

(March 15, 2006)

Invariant differential operators

Paul Garrett garrett@math.umn.edu <http://www.math.umn.edu/~garrett/>

- Derivatives of group actions: Lie algebras
- Laplacians and Casimir operators
- Descending to G/K
- Example computation: $SL_2(\mathbb{R})$
- Enveloping algebras and adjoint functors
- Appendix: brackets
- Appendix: proof of Poincaré-Birkhoff-Witt

We want an intrinsic approach to proving existence of differential operators invariant under group actions.

The translation-invariant operators $\frac{\partial}{\partial x_i}$ on \mathbb{R}^n , and the rotation-invariant Laplacian on \mathbb{R}^n are *deceptively easily* proven invariant, in the sense that these examples provide few clues about more complicated situations.

For example, we expect rotation-invariant *Laplacians* (second-order operators) on spheres, and we do *not* want to write a formula in generalized spherical coordinates and verify computationally that it is invariant. Nor do we want to be constrained to imbedding spheres in Euclidean spaces and using the ambient spaces' geometries, even though by chance this does succeed for *spheres* themselves.

As another basic example, it is often asserted that the operator

$$y^2 \left(\left(\frac{\partial}{\partial x} \right)^2 + \left(\frac{\partial}{\partial y} \right)^2 \right)$$

on the complex upper half-plane \mathfrak{H} is invariant under the linear fractional action of $SL_2(\mathbb{R})$, but it is oppressive to verify this directly. Worse, our goals ought not be merely to *verify* an expression presented as a *deus ex machina*, but, rather, be able to systematically *generate* suitable expressions. An important part of this intention is that we understand *reasons* for the existence of invariant operators, and derivability of expressions in coordinates should be a justifiably foregone conclusion.

(No prior acquaintance with *Lie groups* or *Lie algebras* is assumed.)

1. Derivatives of group actions: Lie algebras

For example, as usual we let

$$SO_n(\mathbb{R}) = \{k \in GL_n(\mathbb{R}) : k^\top k = 1_n, \det k = 1\}$$

act on functions f on the sphere $S^{n-1} \subset \mathbb{R}^n$, by

$$(k \cdot f)(m) = f(mk)$$

with $m \times k \rightarrow mk$ being *right* matrix multiplication of the *row* vector $m \in \mathbb{R}^n$. This action is relatively easy to understand because it is *linear*.

The linear fractional transformation action

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} (z) = \frac{az + b}{cz + d} \quad (\text{for } z \in \mathfrak{H} \text{ and } \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in SL_2(\mathbb{R}))$$

of $SL_2(\mathbb{R})$ ^[1] on the complex upper half-plane \mathfrak{H} is superficially more complicated, but, as we saw earlier, is descended from the *linear* action of $GL_2(\mathbb{C})$ on \mathbb{C}^2 , which induces an action on $\mathbb{P}^1 \supset \mathbb{C} \supset \mathfrak{H}$.

[1] Recall the standard notation that $GL_n(R)$ is n -by- n invertible matrices with entries in a commutative ring R , and $SL_n(R)$ is the subgroup of $GL_n(R)$ consisting of matrices with determinant 1.

Abstracting this a little,^[2] let G be a subgroup of $GL(n, \mathbb{R})$ acting *differentiably*^[3] on the *right* on a subset M of \mathbb{R}^n ,^[4] thereby acting on *functions* f on M by

$$(g \cdot f)(m) = f(mg)$$

Define the (real) **Lie algebra**^[5] \mathfrak{g} of G by

$$\mathfrak{g} = \{\text{real } n\text{-by-}n \text{ real matrices } x : e^{tx} \in G \text{ for all real } t\}$$

where

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

is the usual exponential,^[6] now applied to matrices.^[7]

Remark: A moment's reflection yields the fact that (with this definition) Lie algebras are closed under scalar multiplication. But at this point, it is not clear that the Lie algebra is closed under addition. When x and y are n -by- n real or complex matrices which *commute*, that is, such that $xy = yx$, then

$$e^{x+y} = e^x \cdot e^y \quad (\text{when } xy = yx)$$

from which we could conclude that $x+y$ is again in a Lie algebra containing x and y . But the general case of closed-ness under addition is much less obvious. We will prove it as a side effect of proving (in an appendix) that the Lie algebra is closed under *brackets*. In any particular example the vector space property is readily verified, as just below.

Remark: These Lie algebras will prove to be \mathbb{R} -vectorspaces with a \mathbb{R} -bilinear operation, $x \times y \rightarrow [x, y]$, which is why they are called *algebras*. However, this binary operation is different from more typical ring or algebra multiplications, at least in that it is not *associative*.

Example: The condition $e^{tx} \in SO_n(\mathbb{R})$ ^[8] for all real t is that

$$1_n = (e^{tx})^\top e^{tx}$$

-
- [2] A fuller abstraction, which is not strictly necessary for illustration of our construction of invariant operators, is that G should be a *Lie group* acting *smoothly* and transitively on a *smooth manifold* M . And for the later parts of this discussion G should be *semi-simple*, or *reductive*. Happily, our introductory discussion of invariant differential operators does not require immediate concern for these notions.
- [3] When the group G and the set M are subsets of Euclidean spaces defined as zero sets or level sets of some differentiable functions, the sense of *differentiability* of the action can be posed in terms of the ambient Euclidean coordinates and the Implicit Function Theorem. In any particular example, even less is usually required to make sense of this requirement.
- [4] As in previous situations where a group acts transitively on a set with additional structure, with modest hypotheses we can view M as a quotient $G_o \backslash G$ of G by the isotropy group G_o of a chosen point in M .
- [5] Named after Sophus Lie, pronounced in English *lee*, not *lie*.
- [6] Even for *Lie groups* not imbedded in matrix groups, there is an *intrinsic* notion of *exponential map*. However, it is more technically expensive than we can justify, since the matrix exponential is essentially all we need.
- [7] We are not attempting to precisely define the class of groups G for which it is appropriate to care about the Lie algebra in this sense. For groups G inside $GL_n(\mathbb{R})$, we would surely require that the exponential $\mathfrak{g} \rightarrow G$ is a surjection to a neighborhood of 1_n in G . This is essentially the same thing as requiring that G be a *smooth* submanifold of $GL_n(\mathbb{R})$.
- [8] One may observe that the determinant-one condition does not play a role. That is, the Lie algebra of $O_n(\mathbb{R})$ is the same as that of $SO_n(\mathbb{R})$. More structurally, the reason for this is that $SO_n(\mathbb{R})$ is the (topologically) connected component of $O_n(\mathbb{R})$ containing the identity, so *any* definition of Lie algebra will give the same outcome for both.

Taking derivatives of both sides with respect to t , this is

$$0 = x^\top (e^{tx})^\top e^{tx} + (e^{tx})^\top x e^{tx}$$

Evaluating at $t = 0$ gives

$$0 = x^\top + x$$

so it is *necessary* that $x^\top = -x$. In fact, assuming this,

$$\begin{aligned} (e^{tx})^\top e^{tx} &= (1 + tx + t^2 x^2/2 + \dots)^\top e^{tx} = (1 + tx^\top + t^2 (x^\top)^2/2 + \dots) e^{tx} \\ &= (1 - tx + t^2 (-x)^2/2 + \dots) e^{tx} = e^{-tx} e^{tx} = e^0 = 1 \end{aligned}$$

That is, the condition $x^\top = -x$ is necessary and sufficient, and

$$\text{Lie algebra of } SO_2(\mathbb{R}) = \{x : x^\top = -x\} \quad (\text{denoted } \mathfrak{so}(n))$$

Example: The condition $e^{tx} \in SL_n(\mathbb{R})$ for all real t is that

$$\det(e^{tx}) = 1$$

To see what this requires of x , we observe that for n -by- n (real or complex) matrices x

$$\det(e^x) = e^{\text{tr } x} \quad (\text{where tr is trace})$$

To see why this is so, note that both determinant and trace are invariant under conjugation $x \rightarrow gxg^{-1}$, so we can suppose without loss of generality that x is *upper-triangular*.^[9] Then e^x is still upper-triangular, with diagonal entries $e^{x_{ii}}$, where the x_{ii} are the diagonal entries of x . Thus,

$$\det(e^x) = e^{x_{11}} \dots e^{x_{nn}} = e^{x_{11} + \dots + x_{nn}} = e^{\text{tr } x}$$

Using this claim, the determinant-one condition is

$$1 = \det(e^{tx}) = e^{t \cdot \text{tr } x} = 1 + t \cdot \text{tr } x + \frac{(t \cdot \text{tr } x)^2}{2!} + \dots$$

Intuitively, since t is arbitrary, *surely* $\text{tr } x = 0$ is the necessary and sufficient condition.^[10] Thus,

$$\text{Lie algebra of } SL_n(\mathbb{R}) = \{x \text{ } n\text{-by-}n \text{ real} : \text{tr } x = 0\} \quad (\text{denoted } \mathfrak{sl}_n(\mathbb{R}))$$

Example: From the identity $\det(e^x) = e^{\text{tr } x}$, any matrix e^x is invertible. Thus,

$$\text{Lie algebra of } GL_n(\mathbb{R}) = \{\text{all real } n\text{-by-}n \text{ matrices}\} \quad (\text{denoted } \mathfrak{gl}_n(\mathbb{R}))$$

[9] The existence of *Jordan normal form* of a matrix over an algebraically closed field more than suffices to show that any matrix can be conjugated (over the algebraic closure) to an upper-triangular matrix. But the assertion that a matrix x can be conjugated (over an algebraic closure) to an upper-triangular matrix is much weaker than the assertion of Jordan normal form. Indeed, it only requires that we know that there is a basis v_1, \dots, v_n for \mathbb{C}^n such that $x \cdot v_i \in \sum_{j \leq i} \mathbb{C} v_j$. That this is so follows from the fact that \mathbb{C} is algebraically closed, so there is an eigenvector v_1 . Then x induces an endomorphism of $\mathbb{C}^n / \mathbb{C} \cdot v_1$, which has an eigenvector w_2 . Let v_2 be any inverse image of w_2 in \mathbb{C}^n . Continue inductively.

