Essays on representations of real groups

Introduction to Lie algebras

Bill Casselman
University of British Columbia
cass@math.ubc.ca

This essay is intended to be a self-contained if rather brief introduction to Lie algebras. I begin with an extremely brief introduction to differential geometry, which is mainly concerned with different ways to consider vector fields. The first several sections after that deal with Lie algebras as spaces of invariant vector fields on a Lie group. This early treatment is intended, among other things, to motivate the eventual introduction of abstract Lie algebras, with which all the later sections are concerned.

The first and second parts are not intended to be complete in all minor points, but just to give a rough idea of how things go. The later parts, on the contrary, are intended to be as complete as possible, but only treat results that lead without too much digression to the final goal of understanding semi-simple and reductive Lie algebras. In the third part I sketch the proof that semi-simple Lie algebras give rise to root systems, although I do not discuss the consequences of this. In particular, I do not say anything about the structure of root systems or results about semi-simple Lie algebras that depend on it, such as the structure of the centre of the enveloping algebra or the classification of irreducible representations. That is another story.

In much of this account, I have followed the standard references [Jacobson:1962], [Serre:1965], and [Serre:1966]. However, although nothing I do is without some basis already in the literature, I believe that some aspects of the proofs in the third part are new. One thing I'd like to point out is my use of primary decompositions of finite-dimensional modules over nilpotent Lie algebras, which I found in [Jacobson:1962]. I find it to be both interesting and illuminating, especially in the proof of Cartan's criterion for solvability, where I find it to be the natural tool. I'm surprised one doesn't come across it more often. Because of the role of the Jacquet and Whittaker functors in the theory of representations of real reductive Lie algebras, it deserves to be better known.

The amount of space allotted to various topics will probably look idiosyncratic to some. But for this I offer no explanation other than that I became interested in pursuing things often left out in other expositions of this material. The overall length of this essay is still rather less than that of most other introductions, so I probably need to make no apology for what I hope to be at least amiable *bavardage*.

Contents

Throughout, I'll use multi-indices k in $\mathbb{N}^n$:

$$|k| = \sum k_i, \quad x^k = \prod x_i^{k_i}, \quad k! = \prod k_i! \quad \text{and } f^{(k)} = \frac{\partial^{|k|} f}{\partial x^k} = \left(\prod \frac{\partial^{k_i}}{\partial x^{k_i}}\right) f .$$

Part I. Differential geometry

1. Manifolds

A **topological manifold** of dimension n is a topological space that is locally homeomorphic to an open subset of $\mathbb{R}^n$. I impose further the requirement that the space be Hausdorff, and also a requirement of 'second countability'. This means that there exists a countable set of pairs (U_i, φ_i) in which (a) each U_i is an open subset of M and φ_i a homeomorphism onto an open subset of $\mathbb{R}^n$; (b) the U_i cover M. As a consequence:

1.1. Lemma. *Every open subset of a manifold is the union of a countable set of relatively compact open subsets U_i that can be embedded in $\mathbb{R}^n$.*

In particular, the manifold itself is a countable union of relatively compact open subsets.

Proof. It suffices to find a countable basis for the topology of $\mathbb{R}^n$. For this, one can take the disks with rational centre and rational radii. ∎

1.2. Lemma. *There exists a countable sequence of compact sets*

$$C_1 \subset \ldots \subset C_i \subset C_{i+1} \subset \ldots$$

with $C_i \subset C_{i+1}^\circ$ whose union is all of M.

Proof. We can find a countable basis of open sets $\{K_i\}$ for the topology on M such that each $\overline{K}_i$ is compact. Using these, we can define by induction the increasing union $\{C_i\}$ of compact subsets whose union is all of M:

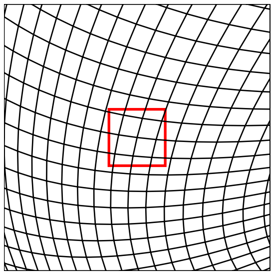
$$C_m = \begin{cases} \emptyset & \text{if } m \le 0 \\ \overline{K_1} & \text{if } m = 1 \\ \overline{K_1 \cup \ldots \cup K_r} & \text{if } m > 1 \text{ and } r \text{ is least with } C_{m-1} \subset \cup_1^r K_i \end{cases}$$

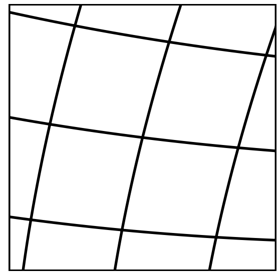
Sure enough, we have

$$C_i^\circ \subset C_i \subset C_{i+1}^\circ \subset C_{i+1} \subset C_{i+2}^\circ .$$ ∎

If (U_i, φ) and (U_j, φ_j) are two pairs in a cover of M, each of the maps φ_i and φ_j determines an embedding of the overlap $U_i \cap U_j$ into $\mathbb{R}^n$, and $\varphi_i \varphi_j^{-1}$ is a homeomorphism of $\varphi_j(U_i \cap U_j)$ with $\varphi_i(U_i \cap U_j)$. The manifold is **smooth** if these homeomorphisms are diffeomorphisms—i.e. smoothly invertible smooth maps—of $\varphi_j(U_i \cap U_j)$ with $\varphi_i(U_i \cap U_j)$. I shall call such a set of pairs (U_i, φ_i) a **defining** or **Euclidean** cover of M.

