using polynomials. We won't need the details here, but this will imply that there is only way to go from the Poisson bracket on polynomials to a Poisson bracket on all smooth functions via some type of "limiting" procedure.

Exercise. Verify that there is only one way to define the Poisson bracket of arbitrary smooth functions in a way that extends the bracket on polynomial functions described above.

In the case of $\mathfrak{so}_3(\mathbb{R})$, which we now identity with $\mathbb{R}^3$ with the cross product as the Lie bracket, we have

$$\mathbf{e}_1 \times \mathbf{e}_2 = \mathbf{e}_3, \ \mathbf{e}_2 \times \mathbf{e}_3 = \mathbf{e}_1, \ \text{and } \mathbf{e}_3 \times \mathbf{e}_1 = \mathbf{e}_2.$$

In terms of the coordinate functions x, y, z on $\mathbb{R}^3$, this translates into the following Poisson bracket relations:

$$\{x, y\} = z, \ \{y, z\} = x, \ \text{and } \{z, x\} = y.$$

The Poisson bracket for general smooth functions on $\mathbb{R}^3$ is then given by

$$\{f, g\} = z\left(\frac{\partial f}{\partial x}\frac{\partial g}{\partial y} - \frac{\partial f}{\partial y}\frac{\partial g}{\partial x}\right) + x\left(\frac{\partial f}{\partial y}\frac{\partial g}{\partial z} - \frac{\partial f}{\partial z}\frac{\partial g}{\partial y}\right) + y\left(\frac{\partial f}{\partial z}\frac{\partial g}{\partial x} - \frac{\partial f}{\partial x}\frac{\partial g}{\partial z}\right),$$

as you can verify.

Exercise. Verify this formula for the Poisson bracket of functions on $\mathfrak{so}_3(\mathbb{R}) \cong \mathbb{R}^3$.

Exercise. A 2×2 matrix of trace zero is of the form $\left[\begin{smallmatrix} x & y \\ z & -x \end{smallmatrix}\right]$, and hence can be characterized by a vector $\left[\begin{smallmatrix} x \\ y \\ z \end{smallmatrix}\right]$ in $\mathbb{R}^3$. The Lie bracket on $\mathfrak{sl}_2(\mathbb{R})$ can then be turned into a Lie bracket on $\mathbb{R}^3$ by declaring the bracket of two vectors $\mathbf{x}$ and $\mathbf{y}$ to be the vector which corresponds to the Lie bracket of the matrices that correspond to $\mathbf{x}$ and $\mathbf{y}$: if $T : \mathfrak{sl}_2(\mathbb{R}) \to \mathbb{R}^3$ is the linear isomorphism sending a matrix to a vector described above, we define $[\mathbf{x}, \mathbf{y}]$ by

$$[\mathbf{x}, \mathbf{y}] = T([T^{-1}(\mathbf{x}), T^{-1}(\mathbf{y})]),$$

where $T^{-1} : \mathbb{R}^3 \to \mathfrak{sl}_2(\mathbb{R})$ sends a vector back to a matrix.

Compute this Lie bracket on $\mathbb{R}^3$ explicitly, in particular by computing the Lie bracket between the standard basis vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$. Then determine the corresponding Poisson bracket of functions on $\mathbb{R}^3$ explicitly, in particular by computing the Poisson bracket of the coordinate functions x, y, z.

Exercise. A matrix in the Heisenberg Lie algebra is of the form $\left[\begin{smallmatrix} 0 & x & y \\ 0 & 0 & z \\ 0 & 0 & 0 \end{smallmatrix}\right]$, and hence can be characterized by a vector $\left[\begin{smallmatrix} x \\ y \\ z \end{smallmatrix}\right]$ in $\mathbb{R}^3$. Thus the Lie bracket on this Lie algebra can be transformed into a Lie bracket on $\mathbb{R}^3$. Compute this Lie bracket explicitly, and describe the corresponding Poisson bracket on $C^\infty(\mathbb{R}^3)$.

Given a degenerate Poisson bracket on $C^\infty(\mathbb{R}^n)$, meaning one for which there are non-constant functions C satisfying $\{C, f\} = 0$ for all f, we can look for the regions in $\mathbb{R}^n$ on which it becomes non-degenerate. For a function C as above satisfying $\{C, f\} = 0$ for all f, the level sets of C give such regions since on such regions C becomes constant, so "$\{C, f\} = 0$ for all f implies that C is constant" is true when all functions considered are restricted to these level sets. The maximal such regions are called the *symplectic leaves* of the Poisson bracket.

For example, recall that the Poisson bracket of functions on $\mathfrak{so}_3(\mathbb{R}) \cong \mathbb{R}^3$ equipped with the cross product is given by

$$\{f, g\} = z\left(\frac{\partial f}{\partial x}\frac{\partial g}{\partial y} - \frac{\partial f}{\partial y}\frac{\partial g}{\partial x}\right) + x\left(\frac{\partial f}{\partial y}\frac{\partial g}{\partial z} - \frac{\partial f}{\partial z}\frac{\partial g}{\partial y}\right) + y\left(\frac{\partial f}{\partial z}\frac{\partial g}{\partial x} - \frac{\partial f}{\partial x}\frac{\partial g}{\partial z}\right).$$

If a function C is to satisfy $\{C, f\} = 0$ for all f, then in particular it will satisfy this when f is a coordinate function, so

$$\{C, x\} = -z\frac{\partial C}{\partial y} + y\frac{\partial C}{\partial z} = 0, \ \{C, y\} = z\frac{\partial C}{\partial x} - x\frac{\partial C}{\partial z} = 0, \ \text{ and } \{C, z\} = x\frac{\partial C}{\partial y} - y\frac{\partial C}{\partial x} = 0.$$

In turn, if these equations all hold, then via the Leibniz rule we get that $\{C, f\} = 0$ for all polynomial functions f, and this implies that $\{C, f\} = 0$ for all smooth functions. Thus, functions C satisfying $\{C, f\} = 0$ for all f are characterized by the three equations

$$y\frac{\partial C}{\partial z} = z\frac{\partial C}{\partial y}, \ z\frac{\partial C}{\partial x} = x\frac{\partial C}{\partial z}, \text{ and } x\frac{\partial C}{\partial y} = y\frac{\partial C}{\partial x}.$$

The first equation suggests that differentiating C with respect to z should should result in an expression with z in it still, and differentiating with respect to y should result in an expression with y in it still. Similarly, we can derive information about what C should look like based on the remaining equations as well.

Exercise. Show that $C = x^2 + y^2 + z^2$ satisfies $\{C, f\} = 0$ for all f in the example above.

The level sets of $C = x^2 + y^2 + z^2$ (spheres!) thus give regions of $\mathbb{R}^3$ on which the Poisson bracket in the $\mathfrak{so}_3(\mathbb{R})$ case becomes non-degenerate, so the symplectic leaves of this Poisson bracket are spheres in $\mathbb{R}^3$. Note that each sphere is indeed a "maximal" region on which the Poisson bracket becomes non-degenerate, since including any more points not that on specific sphere makes is it so that $C = x^2 + y^2 + z^2$ is no longer constant on that region. This function is also constant on, say, the upper half of a given sphere, but such upper-halves are not "maximal" since they are contained on the full sphere on which C is still constant.