[10] One choice of making the conclusion $\text{tr } x = 0$ precise is as follows. Taking the derivative in t and setting $t = 0$ gives a *necessary* condition for $\det(e^{tx}) = 1$, namely $0 = \text{tr } x$. But, looking at the right-hand side of the expanded $1 = \det(e^{tx})$, this condition is also *sufficient* for $\det(e^{tx}) = 1$.

For each $x \in \mathfrak{g}$ we have a **differentiation** X_x of functions f on M in the *direction* of x , by

$$(X_x f)(m) = \left. \frac{d}{dt} \right|_{t=0} f(m \cdot e^{tx})$$

Note that this definition applies uniformly to *any* space M on which G acts (differentiably).

These differential operators X_x for $x \in \mathfrak{g}$ do *not* typically commute with the action of $g \in G$, although the relation between the two is reasonable. ^[11]

Remark: In the extreme and simple case that the space M is G itself, there is a second action of G on itself in addition to right multiplication, namely *left* multiplication. The *right* differentiation by elements of \mathfrak{g} *does* commute with the *left* multiplication by G , for the simple reason that

$$F(h \cdot (g e^{tx})) = F((h \cdot g) \cdot e^{tx}) \quad (\text{for } g, h \in G, x \in \mathfrak{g})$$

That is, \mathfrak{g} gives left G -invariant differential operators on G . ^[12]

Claim: For $g \in G$ and $x \in \mathfrak{g}$

$$g \cdot X_x \cdot g^{-1} = X_{g x g^{-1}}$$

Proof: This is essentially a direct computation. For a smooth function f on M ,

$$(g \cdot X_x \cdot g^{-1} \cdot f)(m) = (g \cdot X_x \cdot f)(m g^{-1}) = \left. \frac{d}{dt} \right|_{t=0} (g \cdot f)(m e^{tx} g^{-1}) = \left. \frac{d}{dt} \right|_{t=0} f(m g e^{tx} g^{-1})$$

Then observe that conjugation and exponentiation interact well, namely

$$\begin{aligned} g e^{tx} g^{-1} &= g \left(1 + tx + \frac{(tx)^2}{2!} + \frac{(tx)^3}{3!} + \dots \right) g^{-1} \\ &= 1 + t g x g^{-1} + \frac{(t g x g^{-1})^2}{2!} + \frac{(t g x g^{-1})^3}{3!} + \dots = e^{t g x g^{-1}} \end{aligned}$$

Thus,

$$(g \cdot X_x \cdot g^{-1} \cdot f)(m) = \left. \frac{d}{dt} \right|_{t=0} f(m g e^{tx} g^{-1}) = \left. \frac{d}{dt} \right|_{t=0} f(m e^{t g x g^{-1}}) = (X_{g x g^{-1}} f)(m)$$

as claimed. ///

Note: This computation also shows that \mathfrak{g} is *stable* under conjugation ^[13] by G .

But since G is *non-abelian* in most cases of interest,

$$e^x \cdot e^y \neq e^y \cdot e^x \quad (\text{typically, for } x, y \in \mathfrak{g})$$

Specifically,

-
- [11] The conjugation action of G on \mathfrak{g} in the *claim* is an instance of what is called the **adjoint action** Ad of G on \mathfrak{g} . But in our examples it is literal conjugation.
- [12] In fact, the argument in the appendix on the closure of Lie algebras under brackets characterizes \mathfrak{g} as the collection of *all* left G -invariant first-order differential operators annihilating constants.
- [13] Again, this literal *conjugation* of matrices has an *intrinsic* description, and is more properly called the *adjoint action* of G on \mathfrak{g} .

Claim: For $x, y \in \mathfrak{g}$

$$e^{tx} e^{ty} e^{-tx} e^{-ty} = 1 + t^2[x, y] + (\text{higher-order terms}) \quad (\text{where } [x, y] = xy - yx)$$

In this context, the commutant expression $[x, y] = xy - yx$ is called the **Lie bracket**.

Proof: This is a direct and unsurprising computation, and easy if we think to drop cubic and higher-order terms.

$$\begin{aligned} e^{tx} e^{ty} e^{-tx} e^{-ty} &= (1 + tx + t^2x^2/2)(1 + ty + t^2y^2/2)(1 - tx + t^2x^2/2)(1 - ty + t^2y^2/2) \\ &= (1 + t(x + y) + \frac{t^2}{2}(x^2 + 2xy + y^2))(1 - t(x + y) + \frac{t^2}{2}(x^2 + 2xy + y^2)) \\ &= 1 + t^2(x^2 + 2xy + y^2 - (x + y)(x + y)) = (1 + t^2(2xy - xy - yx)) = 1 + t^2[x, y] \end{aligned}$$

as claimed. ///

A seemingly similar result about composition of these derivatives is

Theorem:

$$X_x X_y - X_y X_x = X_{[x, y]}$$

In fact, the proof of this theorem is non-trivial, and is given in an appendix.

The point is that the map $x \rightarrow X_x$ is a **Lie algebra homomorphism**, meaning what that it respects these commutants (brackets).

Here is a *heuristic* for the correctness of the assertion of the theorem. For simplicity, just have the group G act on *itself* on the right. ^[14] First, just computing in *matrices* (writing expressions modulo s^2 and t^2 terms, which will vanish upon application of $\frac{d}{dt}\big|_{t=0}$ and $\frac{d}{ds}\big|_{s=0}$),

$$\begin{aligned} &\frac{d}{dt}\bigg|_{t=0} \frac{d}{ds}\bigg|_{s=0} (e^{tx} e^{sy} - e^{sy} e^{tx}) \\ &= \frac{d}{dt}\bigg|_{t=0} \frac{d}{ds}\bigg|_{s=0} (1 + sy + tx + stxy + \dots) - (1 + sy + tx + styx + \dots) = xy - yx \end{aligned}$$

Now we *imagine* that it is legitimate to write something like

$$\frac{d}{dt}\bigg|_{t=0} f(m \cdot e^{tx}) = \frac{d}{dt}\bigg|_{t=0} f(m \cdot (1 + tx + O(t^2))) = \frac{d}{dt}\bigg|_{t=0} (f(m) + \nabla f(m) \cdot (tmx + O(t^2))) = \nabla f(m) \cdot mx$$

From the definition of the differential operators X_x and X_y ^[15]

$$(X_x \circ X_y - X_y \circ X_x)f(m) = \frac{\partial}{\partial t}\bigg|_{t=0} \frac{\partial}{\partial s}\bigg|_{s=0} (f(m e^{tx} e^{sy}) - f(m e^{sy} e^{tx}))$$

^[14] Indeed, when G acts *transitively* on a space M , we should expect (with some additional mild hypotheses in general) that $M \approx G_o \backslash G$ where G_o is the isotropy group of some chosen point in M . Thus, all functions on M give rise to functions on G , and any reasonable notion of invariant differential operator on the quotient should lift to G via the quotient map.

^[15] And we do presume that we can interchange the partial derivatives.