There is no condition on the smooth function $\varphi_i \varphi_j^{-1}$ other than that it be a diffeomorphism. The figure below shows a possibility. The basic fact of calculus is that although these maps in the large can be quite exotic, they become better and better approximated by linear maps at small scale:





A continuous function on M is called smooth when its restriction to each defining U_i is smooth with respect to the coordinates associated to φ_i. The definition suggests a natural way to construct smooth functions on M. Suppose we are given a defining cover (U_i, φ_i), and for each i a smooth function f_i on $X_i = \varphi(U_i)$ in $\mathbb{R}^n$, to which is associated a function

$$[\varphi^* f_i](u) = f_i(\varphi_i(u))$$

on U_i. This will define a smooth function on all of M if

$$f_i(\varphi_i(\varphi_j^{-1}(u)) = f_j(u)$$

on U_j.

Two defining covers (U_i, φ_i), (V_j, ψ_j) are equivalent if and only if the set of smooth functions determined by these on overlaps are the same.

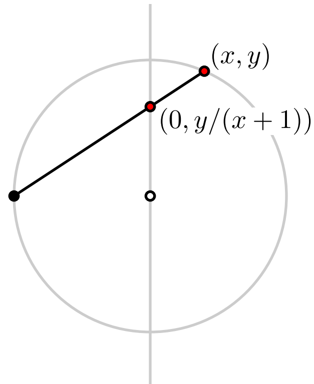
Suppose

$$\varphi: M \to N$$

to be a continuous map from one manifold to another. If f is a continuous function on N, then $[\varphi^* f](x)$ is a continuous function on M. The map φ itself is smooth if $\varphi^* f$ is smooth whenever f is. In practice, such maps can be obtained from pasting together maps defined on a cover of M.

Remark. In general, it is the diffeomorphisms on overlaps that define the structure of M. All manifolds are locally the same, but it has been known since the mid twentieth century that some topological manifolds possess globally inequivalent smooth structures! To gets some idea of what this means, look at the Wikipedia article on exotic spheres listed among the references.

Example. Let $\mathbb{S}$ be the unit circle in $\mathbb{C}$. Define V_+ to be the complement of -1, and V_- the complement of 1. Each of these is parametrized by $\mathbb{R}$, through stereographic projection, for example V_+ in the following figure. The point (x, y) on the circle corresponds to $z_+ = y/(1+x)$ in $\mathbb{R}$. And on V_- it corresponds to $z_- = y/(1-x) = 1/z_+$.



Example. The space $\mathbb{P}^1(\mathbb{R})$ is the set of lines through the origin in $\mathbb{R}^2$. The open sets are determined by the $\mathbb{R}^\times$-invariant open sets in the complement of the origin. This space is the union of two open sets U_0 and U_∞, each isomorphic to $\mathbb{R}$. The first is the set of lines intersecting the line $y = 1$, the second the set of lines intersecting $x = 1$. Each is parametrized by its point of intersection, either $(x, 1)$ or $(1, y)$. Since

$$[\,x \quad 1\,] = x\,[\,1 \quad 1/x\,]$$

we see that $y = 1/x$ on $U_0 \cap U_\infty$.

For example, the function $1/(x^2 + 1)$ on U_0 extends to a smooth function on all of $\mathbb{P}^1(\mathbb{R})$ that is equal to $y/(y^2 + 1)$ on U_∞.

Similarly, the space $\mathbb{P}^{n-1}(\mathbb{R})$ is the set of lines through the origin in $\mathbb{R}^n$. It is covered by the open sets U_i $(1 \le i \le n)$ of lines that intersect the hyperplane $x_i = 1$. The point of intersection determines a bijection of U_i with $\mathbb{R}^{n-1}$.

Remark. The space $\mathbb{P}^1(\mathbb{R})$ may in fact be identified with the unit circle $\mathbb{S}$ in $\mathbb{C}$. If z is a point in $\mathbb{S}$, define $\ell(z)$ to be the line through $\sqrt{z}$ and $-\sqrt{z}$. This is certainly well defined, and it can be seen to be smooth by looking at what it does on each of V_0 and V_∞. What's slightly peculiar about this map is that it is a diffeomorphism of the two manifolds, but not an isomorphism of algebraic varieties. (This is related to results of John Nash about smooth manifolds and real algebraic manifolds.)

∘ ———— ∘

REGULAR SUBMANIFOLDS. Recall that if f is a smooth function on $\mathbb{R}^n$, then $\operatorname{grad}(f)$ is the array $(\partial f/\partial x_i)$. Many smooth manifolds arise as a consequence of this:

1.3. Proposition. *Suppose f to be a smooth function on $\mathbb{R}^n$, and let M be the hypersurface $f(x) = 0$. If $\operatorname{grad}(f)$ does not vanish anywhere on M, it is a smooth manifold.*

For example, spheres of non-zero radius are manifolds.