Exercise. Find the simplest non-constant function C you can in the $\mathfrak{sl}_2(\mathbb{R}) \cong \mathbb{R}^3$ case satisfying $\{C, f\} = 0$ for all f, and hence describe the symplectic leaves of this Poisson bracket in $\mathbb{R}^3$.

Exercise. Find the simplest non-constant function C you can in the Heisenberg case satisfying $\{C, f\} = 0$ for all f, and hence describe the symplectic leaves of this Poisson bracket in $\mathbb{R}^3$.

In each exercise above, the symplectic leaves should be recognizable geometric subsets of $\mathbb{R}^3$. (Spoiler alert: in the Heisenberg case the symplectic leaves are planes!)

Adjoint Orbits

Back to groups. There are certain actions associated with a matrix group G that are of main interest. First, we consider the action of G on itself by conjugation:

$$A \cdot B := ABA^{-1}.$$

This is the simplest action we can write down of G on itself that makes use of the group multiplication which *preserves* the identity element.

The fact that this preserves the identity element guarantees that we can differentiate this action to get an action of G on its Lie algebra $\mathfrak{g}$. To be clear, given a Lie algebra element $B'(0) \in \mathfrak{g}$ which arises from a path $B(t)$ in G starting at $B(0) = I$, we define the *adjoint* action of $A \in G$ on $B'(0) \in \mathfrak{g}$ by

$$A \cdot B'(0) := \frac{d}{dt}\Big|_{t=0} A \cdot B(t).$$

That is, we act by A on each element of the path $B(t)$, and then differentiate at time 0 to get a Lie algebra element as the result. Since $A \cdot B(t)$ is just conjugation and $AB(0)A^{-1} = AIA^{-1} = I$, $A \cdot B(t)$ is still a path in G starting at the identity, which is why differentiating at time 0 is guaranteed to give an element of $\mathfrak{g}$. (This would not be true if we simply took the initial action of G on itself to be just multiplication, since the path $AB(t)$ in G would start at $AB(0) = AI = A$ instead of I. Differentiating this does not give an element of the Lie algebra.) Moreover,

$$\frac{d}{dt}\Big|_{t=0} A \cdot B(t) = \frac{d}{dt}\Big|_{t=0} AB(t)A^{-1} = AB'(0)A^{-1}$$

by the product rule since A is constant with respect to t, so the adjoint action of G on $\mathfrak{g}$ just ends up being conjugation again. The only difference is that now we are conjugating arbitrary matrices by $A \in G$, not just invertible ones as was the case with the initial action of G on itself.

We denote the adjoint action of G on $\mathfrak{g}$ by Ad, so that for $A \in G$ and $X \in \mathfrak{g}$, the adjoint action of A on X is

$$Ad_A X = AXA^{-1}.$$

This is an action by linear transformations, meaning that for each $A \in G$ the map $Ad_A : \mathfrak{g} \to \mathfrak{g}$ given by acting by A is linear. Moreover, these linear transformations all preserve the Lie bracket.

Exercise. Show that for any $A \in G$ and $X, Y \in \mathfrak{g}$, $Ad_A[X, Y] = [Ad_A X, Ad_A Y]$.

To get the adjoint action of G on $\mathfrak{g}$ we differentiated the initial action of G on G in the term being acted upon, meaning that in $A \cdot B(t)$, A was fixed and $B(t)$ varied. Now we can vary the element A which is acting and differentiate in *this* term instead. To be clear, fix a Lie algebra element $B'(0) \in \mathfrak{g}$ as above, and now take a path $A(t) \in G$ starting at the identity. Acting on $B'(0)$ by each element in the path $A(t)$ gives a path in $\mathfrak{g}$, and we can differentiate this path with respect to t at $t = 0$, giving what we interpret as an action of $A'(0) \in \mathfrak{g}$ on $B'(0)$:

$$A'(0) \cdot B'(0) := \frac{d}{dt}\Big|_{t=0} A(t) \cdot B'(0).$$

We call this the *adjoint* action of $\mathfrak{g}$ on itself and denote it by a lowercase *ad*. Since the adjoint action of G on $\mathfrak{g}$ is just given by conjugation $A(t) \cdot B'(0) = A(t)B'(0)A(t)^{-1}$, we have

$$ad_{A'(0)} B'(0) = \frac{d}{dt}\Big|_{t=0} A(t) \cdot B'(0) = \frac{d}{dt}\Big|_{t=0} A(t)B'(0)A(t)^{-1}.$$

By the product rule, this becomes

$$ad_{A'(0)} B'(0) = A'(0)B'(0)A(0)^{-1} + A(0)B'(0)\left(\frac{d}{dt}\Big|_{t=0} A(t)^{-1}\right).$$

Using the formula for the derivative an inverse derived on a previous exercise, we get

$$ad_{A'(0)} B'(0) = A'(0)B'(0)A(0)^{-1} + A(0)B'(0)\left(-A(0)^{-1}A'(0)A(0)^{-1}\right),$$

which finally after using the fact that our paths begin at the identity becomes

$$ad_{A'(0)}B'(0) = A'(0)B'(0) - B'(0)A(0).$$

The conclusion is that the adjoint of action of $\mathfrak{g}$ on itself is nothing but the Lie bracket $[A'(0), B'(0)]$ we defined for matrices. Indeed, this is where this definition of the Lie bracket of matrices comes from: it arises by differentiating adjoint actions and ultimately from conjugations. That is, the Lie bracket of matrices is the "infinitesimal" analog of conjugation and is thus the simplest way to infinitesimally capture the group multiplication.

Exercise. Derive the Jacobi identity via differentiation. (A PRECISE STATEMENT OF THIS EXERCISE IS STILL TO COME.)

For now we are interested in understanding the *orbits* of the adjoint action of G on $\mathfrak{g}$. That is, for a fixed $X \in \mathfrak{g}$, we want to determine the set of all possible $Ad_A X \in \mathfrak{g}$ as $A \in G$ varies. In the example of $SO_3(\mathbb{R})$, we can be very explicit. First, recalling that $\mathfrak{so}_3(\mathbb{R})$ is isomorphic as a Lie algebra to $\mathbb{R}^3$ equipped with the cross product, we wish to determine what the adjoint action of $SO_3(\mathbb{R})$ on $\mathfrak{so}_3(\mathbb{R})$ becomes when thought of as an action of $SO_3(\mathbb{R})$ on $\mathbb{R}^3$ instead.

Exercise. Show that there exists an isomorphism $\mathfrak{so}_3(\mathbb{R}) \to \mathbb{R}^3$ of Lie algebras such that the adjoint action $Ad_A X = AXA^{-1}$ of $SO_3(\mathbb{R})$ on $\mathfrak{so}_3(\mathbb{R})$ becomes the usual action $A \cdot \mathbf{x} = A\mathbf{x}$ of $SO_3(\mathbb{R})$ on $\mathbb{R}^3$ by matrix multiplication. At some point it will likely help to recall (or look up) the formula for the inverse of a matrix in terms of its *adjugate* matrix.

