(*May 13, 2014*)

Invariant Laplacians

Paul Garrett garrett@math.umn.edu http://www.math.umn.edu/˜garrett/

[This document is
http://www.math.umn.edu/˜garrett/m/mfms/notes_2013-14/11d_diffops.pdf]

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

The translation-invariant operators $\partial/\partial x_i$ on $\mathbb{R}^n$, and the rotation-invariant Laplacian on $\mathbb{R}^n$ are *deceptively-easily* proven invariant, providing few clues about more complicated situations.

For example, we expect rotation-invariant *Laplacians* (second-order operators) on spheres, and do *not* want to write a formula in generalized spherical coordinates and *verify* invariance computationally. Nor do we want to be constrained to imbedding spheres in Euclidean spaces and using the ambient geometry, even though this can succeed in some examples.

Another important example is the operator

$$\Delta^{\mathfrak{H}} \;=\; y^2\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)$$

on the complex upper half-plane $\mathfrak{H}$. Although this operator is provably *invariant* under the linear fractional action of $SL_2(\mathbb{R})$, it is oppressive and unenlightening to verify invariance directly. Worse, it is *misguided* to think in terms of such verification. The more relevant issue is the *origin* of the expression in coordinates. We do not want to *verify* invariance of expressions in coordinates, but systematically *obtain* suitable expressions. We want *intrinsic* descriptions of invariant operators, so that expressions in coordinates are *a priori* invariant.

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

---

1. *Derivatives of group actions: Lie algebras*

The usual orthogonal groups

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

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

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

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

The linear fractional transformation action

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

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

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

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

Our operational definition of 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 the matrix exponential is [6]

$$\exp(x) \;=\; e^x \;=\; 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots$$

[1.0.1] Remark: With this definition, Lie algebras are closed under scalar multiplication. But it is not clear that Lie algebras are closed under *addition*. When x and y are n-by-n real or complex matrices which *commute*, that is, such that $xy = yx$, then

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

from which we conclude that $x + y$ is again in a Lie algebra containing x and y. The general case of closedness under addition is less obvious. We will prove it as a side effect of proof (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.

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

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

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

---

[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.

[2] A fuller abstraction, not strictly necessary for illustration of construction of invariant operators, is that G should be a *Lie group* acting *smoothly* and transitively on a *smooth manifold* M. For the later parts, G should be *semi-simple*, or *reductive*. Our introductory discussion of invariant differential operators does not require 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 differentiable functions, *differentiability* of the action can be posed in 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] Probably M should be a *smooth manifold*, to make sense of the differentiability condition. As in other instances of a group acting transitively on a set with additional structure, under modest hypotheses M is 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] Many different arguments show convergence. Perhaps the clearest is to use submultiplicativity of the *operator norm*, immediately reducing to the scalar situation. For *Lie groups* not imbedded in matrix groups, there is an *intrinsic* notion of *exponential map*. However, it is technically expensive. The matrix exponential is all we need.

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

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

To see why, note that both determinant and trace are invariant under conjugation $x \to gxg^{-1}$, so we can suppose without loss of generality that x is *upper-triangular*. [7] Then e^x is still upper-triangular, with diagonal entries $e^{x_{ii}}$, where the x_{ii} are the diagonal entries of x. Thus,

$$\det(e^x) \;=\; e^{x_{11}} \cdots e^{x_{nn}} \;=\; e^{x_{11}+\ldots+x_{nn}} \;=\; e^{\mathrm{tr} x}$$

Using this, the determinant-one condition is

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

Taking the derivative with respect to t and setting $t = 0$ gives $0 = \mathrm{tr} x$. Looking at the right-hand side of the expanded $1 = \det(e^{tx})$, this condition is also *sufficient* for $\det(e^{tx}) = 1$. Thus,

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

[1.0.4] Example: From $\det(e^x) = e^{\mathrm{tr} x}$, any matrix e^x is invertible. Thus,

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

[1.0.5] Example: For $G = O(n,\mathbb{R}) = \{g \in GL_n(\mathbb{R}) : g^\top \cdot g = 1_n\}$, using $(e^{tx})^\top = e^{tx^\top}$,

$$1 \;=\; (e^{tx})^\top \cdot e^{tx} \;=\; (1 + tx^\top + \ldots)\cdot(1 + tx + \ldots) \;=\; 1 + t(x + x^\top) + \ldots$$

Thus, necessarily $x^\top + x = 0$. On the other hand, when $x^\top + x = 0$ we have $x^\top = -x$, so

$$(e^{tx})^\top \cdot e^{tx} \;=\; e^{-tx}\cdot e^{tx} \;=\; (e^{tx})^{-1}\cdot e^{tx} \;=\; 1$$

This shows that the Lie algebra of $O(n,\mathbb{R})$ is skew-symmetric matrices.

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

$$(X_x\, f)(m) \;=\; \frac{d}{dt}\Big|_{t=0} f(m\cdot e^{tx})$$

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. [8]

[7] The existence of *Jordan normal form* of a matrix over an algebraically closed field shows 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 to an upper-triangular matrix is weaker than the assertion of Jordan normal form, only requiring that there is a basis $v_1, \ldots, v_n$ for $\mathbb{C}^n$ such that $x \cdot v_i \in \Sigma_{j\le i}\mathbb{C}\, v_j$. This 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.