Proof. We must define coordinate patches and verify smooth overlap compatibility.

Suppose x to be a point of M, so that $f(x) = 0$. By assumption, some $\partial f/\partial x_i \ne 0$ at x. There exists some affine change of coordinates such that x is the origin and $\partial f/\partial x_i = 0$ if $i < n$ and 1 if $i = n$. The hyperplane $x_n = 0$ is presumably the hyperplane tangent to the variety $f = 0$ passing through the origin. The basic idea is to use the approximation by this hyperplane to parametrize the hypersurface locally.

Let τ be the smooth map taking

$$x = (x_1, \ldots, x_{n-1}, x_n) \longmapsto (x_1, \ldots, x_{n-1}, f(x))\,.$$

Its Jacobian is

$$\begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ \partial f/\partial x_1 & \ldots & \ldots & \partial f/\partial x_n \end{bmatrix}$$

which is I at $x = 0$. By the inversion theorem (§III.6 in [Courant:1936]) it possesses a local inverse in some open neighbourhood U of the origin. Then $U \cap \tau^{-1}(\mathbb{R}^{n-1})$ is equal to $U \cap \{f = 0\}$, hence τ parametrizes it.

Smooth compatibility on overlaps is a straightforward consequence of this construction.

Example. Consider the unit sphere $\mathbb{S}^{n-1}$, embedded in $\mathbb{R}^n$. It is covered by the open hemispheres

$$U_i^+ = \{x \,|\, 0 < x_i \le 1\}, \quad U_i^- = \{x \,|\, -1 \le x_i < 0\}\,.$$

Projection from either of these onto the hyperplane where $x_i = 0$ is a bijection with the interior of the open unit disk in $\mathbb{R}^{n-1}$. Explicitly, on $U_i^\pm$

$$x_i = \pm\sqrt{1 - \sum\nolimits_{j \ne i} x_j^2}\,.$$

Example. The orthogonal group is made up of real matrices X such that

$$\mathrm{O}_n = \{X \in M_n(\mathbb{R}) \,|\, {}^tXX = I\}\,.$$

It too is a manifold. I'll not give full details here, but just show what happens near the identity I.

The points of the tangent hyperplane at I are those of the form $I + X$ in which ${}^t(I+X)(I+X) = I$ up to order two. This means that ${}^tX + X = 0$, or that X lies in the vector space Λ^n of anti-symmetric matrices. The role played by the variable x_n above is now played by the subspace of symmetric matrices S^n, since $M_n = \Lambda^n \oplus S^n$. We expect that, for X anti-symmetric and small, the map taking $X + Y$ to $(X, {}^t(I+X+Y)(I+X+Y) - I)$ is a locally invertible map to M_n. But

$$\begin{aligned} {}^t(I+X+Y)(I+X+Y) - I &= {}^tX + X + {}^tXX + {}^tY + Y + {}^tXY + {}^tYX + {}^tYY \\ &= \big(YX - XY\big) + {}^tXX + {}^tY + Y + {}^tYY\,. \end{aligned}$$

The matrix $YX - XY$ is anti-symmetric, so what we want to show is that for small X the map

$$Y \longmapsto -XX + {}^tY + Y + {}^tYY$$

is invertible—i.e. that for X small there exists Y making this vanish. But since ${}^tX = -X$ and ${}^tY = Y$, this amounts to

$$I + X^2 = (I+Y)^2$$

which can be solved by $I + Y = \sqrt{I + X^2}$. This makes sense because $I + X^2$ is symmetric, and positive if X is small.

∘——∘

VECTOR BUNDLES. If M and N are smooth manifolds, there is an obvious way in which to define the direct product $M \times N$ as one. But there is also a more subtle, if similar, way to make new manifolds from old ones, that generalizes the construction of products. Suppose B, M to be manifolds, $\pi\colon B \to M$ a surjective continuous map. Suppose that there exists a cover of M by open sets U_i and on each $\pi^{-1}(U_i)$ a bijection φ_i with $U_i \times \mathbb{R}^n$ compatible with projection to U_i. The **fibre** over u is the inverse image $\pi^{-1}(u)$, isomorphic to $\mathbb{R}^n$. On the fibre over a point u on $\varphi_j(U_i \cap U_j)$ the map $\varphi_i\varphi_j^{-1}$ is required to be an invertible linear transformation, smoothly varying with u. There exists a unique manifold structure on B compatible with these maps, and also a unique compatible vector space structure on each fibre $\pi^{-1}(u)$. The set (B, π, M) and the covers (U_i, φ_i) make up a **vector bundle** over M of fibre dimension n.

A **smooth section** s of a vector bundle (B, π, M) is a smooth map s from M to B such that $\pi(s(m)) = m$. The set of $\Gamma, M, B)$ of all smooth sections is a module over $C^\infty(M)$. A vector bundle is completely characterized by this module. Better, if M is connected, any module over $C^\infty(M)$ that is finitely generated and projective is the space of sections of some unique vector bundle. (These things are proved in Chapter 11 of [Nestruev:2003].)