Since the adjoint action of $SO_3(\mathbb{R})$ on $\mathfrak{so}_3(\mathbb{R}) \cong \mathbb{R}^3$ is just usual matrix multiplication, the orbits are simple to describe: for a fixed nonzero $\mathbf{x} \in \mathbb{R}^3$, the orbit $\{A\mathbf{x} \mid A \in SO_3(\mathbb{R})\}$ consists of all vectors obtained by rotating $\mathbf{x}$ in all possible ways, and is thus a sphere of radius given by the length of $\mathbf{x}$. The key observation is that these orbits are (drum roll please) precisely the symplectic leaves of the corresponding Poisson bracket on $\mathfrak{so}_3(\mathbb{R}) \cong \mathbb{R}^3$ described previously! It is no accident that this is true in this case, as we will see.

The same phenomenon occurs in the $SL_2(\mathbb{R})$ case, if we are careful about how to interpret the claim. A previous exercise asked to determine the symplectic leaves of the Poisson bracket derived from $\mathfrak{sl}_2(\mathbb{R})$, and it turns out that these symplectic leaves are hyperboloids (some one-sheeted, some two-sheeted), or a (double) cone. The orbits of the adjoint action of $SL_2(\mathbb{R})$ on $\mathfrak{sl}_2(\mathbb{R}) \cong \mathbb{R}^3$ (using a previous exercise to write down a Lie bracket on $\mathbb{R}^3$ that is isomorphic to the one on $\mathfrak{sl}_2(\mathbb{R})$) turn out to also be hyperboloids/cones, although not of the exact same shape as the symplectic leaves. Let us dive into this a bit more.

The adjoint action of $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \in SL_2(\mathbb{R})$ on $\left[\begin{smallmatrix} x & y \\ z & -x \end{smallmatrix}\right] \in \mathfrak{sl}_2(\mathbb{R})$ is given by

$$\begin{aligned} Ad_{\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]} \left[\begin{smallmatrix} x & y \\ z & -x \end{smallmatrix}\right] &= \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \left[\begin{smallmatrix} x & y \\ z & -x \end{smallmatrix}\right] \left[\begin{smallmatrix} d & -b \\ -c & a \end{smallmatrix}\right] \\ &= \begin{bmatrix} adx + bdz - acy + cbx & -abx - b^2z + a^2y - abx \\ cdx + d^2z - c^2y + cdx & -adx - bdz + acy - cbx \end{bmatrix}, \end{aligned}$$

where we use the fact that $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]^{-1} = \left[\begin{smallmatrix} d & -b \\ -c & a \end{smallmatrix}\right]$ since elements of $SL_2(\mathbb{R})$ have determinant 1. Under the isomorphism $\mathfrak{sl}_2(\mathbb{R}) \cong \mathbb{R}^3$ where $\left[\begin{smallmatrix} x & y \\ z & -x \end{smallmatrix}\right]$ corresponds to $\left[\begin{smallmatrix} x \\ y \\ z \end{smallmatrix}\right]$, the adjoint action of $SL_2(\mathbb{R})$ on $\mathfrak{sl}_2(\mathbb{R}) \cong \mathbb{R}^3$ thus looks like

$$Ad_{\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]} \left[\begin{smallmatrix} x \\ y \\ z \end{smallmatrix}\right] = \left[\begin{smallmatrix} adx+bdz-acy+cbx \\ -abx-b^2z+a^2y-abx \\ cdx+d^2z-c^2y+cdx \end{smallmatrix}\right].$$

Let us take, for instance, the action on the vector $\left[\begin{smallmatrix}1\\0\\0\end{smallmatrix}\right]$:

$$Ad_{\left[\begin{smallmatrix}a & b\\ c & d\end{smallmatrix}\right]}\left[\begin{smallmatrix}1\\0\\0\end{smallmatrix}\right] = \left[\begin{smallmatrix}ad+bc\\-2ab\\2cd\end{smallmatrix}\right].$$

Note that $\left[\begin{smallmatrix}1\\0\\0\end{smallmatrix}\right]$ satisfies $yz + x^2 = 1$. We can check that the vector $\left[\begin{smallmatrix}x\\y\\z\end{smallmatrix}\right] = \left[\begin{smallmatrix}ad+bc\\-2ab\\2cd\end{smallmatrix}\right]$ in the orbit of $\left[\begin{smallmatrix}1\\0\\0\end{smallmatrix}\right]$ also satisfies this same equation:

$$\begin{aligned}(-2ab)(2cd) + (ad+bc)^2 &= -4abcd + a^2d^2 + 2adbc + b^2c^2\\ &= a^2d^2 - 2abcd + b^2c^2\\ &= (ad-bc)^2 = 1,\end{aligned}$$

where again we use the fact that $\left[\begin{smallmatrix}a & b\\ c & d\end{smallmatrix}\right] \in SL_2(\mathbb{R})$ has determinant $ad - bc$ equal to 1. This shows that the entire orbit containing $\left[\begin{smallmatrix}1\\0\\0\end{smallmatrix}\right]$ lies on the surface $yz + x^2 = 1$, which is indeed a (one-sheeted) hyperboloid. The orbits through points other than $\left[\begin{smallmatrix}1\\0\\0\end{smallmatrix}\right]$ will make up other hyperboloids, or a double cone in one case. (Actually, the orbit through the zero vector consists of only the zero vector, so this orbit is just a point. In the discussion above we really only care about non-origin orbits. In the "double cone" case we technically get a double cone with the origin removed.)

Exercise. Verify that the orbits of the adjoint action of $SL_2(\mathbb{R})$ on $\mathfrak{sl}_2(\mathbb{R}) \cong \mathbb{R}^3$ are hyperboloids or a double cone, by showing that the value of $yz + x^2$ is the same for all points in an orbit.

So we do get hyperboloids/cones as orbit for $SL_2(\mathbb{R})$, although not the same hyperboloids/cone as the symplectic leaves. (In other words, the symplectic leaves are given by an equation similar to $yz + x^2 = k$, but not literally this exact equation.) However, the hyperboloids/orbits we get in each case can easily be related by some map $\mathbb{R}^3 \to \mathbb{R}^3$ which transforms the one type of hyperboloid/cone to the other. The orbits of the adjoint action of $SL_2(\mathbb{R})$ are thus, as we say, *diffeomorphic* (a "diffeomorphism" is a bijective smooth map whose inverse is also smooth) to the symplectic leaves of the corresponding Poisson bracket, so the relation between adjoint orbits and symplectic leaves still holds here as it did for $SO_3(\mathbb{R})$.

Exercise. Find the (bijective) map $\mathbb{R}^3 \to \mathbb{R}^3$ which sends the orbits of the adjoint action of $SL_2(\mathbb{R})$ to the symplectic leaves of the Poisson bracket derived from $\mathfrak{sl}_2(\mathbb{R})$.

Alas, the relation between adjoint orbits and symplectic leaves breaks down when we get to the Heisenberg group and its Lie algebra, even if we allow for spaces that are only "diffeomorphic" and not literally the same.

Exercise. Use the identification of the Heisenberg Lie algebra with $\mathbb{R}^3$ given in a previous exercise to determine the orbits of the adjoint action of the Heisenberg group. You should find that these orbits are actually *lines*, and not planes as the symplectic leaves were.