Writing the exponentials out, modulo s^2 and t^2 terms,

$$\begin{aligned} & f(m \cdot (1 + sy + tx + stxy + \dots)) - f(m \cdot (1 + sy + tx + styx + \dots)) \\ & \sim (f(m) + \nabla f(m) \cdot m(tx + sy + stxy + \dots)) - (f(m) + \nabla f(m) \cdot m(tx + sy + styx + \dots)) \\ & = \nabla f(m) \cdot m(st(xy - yx) + \dots) \end{aligned}$$

When the operator $\frac{\partial}{\partial t} \Big|_{t=0} \frac{\partial}{\partial s} \Big|_{s=0}$ is applied, this gives

$$\nabla f(m) \cdot m(xy - yx)$$

as claimed. ^[16]

2. Laplacians and Casimir operators

The theorem of the last section notes that *commutants* of differential operators coming from Lie algebras \mathfrak{g} are again differential operators coming from the Lie algebra, namely

$$X_x X_y - X_y X_x = [X_x, X_y] = X_{[x, y]} = X_{xy - yx}$$

Indeed, closure under a bracket operation is a defining attribute of a Lie algebra. ^[17]

However, the *composition* of differential operators has no analogue inside the Lie algebra. That is, typically,

$$X_x X_y \neq X_\varepsilon \quad (\text{for any } \varepsilon \in \mathfrak{g})$$

But we *do* want to create something from the Lie algebra that allows us to compose in this fashion.

Remark: For Lie algebras \mathfrak{g} such as $\mathfrak{so}(n)$, \mathfrak{sl}_n , or \mathfrak{gl}_n lying inside matrix rings, typically

$$X_x X_y \neq X_{xy}$$

That is, multiplication of *matrices* is *not* multiplication in any sense that will match multiplication (composition) of *differential operators*. ^[18]

^[16] One problem with this heuristic is that there is an implicit assumption that f has an extension to the ambient space of matrices, and the computation depends on this extension, at least superficially. In fact, such extensions do exist, but that's not the point. This sort of *extrinsic* argument will cause trouble, since (for example) we cannot easily prove that it is compatible with *mappings* to other groups. See the appendix for a compelling and not-too-expensive argument.

^[17] We are cheating a little in our definition of Lie algebra by restricting our scope to matrix groups G , and by defining the Lie bracket via matrix multiplication, $[x, y] = xy - yx$. This implicitly engenders further properties which would otherwise need to be explicitly declared, such as the *Jacobi identity* $[x, [y, z]] - [y, [x, z]] = [[x, y], z]$. For matrices x, y, z this can be verified directly by expanding the brackets. The general definition of *Lie algebra* explicitly requires this relation. The *content* of this identity is that the map $\text{ad} : \mathfrak{g} \rightarrow \text{End}(\mathfrak{g})$ by $(\text{ad}x)(y) = [x, y]$ is a Lie algebra *homomorphism*. That is, $[\text{ad}x, \text{ad}y] = \text{ad}[x, y]$.

^[18] Many years ago, it was disturbing to me that that matrix multiplication is not the correct multiplication to match composition of associated differential operators!

What we *want* is an *associative algebra* ^[19] $U(\mathfrak{g})$ which is *universal* in the sense that any linear map $\varphi : \mathfrak{g} \rightarrow A$ to an associative algebra A respecting brackets

$$\varphi([x, y]) = \varphi(x)\varphi(y) - \varphi(y)\varphi(x) \quad (\text{for } x, y \in \mathfrak{g})$$

should give a (unique) *associative algebra homomorphism*

$$\Phi : U(\mathfrak{g}) \rightarrow A$$

In fact, we realize that there must be a connection to the original $\varphi : \mathfrak{g} \rightarrow A$, so we should require existence of a fixed map $i : \mathfrak{g} \rightarrow U(\mathfrak{g})$ *respecting brackets* and commutativity of a diagram

$$\begin{array}{ccc} & U(\mathfrak{g}) & \\ & \uparrow i \text{ (Lie)} & \searrow \Phi \text{ (assoc)} \\ \mathfrak{g} & \xrightarrow{\varphi \text{ (Lie)}} & A \end{array}$$

where the labels tell the *type* of the maps.

A related construction is the **tensor algebra** $\otimes^\bullet V$ of a vector space V over a field k , with a specified linear $j : V \rightarrow \otimes^\bullet V$. (The name is a description of the construction, not of its properties!) The defining property is that any linear map $V \rightarrow A$ to an (associative) algebra A extends to a unique (associative) algebra map $\otimes^\bullet V \rightarrow A$. That is, there is a diagram

$$\begin{array}{ccc} & \otimes^\bullet V & \\ & \uparrow j \text{ (linear)} & \searrow \Phi \text{ (assoc)} \\ V & \xrightarrow{\varphi \text{ (linear)}} & A \end{array}$$

The *construction* of $\otimes^\bullet V$ is

$$\otimes^\bullet V = k \oplus V \oplus (V \otimes V) \oplus (V \otimes V \otimes V) \oplus \dots$$

with multiplication given by (the bilinear extension of) the obvious

$$(v_1 \otimes \dots \otimes v_m) \cdot (w_1 \otimes \dots \otimes w_n) = v_1 \otimes \dots \otimes v_m \otimes w_1 \otimes \dots \otimes w_n$$

Since the tensor algebra $\otimes^\bullet \mathfrak{g}$ is universal with respect to maps $\mathfrak{g} \rightarrow A$ that are merely *linear*, not necessarily preserving the Lie brackets, we expect that there is a (unique) natural (quotient) map $q : \otimes^\bullet \mathfrak{g} \rightarrow U(\mathfrak{g})$. Indeed, the fixed $i : \mathfrak{g} \rightarrow U(\mathfrak{g})$ is a map to an associative algebra, so there is induced a (unique) associative algebra map q giving a commutative diagram

$$\begin{array}{ccc} & \otimes^\bullet \mathfrak{g} & \\ & \uparrow j \text{ (linear)} & \searrow q \text{ (assoc)} \\ \mathfrak{g} & \xrightarrow{i \text{ (linear)}} & U(\mathfrak{g}) \end{array}$$

^[19] For present purposes, all *algebras* are either \mathbb{R} - or \mathbb{C} -algebras, as opposed to using some more general field, or a ring. An *associative algebra* is what would often be called simply an *algebra*, but since Lie algebras are *not* associative, we have to adjust the terminology to enable ourselves to talk about them. So an associative algebra is one whose multiplication is associative, namely $a(bc) = (ab)c$. Addition is associative and commutative, and multiplication distributes over addition, both on the left and on the right.

Remark: *Existence* of the universal enveloping algebra (and comparable treatment of the tensor algebra) will be given in the next section. For the moment, we want to see the *application* to construction of invariant operators.

Likewise, the conjugation action $x \longrightarrow gxg^{-1}$ should extend to an action of G on $U(\mathfrak{g})$ (which we'll still write as conjugation) compatible with the multiplication in $U(\mathfrak{g})$. That is, we *require*

$$\begin{aligned} g(\alpha) &= g\alpha g^{-1} && (\text{for } \alpha \in \mathfrak{g} \text{ and } g \in G) \\ g(\alpha\beta) &= g(\alpha) \cdot g(\beta) && (\text{for } \alpha, \beta \in U(\mathfrak{g}) \text{ and } g \in G) \end{aligned}$$

This condition could be met (as we'll see in the next section) by taking the *obvious* G -conjugation^[20] on $\otimes^{\bullet} \mathfrak{g}$ given by

$$g(x_1 \otimes \dots \otimes x_m)g^{-1} = gx_1g^{-1} \otimes \dots \otimes gx_mg^{-1}$$

showing that the kernel of $\otimes^{\bullet} \mathfrak{g} \longrightarrow U(\mathfrak{g})$ is G -stable, thus inducing a natural action of G on $U(\mathfrak{g})$.

The last item we need is more special, and is not possessed by all Lie algebras. We want a *non-degenerate* symmetric \mathbb{R} -bilinear map

$$\langle \cdot, \cdot \rangle : \mathfrak{g} \times \mathfrak{g} \longrightarrow \mathbb{R}$$

which is G -equivariant in the sense that

$$\langle gxg^{-1}, gxg^{-1} \rangle = \langle x, y \rangle$$

Happily, for $\mathfrak{so}(n)$, $\mathfrak{sl}_n(\mathbb{R})$, and $\mathfrak{gl}_n(\mathbb{R})$, the obvious guess

$$\langle x, y \rangle = \text{tr}(xy)$$

suffices.^[21] The non-degeneracy and G -equivariance of $\langle \cdot, \cdot \rangle$ give a natural G -equivariant isomorphism $\mathfrak{g} \longrightarrow \mathfrak{g}^*$ by

$$x \longrightarrow \lambda_x \quad \text{by} \quad \lambda_x(y) = \langle x, y \rangle \quad (\text{for } x, y \in \mathfrak{g})$$

Note that when a group G acts on a vector space V the action on the dual V^* is by

$$(g \cdot \lambda)(v) = \lambda(g^{-1} \cdot v) \quad (\text{for } v \in V \text{ and } \lambda \in V^*)$$

where the inverse appears (as usual!) to preserve *associativity*. Then the equivariance of $\langle \cdot, \cdot \rangle$ gives

$$\lambda_{g \cdot x}(y) = \lambda_{gxg^{-1}}(y) = \langle gxg^{-1}, y \rangle = \langle x, g^{-1}yg \rangle = \lambda_x(g^{-1}yg) = \lambda_x(g^{-1} \cdot y) = (g \cdot \lambda_x)(y)$$

proving that the map $x \longrightarrow \lambda_x$ is a G -isomorphism.