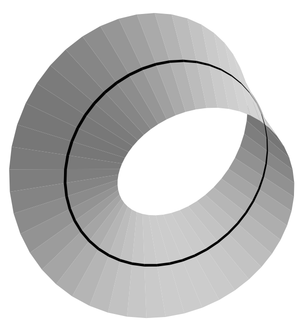
The simplest example is the product $M \times \mathbb{R}^n$, the **trivial bundle** of fibre dimension n. Others are more interesting. Every vector bundle over an open disk is trivial, as is shown nicely in the note by Dan Freed. It uses the parallel transport associated to a connection.

Example. Let $M = \mathbb{P}^1(\mathbb{R})$. Let B be the set of pairs (ℓ, x) where ℓ is a line through the origin in $\mathbb{R}^2$—i.e. a point of M—and x is a point on ℓ. Let π be the obvious projection from this onto M.

Let $V_0 = \pi^{-1}(U_0)$, $V_\infty = \pi^{-1}(U_\infty)$. A point in V_0 amounts to a line through some $(x, 1)$ together with a point on that line. We can parametrize that line by $\mathbb{R}$, taking t to the point (tx, t). Similarly, we can parametrize points on a line through $(1, y)$ by taking s to (s, sy). Thus we have a bijection of $\mathbb{R} \times \mathbb{R}$ with V_0, taking (x, s) to the pair $(\langle\!\langle x, 1\rangle\!\rangle, (sx, s)$, and similarly one of V_∞. Here $\langle\!\langle x, y\rangle\!\rangle$ is the line through (x, y). On the overlap $U_0 \cap U_\infty$ we can define a bijection of V_0 with V_∞. These define coordinate patches for B.

1.4. Proposition. *Every smooth section of B over M vanishes somewhere.*

Proof. The manifold M is the quotient of the unit circle $\mathbb{S}$ in $\mathbb{R}^2$ by ± 1, since every line through the origin in $\mathbb{R}^2$ passes through opposite points on the unit circle. The map $z \mapsto z^2$ identifies this quotient with $\mathbb{S}$ itself, so in fact $\mathbb{P}^1(\mathbb{R})$ with $\mathbb{S}$. Suppose we had a section $\ell \mapsto (\ell, x(\ell))$. If $x(\ell)$ is a non-zero point on ℓ then $x/\|x\|$ is a point on $\mathbb{S}$, so a non-zero section of B would give rise to a map back from M to $\mathbb{S}$. However, there are no such maps, as one can verify by considering what such a map has to look like on each U_i. ∎



Example. Let M be the two-sphere $\mathbb{S}^2$ in $\mathbb{R}^3$. At a point $\alpha = (a, b, c)$ of M the tangent plane is the surface

$$T_\alpha = \alpha + \alpha^\perp ,$$

where $\alpha^\perp$ is the plane through the origin perpendicular to α. Let B be the set of all pairs (α, v) with α in M and and v in $\alpha^\perp$. One can verify easily the definition of a vector bundle, applying the covering $\{U_i^\pm\}$ discussed earlier. Here, too, this is not a product bundle:

1.5. Proposition. *Every continuous vector field on the unit sphere in $\mathbb{R}^3$ must vanish somewhere.*

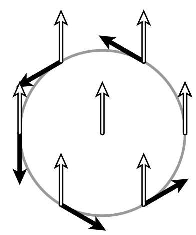
Proof sketch. Suppose φ to be a map from the unit circle to itself. Its **index** is the number of times it wraps around, which is also the integral

$$\frac{1}{2\pi} \int_0^{2\pi} \varphi^*(d\theta) .$$

It varies continuously with φ, but it also takes dsicrete values, which means that it remains constant.

Suppose x to be a point in $\mathbb{R}^2$, C a positively oriented circle with x in its interior, v a non-zero vector field defined throughout the interior of C, including C itself.

For each c in C, let $\theta(c)$ be the angle of rotation from $v(c)$ to $\tau(c)$, the positively oriented tangent vector at c. As c travels around C, the angle $\theta(c)$ describes a closed path in the unit circle. As the following figure shows, its index is 1 near x, but as C expands it does not change.



If a vector field on the sphere never vanishes, this index near at every point is 1. But this contradicts the fact that the orientation of the equator is different with respect to the two poles.

2. Advanced calculus

My principal reference for this section is [de Rham:1969].

EUCLIDEAN SPACE. The following is a familiar assertion about the Taylor series of a smooth function:

2.1. Lemma. *Suppose U to be a convex open neighbourhood of the origin in $\mathbb{R}^n$. If f is a smooth function on U then for any m it may be expressed as*

$$f(x) = \sum_{|k|<m} f^{(k)}(0)\,\frac{x^k}{k!} + \sum_{|k|=m} x^k f_k(x)$$

where each f_k is a smooth function on U.