The reason for this discrepancy has to do with a difference in the structure of the Heisenberg Lie algebra as compared to either $\mathfrak{so}_3(\mathbb{R})$ and $\mathfrak{sl}_2(\mathbb{R})$: the Heisenberg Lie algebra is not what's called *semisimple*, whereas $\mathfrak{so}_3(\mathbb{R})$ and $\mathfrak{sl}_2(\mathbb{R})$ are semisimple. We will discuss semisimplicity and see what it has to do with describing the relation between adjoint orbits and symplectic leaves later. But first we must recast everything we've done so far in the "correct settting", where the relation between orbits and symplectic leaves *always* holds, even for the Heisenberg case. This correct setting comes from considering the *coadjoint action* of a group instead of its adjoint action, and to discuss coadjoint actions we need to first discuss “dual spaces”.

Dual Spaces

The *dual space* of a vector space V is the space V^* of all linear transformations from V to $\mathbb{R}$:

$$V^* := \{f : V \to \mathbb{R} \mid f \text{ is linear}\}.$$

(Such linear transformations which map into $\mathbb{R}$ are often called *linear functionals*, so V^* is the space of all linear functionals on V.) The dual space is itself a vector space under the operations of addition of functions and scalar multiplication of functions. If V is finite-dimensional, then V^* is also finite dimensional and of the same dimension as V.

Exercise. Let $e_1, \ldots, e_n$ be a basis for V. Define the linear functionals $f_1, \ldots, f_n : V \to \mathbb{R}$ by specifying that

$$f_i(e_j) = 1 \text{ when } i = j \text{ and } 0 \text{ when } i \neq j,$$

which by linearity determines the value of f_i on any element of V. Show that these $f_1, \ldots, f_n$ form a basis for V^*, which we call the *dual basis* of $e_1, \ldots, e_n$.

For example, take $V = \mathbb{R}^n$ with elements thought of as column vectors. A linear functional $\mathbb{R}^n \to \mathbb{R}$ is then given by a $1 \times n$ matrix: for instance, the 1×3 matrix $\begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$ gives the linear functional $\mathbb{R}^3 \to \mathbb{R}$ defined by

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} \mapsto \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = x + 2y + 3z.$$

(Technically the result of the product in the middle is a 1×1 matrix, but we just think of this is an element in $\mathbb{R}$.) So, the dual space of $\mathbb{R}^n$ is the space of all $1 \times n$ matrices, which we can also think of as $\mathbb{R}^n$ only with elements viewed as *row* vectors instead of column vectors. The standard basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ of $\mathbb{R}^n$ then gives the dual basis $\mathbf{e}_1^T, \ldots, \mathbf{e}_n^T$ (the transpose turns each column vector into a row vector) of $(\mathbb{R}^n)^* \cong \mathbb{R}^n$. In this case, the linear functional $\mathbf{e}_i^T$ picks out the i-th coordinate of a vector: if $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, then $\mathbf{e}_i^T(\mathbf{x}) = x_i$.

The same is true more generally for any dual basis: if $c_1, \ldots, c_n$ are the coordinates of $v \in V$ relative to the basis $e_1, \ldots, e_n$, so that

$$v = c_1 e_1 + \cdots + c_n e_n,$$

then the corresponding dual basis vector f_i picks out the coordinate c_i corresponding to e_i:

$$f_i(v) = f_i(c_1 e_1 + \cdots + c_n e_n) = c_1 f_i(e_1) + \cdots + c_n f_i(e_n) = c_i$$

where we use the defining property of f_i as evaluating to 1 at e_i and to 0 at any other e_j. In terms of the dual basis, the expression for v above can thus be written as

$$v = f_1(v) e_1 + \cdots + f_n(v) e_n.$$

Elements in the dual basis can hence be thought of as "coordinate functions", and indeed elements of V^* in general should be viewed as types of "coordinate functions". The upshot is that the dual space of V in general gives a way to work with all possible coordinate functions on V without having to pick out any one specific basis of V. For various reasons this is desirable, since it leads to constructions which are "coordinate independent".

Given a linear map $T : V \to W$, we can define a linear map $T^* : W^* \to V^*$ going between the dual spaces of V and W (but in reverse order!) in the following way: for $g \in W^*$, we define T^*g to be the composition $g \circ T$. To be clear, $g \in W^*$ is a map $g : W \to \mathbb{R}$, so the composition $g \circ T$ maps from V to $\mathbb{R}$, just as an element of V^* should be. As an element of V^*, T^*g takes as input $v \in V$ and outputs the scalar $(T^*g)v = g(T(v))$, which makes sense since $T(w)$ is an element in the domain W of $g \in W^*$. We call the map $T^* : W^* \to V^*$ defined in this way the *dual map* of T.

Exercise. Show that the dual map $T^* : W^* \to V^*$ of $T : V \to W$ is indeed linear, and that $(ST)^* = T^*S^*$ where $S : W \to Z$ is a linear map and ST denotes the composition of S and T. Note that the two dual maps on the right T^* and S^* are being composed in the order opposite to that of S and T.

Here's an example. Take $T : \mathbb{R}^3 \to \mathbb{R}^2$ to be the linear map defined by the matrix

$$A = \begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & 1 \end{bmatrix}.$$

The dual map T^* should take an element of $(\mathbb{R}^2)^*$ as input and output an element of $(\mathbb{R}^3)^*$. If we view these two dual spaces as consisting of row vectors, then

$$T^* \begin{bmatrix} a & b \end{bmatrix} \text{ should be a } 1 \times 3 \text{ row vector.}$$

As an element of $(\mathbb{R}^3)^*$, this $T^* \begin{bmatrix} a & b \end{bmatrix}$ takes $\begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3$ as input and outputs the scalar

$$\begin{aligned}
\left(T^* \begin{bmatrix} a & b \end{bmatrix}\right) \begin{bmatrix} x \\ y \\ z \end{bmatrix} &= \begin{bmatrix} a & b \end{bmatrix} \left(T \begin{bmatrix} x \\ y \\ z \end{bmatrix} \right) \\
&= \begin{bmatrix} a & b \end{bmatrix} \left(\begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \right) \\
&= \begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} x + 2y + 3z \\ -x + z \end{bmatrix} \\
&= ax + 2ay + 3az - bx + bz.
\end{aligned}$$

This scalar can be written as

$$\left(T^* \begin{bmatrix} a & b \end{bmatrix}\right) \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} a - b & 2a & 3a + b \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix},$$

so concretely the dual map T^* sends $\begin{bmatrix} a & b \end{bmatrix}$ to $\begin{bmatrix} a - b & 2a & 3a + b \end{bmatrix}$. This can also be seen by noting that the composition gT, where g is the linear functional corresponding to $\begin{bmatrix} a & b \end{bmatrix}$, defining T^*g is given by the matrix

$$gT = \begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} a - b & 2a & 3a + b \end{bmatrix}.$$

Continuing with this same example, we now seek to the represent $T^* : (\mathbb{R}^2)^* \to (\mathbb{R}^3)^*$ as a matrix relative to the dual bases of $(\mathbb{R}^2)^*$ and $(\mathbb{R}^3)^*$ which correspond to the standard bases of $\mathbb{R}^2$ and $\mathbb{R}^3$. Let's unwind what this means. The standard basis of $\mathbb{R}^2$ and $\mathbb{R}^3$ give the dual bases

$$\begin{bmatrix}1 & 0\end{bmatrix}, \begin{bmatrix}0 & 1\end{bmatrix} \text{ for } (\mathbb{R}^2)^* \text{ and } \begin{bmatrix}1 & 0 & 0\end{bmatrix}, \begin{bmatrix}0 & 1 & 0\end{bmatrix}, \begin{bmatrix}0 & 0 & 1\end{bmatrix} \text{ for } (\mathbb{R}^3)^*.$$