We should recall one more thing, the natural isomorphism

$$V \otimes_k V^* \xrightarrow{\text{isom}} \text{End}_k V \quad (V \text{ a finite-dimensional vector space over a field } k)$$

[20] Again, G -conjugation on \mathfrak{g} is really a more intrinsic thing, and is called the *adjoint* action.

[21] There is an *intrinsic* way to construct such a G -equivariant symmetric bilinear form on any (real) Lie algebra $\mathfrak{g} = \text{Lie}(G)$, by first defining $(\text{adx})(y) = [x, y]$ (this is the *adjoint action* of \mathfrak{g} on itself) and taking $\langle x, y \rangle = \text{tr}(\text{adx} \circ \text{ady})$. This is the **Killing form**, named after Wilhelm Killing (*not* because it *kills* anything). Up to a normalization, the more direct trace-of-matrix definition we've given here is the same, though there's little reason to worry about verifying this. The fact that this bilinear form is *non-degenerate* on Lie algebras of interest to us capsulizes in an important technical fashion some virtues of these Lie algebras. Indeed, **Cartan's criterion** for *semi-simplicity* (whose definition we'll postpone) of \mathfrak{g} is exactly that the Killing form be non-degenerate. That is, we are indirectly and covertly using an intrinsic and abstract-able aspect of our tangible Lie algebras. This is good, since it means that the ideas here are applicable very broadly.

given by the k -linear extension of the map

$$(v \otimes \lambda)(w) = \lambda(w) \cdot v \quad (\text{for } v, w \in V \text{ and } \lambda \in V^*)$$

The fact that the map is an isomorphism follows by dimension counting, using the finite-dimensionality. [22]

Now we can construct the simplest non-trivial G -invariant element in $U(\mathfrak{g})$, the **Casimir element**. As above, under any (smooth) action of G on a smooth manifold the Casimir element gives rise to a G -invariant differential operator, a **Casimir operator**. In many familiar situations this differential operator is the suitable notion of **invariant Laplacian**.

Consider the map $\zeta : \text{End}_{\mathbb{C}}(\mathfrak{g}) \rightarrow U(\mathfrak{g})$ defined by

$$\begin{array}{ccccccc} \text{End}_{\mathbb{C}}(\mathfrak{g}) & \xrightarrow{\text{natural } \approx} & \mathfrak{g} \otimes \mathfrak{g}^* & \xrightarrow{\approx \text{ via } \langle, \rangle} & \mathfrak{g} \otimes \mathfrak{g} & \xrightarrow{\text{inclusion}} & \bigotimes^{\bullet} \mathfrak{g} & \xrightarrow{\text{quotient}} & U(\mathfrak{g}) \\ & & & & & & & & \nearrow \zeta \end{array}$$

An obvious choice of endomorphism of \mathfrak{g} commuting with the action of G on \mathfrak{g} is the *identity map* $\text{id}_{\mathfrak{g}}$.

Claim: The **Casimir element** $\Omega = \zeta(\text{id}_{\mathfrak{g}})$ is a G -invariant element of $U(\mathfrak{g})$.

Proof: Since ζ is G -equivariant by construction,

$$g\zeta(\text{id}_{\mathfrak{g}})g^{-1} = \zeta(g \text{id}_{\mathfrak{g}} g^{-1}) = \zeta(g g^{-1} \text{id}_{\mathfrak{g}}) = \zeta(\text{id}_{\mathfrak{g}})$$

since $\text{id}_{\mathfrak{g}}$ commutes with *anything*. Thus, $\zeta(\text{id}_{\mathfrak{g}})$ is a G -invariant element of $U(\mathfrak{g})$. ///

Remark: The slight hitch is that we don't have a simple way to show that $\zeta(\text{id}_{\mathfrak{g}}) \neq 0$. This is a corollary of a surprisingly serious result, the *Poincaré-Birkhoff-Witt* theorem, proven in an appendix.

This prescription *does* tell us how to compute the Casimir element $\Omega = \zeta(\text{id}_{\mathfrak{g}})$ in *various* coordinates. Namely, for any basis x_1, \dots, x_n of \mathfrak{g} , let x_1^*, \dots, x_n^* be the corresponding *dual basis*, meaning as usual that

$$\langle x_i, x_j^* \rangle = \begin{cases} 1 & (\text{for } i = j) \\ 0 & (\text{for } i \neq j) \end{cases}$$

Then $\text{id}_{\mathfrak{g}}$ maps to $\sum_i x_i \otimes x_i^*$ in $\mathfrak{g} \otimes \mathfrak{g}$, which imbeds in $\bigotimes^{\bullet} \mathfrak{g}$, and then by the quotient map is sent to $U(\mathfrak{g})$.

Remark: The *intrinsic* description of the Casimir element as $\zeta(\text{id}_{\mathfrak{g}})$ shows that it does not depend upon the choice of basis x_1, \dots, x_n . [23]

3. Descending to G/K

[22] The dimension of $V \otimes_k V^*$ is $(\dim_k V)(\dim_k V^*)$, which is $(\dim_k V)^2$, the same as the dimension of $\text{End}_k V$. To see that the map is *injective*, suppose $\sum_i (v_i \otimes \lambda_i)(w) = 0$ for all $w \in V$, with the v_i linearly independent (without loss of generality), and none of the λ_i the 0 functional. Then, by the definition, $\sum_i \lambda_i(w) \cdot v_i = 0$. This vanishing for all w would assert linear dependence relation(s) among the v_i , since none of the λ_i is the 0 functional. Since the spaces are finite-dimensional and of the same dimension, a linear injection is an isomorphism. This argument fails for infinite-dimensional spaces, and the conclusion is false for infinite-dimensional spaces.

[23] Some sources *define* the Casimir element as the element $\sum_i x_i x_i^*$ in the universal enveloping algebra, show by computation that it is G -invariant, and show by change-of-basis that the defined object is independent of the choice of basis. That element $\sum_i x_i x_i^*$ is of course the image in $U(\mathfrak{g})$ of the tensor $\sum_i x_i \otimes x_i^*$ (discussed here) which is simply the image, in coordinates, of $\text{id}_{\mathfrak{g}}$.

Now we see how the Casimir operator Ω on G gives G -invariant Laplacian-like differential operators on quotients G/K , such as $SL_2(\mathbb{R})/SO_2(\mathbb{R}) \approx \mathfrak{H}$. The pair $G = SL_n(\mathbb{R})$ and $K = SO_n(\mathbb{R})$ is a prototypical example. Let $\mathfrak{k} \subset \mathfrak{g}$ be the Lie algebra of K .^[24]

Again, the action of $x \in \mathfrak{g}$ on the *right* on functions F on G , by

$$(x \cdot f)(g) = \left. \frac{d}{dt} \right|_{t=0} F(g e^{tx})$$

is *left* G -invariant for the straightforward reason that

$$F(h \cdot (g e^{tx})) = F((h \cdot g) \cdot e^{tx}) \quad (\text{for } g, h \in G, x \in \mathfrak{g})$$

For a (closed) subgroup K of G let $q : G \rightarrow G/K$ be the quotient map. A function f on G/K gives the right K -invariant function $F = f \circ q$ on G . Given $x \in \mathfrak{g}$, the differentiation

$$(x \cdot (f \circ q))(g) = \left. \frac{d}{dt} \right|_{t=0} (f \circ q)(g e^{tx})$$

makes sense. *However*, $x \cdot (f \circ q)$ is *not* usually right K -invariant. Indeed, the condition for right K -invariance is

$$\left. \frac{d}{dt} \right|_{t=0} F(g e^{tx}) = (x \cdot F)(g) = (x \cdot F)(gk) = \left. \frac{d}{dt} \right|_{t=0} F(gk e^{tx}) \quad (k \in \mathfrak{k})$$

Using the right K -invariance of $F = f \circ q$,

$$F(gk e^{tx}) = F(gk e^{tx} k^{-1} k) = F(g e^{t \cdot kxk^{-1}})$$

Thus, unless $kxk^{-1} = x$ for all $k \in K$ it is unlikely that $x \cdot F$ is still right K -invariant. That is, the left G -invariant differential operators coming from \mathfrak{g} usually do *not* descend to differential operators on G/K .

The differential operators in

$$Z(\mathfrak{g}) = \{\alpha \in U(\mathfrak{g}) : g\alpha g^{-1}\}$$

do descend to G/K , exactly because of the commutation property, as follows. For any function φ on G let $(k \cdot \varphi)(g) = \varphi(gk)$. For F right K -invariant on G , for $\alpha \in Z(\mathfrak{g})$ we compute directly

$$k \cdot (\alpha \cdot F) = \alpha \cdot (k \cdot F) = \alpha \cdot F$$

showing the right K -invariance of $\alpha \cdot F$. Thus, $\alpha \cdot F$ gives a well-defined function on G/K .