Proof. I follow [Courant:1936] and [Courant:1937], and I first look at the case of dimension one. The fundamental theorem of calculus tells us that

$$f(x) - f(0) = \int_0^x f'(s)\,ds\,. \tag{2.2}$$

An easy estimate tells us that the integral is $O(x)$, but a simple trick will do better. If we set $s = tx$ this equation becomes

$$f(x) = f(0) + x\int_0^1 f'(tx)\,dt\,,$$

and the integral

$$f_1(x) = \int_0^1 f'(tx)\,dt$$

is a smooth function of x. Applying induction to f' in (2.2) gives us

$$f(x) = \sum_{k<m} c_k x^k + x^m f_m(x)$$

with $f_m(x)$ smooth. An easy calculation tells us that $c_k = f^{(k)}(0)/k!$.

In any dimension, for x in U define

$$\varphi(t) = f(tx)\,.$$

Then

$$\varphi(1) - \varphi(0) = \int_0^1 \varphi'(t)\,dt\,,$$

which because of our assumption and the chain rule translates to

$$f(x) - f(0) = \sum_{i=1}^{m} x_i \int_0^1 \frac{\partial f}{\partial x_i}\,(tx)\,dt = \sum_{i=1}^{n} x_i f_i(x)$$

with $f_i(x)$ smooth. Again apply induction.

If U is any open neighbourhood of p, let $\mathfrak{m}_0$ be the maximal ideal of f in $C^\infty(U)$ vanishing at p. The embedding of $\mathbb{C}[x]$ into $C^\infty(U)$ induces a ring homomorphism from $\mathbb{C}[x]/(x-p)^m$ in $C^\infty(U)/\mathfrak{m}_p^m$

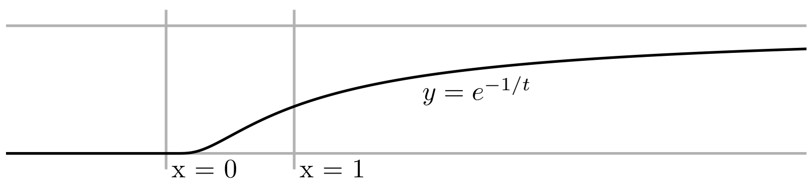
2.3. Corollary. *This ring homomorphism is an isomorphism.*

A well known result of Émile Borel asserts that this remains true in $m = \infty$, but I shan't need that.

BUMP FUNCTIONS. Since e^t grows more rapidly than any power of t as $t \to \infty$, the Taylor series of $e^{-1/t}$ (for $t > 0$) at the origin vanishes identically, so the function

$$\varphi(x) = \begin{cases} 0 & \text{if } t \le 0 \\ e^{-1/t} & \text{otherwise} \end{cases}$$

is smooth on all of $\mathbb{R}$. Its graph looks like this:



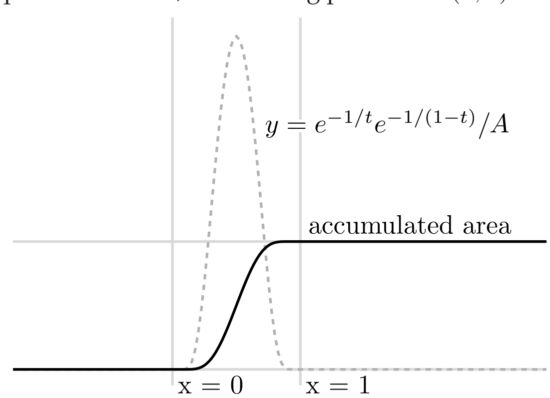
The function $e^{-1/t}e^{-1/(1-t)}$ then vanishes of infinite order at both 0 and 1, so the function specified to be 0 outside the interval $[0,1]$ and $e^{-1/t}e^{-1/(1-t)}$ inside it is also smooth on all of $\mathbb{R}$. If

$$A = \int_0^1 \varphi(t)\,dt$$

then

$$\psi(x) = \frac{1}{A}\int_{-\infty}^{x} \varphi(t)\,dt$$

vanishes for $t \le 0$ and is equal to 1 for $x\mathfrak{e}1$, while being positive in $(0,1)$. Its graph looks like this:



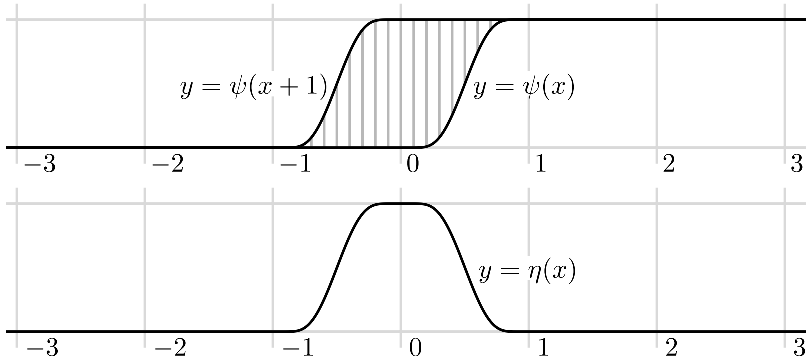
Let

$$\eta(x) = \psi(x+1) - \psi(x)\,.$$

It measures the distance between the function ψ shifted to the left by 1 and ψ itself. It is also characterized by

$$\eta(x) = \begin{cases} \psi(x+1) & \text{if } x \le 0 \\ 1 - \psi(x) & \text{otherwise.} \end{cases}$$