[8] The conjugation action of G on $\mathfrak{g}$ in the *claim* is an instance of the **adjoint action** Ad of G on $\mathfrak{g}$. In our examples it is literal conjugation.

[1.0.6] Remark: In the extreme, 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

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

That is, $\mathfrak{g}$ gives left G-invariant differential operators on G. [9]

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

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

The conjugation action of G on $\mathfrak{g}$ stabilizes $\mathfrak{g}$.

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

$$(g \cdot X_x \cdot g^{-1} \cdot f)(m) \;=\; (g(X_x(g^{-1}f)))(m) \;=\; (X_x(g^{-1}f))(mg)$$

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

Conjugation and exponentiation interact well, namely

$$g\, e^{tx}\, g^{-1} \;=\; g\left(1 + tx + \frac{(tx)^2}{2!} + \frac{(tx)^3}{3!} + \ldots\right) g^{-1}$$

$$= 1 + tgxg^{-1} + \frac{(tgxg^{-1})^2}{2!} + \frac{(tgxg^{-1})^3}{3!} + \ldots \;=\; e^{tgxg^{-1}}$$

Thus,

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

as claimed. The same argument shows that $gxg^{-1} \in \mathfrak{g}$. ///

[1.0.8] Remark: Again, this literal *conjugation* of matrices is more properly called the *adjoint action* of G on $\mathfrak{g}$.

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

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

Specifically,

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

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

The commutant expression $[x,y] = xy - yx$ is the **Lie bracket**.

[9] 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.

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

$$e^{tx}\,e^{ty}\,e^{-tx}\,e^{-ty} \;=\; (1+tx+t^2x^2/2)(1+ty+t^2y^2/2)(1-tx+t^2x^2/2)(1-ty+t^2y^2/2)$$

$$=\; (1+t(x+y)+\frac{t^2}{2}(x^2+2xy+y^2))\;(1-t(x+y)+\frac{t^2}{2}(x^2+2xy+y^2))$$

$$=\; 1+t^2\big(x^2+2xy+y^2-(x+y)(x+y)\big) \;=\; \big(1+t^2(2xy-xy-yx)\big) \;=\; 1+t^2[x,y]$$

as claimed. ///

[1.0.10] Claim: The conjugation/adjoint action of G on $\mathfrak{g}$ respects brackets:

$$[gxg^{-1},\, gyg^{-1}] \;=\; g[x,y]g^{-1} \qquad\qquad (\text{for } x,y\in\mathfrak{g} \text{ and } g\in G)$$

Proof: For Lie brackets expressed in terms of matrix operations, this is straightforward:

$$[gxg^{-1},\, gyg^{-1}] \;=\; gxg^{-1}\,gyg^{-1} - gyg^{-1}\,gxg^{-1} \;=\; gxyg^{-1} - gyxg^{-1} \;=\; g(xy-yx)g^{-1} \;=\; g[x,y]g^{-1}$$

as claimed. ///

Composition of the derivatives X_x operators mirrors the bracket in the Lie algebra:

[1.0.11] Theorem:

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

The proof of this theorem is non-trivial, given in an appendix.

The point is that the map $x\to X_x$ is a **Lie algebra homomorphism**, meaning 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. [10] First, 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 $\frac{d}{ds}\big|_{s=0}$),

$$\frac{d}{dt}\bigg|_{t=0}\frac{d}{ds}\bigg|_{s=0}\big(e^{tx}\,e^{sy}-e^{sy}\,e^{tx}\big)$$

$$=\; \frac{d}{dt}\bigg|_{t=0}\frac{d}{ds}\bigg|_{s=0}(1+sy+tx+stxy+\ldots)-(1+sy+tx+styx+\ldots)) \;=\; xy-yx$$

Slightly more rigorously, *imagine* that it is legitimate to write something like

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

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

On one hand, replacing x by $[x,y]$ in the previous gives

$$(X_{[x,y]}\,f)(m) \;=\; (\nabla f)(m)\cdot m[x,y]$$

[10] When G acts *transitively* on a space M, we should expect (under mild hypotheses) that $M\approx G_o\backslash G$ where G_o is the isotropy group of a 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.

On the other hand, from the definition of the differential operators X_x and X_y, [11]

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

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

$$f(m\cdot(1+sy+tx+stxy+\ldots)) - f(m\cdot(1+sy+tx+styx+\ldots))$$

$$\sim (f(m) + \nabla f(m)\cdot m(tx+sy+stxy+\ldots)) - (f(m)+\nabla f(m)\cdot m(tx+sy+styx+\ldots))$$
$$= \nabla f(m)\cdot m(st\,(xy-yx)+\ldots)$$

Applying the operator $\frac{\partial}{\partial t}\big|_{t=0}\ \frac{\partial}{\partial s}\big|_{s=0}$,

$$(X_{[x,y]}\, f)(m) \;=\; \nabla f(m)\cdot m(xy-yx)$$

as claimed. [12] We'll do this computation legitimately in the appendix.