4. Example computation: $SL_2(\mathbb{R})$

Here we compute Casimir operators in coordinates in the simplest examples.

Let $\mathfrak{g} = \mathfrak{sl}_2(\mathbb{R})$, the Lie algebra of the group $G = SL_2(\mathbb{R})$. A typical choice of basis for \mathfrak{g} is

$$H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad Y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

These have the easily verified relations

$$[H, X] = HX - XH = 2X \quad [H, Y] = HY - YH = -2Y \quad [X, Y] = XY - YX = H$$

^[24] It is implicit that K is a Lie group in the sense that it has a Lie algebra. This is visibly verifiable for the explicit examples mentioned.

Use the pairing

$$\langle v, w \rangle = \text{tr}(vw) \quad (\text{for } v, w \in \mathfrak{g})$$

To prove that this is *non-degenerate*, use the fact that \mathfrak{g} is stable under *transpose* $v \longrightarrow v^\top$, and then

$$\langle v, v^\top \rangle = \text{tr}(vv^\top) = 2a^2 + b^2 + c^2 \quad (\text{for } v = \begin{bmatrix} a & b \\ c & -a \end{bmatrix})$$

We easily compute that

$$\langle H, H \rangle = 2 \quad \langle H, X \rangle = 0 \quad \langle H, Y \rangle = 0 \quad \langle X, Y \rangle = 1$$

Thus, for the basis H, X, Y we have *dual* basis $H^* = H/2$, $X^* = Y$, and $Y^* = X$, and in these coordinates the Casimir operator is

$$\Omega = HH^* + XX^* + YY^* = \frac{1}{2}H^2 + XY + YX \quad (\text{now inside } U(\mathfrak{g}))$$

Since $XY - YX = H$ ^[25] the expression for Ω can be rewritten in various useful forms, such as

$$\Omega = \frac{1}{2}H^2 + XY + YX = \frac{1}{2}H^2 + XY - YX + 2YX = \frac{1}{2}H^2 + H + 2YX$$

and, similarly,

$$\Omega = \frac{1}{2}H^2 + XY + YX = \frac{1}{2}H^2 + XY - (-YX) = \frac{1}{2}H^2 + 2XY - (XY - YX) = \frac{1}{2}H^2 + 2XY - H$$

To make a G -invariant differential operator on the upper half-plane \mathfrak{H} , we use the G -space isomorphism $\mathfrak{H} \approx G/K$ where $K = SO_2(\mathbb{R})$ is the isotropy group of the point $i \in \mathfrak{H}$. Let $q : G \longrightarrow G/K$ be the quotient map

$$q(g) = gK \longleftrightarrow g(i)$$

A function f on \mathfrak{H} naturally yields the right K -invariant function $f \circ q$

$$(f \circ q)(g) = f(g(i)) \quad (\text{for } g \in G)$$

As above, for any $z \in \mathfrak{g}$ there is the corresponding left G -invariant differential operator on a function F on G by

$$(z \cdot F)(g) = \left. \frac{d}{dt} \right|_{t=0} F(g e^{tz})$$

but these linear operators should not be expected to descend to operators on G/K . But elements such as the Casimir operator Ω in $Z(\mathfrak{g})$ *do* descend.

We can simplify the computation of Ω on $f \circ q$ by using the right K -invariance of $f \circ q$, which implies that $f \circ q$ is annihilated by

$$\mathfrak{so}_2(\mathbb{R}) = \text{Lie algebra of } SO_2(\mathbb{R}) = \text{skew-symmetric 2-by-2 real matrices} = \left\{ \begin{bmatrix} 0 & t \\ -t & 0 \end{bmatrix} : t \in \mathbb{R} \right\}$$

In terms of the basis H, X, Y above, this says that $X - Y$ annihilates $f \circ q$.

[25] The identity $XY - YX = H$ holds in *both* the universal enveloping algebra *and* as matrices.

Among other possibilities, a point $z = x + iy \in \mathfrak{H}$ is the image

$$x + iy = (n \cdot m)(i) \quad \text{where} \quad n_x = \begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix} \quad m_y = \begin{bmatrix} \sqrt{y} & 0 \\ 0 & \frac{1}{\sqrt{y}} \end{bmatrix}$$

These are convenient group elements because they fit well with the exponentiated Lie algebra elements:

$$e^{tX} = n_t \quad e^{tH} = m_{e^{2t}}$$

By contrast, the exponentiated Y has a more complicated action on \mathfrak{H} . This suggests invocation of the fact that $X - Y$ acts trivially on right K -invariant functions on G . That is, the action of Y is the same as the action of X on right K -invariant functions. Then for right K -invariant F on G we compute

$$\begin{aligned} (\Omega F)(n_x m_y) &= \left(\frac{H^2}{2} + XY + YX\right)F(n_x m_y) = \left(\frac{H^2}{2} + XY + YX\right)F(n_x m_y) \\ &= \left(\frac{H^2}{2} + 2XY - H\right)F(n_x m_y) = \left(\frac{H^2}{2} + 2X^2 - H\right)F(n_x m_y) \end{aligned}$$

Compute the pieces separately. First, using the identity

$$m_y n_t = (m_y n_t m_y^{-1}) m_y = \begin{bmatrix} \sqrt{y} & 0 \\ 0 & \frac{1}{\sqrt{y}} \end{bmatrix} \begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{y} & 0 \\ 0 & \frac{1}{\sqrt{y}} \end{bmatrix}^{-1} m_y = n_{yt} m_y$$

we compute the effect of X

$$(X \cdot F)(n_x m_y) = \left. \frac{d}{dt} \right|_{t=0} F(n_x m_y n_t) = \left. \frac{d}{dt} \right|_{t=0} F(n_x n_{yt} m_y) = \left. \frac{d}{dt} \right|_{t=0} F(n_{x+yt} m_y) = y \frac{\partial}{\partial x} F(n_x m_y)$$

Thus, the term $2X^2$ gives

$$2X^2 \longrightarrow 2\left(y \frac{\partial}{\partial x}\right)^2 = 2y^2 \left(\frac{\partial}{\partial x}\right)^2$$

The action of H is

$$(H \cdot F)(n_x m_y) = \left. \frac{d}{dt} \right|_{t=0} F(n_x m_y m_{e^{2t}}) = \left. \frac{d}{dt} \right|_{t=0} F(n_x m_y e^{2t}) = 2y \frac{\partial}{\partial y} F(n_x m_y)$$

Then

$$\frac{H^2}{2} - H = \frac{1}{2}(2y \frac{\partial}{\partial y})^2 - (2y \frac{\partial}{\partial y}) = 2y^2 \left(\frac{\partial}{\partial y}\right)^2 + 2y \frac{\partial}{\partial y} - 2y \frac{\partial}{\partial y} = 2y^2 \left(\frac{\partial}{\partial y}\right)^2$$

Thus, altogether, on right K -invariant functions F ,

$$(\Omega F)(n_x m_y) = 2y^2 \left(\left(\frac{\partial}{\partial x}\right)^2 + \left(\frac{\partial}{\partial y}\right)^2 \right) F(n_x m_y)$$

That is, in the usual coordinates $z = x + iy$ on \mathfrak{H} ,

$$\Omega = y^2 \left(\left(\frac{\partial}{\partial x}\right)^2 + \left(\frac{\partial}{\partial y}\right)^2 \right)$$

The factor of 2 in the front does not matter much.

5. Enveloping algebras and other adjoint functors

To *construct* universal enveloping algebras, that is, to prove *existence*, we want a broader perspective.

There is a useful, more formulaic, version of the defining property of universal enveloping algebras. First, given an *associative* algebra A , there is an associated *Lie* algebra

$$\text{Lie}(A) = \text{Lie algebra with underlying vector space } A, \text{ with bracket } [a, b] = ab - ba$$

where the expression $ab - ba$ uses the original multiplication in A . Thus, a Lie homomorphism $\mathfrak{g} \rightarrow A$ is really a Lie homomorphism $\mathfrak{g} \rightarrow \text{Lie}(A)$. Again suppressing the fact that the object $U(\mathfrak{g})$ also carries along the map $i : \mathfrak{g} \rightarrow U(\mathfrak{g})$, there is a *natural* isomorphism

$$\text{Hom}_{\text{assoc}}(U(\mathfrak{g}), A) \approx \text{Hom}_{\text{Lie}}(\mathfrak{g}, \text{Lie}(A)) \quad (\text{of complex vectorspaces})$$

This behavior of the two mappings $\mathfrak{g} \rightarrow U(\mathfrak{g})$ and $A \rightarrow \text{Lie}(A)$ as arguments to $\text{Hom}(\cdot, \cdot)$'s is an example of an **adjunction relation**, or, equivalently, that $\mathfrak{g} \rightarrow U(\mathfrak{g})$ is a **left adjoint (functor)** to its **right adjoint (functor)** $A \rightarrow \text{Lie}(A)$. There are more elementary and familiar examples of adjoint functors arising when one of the two functors is *forgetful*, typically mapping to underlying *sets*. For example,

$$\text{Hom}_{\mathbb{Z}\text{-mods}}(F(S), M) \approx \text{Hom}_{\text{sets}}(S, M) \quad (F(S) \text{ is free module on set } S)$$