In terms of this bases, $\begin{bmatrix}a & b\end{bmatrix}$ is written as

$$\begin{bmatrix}a & b\end{bmatrix} = a\begin{bmatrix}1 & 0\end{bmatrix} + b\begin{bmatrix}0 & 1\end{bmatrix}$$

and $T^*\begin{bmatrix}a & b\end{bmatrix} = \begin{bmatrix}a-b & 2a & 3a+b\end{bmatrix}$ is

$$\begin{bmatrix}a-b & 2a & 3a+b\end{bmatrix} = (a-b)\begin{bmatrix}1 & 0 & 0\end{bmatrix} + 2a\begin{bmatrix}0 & 1 & 0\end{bmatrix} + (3a+b)\begin{bmatrix}0 & 0 & 1\end{bmatrix}.$$

We want the matrix that has the effect of sending the *coordinate vector* of $\begin{bmatrix}a & b\end{bmatrix}$ relative to the dual basis to the coordinate vector of $T^*\begin{bmatrix}a & b\end{bmatrix}$ relative to the dual basis, which means the matrix satisfying

$$(\text{matrix})\begin{bmatrix}a \\ b\end{bmatrix} = \begin{bmatrix}a-b \\ 2a \\ 3a+b\end{bmatrix}.$$

The matrix which does this is

$$\begin{bmatrix}1 & -1 \\ 2 & 0 \\ 3 & 1\end{bmatrix},$$

so this is the matrix which represents T^* relative to the dual bases. The key observation is that this is nothing but the transpose of the matrix

$$A = \begin{bmatrix}1 & 2 & 3 \\ -1 & 0 & 1\end{bmatrix}$$

which originally defined T. Indeed, this is true in general.

Exercise. Suppose $e_1, \ldots, e_n$ is a basis for V and $u_1, \ldots, u_m$ is a basis for W. Let $T : V \to W$ be a linear map, whose matrix relative to these basis is A. Show that the matrix of the dual map $T^* : W^* \to V^*$ relative to the corresponding dual bases is A^T.

Thus, when everything is expressed in terms of matrices relative to chosen bases, taking the dual map just corresponds to taking the transpose. The fact that dual maps satisfy $(ST)^* = T^*S^*$ then just corresponds to the analogous fact $(AB)^T = B^TA^T$ for transposes.

Finally, we return to the construction of the Poisson bracket, which as we said when we outlined it previously was fairly ad-hoc. Now we have the tools needed to give the "correct" definition. In actuality, given a Lie algebra $\mathfrak{g}$, the corresponding linear Poisson bracket should in fact be a bracket on $C^\infty(\mathfrak{g}^*)$. (The definition we gave previously as a Poisson bracket on $C^\infty(\mathbb{R}^n)$ comes from picking a basis with which to identify $\mathfrak{g}^*$ and $\mathbb{R}^n$, and indeed this is how we define "smooth" for a function on $\mathfrak{g}^*$: take an isomorphism $\mathfrak{g}^* \cong \mathbb{R}^n$ and say that $\mathfrak{g}^* \to \mathbb{R}$ is smooth if the corresponding $\mathbb{R}^n \to \mathbb{R}$ is smooth.) A Poisson bracket on $C^\infty(\mathfrak{g}^*)$ should take two smooth functions $\mathfrak{g}^* \to \mathbb{R}$ as inputs and output a third. Note first that any element of $\mathfrak{g}$ can be viewed as defining a function $\mathfrak{g}^* \to \mathbb{R}$: for $v \in \mathfrak{g}$, the corresponding function on $\mathfrak{g}^*$ takes $f \in \mathfrak{g}^*$ as input (recall that f is itself then a linear function $f : \mathfrak{g} \to \mathbb{R}$) and outputs $f(v)$. In other words, in the expression $f(v)$ where $v \in \mathfrak{g}$ and $f \in \mathfrak{g}^*$, where we normally think about f as the fixed function and v the thing that varies, we instead think about v as the fixed thing and f as the thing that varies. In this way, we can thus think about elements of $\mathfrak{g}$ as giving specific elements of $C^\infty(\mathfrak{g}^*)$.

The Poisson bracket on $C^\infty(\mathfrak{g}^*)$ is then defined by taking the Lie bracket on $\mathfrak{g} \subseteq C^\infty(\mathfrak{g}^*)$ as a starting point, so that for $v_1, v_2 \in \mathfrak{g}$ viewed as elements of $C^\infty(\mathfrak{g}^*)$, we define their Poisson bracket to be their Lie bracket:

$$\{v_1, v_2\} := [v_1, v_2].$$

The result is an element of $\mathfrak{g}$, which can still be viewed as a function on $\mathfrak{g}^*$. Using the Leibniz rule we then we get something like

$$\{v_1^2, v_2\} = \{v_1 v_1, v_2\} = 2v_1\{v_1, v_2\}$$

where $v_1^2 = v_1 v_1$ is now an element of $C^\infty(\mathfrak{g})$ that is not in $\mathfrak{g}$. (As a function on $\mathfrak{g}^*$, v_1^2 is the square of the function v_1, so it sends $f \in \mathfrak{g}^*$ to $f(v_1)^2$.) In this way we can extend the Poison bracket on $\mathfrak{g} \subseteq C^\infty(\mathfrak{g}^*)$ to polynomials, and then to all smooth functions using denseness. The symplectic leaves we defined previously are thus really subsets $\mathfrak{g}^*$, where we go from here to subsets of $\mathbb{R}^3$ as in the $\mathfrak{sl}_2, \mathfrak{so}_3$, and Heisenberg examples by using an isomorphism $\mathfrak{g}^* \cong \mathbb{R}^3$.

Coadjoint Orbits

The *coadjoint action* of a group G is an action of G on the dual $\mathfrak{g}^*$ of its Lie algebra. To define this, for each $g \in G$ take the linear map $Ad_g : \mathfrak{g} \to \mathfrak{g}$ defined by the adjoint action of G on $\mathfrak{g}$. The dual map is then a linear map $Ad_g^* : \mathfrak{g}^* \to \mathfrak{g}^*$, so these maps can be taken as defining an action of G on $\mathfrak{g}^*$. But actually it is typical to take a slightly different approach, and define the coadjoint action by $g \in G$ on $\mathfrak{g}^*$ to be the dual of the adjoint action by the *inverse* of g on $\mathfrak{g}$:

$$Ad_g^* := \text{the dual of } Ad_{g^{-1}}.$$

Explicitly, this gives

$$(Ad_g^* f)v = f(Ad_{g^{-1}}v)$$

for $f \in \mathfrak{g}^*$ and $v \in \mathfrak{g}$. To unpack this, $Ad_g^* f$ should be an element of $\mathfrak{g}$, and this says that the value of this functional on v is the value of the functional f on $Ad_{g^{-1}}v \in V$. The *coadjoint orbits* of G are the orbits in $\mathfrak{g}^*$ of the coadjoint action.