---

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

$$X_xX_y - X_yX_x \;=\; [X_x, X_y] \;=\; X_{[x,y]} \;=\; X_{xy-yx}$$

Indeed, closure under a bracket operation is a defining attribute of a Lie algebra. [13] However, the *composition* of differential operators has no analogue inside the Lie algebra. That is, typically,

$$X_xX_y \;\neq\; X_\varepsilon \qquad (\text{for any } \varepsilon\in\mathfrak{g})$$

We want to create something from the Lie algebra that allows composition in this fashion. [14]

---

[11] We presume that we can interchange the partial derivatives. This would be *Clairault's theorem*, but we have sufficiently strong hypothesis that there's no issue.

[12] One problem with this heuristic is the implicit assumption that f extends to the ambient space of matrices. The computation depends on this extension, at least superficially. 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 compatibility with *mappings* to other groups. See the appendix for a better argument.

[13] Restricting our scope to matrix groups G, defining the Lie bracket via matrix multiplication, $[x,y] = xy - yx$, conveniently entails 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 $\mathrm{ad}:\mathfrak{g}\to\mathrm{End}(\mathfrak{g})$ by $(\mathrm{ad}x)(y) = [x,y]$ is a Lie algebra *homomorphism*. That is, $[\mathrm{ad}x, \mathrm{ad}y] = \mathrm{ad}[x,y]$.

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

$$X_xX_y \;\neq\; X_{xy}$$

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

To accomplish this construction, and to see the efficacy of the approach we take, we give as few details as possible in a first pass. The details are filled in carefully a little later.

[2.1] Intrinsic description of Casimir We want an *associative* algebra[15] $U\mathfrak{g}$ which is *universal* in the sense that any linear map $\varphi : \mathfrak{g} \to B$ to an associative algebra B respecting brackets

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

should give a unique *associative algebra* homomorphism

$$\Phi : U\mathfrak{g} \;\longrightarrow\; B$$

There must be a connection to the original $\varphi : \mathfrak{g} \to B$, so we require existence of a fixed map $i : \mathfrak{g} \to U\mathfrak{g}$ *respecting brackets* and commutativity of a diagram

$$\begin{tikzcd} U\mathfrak{g} \arrow[dr, dashed, "\Phi\ (\text{assoc})"] & \\ \mathfrak{g} \arrow[u, "i\ (\text{Lie})"] \arrow[r, "\varphi\ (\text{Lie})"'] & B \end{tikzcd}$$

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

A related object is the *universal associative algebra*[16] AV of a vector space V over a field k, with a specified linear $j : V \to AV$. The characterizing property is that any linear map $V \to B$ to an (associative) algebra B extends to a unique associative algebra map $AV \to B$. That is, there is a diagram

$$\begin{tikzcd} AV \arrow[dr, dashed, "\Phi\ (\text{assoc})"] & \\ V \arrow[u, "j\ (\text{linear})"] \arrow[r, "\varphi\ (\text{linear})"'] & B \end{tikzcd}$$

Since the universal associative algebra $j\mathfrak{g} \to A\mathfrak{g}$ is universal with respect to maps $\mathfrak{g} \to B$ that are merely *linear*, not necessarily preserving the Lie brackets, there should be (unique) natural (quotient) map $q : A\mathfrak{g} \to U\mathfrak{g}$.

The conjugation (Adjoint) action $x \to gxg^{-1}$ of G on $\mathfrak{g}$ should extend to an action of G on $U\mathfrak{g}$ (which we may still write as conjugation) compatible with the multiplication in $U\mathfrak{g}$. That is, we expect

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

The action of G on $\mathfrak{g}$ should extend to $A\mathfrak{g}$, too, and the quotient map $q : A\mathfrak{g} \to U\mathfrak{g}$ should respect that action.

[15] 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 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. The associative algebras here also have a *unit* 1.

[16] The universal associative algebra AV of a vector space V is very often called the *tensor algebra*, although, unhelpfully, this name refers to a specific *construction*, rather than to the characterizing property of the algebra.

We also *assume* for the moment that we have a non-degenerate symmetric bilinear form $\langle,\rangle$ on $\mathfrak{g}$, and that this form has the G-action property

$$\langle gxg^{-1}, gyg^{-1}\rangle \;=\; \langle x, y\rangle \qquad\qquad (\text{for } x, y \in \mathfrak{g} \text{ and } g \in G)$$

Granting these things, we can *intrinsically* describe the simplest non-trivial G-invariant element in $U\mathfrak{g}$. As earlier, 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 situations this differential operator is the suitable notion of **invariant Laplacian**.