Extension of scalars functors are adjoints to *slightly*-forgetful functors, such as $V \rightarrow \text{Res}_k^K V$ where V is a K -vector space, and K is an overfield of a field k . Then we have two *different* adjunction relations

$$\begin{aligned} \text{Hom}_K(K \otimes_k W, V) &\approx \text{Hom}_k(W, \text{Res}_k^K V) && (V \text{ over } K, W \text{ over } k) \\ \text{Hom}_k(\text{Res}_k^K V, W) &\approx \text{Hom}_K(V, \text{Hom}_k(K, W)) && (V \text{ over } K, W \text{ over } k) \end{aligned}$$

As usual, the universal mapping property characterization of $U(\mathfrak{g})$ proves its uniqueness, if it exists at all. To prove *existence*, we begin by asking for a simpler universal object, namely the so-called **tensor algebra** attached to a vector space V (over a fixed field k). That is, letting $A \rightarrow F(A)$ be the forgetful map which takes (associative) k -algebras to the underlying k -vectorspaces, we want a left adjoint $V \rightarrow T(V)$ (and a linear map $V \rightarrow T(V)$) such that

$$\text{Hom}_{k\text{-alg}}(T(V), A) \approx \text{Hom}_{k\text{-vs}}(V, F(A)) \quad (\text{algebra and vector space homs, respectively})$$

Claim: The tensor algebra $T(V)$ of a k -vector space exists.

Proof: ///

Claim: The universal enveloping algebra $U(\mathfrak{g})$ of a Lie algebra \mathfrak{g} exists.

Proof: [... iou ...] ///

Claim: The kernel of the quotient map $\otimes^\bullet \mathfrak{g} \rightarrow U(\mathfrak{g})$ is G -stable.

Proof: [... iou ...] ///

So far, there is nothing subtle or substantial here. But the following result, proven in an appendix, is both essential and non-trivial.

Theorem: (*Poincaré-Birkhoff-Witt*) For any basis $\{x_i : i \in I\}$ of a Lie algebra \mathfrak{g} with *ordered* index set I , the monomials

$$x_{i_1}^{e_1} \dots x_{i_n}^{e_n} \quad (\text{with } i_1 < \dots < i_n, \text{ and integers } e_i > 0)$$

form a basis for the enveloping algebra $U(\mathfrak{g})$.

(*Proof in appendix.*)

Corollary: The map $\mathfrak{g} \rightarrow U(\mathfrak{g})$ is an *injection*. ///

6. Appendix: brackets

Here we prove the basic but non-trivial result about intrinsic derivatives. Let G act on itself by *right translations*, and thus an operation on functions on G by

$$(g \cdot f)(h) = f(hg) \quad (\text{for } g, h \in G)$$

For $x \in \mathfrak{g}$, define a differential operator X_x on smooth functions f on G by

$$(X_x f)(h) = \left. \frac{d}{dt} \right|_{t=0} f(h \cdot e^{tx})$$

Theorem:

$$X_x X_y - X_y X_x = X_{[x,y]} \quad (\text{for } x, y \in \mathfrak{g})$$

Remark: Actually, if we had set things up differently, the assertion about brackets would *define* $[x, y]$. (That would still leave the issue of computations in more practical terms.)

Proof: We first *re-characterize* the Lie algebra \mathfrak{g} in a less formulaic but more useful form.

The **tangent space** $T_m M$ to a smooth manifold M at a point $m \in M$ is *intended* to be the collection of first-order (homogeneous) differential operators, on functions near m , followed by *evaluation* of the resulting functions at the point m .

One way to make the description of the tangent space precise is as follows. Let \mathcal{O} be the ring of germs^[26] of smooth functions at m . Let $e_m : f \rightarrow f(m)$ be the evaluation-at- m map $\mathcal{O} \rightarrow \mathbb{R}$ on (germs of) functions in \mathcal{O} . Since evaluation is a ring homomorphism, (and \mathbb{R} is a field) the kernel \mathfrak{m} of e_m is a *maximal* ideal in \mathcal{O} . A first-order homogeneous differential operator D might be characterized by the *Leibniz rule*

$$D(f \cdot F) = Df \cdot F + f \cdot DF$$

Then $e_m \circ D$ vanishes on \mathfrak{m}^2 , since

$$(e_m \circ D)(f \cdot F) = f(m) \cdot DF(m) + Df(m) \cdot F(m) = 0 \cdot DF(m) + Df(m) \cdot 0 = 0 \quad (\text{for } f, F \in \mathfrak{m})$$

Thus, D gives a linear functional on \mathfrak{m} that factors through $\mathfrak{m}/\mathfrak{m}^2$. We *define*

$$\text{tangent space to } M \text{ at } m = T_m M = (\mathfrak{m}/\mathfrak{m}^2)^* = \text{Hom}_{\mathbb{C}}(\mathfrak{m}/\mathfrak{m}^2, \mathbb{C})$$

To see that we have included exactly what we want, and nothing more, use the defining fact (for *manifold*) that m has a neighborhood U and a homeomorphism-to-image $\varphi : U \rightarrow \mathbb{R}^n$.^[27] The precise definition

[26] The **germ** of a smooth function f near a point x_o on a smooth manifold M is the equivalence class of f under the equivalence relation \sim , where $f \sim g$ if f, g are smooth functions defined on some neighborhoods of x_o , and which *agree* on *some* neighborhood of x_o . But this is a *construction*, which does admit a more functional reformulation. That is, for each neighborhood U of x_o , let $\mathcal{O}(U)$ be the ring of smooth functions on U , and for $U \supset V$ neighborhoods of x_o let $\rho_{UV} : \mathcal{O}(U) \rightarrow \mathcal{O}(V)$ be the restriction map. Then the *colimit* $\text{colim}_U \mathcal{O}(U)$ is exactly the ring of germs of smooth functions at x_o .

[27] This map φ is presumably part of an *atlas*, meaning a maximal family of *charts* (homeomorphisms-to-image) φ_i of opens U_i in M to subsets of a fixed \mathbb{R}^n , with the *smooth manifold* property that on *overlaps* things fit together smoothly, in the sense that

$$\varphi_i \circ \varphi_j^{-1} : \varphi_j(U_i \cap U_j) \rightarrow U_i \cap U_j \rightarrow \varphi_i(U_i \cap U_j)$$

is a *smooth* map from the subset $\varphi_j(U_i \cap U_j)$ of \mathbb{R}^n to the subset $\varphi_i(U_i \cap U_j)$.

of *smoothness* of a function f near m is that $f \circ \varphi^{-1}$ be smooth on some subset of $\varphi(U)$.^[28] Thus, in brief, the nature of $\mathfrak{m}/\mathfrak{m}^2$ and $(\mathfrak{m}/\mathfrak{m}^2)^*$ can be immediately transported to an open subset of \mathbb{R}^n . From Maclaurin-Taylor expansions we know that the pairing

$$v \times f \longrightarrow (\nabla f)(m) \cdot v \quad (\text{for } v \in \mathbb{R}^n \text{ and } f \text{ smooth at } m \in \mathbb{R}^n)$$

induces an isomorphism $\mathbb{R}^n \longrightarrow (\mathfrak{m}/\mathfrak{m}^2)^*$. Thus, $(\mathfrak{m}/\mathfrak{m}^2)^*$ is a good notion of *tangent space*.

Claim: The Lie algebra \mathfrak{g} of G is naturally identifiable with the tangent space to G at 1, via

$$x \times f \longrightarrow \left. \frac{d}{dt} \right|_{t=0} f(e^{tx}) \quad (\text{for } x \in \mathfrak{g} \text{ and } f \text{ smooth near } 1)$$

Proof: [... iou ...]

///

Define the *left* translation action of G on functions on G by

$$(L_g f)(h) = f(g^{-1}h) \quad (g, h \in G)$$

with the inverse for associativity, as usual.

Claim: The map

$$x \longrightarrow X_x$$

gives a \mathbb{C} -linear isomorphism

$$\mathfrak{g} \longrightarrow \text{left } G\text{-invariant vector fields on } G$$

Proof: (of claim) On one hand, since the action of x is on the *right*, it is not surprising that X_x is invariant under the *left* action of G , namely

$$(X_x \circ L_g)f(h) = X_x f(g^{-1}h) = \left. \frac{d}{dt} \right|_{t=0} f(g^{-1}he^{tx}) = L_g \left. \frac{d}{dt} \right|_{t=0} f(he^{tx}) = (L_g \circ X_x)f(h)$$

On the other hand, let X be a left-invariant vector field. Then

$$(Xf)(h) = (L_h^{-1} \circ X)f(1) = (X \circ L_h^{-1})f(1) = X(L_h^{-1}f)(1)$$

That is, X is completely determined by what it does to functions at 1.