Here is its graph:

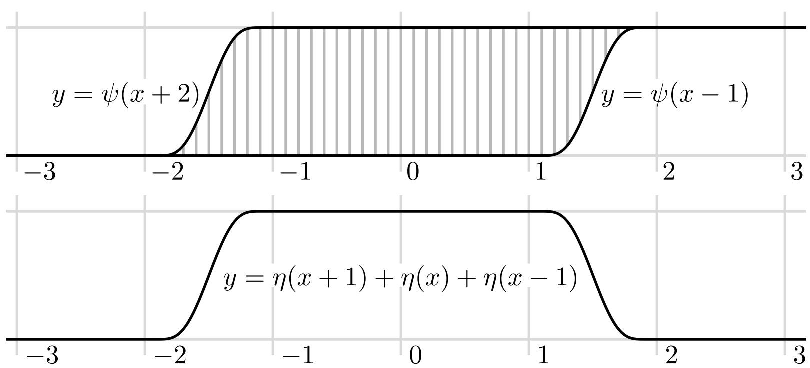


The function $\eta(x)$ is even. It vanishes for $|x|\mathfrak{e}1$ or, equivalently, $\eta(x) \neq 0$ if and only if $|x| < 1$. It takes value 1 at 0 and takes values everywhere else in $(0, 1)$. It is a **bump function**.

The special form of η gives it a very useful property. For $k \le \ell$ in $\mathbb{N}$ we have

$$\begin{aligned}
\eta(x+\ell) &+ \eta(x+\ell-1)) + \cdots + \eta(x+k) \\
&= \big(\psi(x+\ell+1) - \psi(x+\ell)\big) + \big(\psi(x+\ell) - \psi(x+\ell-1)\big) + \cdots + \big(\psi(x+k+1) - \psi(x+k)\big) \\
&= \psi(x+\ell+1) - \psi(x+k)\,.
\end{aligned}$$

It vanishes for $x \le -(\ell+1)$, $x\mathfrak{e}1-k$, and in the range $[-\ell, -k]$ is identically equal to 1. It is an approximation to the characteristic function of the interval $[-\ell, -k]$. For example, with $\ell = 1$, $k = -1$:



Taking limits as $k \to -\infty$, $\ell \to \infty$:

2.4. Lemma. *For any x in $\mathbb{R}$ the sum*

$$\sum\nolimits_{\lambda \in \mathbb{Z}} \eta(x-\lambda)$$

is locally finite and equal to 1.

Something similar happens in all dimensions. On $\mathbb{R}^n$ define the norm

$$|x| = \sup_i |x_i|\,,$$

and set

$$\eta_n(x) = \prod\nolimits_1^n \eta(x_i)\,.$$

This is a smooth function on $\mathbb{R}^n$, centred at 0. It is equal to 1 at 0, and takes values in $[0, 1]$ everywhere. It is worth noting formally:

2.5. Lemma. *Suppose x in $\mathbb{R}^n$. Then $\eta_n(x) \neq 0$ if and only if $|x| < 1$.*

In other words, the support of η_n is a cube centred at the origin.

2.6. Proposition. *The infinite sum*

$$\sum_{\lambda\in\mathbb{Z}^n}\eta_n(x-\lambda)$$

is locally finite and identically equal to 1*.*

Proof. Because it is equal to

$$\prod_1^n\Big(\sum_{\lambda_i\in\mathbb{Z}}\eta(x_i-\lambda_i)\Big)$$

which according to Lemma 2.4 is identically 1.

Suppose A to be a compact subset of $\mathbb{R}^n$ contained in an open set U. Let B be the closed complement of U, and for x in $\mathbb{R}^n$ let

$$\operatorname{dist}(x,B)=\inf_{b\in B}|a-b|\,.$$

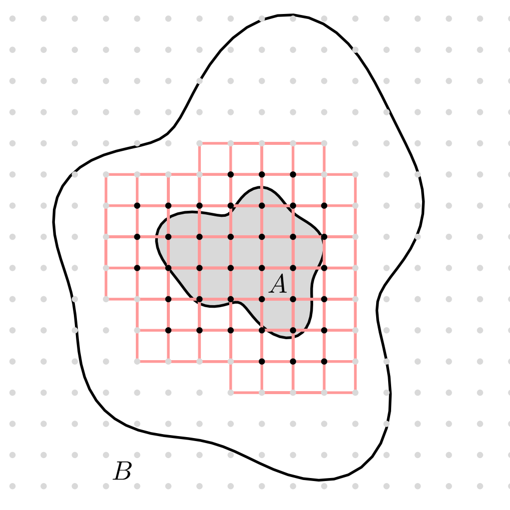
The function $\operatorname{dist}(x,B)$ is continuous, and vanishes only on B. On the compact set A it therefore takes a minimum value $\delta=\operatorname{dist}(A,B)>0$. For any a in A the disk $|x-a|<\delta$ is contained in U.