Map $\zeta : \mathrm{End}_{\mathbb{C}}(\mathfrak{g}) \to U\mathfrak{g}$ by

$$\begin{tikzcd} \mathrm{End}_{\mathbb{C}}(\mathfrak{g}) \arrow[r, "\text{natural} \approx"] \arrow[rrrr, bend right, "\zeta"'] & \mathfrak{g} \otimes \mathfrak{g}^* \arrow[r, "\approx \text{ via } \langle,\rangle"] & \mathfrak{g} \otimes \mathfrak{g} \arrow[r, "\text{inclusion}"] & A\mathfrak{g} \arrow[r, "\text{quotient}"] & U\mathfrak{g} \end{tikzcd}$$

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

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

Proof: Since ζ is G-equivariant by construction,

$$g\zeta(\mathrm{id}_{\mathfrak{g}})g^{-1} \;=\; \zeta(g\,\mathrm{id}_{\mathfrak{g}}\,g^{-1}) \;=\; \zeta(g\,g^{-1}\,\mathrm{id}_{\mathfrak{g}}) \;=\; \zeta(\mathrm{id}_{\mathfrak{g}})$$

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

[2.1.2] Remark: The possible hitch is that we don't have a simple way to show that $\zeta(\mathrm{id}_{\mathfrak{g}}) \neq 0$. This non-vanishing can be proven by demonstrating at least one associative algebra B and $\mathfrak{g} \to B$ so that the induced image of Casimir is non-zero in B. The non-vanishing is also a corollary of a surprisingly serious result, the *Poincaré-Birkhoff-Witt* theorem, proven in an appendix.

[2.2] Casimir in coordinates

The above prescription *does* tell how to express the Casimir element $\Omega = \zeta(\mathrm{id}_{\mathfrak{g}})$ in *various* coordinates. Namely, for any basis $x_1, \ldots, x_n$ of $\mathfrak{g}$, let $x_1^*, \ldots, x_n^*$ 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 $\mathrm{id}_{\mathfrak{g}}$ maps to $\sum_i x_i \otimes x_i^*$ in $\mathfrak{g}\otimes\mathfrak{g}^*$, then to $\sum_i x_i \otimes x_i$ in $\mathfrak{g}\otimes\mathfrak{g}$ for x_i an *orthonormal* basis, which imbeds in $\bigotimes^\bullet \mathfrak{g}$, and by the quotient map is sent to $U\mathfrak{g}$.

[2.2.1] Remark: The *intrinsic* description of the Casimir element as $\zeta(\mathrm{id}_{\mathfrak{g}})$ shows that it does not depend upon the choice of basis $x_1, \ldots, x_n$. [17]

[2.2.2] Remark: Example computations will be given below.

[17] 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 of $\mathrm{id}_{\mathfrak{g}}$ in coordinates.

3. *Details about universal algebras*

We fill in details about $U\mathfrak{g}$ and $A\mathfrak{g}$, including constructions.

[3.1] Universal algebras Again, we want an *associative* algebra $U\mathfrak{g}$ such that any Lie algebra map $\varphi : \mathfrak{g} \to B$ to an associative algebra B with the property

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

gives a unique *associative algebra* homomorphism

$$\Phi : U\mathfrak{g} \;\longrightarrow\; B$$

fitting into a commutative diagram

$$\begin{tikzcd}
U\mathfrak{g} \arrow[dr, dashed, "\Phi\ (\text{assoc})"] & \\
\mathfrak{g} \arrow[u, "i\ (\text{Lie})"] \arrow[r, "\varphi\ (\text{Lie})"'] & B
\end{tikzcd}$$

Similarly, we want a *universal associative algebra* AV of a vector space V over a field k, with a specified linear $j : V \to AV$, such that any linear map $V \to B$ to an associative algebra B extends to a unique associative algebra map $AV \to B$ fitting into a commutative diagram

$$\begin{tikzcd}
AV \arrow[dr, dashed, "\Phi\ (\text{assoc})"] & \\
V \arrow[u, "j\ (\text{linear})"] \arrow[r, "\varphi\ (\text{linear})"'] & B
\end{tikzcd}$$

Granting for a moment the *existence* of $A\mathfrak{g}$, construct $U\mathfrak{g}$ as the quotient of $A\mathfrak{g}$ by the two-sided ideal generated by all elements

$$\big(jx \otimes jy - jy \otimes jx\big) - j[x,y] \qquad\qquad (\text{where } x,y \in \mathfrak{g})$$

The map $i : \mathfrak{g} \to U\mathfrak{g}$ is the obvious composite $q \circ j$. Given a Lie algebra map $\varphi : \mathfrak{g} \to B$ from $\mathfrak{g}$ to an associative algebra, we show that the induced map $\Phi : A\mathfrak{g} \longrightarrow B$ factors through $q : A\mathfrak{g} \longrightarrow U\mathfrak{g}$. Diagrammatically, we claim the existence of an arrow to fill in a commutative diagram

$$\begin{tikzcd}
A\mathfrak{g} \arrow[dr, "q"] \arrow[ddrr, bend left, "\Phi"] & & \\
& U\mathfrak{g} \arrow[dr, dashed] & \\
\mathfrak{g} \arrow[uu, "j"] \arrow[ur, "i"] \arrow[rr, "\varphi"] & & B
\end{tikzcd}$$

Indeed, the the Lie algebra homomorphism property $\varphi(x)\varphi(y) - \varphi(y)\varphi(x) - \varphi[x,y] = 0$ and the commutativity imply that

$$\Phi\Big(\big(jx \otimes jy - jy \otimes jx\big) - j[x,y]\Big) \;=\; 0 \qquad\qquad (\text{for all } x,y \in \mathfrak{g})$$

That is, Φ vanishes on the kernel of the quotient map $q : A\mathfrak{g} \to U\mathfrak{g}$, so factors through this quotient map. This proves the existence of $U\mathfrak{g}$ in terms of $A\mathfrak{g}$.