Exercise. The reason for using g^{-1} instead of just g when defining the coadjoint action stems from the fact that taking the dual of a product swaps order in the sense that $(ST)^* = T^*S^*$, whereas an action of G on $\mathfrak{g}^*$ should satisfy the typical property

$$g_1 \cdot (g_2 \cdot x) = (g_1 g_2) \cdot v$$

required of groups actions. Show that if we try to define $Ad_g^* : \mathfrak{g}^* \to \mathfrak{g}^*$ to be the dual map of $Ad_g : \mathfrak{g} \to \mathfrak{g}$, the property above is not satisfied, but that it is satisfied if we take Ad_g^* to be the dual map of $Ad_{g^{-1}} : \mathfrak{g} \to \mathfrak{g}$ instead as in the actual definition of the coadjoint action.

Let us compute some examples. Recall that for $SO_3(\mathbb{R})$ the adjoint action on $\mathfrak{so}_3(\mathbb{R}) \cong \mathbb{R}^3$ (using the right isomorphism) is just matrix multiplication: for $A \in SO_3(\mathbb{R})$ and $\mathbf{x} \in \mathbb{R}^3$, we have

$$Ad_A \mathbf{x} = A\mathbf{x},$$

so that the matrix of Ad_A relative to the standard basis is just A. We identify $\mathbb{R}^3$ with its dual space using the dual standard basis, so that the coadjoint action of $SO_3(\mathbb{R})$ on $\mathfrak{so}_3(\mathbb{R})^*$ can also be viewed as an action on $\mathbb{R}^3$. The matrix of this coadjoint transformation Ad^*_A is the matrix of the dual map of $Ad_{A^{-1}}$ by the definition of the coadjoint action, and the matrix of $Ad_{A^{-1}}$ is A^{-1}, so by what we learned about dual maps and transposes earlier we get that the matrix of the dual map is $(A^{-1})^T$. Thus Ad^*_A is the linear map corresponding to the matrix $(A^{-1})^T$, so the coadjoint action of $SO_3(\mathbb{R})$ on $\mathfrak{so}_3(\mathbb{R})^* \cong \mathbb{R}^3$ is

$$Ad^*_A\mathbf{x} = (A^{-1})^T\mathbf{x}.$$

But in this case $A \in SO_3(\mathbb{R})$ is an orthogonal matrix, so $A^{-1} = A^T$ and hence $(A^{-1})^T = A$. Thus $Ad^*_A\mathbf{x} = A\mathbf{x}$, meaning that the adjoint and coadjoint actions are the same in the $SO_3(\mathbb{R})$ case. Hence the coadjoint orbits of $SO_3(\mathbb{R})$ are the same spheres we saw before when computing the adjoint orbits, and thus agree with the symplectic leaves in $\mathfrak{g}^* \cong \mathbb{R}^3$.

We computed previously that the adjoint action of $SL_2(\mathbb{R})$ on $\mathfrak{sl}_2(\mathbb{R}) \cong \mathbb{R}^3$ was given by

$$Ad_{\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]} \left[\begin{smallmatrix} x \\ y \\ z \end{smallmatrix}\right] = \left[\begin{smallmatrix} adx+bdz-acy+cbx \\ -abx-b^2z+a^2y-abx \\ cdx+d^2z-c^2y+cdx \end{smallmatrix}\right].$$

The matrix of the adjoint action by $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ relative to the standard basis of $\mathbb{R}^3$ is thus

$$Ad_{\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]} = \begin{bmatrix} ad+bc & -ac & bd \\ -2ab & a^2 & -b^2 \\ 2cd & -c^2 & d^2 \end{bmatrix}.$$

The coadjoint action by $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ is then given by multiplication by the transpose of the inverse of this matrix. Instead of working with this inverse transpose, we will instead stick with the transpose of the matrix above. This transpose thus actually gives the coadjoint action by the matrix $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]^{-1}$ instead of by $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$, so that what we are applying is

$$Ad^*_{\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]^{-1}} = \begin{bmatrix} ad+bc & -2ab & 2cd \\ -ac & a^2 & -c^2 \\ bd & -b^2 & d^2 \end{bmatrix} \quad \text{instead of } Ad^*_{\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]},$$

but this does not matter when determining the orbits since we run though all elements of $SO_3(\mathbb{R})$ when compute $Ad^*_A\mathbf{x}$ anyway, and both A and A^{-1} show up as such elements. (In other words, computing all possible $Ad^*_{A^{-1}}\mathbf{x}$ for all possible A still gives the entire orbit.) Elements in the orbit of $\mathbf{x} = \left[\begin{smallmatrix} x \\ y \\ z \end{smallmatrix}\right]$ are thus of the form

$$Ad^*_{\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]^{-1}}\mathbf{x} = \begin{bmatrix} ad+bc & -2ab & 2cd \\ -ac & a^2 & -c^2 \\ bd & -b^2 & d^2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} adx+bcx-2aby+2cdz \\ -acx+a^2y-c^2z \\ bdx-b^2y+d^2z \end{bmatrix}.$$

Take for example $\mathbf{x} = \left[\begin{smallmatrix} 1 \\ 0 \\ 0 \end{smallmatrix}\right]$. The orbit through point consists of all things of the form

$$Ad^*_{\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]^{-1}} \left[\begin{smallmatrix} 1 \\ 0 \\ 0 \end{smallmatrix}\right] = \left[\begin{smallmatrix} ad+bc \\ -ac \\ bd \end{smallmatrix}\right].$$

The point $\left[\begin{smallmatrix} 1 \\ 0 \\ 0 \end{smallmatrix}\right]$ satisfies $4yz + x^2 = 1$, and indeed so does any point in this specific coadjoint orbit:

$$4(-ac)(bd) + (ad+bc)^2 = a^2d^2 + b^2c^2 - 2abcd = (ad-bc)^2 = 1$$

since $ad - bc = 1$. Just as we saw for the adjoint orbits, we get that the coadjoint orbit through $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ is a hyperboloid, only this time with equation $4yz + x^2 = 1$ as opposed to $yx + x^2 = 1$ as was the case for the adjoint orbit before.

Exercise. Show that all coadjoint orbits are given by equations of the form $4yz + x^2 = k$, which are hyperboloids or in one case a cone. (Exclude the case where we take the coadjoint orbit of the zero vector, which is just a point.)

The difference between this coadjoint orbit example and the previous adjoint orbit example for $SL_2(\mathbb{R})$ is that now the coadjoint orbits $4yz + x^2 = k$ are *exactly* the same as the symplectic leaves of the corresponding Poisson bracket on $\mathfrak{sl}_2(\mathbb{R})^* \cong \mathbb{R}^3$! A previous exercise asked to determine these symplectic leaves, and the answer turns out to be precisely the surfaces of the form $4yz + x^2 = k$, which are also the coadjoint orbits. For the adjoint orbits we still saw that the adjoint orbits—described by equations $yz + x^2 = k$—were diffeomorphic to the symplectic leaves, but for the coadjoint orbits they are literally the same.

Exercise. Compute the orbits of the coadjoint action of the Heisenberg group on the dual of its Lie algebra, which we identify with $\mathbb{R}^3$ as in previous exercises. The answer should be exactly the same as the symplectic leaves computed in a previous example.

Thus we have the main theorem in this subject: for any matrix group G with Lie algebra $\mathfrak{g}$, the orbits of the coadjoint action of G on $\mathfrak{g}^*$ are the symplectic leaves of the Poisson bracket on $C^\infty(\mathfrak{g}^*)$ induced from the Lie bracket.