2.7. Corollary. *Suppose* $A\subset\mathbb{R}^n$ *to be a compact subset,* $\varepsilon>0$*. Let* Λ *be the set of all* λ *in* $\mathbb{Z}^n$ *such that there exists some point* a *of* A *such that* $|a-\varepsilon\lambda|<1$*. The sum*

$$\sum_{\lambda\in\Lambda}\eta(x-\varepsilon\lambda)\,.$$

is equal to 1 *for every* x *in* A*, and vanishes if* $|x-a|$¢2ε *for all* a *in* A*.*

2.8. Corollary. *Suppose* $A\subset\mathbb{R}^n$ *to be compact,* B *a closed set in the complement of* A*. There exists a smooth function equal to* 1 *on* A *that vanishes on* B*.*



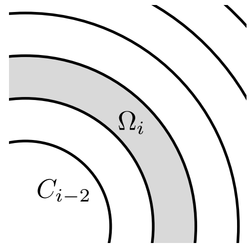
Proof. Apply Corollary 2.7 with $\varepsilon<\operatorname{dist}(A,B)/2$.

PARTITIONS OF UNITY. Let M be a smooth manifold.

2.9. Proposition. *Suppose* $\{A_\alpha\}$ *to be a covering of* M*. There exists a countable, locally finite covering by open sets* $\{V_i\}$ *such that each* $\overline{V}_i$ *is compact and contained in some open* U_i *that is in turn contained in some* A_α*.*

Proof. Apply Lemma 1.2 to obtain a sequence of relatively compact open sets C_i that cover M, such that $C_i\subset C_{i+1}^\circ$.

Each $\Omega_i=C_i-C_{i-1}^\circ$ is compact, M is the union of the Ω_i, and Ω_i is disjoint from C_{i-2}.



If x is in M, it will lie in some Ω_i, and there will exist (i) some open coordinate patch U containing it which is in turn contained in some A_α as well as in $C_{i+1} - C^\circ_{i-1}$ and (ii) some open V with $\overline{V} \subset U$. The compact set Ω_i will be covered by a finite number of these V. The union of all of these as i varies provides the cover we want.

2.10. Corollary. *Suppose given an open covering $\{A_\alpha\}$ of M.. There exists a countable family of smooth functions $\{f_i\}$ on M such that (a) $f_i(x)\mathfrak{e}0$ for all x; (b) the support of f_i is contained in some A_α; (c) the colection of sets $\mathrm{supp}(f_i)$ is locally finite; $\sum f_i(x) = 1$ for all x.*

Proof. Replace the A_α if necessary by a refinement of coordinate patches. Find a covering by sets $V_i \subset U_i$ as in the previous result. By Corollary 2.8 there exists a smooth function $\varphi_i\mathfrak{e}0$ equal to 1 on V_i with support in U_i. Define

$$f_i(x) = \frac{\varphi_i(x)}{\sum \varphi_i(x)} .$$

2.11. Proposition. *If A is a compact subset of M and B a closed subset of its complement, then there exists a smooth function that is 1 on A and vanishes on B.*

Proof. Combine Corollary 2.8 with Corollary 2.10.

2.12. Corollary. *Suppose A to be a compact set contained in the open set U, f a smooth function on U. There exists a smooth function on all of M whose restriction to A is the same as f.*

Proof. Let χ have support in U, $\chi = 1$ on A. Then $f\chi$ is the function we want.

MAXIMAL IDEALS ON M. Suppose x to be a point in M, and let $\mathfrak{m}_x$ be the maximal ideal of $C^\infty(M)$ consisting of functions on M vanishing at x. Suppose U to be any neighbourhood of x diffeomorphic to an open subset of $\mathbb{R}^n$, and let $\mathfrak{u}_x$ be the maximal ideal of functions in $C^\infty(U)$ vanishing at x.

2.13. Proposition. *Suppose M to be a smooth manifold, x a point on M, and suppose U to be a neighbourhood of x. The canonical homomorphism from $C^\infty(M)/\mathfrak{m}_x^m$ to $C^\infty(U)/\mathfrak{u}_x^m$ is an isomorphism.*

Proof. Suppose f to be in $C^\infty(U)$. Choose V a neighbourhood of x with $\overline{V}$ compact and included in U. According to Proposition 2.11, there exists φ in $C^\infty(M)$ equal to f on $\overline{V}$. Then $f - \varphi$ is identically 0 in a neighbourhood of x. Apply:

2.14. Lemma. *Any function in $C^\infty(M)$ vanishing identically in the neighbourhood of x is in every $\mathfrak{m}_x^m$.*

Proof. Let F be such a function. Choose χ in $C^\infty(M)$ vanishing identically near x and identically 1 on the support of F. Then $\chi^m F = F$ for all m.

RECOVERING *M* FROM ITS SMOOTH FUNCTIONS. Any point p of M gives rise to a ring homomorphism

$$\mathfrak{e}_p\colon\ C^\infty(M) \longrightarrow \mathbb{R}, \qquad f \longmapsto f(p) .$$

2.15. Proposition. *Any homomorphism of $\mathbb{R}$-algebras from $C^\infty(M)$ to $\mathbb{R}$ is equal to $\mathfrak{e}_p$ for some p in M.*

Proof. I take this from Chapter 7 of [Nestruev:2003]. Let

$$C_1 \subset \ldots \subset C_i \subset C^\circ_{i+1} \subset C_{i+1} \subset \ldots$$

be the sequence of compact sets guaranteed by Lemma 1.2. Let σ_i be a function equal to 0 in C_{i-1} and 1 in the complement of C°_i. The series $\sigma = \sum \sigma_i$ is locally finite and $\sigma^{-1}(I)$ is compact if and only if I is.