Let \mathfrak{m} be the maximal ideal of functions vanishing at 1, in the ring \mathcal{O} of germs of smooth functions at 1 on G . The first-order nature of *vector fields* is captured by the Leibniz rule

$$X(f \cdot F) = f \cdot XF + Xf \cdot F$$

As above, the Leibniz rule implies that $e_1 \circ X$ *vanishes* on \mathfrak{m}^2 . Thus, we can identify $e_1 \circ X$ with an element of

$$(\mathfrak{m}/\mathfrak{m}^2)^* = \text{Hom}_{\mathbb{C}}(\mathfrak{m}/\mathfrak{m}^2, \mathbb{C}) = \text{tangent space to } G \text{ at } 1 = \mathfrak{g}$$

Thus, the map $x \longrightarrow X_x$ is an isomorphism from \mathfrak{g} to left invariant vector fields, proving the claim. ///

Now we use the re-characterized \mathfrak{g} to prove

$$[X_x, X_y] = X_z$$

^[28] The well-definedness of this definition depends on the maximal property of an *atlas*.

for *some* $z \in \mathfrak{g}$. Consider $[X_x, X_y]$ for $x, y \in \mathfrak{g}$. That this operator is *left* G -invariant is clear, since it is a difference of composites of such. It is *less* clear that it satisfies Leibniz' rule (and thus is *first-order*). But, indeed, for *any* two vector fields X, Y we compute

$$\begin{aligned} [X, Y](fF) &= XY(fF) - YX(Ff) = X(Yf \cdot F + f \cdot YF) - Y(Xf \cdot F + f \cdot XF) \\ &= (XYf \cdot F + Yf \cdot XF + Xf \cdot YF + f \cdot XYF) - (YXf \cdot F + Xf \cdot YF + Yf \cdot XF + f \cdot YXF) \\ &= [X, Y]f \cdot F + f \cdot [X, Y]F \end{aligned}$$

so $[X, Y]$ *does* satisfy the Leibniz rule. In particular, $[X_x, X_y]$ is again a left- G -invariant vector field, so is of the form $[X_x, X_y] = X_z$ for *some* $z \in \mathfrak{g}$.

In fact, the relation $[X_x, X_y] = X_z$ is the *intrinsic definition* of the Lie bracket on \mathfrak{g} , since we could *define* the element $z = [x, y]$ by the relation $[X_x, X_y] = X_{[x, y]}$. However, we are burdened by having the *ad hoc* but elementary definition

$$[x, y] = xy - yx \quad (\text{matrix multiplication})$$

On the other hand, this burden is off-set by the commensurate fact that our group G is inside some $GL_n(\mathbb{R})$, which sits inside a familiar Euclidean space, on which we can exhibit many smooth functions explicitly.

Indeed, consider linear functions on the ambient Euclidean space

$$f(g) = \text{tr}(\lambda \cdot g)$$

where λ is an n -by- n real matrix. Unsurprisingly, after the above reduction and clarification in intrinsic terms, our earlier computational heuristic about brackets is sufficient to conclude our discussion. That is,

$$\begin{aligned} [X_x, X_y]f(1) &= \left. \frac{d}{dt} \right|_{t=0} \left. \frac{d}{ds} \right|_{s=0} (f(e^{tx}e^{sy}) - f(e^{sy}e^{tx})) \\ &= \left. \frac{d}{dt} \right|_{t=0} \left. \frac{d}{ds} \right|_{s=0} (f(1 + sy + tx + stxy + \dots) - f(1 + sy + tx + styx + \dots)) \\ &= \text{tr}(\lambda \cdot (xy - yx)) = X_{[x, y]}f(1) \end{aligned}$$

The set of images in $\mathfrak{m}/\mathfrak{m}^2$ of the aggregate of these functions is the whole space, so we have the equality

$$[X_x, X_y] = X_{[x, y]}$$

with the *ad hoc* definition of $[x, y]$. ///

Remark: Again, the *intrinsic* definition of $[x, y]$ is given by first proving that the Lie bracket of (left G -invariant) vector fields is a vector field (as opposed to some higher-order operator), and observing the identification of left-invariant vector fields with the tangent space \mathfrak{g} to G at 1. Our *extrinsic* matrix definition of the Lie bracket is appealing, but does require reconciliation with the more meaningful notion.

7. Appendix: proof of Poincare'-Birkhoff-Witt

The following result does not use any further properties of the Lie algebra \mathfrak{g} , so *must* be very general. The result is constantly invoked, so frequently, in fact, that one might tire of citing it and declare that it is understood that everyone should keep this in mind. Still, it is surprisingly difficult to prove.

Thinking in terms of the universal property of the universal enveloping algebra, we might interpret the *free-ness* assertion of the theorem as an assertion that, in the range of possibilities for abundance or poverty of representations of the Lie algebra \mathfrak{g} , the actuality is at the end of *extreme abundance*.

Theorem: For any basis $\{x_i : i \in I\}$ of a Lie algebra \mathfrak{g} with *ordered* index set I , the monomials

$$x_{i_1}^{e_1} \dots x_{i_n}^{e_n} \quad (\text{with } i_1 < \dots < i_n, \text{ and integers } e_i > 0)$$

form a *basis* for the enveloping algebra $U(\mathfrak{g})$.

Corollary: The natural map of a Lie algebra to its universal enveloping algebra is an *injection*. ///

Proof: Since we do not yet know that \mathfrak{g} *injects* to $U(\mathfrak{g})$, let $i : \mathfrak{g} \rightarrow U(\mathfrak{g})$ be the natural Lie homomorphism. The easy part of the argument is to observe that these monomials *span*. Indeed, whatever unobvious relations may hold in $U(\mathfrak{g})$,

$$U(\mathfrak{g}) = \mathbb{R} + \sum_{n=1}^{\infty} \underbrace{i(\mathfrak{g}) \dots i(\mathfrak{g})}_n$$

though we are not claiming that the sum is direct (it is not). Let

$$U(\mathfrak{g})^{\leq N} = \mathbb{R} + \sum_{n=1}^N \underbrace{i(\mathfrak{g}) \dots i(\mathfrak{g})}_n$$

Start from the fact that $i(x_k)$ and $i(x_\ell)$ commute modulo $i(\mathfrak{g})$, specifically,

$$i(x_k)i(x_\ell) - i(x_\ell)i(x_k) = i[x_k, x_\ell]$$

This reasonably suggests an induction proving that for α, β in $U(\mathfrak{g})^{\leq n}$

$$\alpha\beta - \beta\alpha \in U(\mathfrak{g})^{\leq n-1}$$

This much does not require much insight. We amplify upon this below.

The hard part of the argument here is basically from Jacobson, and applies to not-necessarily finite-dimensional Lie algebras over arbitrary fields k of characteristic 0, using no special properties of \mathbb{R} . The same argument appears later in Varadarajan. There is a different argument given in Bourbaki, and then in Humphreys.

N. Bourbaki, *Groupes et algèbres de Lie, Chap. 1*, Paris: Hermann, 1960.

N. Jacobson, *Lie Algebras*, Dover, 1962.

J. Humphreys, *Introduction to Lie Algebras and Representation Theory*, Springer-Verlag, 1972.

V.S. Varadarajan, *Lie Groups, Lie Algebras, and their Representations*, Springer-Verlag, 1974, 1984.

Remark: It may not be clear at the outset that the *Jacobi identity*

$$[x, [y, z]] - [y, [x, z]] = [[x, y], z]$$

plays an essential role in the argument, but it does. At the same time, apart from Jacobson's device of use of the endomorphism L (below), the argument is natural.

Let T_n be

$$T_n = \underbrace{\mathfrak{g} \otimes \dots \otimes \mathfrak{g}}_n$$

the space of *homogeneous tensors* of degree n , and T the **tensor algebra**

$$T = k \oplus T_1 \oplus T_2 \oplus \dots$$

of \mathfrak{g} . For $x, y \in \mathfrak{g}$ let

$$u_{x,y} = (x \otimes y - y \otimes x) - [x, y] \in T_2 + T_1 \subset T$$

Let J be the two-sided ideal in T generated by the set of all elements $u_{x,y}$. Since $u_{x,y} \in T_1 + T_2$, the ideal J contains no elements of $T_0 \approx k$, so J is a *proper* ideal in T .

Let $U = T/J$ be the quotient, the **universal enveloping algebra** of \mathfrak{g} . Let

$$q : T \rightarrow U$$

be the quotient map.

For any basis $\{x_i : i \in I\}$ of \mathfrak{g} the images $q(x_{i_1} \otimes \dots \otimes x_{i_n})$ in U of tensor monomials $x_{i_1} \otimes \dots \otimes x_{i_n}$ *span* the enveloping algebra over k , since they span the tensor algebra.