The conjugation (Adjoint) action $x \to gxg^{-1}$ of G on $\mathfrak{g}$ should extend to an action of G on $U\mathfrak{g}$ (which we may still write as conjugation) compatible with the multiplication in $U\mathfrak{g}$. That is, we expect

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

The action of G on $\mathfrak{g}$ should extend to $A\mathfrak{g}$, too, and the quotient map $q : A\mathfrak{g} \to U\mathfrak{g}$ should respect that action. Fulfillment of this requirement, or the observation that it is automatically fulfilled, is best understood from further details about $A\mathfrak{g}$, just below.

[3.1.1] Remark: It would be reasonable to approach the G-action on $U\mathfrak{g}$ by looking at the smaller category of associative G-algebras, meaning having a G-action respecting the associative algebra structure, rather than the larger category of *all* associative algebras, with or without G-structure. Thus, we'd define the universal associative G-algebra, and universal enveloping G-algebra. But there is no *a priori* guarantee that these are the same as $A\mathfrak{g}$ and $U\mathfrak{g}$. It turns out that this approach is not necessary, because the universal algebras are sufficiently large, in a sense clarified below.

[3.2] Construction of universal associative algebras

The tensor *construction* of $A\mathfrak{g}$ gives enough further information so that we can see that it inherits an action of G from $\mathfrak{g}$, and that this action is inherited by $U\mathfrak{g}$. The *construction* of AV in terms of tensors is

$$AV \;=\; k \,\oplus\, V \,\oplus\, (V\otimes V) \,\oplus\, (V\otimes V\otimes V)\oplus\ldots$$

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

$$(v_1\otimes\ldots\otimes v_m)\cdot(w_1\otimes\ldots\otimes w_n) \;=\; v_1\otimes\ldots\otimes v_m\otimes w_1\otimes\ldots\otimes w_n$$

The well-definedness of the multiplication follows from noting that there is a unique linear map $\bigotimes^m V\otimes\bigotimes^n V \longrightarrow \bigotimes^{m+n} V$ induced from the bilinear map

$$(v_1\otimes\ldots\otimes v_m)\times(w_1\otimes\ldots\otimes w_n) \;\longrightarrow\; v_1\otimes\ldots\otimes v_m\otimes w_1\otimes\ldots\otimes w_n$$

Distributivity of multiplication over addition follows from the fact that the multiplication maps are induced from bilinear maps.

The map $V \to AV$ is to the summand $V \subset AV$, which shows that this map is *injective*. It is also true that $\mathfrak{g} \to U\mathfrak{g}$ is injective, but the latter fact is less trivial to prove.

To verify that this constructed object has the requisite universal property, let $\varphi : V \to B$ be a linear map to an associative algebra. Then the linear map $\Phi_n : \bigotimes^n V \to B$ defined by

$$\Phi(v_1\otimes\ldots\otimes v_n) \;=\; \varphi(v_1)\ldots\varphi(v_n) \qquad\qquad (\text{latter is multiplication in } B)$$

is *well-defined*, being induced from the n-multilinear map

$$\underbrace{V\times\ldots\times V}_{n} \;\longrightarrow\; B \qquad\text{by}\qquad v_1\times\ldots\times v_n \;\longrightarrow\; \varphi(v_1)\ldots\varphi(v_n)$$

Letting k be the underlying field (probably either $\mathbb{C}$ or $\mathbb{R}$), there is also the map $\Phi_0 : k \to B$ by $a \to 1_B$. The collection of maps Φ_n gives a linear map $\Phi : AV \to B$. It also obviously preserves multiplication. This proves that the tensor construction yields the universal associative algebra.

[3.3] G-action on $A\mathfrak{g}$ and $U\mathfrak{g}$

The notationally obvious G-action on $A\mathfrak{g}$ is

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

This gives a well-defined linear map of each $\bigotimes^n \mathfrak{g}$ to itself, because it is the unique map induced by the multilinear map

$$\underbrace{\mathfrak{g} \times \ldots \times \mathfrak{g}}_{n} \longrightarrow \bigotimes{}^n \mathfrak{g} \qquad \text{by} \qquad v_1 \times \ldots \times v_n \longrightarrow gv_1g^{-1} \otimes \ldots \otimes gv_ng^{-1}$$

The map is visibly compatible with multiplication.