Semisimplicity

We now focus on understanding the relation between the adjoint and coadjoint orbits a bit better, and in particular to understanding why in the $\mathfrak{so}_3(\mathbb{R})$ and $\mathfrak{sl}_2(\mathbb{R})$ cases they were essentially the same (literally the same for $\mathfrak{so}_3(\mathbb{R})$, and "diffeomorphic" for $\mathfrak{sl}_2(\mathbb{R})$) but not in the Heisenberg case. The key notion to grasp is that of semisimplicity.

First, we define a bilinear pairing on a Lie algebra $\mathfrak{g}$ called the *Killing form* as follows. This pairing will take two Lie algebra elements X, Y as inputs and output a real number that we denote by $B(X, Y)$. To define this, consider the linear transformations $ad_X, ad_Y : \mathfrak{g} \to \mathfrak{g}$ that are given by the (lowercse) adjoint actions of X and Y:

$$ad_X Z := [X, Z] \quad \text{and} \quad ad_Y Z := [Y, Z].$$

The composition of these linear transformations is also a linear transformation $ad_X \circ ad_Y : \mathfrak{g} \to \mathfrak{g}$, and we define $B(X, Y)$ to be the *trace* of this composition:

$$B(X, Y) = \operatorname{tr}(ad_X \circ ad_Y).$$

(Recall that the trace of a linear transformation is the trace of the matrix that represents that transformation with respect to any basis.) The Killing form is bilinear: for fixed $X \in \mathfrak{g}$ the mapping $Y \mapsto B(X, Y)$ is linear, and for fixed $Y \in \mathfrak{g}$ the mapping $X \mapsto B(X, Y)$ is linear. When $\mathfrak{g}$ is the Lie algebra of a group G, the Killing form is also Ad-invariant in the sense that

$$B(Ad_g X, Ad_g Y) = B(X, Y)$$

for any $g \in G$ and $X, Y \in \mathfrak{g}$.

Exercise. Justify the Ad-invariance property of the Killing form. You will need to use the fact that similar matrices have the same trace.

For $SO_3(\mathbb{R})$, the adjoint transformation $ad_X : \mathfrak{so}_3(\mathbb{R}) \to \mathfrak{so}_3(\mathbb{R})$ for $X = \left[\begin{smallmatrix} 0 & x & y \\ -x & 0 & z \\ -y & -z & 0 \end{smallmatrix}\right] \in \mathfrak{so}_3(\mathbb{R})$ is given by

$$ad_X \left[\begin{smallmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 0 & x & y \\ -x & 0 & z \\ -y & -z & 0 \end{smallmatrix}\right]\left[\begin{smallmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{smallmatrix}\right] - \left[\begin{smallmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{smallmatrix}\right]\left[\begin{smallmatrix} 0 & x & y \\ -x & 0 & z \\ -y & -z & 0 \end{smallmatrix}\right]$$
$$= \begin{bmatrix} 0 & -yc+bz & xc-az \\ yc-az & 0 & -xb+ay \\ -xc+az & xb-ay & 0 \end{bmatrix}.$$

With respect to the standard basis of $\mathfrak{so}_3(\mathbb{R})$ that gives the isomorphism $\left[\begin{smallmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{smallmatrix}\right] \mapsto \left[\begin{smallmatrix} a \\ b \\ c \end{smallmatrix}\right]$ with $\mathbb{R}^3$, the adjoint transformation ad_X satisfies

$$ad_X \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} -yc+bz \\ xc-az \\ -xb+ay \end{bmatrix}$$

and hence is represented by

$$ad_X = \begin{bmatrix} 0 & z & -y \\ -z & 0 & x \\ y & -x & 0 \end{bmatrix}.$$

For two such adjoint transformations ad_X and ad_Y (take X as above only with entries using 1 as a subscript and Y of the same form but using 2 as a subscript), we have

$$ad_X \circ ad_Y = \begin{bmatrix} 0 & z_1 & -y_1 \\ -z_1 & 0 & x_1 \\ y_1 & -x_1 & 0 \end{bmatrix}\begin{bmatrix} 0 & z_2 & -y_2 \\ -z_2 & 0 & x_2 \\ y_2 & -x_2 & 0 \end{bmatrix}$$
$$= \begin{bmatrix} -z_1z_2 - y_1y_2 & y_1x_2 & z_1x_2 \\ x_1y_2 & -z_1z_2 - x_1x_2 & z_1y_2 \\ x_1z_2 & y_1z_2 & -y_1y_2 - x_1x_2 \end{bmatrix},$$

so $B(X,Y) = \operatorname{tr}(ad_X \circ ad_Y) = -2(x_1x_2 + y_2y_2 + z_2z_2)$ is the Killing form on $\mathfrak{so}_3(\mathbb{R})$.

Exercise. Compute the Killing form on $\mathfrak{sl}_2(\mathbb{R})$.

Exercise. Compute the Killing form on the Heisenberg Lie algebra.

We say that the Lie algebra $\mathfrak{g}$ is *semisimple* if its Killing form is non-degenerate, which means that the only $X \in \mathfrak{g}$ satisfying $B(X,Y) = 0$ for all $Y \in \mathfrak{g}$ is $X = 0$. This is true for the example of $\mathfrak{so}_3(\mathbb{R})$ above: if for fixed x_1, y_1, z_1 we have

$$-2(x_1x_2 + y_1y_2 + z_1z_2) = 0$$

for all x_2, y_2, z_2, then taking $(x_2, y_2, z_2) = (1,0,0), (0,1,0)$, and $(0,0,1)$ will show that $x_1 = 0, y_1 = 0$, and $z_1 = 0$ respectively. Thus $\mathfrak{so}_3(\mathbb{R})$ is semisimple.

Exercise. Show that $\mathfrak{sl}_2(\mathbb{R})$ is semisimple.

It should be clear from the computation of the Killing form on the Heisenberg algebra that it is degenerate, so that the Heisenberg algebra is not semisimple.

The Killing form us to define a linear map $\mathfrak{g} \to \mathfrak{g}^*$ as follows: for each $X \in \mathfrak{g}$, define $B_X \in \mathfrak{g}^*$ to be the linear functional which sends $Y \in \mathfrak{g}$ to $B(X, Y)$:

$$B_X(Y) := B(X, Y) \text{ for all } Y \in \mathfrak{g}.$$

For the Killing form on $\mathfrak{so}_3(\mathbb{R}) \cong \mathbb{R}^3$ computed above for example, for fixed $\mathbf{x} = \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}$ we have

$$B_{\mathbf{x}} \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = B\left(\begin{bmatrix} x_1 \\ y_1 \\ z_2 \end{bmatrix}, \begin{bmatrix} x_2 \\ y_2 \\ z_3 \end{bmatrix}\right) = -2(x_1x_2 + y_1y_2 + z_1z_2),$$

so the linear functional $B_{\mathbf{x}} : \mathbb{R}^3 \to \mathbb{R}$ is represented by the 1×3 matrix

$$B_{\mathbf{x}} = \begin{bmatrix} -2x_1 & -2y_1 & -2z_1 \end{bmatrix} \in \mathfrak{so}_3(\mathbb{R})^* \cong (\mathbb{R}^3)^*$$

and the map $\mathfrak{so}_3(\mathbb{R}) \cong \mathbb{R}^3 \to \mathfrak{so}_3(\mathbb{R})^* \cong (\mathbb{R}^3)^*$ induced by the Killing form is $\mathbf{x} \mapsto -2\mathbf{x}^T$. Note that this map is an isomorphism, which is in fact always true in the semisimple case: the fact that the Killing form is non-degenerate says precisely that the map $X \mapsto B_X$ has trivial kernel, so it is injective and hence surjective since $\mathfrak{g}$ and $\mathfrak{g}^*$ have the same dimension.