With an *ordered* index set I for the basis of \mathfrak{g} , using the Lie bracket $[\cdot, \cdot]$, we can do a certain amount of rearranging of the x_{i_j} in a monomial. We *anticipate* that everything in U can be rewritten to be as sum of monomials $x_{i_1} \dots x_{i_n}$ where

$$i_1 \leq i_2 \leq \dots \leq i_n$$

A monomial in which the indices possess this ordering is a **standard monomial**.

To form the induction that proves that the (images of) standard monomials *span* U , consider a monomial $x_{i_1} \dots x_{i_n}$ with indices not correctly ordered. There must be at least one index j such that

$$i_j > i_{j+1}$$

Since

$$x_{i_j} x_{i_{j+1}} - x_{i_{j+1}} x_{i_j} - [x_{i_j}, x_{i_{j+1}}] \in J$$

we have

$$\begin{aligned} x_{i_1} \dots x_{i_n} &= x_{i_1} \dots x_{i_{j-1}} \cdot (x_{i_j} x_{i_{j+1}} - x_{i_{j+1}} x_{i_j} - [x_{i_j}, x_{i_{j+1}}]) \cdot x_{i_{j+2}} \dots x_{i_n} \\ &\quad + x_{i_1} \dots x_{i_{j-1}} x_{i_{j+1}} x_{i_j} x_{i_{j+2}} \dots x_{i_n} + x_{i_1} \dots x_{i_{j-1}} [x_{i_j}, x_{i_{j+1}}] x_{i_{j+2}} \dots x_{i_n} \end{aligned}$$

The first summand lies inside the ideal J , while the third is a tensor of smaller degree. Thus, do induction on degree of tensors, and for each fixed degree do induction on the number of pairs of indices out of order.

The serious assertion is *linear independence*. For brevity, given a tensor monomial $x_{i_1} \otimes \dots \otimes x_{i_n}$, say that the **defect** of this monomial is the number of pairs of indices $i_j, i_{j'}$ so that $j < j'$ but $i_j > i_{j'}$. *Suppose* that we can define a linear map

$$L : T \rightarrow T$$

such that L is the identity map on a standard monomial, and whenever $i_j > i_{j+1}$

$$\begin{aligned} L(x_{i_1} \otimes \dots \otimes x_{i_n}) &= L(x_{i_1} \otimes \dots \otimes x_{i_{j+1}} \otimes x_{i_j} \otimes \dots \otimes x_{i_n}) \\ &\quad + L(x_{i_1} \otimes \dots \otimes [x_{i_j}, x_{i_{j+1}}] \otimes \dots \otimes x_{i_n}) \end{aligned}$$

If there is such L , then $L(J) = 0$, while L acts as the identity on any linear combination of standard monomials. This would prove that the subspace of T consisting of linear combinations of standard monomials meets the ideal J just at 0, so maps injectively to the enveloping algebra.

Note that, incidentally, L would have the property that

$$\begin{aligned} L(y_{i_1} \otimes \dots \otimes y_{i_n}) &= L(y_{i_1} \otimes \dots \otimes y_{i_{j+1}} \otimes y_{i_j} \otimes \dots \otimes y_{i_n}) \\ &\quad + L(y_{i_1} \otimes \dots \otimes [y_{i_j}, y_{i_{j+1}}] \otimes \dots \otimes y_{i_n}) \end{aligned}$$

for *any* vectors y_{i_j} in \mathfrak{g} .

Thus, the problem reduces to defining such a map L . We do an induction to define L . First, define L to be the identity on $T_0 + T_1$. Note that the second condition on L is vacuous here, and the first condition is met since every monomial tensor of degree 1 or 0 is standard.

Now fix $n \geq 2$, and attempt to define L on monomials in $T_{\leq n}$ inductively by using the second required property: define $L(x_{i_1} \otimes \dots \otimes x_{i_n})$ by

$$\begin{aligned} L(x_{i_1} \otimes \dots \otimes x_{i_n}) &= L(x_{i_1} \otimes \dots \otimes x_{i_{j+1}} \otimes x_{i_j} \otimes \dots \otimes x_{i_n}) \\ &\quad + L(x_{i_1} \otimes \dots \otimes [x_{i_j}, x_{i_{j+1}}] \otimes \dots \otimes x_{i_n}) \end{aligned}$$

where $i_j > i_{j+1}$. One term on the right-hand side is of lower degree, and the other is of smaller defect. Thus, we do induction on degree of tensor monomials, and for each fixed degree do induction on defect.

The potential problem is the well-definedness of this definition. Monomials of degree n and of defect 0 are already standard. For monomials of degree n and of defect 1 the definition is unambiguous, since there is just one pair of indices that are out of order.

So suppose that the defect is at least two. Let $j < j'$ be two indices so that both $i_j > i_{j+1}$ and $i_{j'} > i_{j'+1}$. To prove well-definedness it suffices to show that the two right-hand sides of the defining relation for $L(x_{i_1} \otimes \dots \otimes x_{i_n})$ are actually the same element of T .

Consider the case that $j + 1 < j'$. Necessarily $n \geq 4$. (In this case the two rearrangements do not interact with each other.) Doing the rearrangement specified by the index j , we have

$$L(x_{i_1} \otimes \dots \otimes x_{i_n}) = L(x_{i_1} \otimes \dots \otimes x_{i_{j+1}} \otimes x_{i_j} \otimes \dots \otimes x_{i_n}) \\ + L(x_{i_1} \otimes \dots \otimes [x_{i_j}, x_{i_{j+1}}] \otimes \dots \otimes x_{i_n})$$

The first summand on the right-hand side has smaller defect, and the second has smaller degree, so we can use the inductive definition to evaluate them both. And still has $i_{j'} > i_{j'+1}$. Nothing is lost if we simplify notation by taking $j = 1, j' = 3$, and $n = 4$, since all the other factors in the monomials are inert. Further, to lighten the notation write x for x_{i_1} , y for x_{i_2} , z for x_{i_3} , and w for x_{i_4} . We use the inductive definition to obtain

$$L(x \otimes y \otimes z \otimes w) = L(y \otimes x \otimes z \otimes w) + L([x, y] \otimes z \otimes w) \\ = L(y \otimes x \otimes w \otimes z) + L(y \otimes x \otimes [z, w]) \\ + L([x, y] \otimes w \otimes z) + L([x, y] \otimes [z, w])$$

But then it is clear (or can be computed analogously) that the same expression is obtained when the roles of j and j' are reversed. Thus, the induction step is completed in case $j + 1 < j'$.

Now consider the case that $j + 1 = j'$, that is, the case in which the interchanges *do* interact. Here nothing is lost if we just take $j = 1, j' = 2$, and $n = 3$. And write x for x_{i_1} , y for x_{i_2} , z for x_{i_3} . Thus,

$$i_1 > i_2 > i_3$$

Then, on one hand, applying the inductive definition by first interchanging x and y , and then doing further reshufflings, we have

$$L(x \otimes y \otimes z) = L(y \otimes x \otimes z) + L([x, y] \otimes z) = L(y \otimes z \otimes x) + L(y \otimes [z, x]) + L([x, y] \otimes z) \\ = L(z \otimes y \otimes x) + L([y, z] \otimes x) + L(y \otimes [z, x]) + L([x, y] \otimes z)$$

On the other hand, starting by doing the interchange of y and z gives

$$L(x \otimes y \otimes z) = L(x \otimes z \otimes y) + L(x \otimes [y, z]) = L(z \otimes x \otimes y) + L([x, z] \otimes y) + L(x \otimes [y, z]) \\ = L(z \otimes y \otimes x) + L(z \otimes [x, y]) + L([x, z] \otimes y) + L(x \otimes [y, z])$$

It remains to see that the two right-hand sides are the same.

Since L is already well-defined, by induction, for tensors of degree $n - 1$ (here in effect $n - 1 = 2$), we can invoke the property

$$L(v \otimes w) = L(w \otimes v) + L([v, w])$$

for all $v, w \in \mathfrak{g}$. Apply this to the second, third, and fourth terms in the first of the two previous computations, to obtain

$$L(x \otimes y \otimes z) \\ = L(z \otimes y \otimes x) + \left(L(x \otimes [y, z]) + L([[y, z], x]) \right) + \left(L([x, z] \otimes y) + L([y, [x, z]]) \right) + \left(L(z \otimes [x, y]) + L([[x, y], z]) \right)$$

The latter differs from the right-hand side of the *second* computation just by the expressions involved doubled brackets, namely

$$L([[y, z], x]) + L([y, [x, z]]) + L([[x, y], z])$$

Thus, we wish to prove that the latter is 0. Reversing *both* (so no net sign change) the brackets in the middle summand gives the equivalent requirement

$$L([[y, z], x]) + L([[z, x], y]) + L([[x, y], z])$$

which is the **Jacobi identity**. ///