Since $\mathfrak{g}$ injects to $A\mathfrak{g}$, we can safely suppress the map j in this discussion. The G-action stabilizes the kernel of the kernel of $q : A\mathfrak{g} \to U\mathfrak{g}$, since

$$g(\Big(x \otimes y - y \otimes x) - [x,y]\Big)g^{-1} \;=\; g(x \otimes y)g^{-1} - g(y \otimes x)g^{-1} - g[x,y]g^{-1}$$
$$= \; gxg^{-1} \otimes gyg^{-1} - gyg^{-1} \otimes gxg^{-1} - [gxg^{-1}, gyg^{-1}]$$

This gives a natural action of G on $U\mathfrak{g}$, respecting the quotient $q : A\mathfrak{g} \to U\mathfrak{g}$, and, therefore, respecting the map $\mathfrak{g} \to U\mathfrak{g}$.

[3.3.1] Remark: Thus, the universal associative algebra $A\mathfrak{g}$ is sufficiently large that, roughly, it has no non-trivial relations. Thus, the notationally-obvious apparent definition of the G-action on $A\mathfrak{g}$ is *well-defined*. Then the G-action descends to $U\mathfrak{g}$.

[3.4] Killing form

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

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

which is G-*equivariant* in the sense that

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

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

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

suffices. The behavior under the action of G is clear:

$$\langle gxg^{-1}, \, gyg^{-1} \rangle \;=\; \mathrm{tr}\big(gxg^{-1} \cdot gyg^{-1}\big) \;=\; \mathrm{tr}\big(gxyg^{-1}\big) \;=\; \mathrm{tr}(xy) \;=\; \langle x, y \rangle$$

The non-degeneracy and G-equivariance of $\langle , \rangle$ give a natural G-equivariant isomorphism $\mathfrak{g} \to \mathfrak{g}^*$ by

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

When a group G acts on a vector space V the action on the dual V^* is by

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

The inverse appears (as usual!) to preserve *associativity*. The equivariance of $\langle , \rangle$ 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 $x \to \lambda_x$ is a G-isomorphism.

[3.5] $\mathrm{End}_k V \approx V \otimes V^*$

The natural isomorphism

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

is given by the k-linear extension of the map

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

Indeed, the fact that the map is an isomorphism follows by dimension counting, using the finite-dimensionality. [18]

[18] 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 $\mathrm{End}_k V$. To

4. *Descending to G/K*

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. [19]

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

$$(x \cdot f)(g) \;=\; \frac{d}{dt}\Big|_{t=0} F(g\, e^{tx})$$

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

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

For a (closed) subgroup K of G let $q : G \to G/K$ 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

$$(x \cdot (f \circ q))(g) \;=\; \frac{d}{dt}\Big|_{t=0} (f \circ q)(g\, e^{tx})$$

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

$$\frac{d}{dt}\Big|_{t=0} F(g\, e^{tx}) \;=\; (x \cdot F)(g) \;=\; (x \cdot F)(gk) \;=\; \frac{d}{dt}\Big|_{t=0} F(gk\, e^{tx}) \qquad\qquad (k \in \mathfrak{k})$$

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

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

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 $\alpha \in Z(\mathfrak{g})$ 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.

---

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. 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.

[19] 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.

5. *Example computation:* $SL_2(\mathbb{R})$ *and* $\mathfrak{H}$

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

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

These have the easily verified relations

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

Use the pairing

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

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

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

We easily compute that

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

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

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

Since $XY - YX = H$ [20] the expression for Ω can be rewritten is 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 $q : G \to G/K$ be the quotient map

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

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

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

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

$$(z \cdot F)(g) \;=\; \frac{d}{dt}\Big|_{t=0} F(g\, e^{tz})$$

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

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

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

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

Thus, in terms of the basis H, X, Y above, $X - Y$ annihilates $f \circ q$.

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

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

These are convenient group elements because they match the exponentiated Lie algebra elements:

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

In 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

$$(\Omega F)(n_x\, m_y) \;=\; (\frac{H^2}{2} + XY + YX)F(n_x\, m_y) \;=\; (\frac{H^2}{2} + XY + YX)F(n_x\, m_y)$$

$$= (\frac{H^2}{2} + 2XY - H)F(n_x\, m_y) \;=\; (\frac{H^2}{2} + 2X^2 - H)F(n_x\, m_y)$$

Compute the pieces separately. First, using the identity

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

we compute the effect of X

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

Thus, the term $2X^2$ gives

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

The action of H is

$$(H \cdot F)(n_x\, m_y) \;=\; \frac{d}{dt}\Big|_{t=0} F(n_x\, m_y\, m_{e^{2t}}) \;=\; \frac{d}{dt}\Big|_{t=0} F(n_x\, m_{ye^{2t}}) \;=\; 2y\frac{\partial}{\partial y}F(n_x\, m_y)$$

Then

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

Altogether, on right K-invariant functions F,

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

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

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

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

---

6. *Appendix: brackets*

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

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

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

$$(X_x\, f)(h) \;=\; \frac{d}{dt}\Big|_{t=0} f(h \cdot e^{tx})$$

[6.0.1] Theorem:

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

[6.0.2] Remark: 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: First, *re-characterize* the Lie algebra $\mathfrak{g}$ in a less formulaic, 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 $\mathscr{O}$ be the ring of germs [21] of smooth functions at m. Let $e_m : f \to f(m)$ be the evaluation-at-m map $\mathscr{O} \to \mathbb{R}$ on (germs of) functions in $\mathscr{O}$. Since evaluation is a ring homomorphism, (and $\mathbb{R}$ is a field) the kernel $\mathfrak{m}$ of e_m is a *maximal* ideal in $\mathscr{O}$. 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