When $\mathfrak{g}$ is the Lie algebra of a group G, the map $\mathfrak{g} \to \mathfrak{g}^*$ induced by the Killing form is equivariant for the adjoint and coadjoint actions of G, which means that acting by G on $\mathfrak{g}$ first and then applying this map is the same as applying the map first and then acting by G:

$$B_{Ad_g X} = Ad_g^* B_X$$

for all $g \in G$. If we unwind the definitions, this equality says that

$$B(Ad_X, Y) = B(X, Ad_{g^{-1}}Y)$$

for all $X, Y \in \mathfrak{g}$, where the left side is the functional $B_{Ad_g}X$ evaluated at Y and the right side is the functional $Ad_g^* B_x$ evaluated at Y.

Exercise. Justify this equivariance property, which is actually just a reformulation of the Ad-invariance property of the Killing form mentioned previously.

Equivariance in particular implies that any adjoint orbit in $\mathfrak{g}$ is sent into a coadjoint orbit in $\mathfrak{g}^*$ under the map $B : \mathfrak{g} \to \mathfrak{g}^*$, since the image an element $Ad_g X$ in the adjoint orbit of X is the same as the element $Ad_g^* B_x$ in the coadjoint orbit of B_X. Thus when this map B is invertible, in other words when we are in the semisimple case, this map induces a bijection (diffeomorphism, actually) between the adjoint orbits and coadjoint orbits, which explains what we saw previously for $\mathfrak{so}_3(\mathbb{R})$ and $\mathfrak{sl}_3(\mathbb{R})$. For $\mathfrak{so}_3(\mathbb{R}) \cong \mathbb{R}^3$, the map $\mathbb{R}^3 \to (\mathbb{R}^3)^* \cong \mathbb{R}^3$ (identity the dual of $\mathbb{R}^3$ with $\mathbb{R}^3$ using the dual bases) induced by the Killing form is $\mathbf{x} \mapsto -2\mathbf{x}$, which indeed sends spheres (i.e., adjoint orbits) to spheres (i.e., coadjoints orbits).

Exercise. Using the Killing form on $\mathfrak{sl}_2(\mathbb{R}) \cong \mathbb{R}^3$ computed previously and identifying $\mathfrak{sl}_2(\mathbb{R})^*$ also with $\mathbb{R}^3$ using the dual bases, verify that an adjoint orbit $yz + x^2 = k$ is sent to a coadjoint orbit $4yz + x^2 = \ell$ under the map induced by the Killing form.

In the Heisenberg case, all adjoint orbits end up being sent to a single coadjoint orbit, namely the coadjoint orbit of the zero vector, which consists of only the zero vector. (The result of the exercise asking to compute the Killing form on the Heisenberg algebra is that the Killing form is zero on all vectors.) Hence we should not expect any relation between the adjoint and coadjoint orbits of the Heisenberg algebra, just as we saw previously.

Mechanics

We finish by giving the motivation behind the subject of linear Poisson geometry, and why one might be led to study coadjoint orbits and symplectic leaves in the first place. The starting point is Hamilton's formulation of Newtonian mechanics. We take the standard Poisson bracket on $\mathbb{R}^{2n}$ with coordinates $(x_1, \dots, x_n, y_1, \dots, y_n)$:

$$\{f, g\} = \sum_i \left(\frac{\partial f}{\partial x_1} \frac{\partial g}{\partial y_i} - \frac{\partial f}{\partial y_i} \frac{\partial g}{\partial x_i} \right).$$

We consider $\mathbb{R}^{2n} = \mathbb{R}^n \times \mathbb{R}^n$ as a *phase space* of a mechanical system, where $\mathbf{x} = (x_1, \dots, x_n)$ describes position and $\mathbf{y} = (y_1, \dots, y_n)$ momentum. The basic tenet of Hamiltonian mechanics is that an observable quantity (i.e., function) f of position and momentum evolves in time along the trajectories of our system according to the differential equation

$$\dot{f} = \{f, H\}$$

where $\dot{f}$ denotes the time-derivative of f and where H is the *Hamiltonian function* of the system, which is usually interpreted as being the total energy of the system:

$$H = (\text{kinetic energy}) + (\text{potential energy}).$$

Let's use the standard total energy $H(\mathbf{x}, \mathbf{y}) = \frac{\mathbf{y} \cdot \mathbf{y}}{2m} + V(\mathbf{x})$ where $V(\mathbf{x})$ is some potential energy.

If we want to know, say, how position changes with respect to time, the equation above says that this rate of change is given by Poisson-bracketing with the Hamiltonian function. Similarly for determining how momentum changes, so we get the equations

$$\dot{x}_i = \{x_i, H\} = \frac{\partial H}{\partial y_i} \quad \text{and} \quad \dot{y}_i = \{y_i, H\} = -\frac{\partial H}{\partial x_i},$$

which are known as *Hamilton's equations.* For the specific Hamiltonian $H = \frac{\mathbf{y} \cdot \mathbf{y}}{2m} + V(\mathbf{x})$, we have $\frac{\partial H}{\partial y_i} = \frac{y_i}{m}$ and $\frac{\partial H}{\partial x_i} = \frac{\partial V}{\partial x_i}$, so Hamilton's equations become

$$\dot{x}_i = \frac{y_i}{m} \quad \text{and} \quad \dot{y}_i = \frac{\partial V}{\partial x_i}.$$

From this we can compute acceleration $\ddot{x}_i$ by substituting the second equation into the time-derivative of the first:

$$\ddot{x}_i = \frac{\dot{y}_i}{m} = \frac{1}{m} \frac{\partial V}{\partial x_i}.$$

Rewriting this as $m\ddot{x}_i = \frac{\partial V}{\partial x_i}$ gives Newton's second law of motion, where $\frac{\partial V}{\partial x_i}$ is the force acting on our system. Thus, we see that Hamiltonian mechanics does indeeed reproduce Newtonian mechanics.

We say that f and g *Poisson commute* if $\{f, g\} = 0$. If f Poisson commutes with the Hamiltonian H, then $\dot{f} = \{f, H\} = 0$, which means that f should not change values along the trajectories of our system, so that f should be constant along the trajectories. (We say that f is a *conserved quantity.*) Thus, the underliyng mechanics should take place on a level set of f. Functions which Poisson commute with all other functions (these are called *Casimir functions*) in particular Poisson commute with the Hamiltonian, so the symplectic leaves (i.e., level sets of Casimir functions) we saw earlier describe spaces on which Hamiltonian mechanics occurs. The non-degeneracy of the Poisson bracket we get on such leaves has useful consequences in physics, and the fact that these leaves correspond to the orbits of a certain group action opens up the use of group theory to their study. Such coadjoint orbits/leaves also arise naturally via a type of "reduction" procedure, but that is a story for another time.