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

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

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

---

[21] 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. This is a *construction*, which does admit a more functional reformulation. That is, for each neighborhood U of x_o, let $\mathscr{O}(U)$ be the ring of smooth functions on U, and for $U \supset V$ neighborhoods of x_o let $\rho_{UV} : \mathscr{O}(U) \to \mathscr{O}(V)$ be the restriction map. Then the *colimit* $\mathrm{colim}_U \mathscr{O}(U)$ is exactly the ring of germs of smooth functions at x_o.

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 $\varphi : U \to \mathbb{R}^n$. [22] The precise definition of *smoothness* of a function f near m is that $f \circ \varphi^{-1}$ be smooth on some subset of $\varphi(U)$. [23] 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, the pairing

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

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

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

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

Proof: The exponential map is a diffeomorphism of the Lie algebra $\mathfrak{g}$ to its image, and the image is a neighborhood of the identity in G. For *linear* Lie groups, the invertibility is immediate from existence of an explicit local inverse to the exponential near 1, given by the usual logarithm. ///

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

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

with the inverse for associativity, as usual.

[6.0.4] Claim: The map

$$x \longrightarrow X_x$$

gives an $\mathbb{R}$-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

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

On the other hand, for a left-invariant vector field X,

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

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

---

[22] 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

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

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

[23] The well-definedness of this definition depends on the *maximality* property of an *atlas*.

Let $\mathfrak{m}$ be the maximal ideal of functions vanishing at 1, in the ring $\mathscr{O}$ 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

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

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

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

$$[X_x, X_y] \;=\; X_z$$

for *some* $z \in \mathfrak{g}$. Consider $[X_x, X_y]$ for $x, y \in \mathfrak{g}$. That this differential 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,

$$[X,Y](fF) \;=\; XY(fF) - YX(Ff) \;=\; X(Yf\cdot F + f\cdot YF) - Y(Xf \cdot F + f\cdot XF)$$

$$=\; (XYf\cdot F + Yf\cdot XF + Xf\cdot YF + f\cdot XYF) - (YXf\cdot F + Xf\cdot YF + Yf\cdot XF + f\cdot YXF)$$

$$=\; [X,Y]f\cdot F + f\cdot [X,Y]F$$

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

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

However, our limiting assumption that G is a subgroup of some $GL_n(\mathbb{R})$ or $GL_n(\mathbb{C})$ allows us to use the explicit exponential and a local logarithm inverse to it, to determine the bracket $[X_x, X_y]$ somewhat more intrinsically, as follows.

Consider *linear* functions on $\mathfrak{g}$, *locally* transported to G via locally inverting the exponential near $1 \in G$. Thus, for $\lambda \in \mathfrak{g}^*$, near $1 \in G$, define

$$f(e^x) \;=\; \lambda(x)$$

Then

$$[X_x, X_y] f_\lambda(1) \;=\; \frac{d}{dt}\Big|_{t=0} \frac{d}{ds}\Big|_{s=0} \Big(\lambda\big(\log(e^{sx}e^{ty})\big) - \lambda(\log(e^{ty}e^{sx}))\big)\Big)$$

Dropping $O(s^2)$ and $O(t^2)$ terms, this is

$$=\; \frac{d}{dt}\Big|_{t=0} \frac{d}{ds}\Big|_{s=0} \Big(\lambda\big(\log(1+sx)(1+ty)\big) - \lambda\big(\log(1+ty)(1+sx)\big)\Big)$$

$$=\; \frac{d}{dt}\Big|_{t=0} \frac{d}{ds}\Big|_{s=0} \lambda\big(\log(1+sx+ty+stxy) - \log(1+ty+sx+styx)\big)$$

$$=\; \frac{d}{dt}\Big|_{t=0} \frac{d}{ds}\Big|_{s=0} \lambda\left((sx+ty+stxy - \tfrac{1}{2}(sx+ty)^2) - (ty+sx+styx - \tfrac{1}{2}(ty+sx)^2)\right)$$

$$=\; \frac{d}{dt}\Big|_{t=0} \frac{d}{ds}\Big|_{s=0} \lambda\left((stxy - \tfrac{1}{2}stxy - \tfrac{1}{2}styx) - (styx - \tfrac{1}{2}stxy - \tfrac{1}{2}styx)\right)$$

$$=\; \frac{d}{dt}\Big|_{t=0} \frac{d}{ds}\Big|_{s=0} st\cdot\lambda(xy-yx) \;=\; \lambda(xy-yx)$$

where the multiplication and commutator $xy - yx$ is in the ring of matrices. Thus, since $\mathfrak{g}^*$ separates points on $\mathfrak{g}$, we have the equality

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

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

[6.0.5] 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 requires reconciliation with the more meaningful notion.

---

7. *Appendix: proof of Poincaré-Birkhoff-Witt*

The following result does not use any further properties of the Lie algebra $\mathfrak{g}$, so *must* be 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. It is surprisingly difficult to prove.

Thinking 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 *abundance* rather than *scarcity.*

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

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

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

[7.0.2] 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 $i : \mathfrak{g} \to U\mathfrak{g}$ 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 $U\mathfrak{g}$,

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

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

$$U\mathfrak{g}^{\le N} \;=\; \mathbb{R} + \sum_{n=1}^{N} \underbrace{i(\mathfrak{g}) \ldots i(\mathfrak{g})}_{n}$$

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 $U\mathfrak{g}^{\le n}$

$$\alpha\beta - \beta\alpha \in U\mathfrak{g}^{\le n-1}$$

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

The hard part of the argument 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.

[7.0.3] Remark: It is not 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

$$T_n \;=\; \underbrace{\mathfrak{g}\otimes\ldots\otimes\mathfrak{g}}_{n}$$

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

$$T \;=\; k\oplus T_1\oplus T_2\oplus\ldots$$

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 $T_o\approx k$, so J is a *proper* ideal in T.

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

$$q:T\longrightarrow U$$

be the quotient map.

For any basis $\{x_i : i\in I\}$ of $\mathfrak{g}$ the images $q(x_{i_1}\otimes\ldots\otimes x_{i_n})$ in U of tensor monomials $x_{i_1}\otimes\ldots\otimes x_{i_n}$ *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 $[,]$, we can rearrange the x_{i_j}'s in a monomial. We *anticipate* that everything in U can be rewritten to be as sum of monomials $x_{i_1}\ldots x_{i_n}$ where

$$i_1\le i_2\le\ldots i_n$$

A monomial in with indices so ordered is a **standard monomial**.

To form the induction that proves that the (images of) standard monomials *span* U, consider a monomial $x_{i_1}\ldots x_{i_n}$ 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

$$x_{i_1}\ldots x_{i_n} \;=\; x_{i_1}\ldots x_{i_{j-1}}\cdot(x_{i_j}x_{i_{j+1}} - x_{i_{j+1}}x_{i_j} - [x_{i_j},x_{i_{j+1}}])\cdot x_{i_{j+2}}\ldots x_{i_n}$$
$$+x_{i_1}\ldots x_{i_{j-1}}x_{i_{j+1}}x_{i_j}x_{i_{j+2}}\ldots x_{i_n} + x_{i_1}\ldots x_{i_{j-1}}[x_{i_j},x_{i_{j+1}}]x_{i_{j+2}}\ldots x_{i_n}$$

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*. Given a tensor monomial $x_{i_1} \otimes \ldots \otimes x_{i_n}$, 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

$$L : T \to T$$

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

$$L(x_{i_1} \otimes \ldots \otimes x_{i_n}) \;=\; L(x_{i_1} \otimes \ldots \otimes x_{i_{j+1}} \otimes x_{i_j} \otimes \ldots \otimes x_{i_n})$$

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

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.

Incidentally, L would have the property that

$$L(y_{i_1} \otimes \ldots \otimes y_{i_n}) \;=\; L(y_{i_1} \otimes \ldots \otimes y_{i_{j+1}} \otimes y_{i_j} \otimes \ldots \otimes y_{i_n})$$

$$+L(y_{i_1} \otimes \ldots \otimes [y_{i_j}, y_{i_{j+1}}] \otimes \ldots \otimes y_{i_n})$$

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

Thus, the problem reduces to defining L. Do an induction to define L. First, define L to be the identity on $T_o + T_1$. 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* $L(x_{i_1} \otimes \ldots \otimes x_{i_n})$ by

$$L(x_{i_1} \otimes \ldots \otimes x_{i_n}) \;=\; L(x_{i_1} \otimes \ldots \otimes x_{i_{j+1}} \otimes x_{i_j} \otimes \ldots \otimes x_{i_n})$$

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

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 $L(x_{i_1} \otimes \ldots \otimes x_{i_n})$ are 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,

$$L(x_{i_1} \otimes \ldots \otimes x_{i_n}) \;=\; L(x_{i_1} \otimes \ldots \otimes x_{i_{j+1}} \otimes x_{i_j} \otimes \ldots \otimes x_{i_n})$$

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

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

$$\begin{aligned} L(x\otimes y\otimes z\otimes w) \;&=\; L(y\otimes x\otimes z\otimes w) + L([x,y]\otimes z\otimes w) \\ &=\; L(y\otimes x\otimes w\otimes z) + L(y\otimes x\otimes [z,w]) \\ &+L([x,y]\otimes w\otimes z) + L([x,y]\otimes [z,w]) \end{aligned}$$

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 further reshuffling,

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

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

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

$$L(x\otimes y\otimes z)$$

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

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. Having the Jacobi identity in mind motivates some rearrangement: move $L([[x,y],z])$ to the right-hand side of the equation, multiply through by -1, and reverse the outer bracket in the first summand, to give the equivalent requirement

$$L([x,[y,z]]) - L([y,[x,z]]) \;=\; L([[x,y],z])$$

This equality follows from application of L to the Jacobi identity. ///