Suppose F to be an $\mathbb{R}$-homomorphism from $C^\infty(M)$ to $\mathbb{R}$ that is equal to no $\mathfrak{e}_p$. But two $\mathbb{R}$-algebra homorphisms are equal if and only if their kernels are equal, so this means that $\mathrm{Ker}(F)$ is distinct from all $\mathfrak{m}_p$.

Hence for each p there exists f_p with $f_p(p) \neq 0$ but $F(f_p) = 0$. Let U_p be the set of all x such that $f_p(x) \neq 0$. It is a neighbourhood of p, and these cover M.

The set $\sigma^{-1}(F(\sigma))$ is compact, so covered by a finite number of the U_x, say the U_i corresponding to p_i with $\sigma(p_i) = F(\sigma)$. Let

$$g = (\sigma - F(\sigma))^2 + \sum_i f_{p_i}^2 ,$$

so that $F(g) = 0$. If $\sigma(x) \neq F(\sigma)$ then the first term does not vanish, while if $\sigma(x) = F(\sigma)$ then some term in the second sum does not vanish. Hence $g(x) > 0$ for all x, g is invertible in $C^\infty(M)$, and $F(g) \neq 0$, a contradiction.

3. Tangent vectors

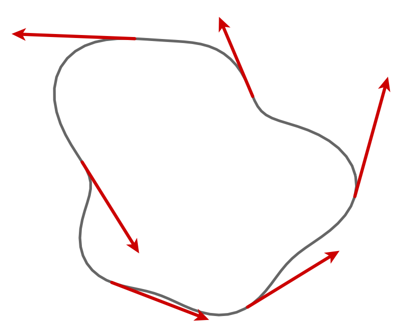
The basic model for tangent vectors in differential geometry is velocity. Suppose

$$\gamma\colon\ t \longmapsto [\, x_1(t) \quad \ldots \quad x_n(t)\,]$$

to be a parametrized path in $\mathbb{R}^n$. Its velocity at t is

$$\lim_{h\to 0} \frac{\gamma(t+h) - \gamma(t)}{h} = [\, x_1'(t) \quad \ldots \quad x_n'(t)\,] \ .$$

In low dimensions the velocities can be plotted graphically:



To each parameter value t is assigned an array determining the *location* of $\gamma(t)$, and attached to this location is a second array determining its *velocity*. The velocity itself does not have a location—a particle moving in space may well have the same velocity at different points on the path. The velocity represents the translation of a parallel vector based at the origin.

Tangent vectors in $\mathbb{R}^n$ are thus simple to characterize in terms of arrays. But how are tangent vectors on an arbitary manifold to be defined?

NOTATION. In the figure above, it should be clear that coordinate arrays play at least two distinct roles, and I want to introduce notation that will take this into account.

Let me begin with a very general situation. If G is any group, a *principal homogeneous space* for G is something on which the group acts transitively and for which all isotropy subgroups are trivial. Of course if you fix any particular point, it becomes a copy of G itself, but in practice there might not be any canonical choice of a particular point, and it can be confusing to try to make one. In the case I am interested in, the group will be the additive group of $\mathbb{R}^n$. In one interpretation of $\mathbb{R}^n$, points record *location*, and in the other *displacement*. I shall try to keep these roles separate by expressing the set of locations by $\mathbb{R}^n$, and the displacements by $\boldsymbol{R}^n$. The latter is a vector space: following one displacement by another amounts to addition of vectors, and you can also multiply a displacement by a scalar, getting $u+v$ or cv. In the other, one can displace a location by a vector: $p \mapsto q = p+v$. Or subtract one location from another by a displacement: $p = q - v$. Or interpolate two locations, since

$$(1-s)p + sq = p + s(q-p)\,.$$

But adding two locations doesn't make sense. So the two different notions are distinguished by what operations you can perform on them.

I'll also distinguish the vector dual $\widehat{\boldsymbol{R}}^n$ from $\boldsymbol{R}^n$. It is made up of linear functions on $\boldsymbol{R}^n$. In contrast with a common convention, for reasons I'll mention later, vectors in $\boldsymbol{R}^n$ will be row vectors, and those in $\widehat{\boldsymbol{R}}^n$ will be columns. Thus an element of $\widehat{\boldsymbol{R}}^n$ is a matrix

$$\widehat{v} = \begin{bmatrix} \widehat{v}_1 \\ \cdots \\ \widehat{v}_n \end{bmatrix},$$

and duality is expressed by a matrix multiplication:

$$\langle v, \widehat{v} \rangle = v \cdot \widehat{v}\,.$$

VECTOR DERIVATIVES. If f is a smooth scalar function defined on an open subset U of $\mathbb{R}^n$ and v is in $\boldsymbol{R}^n$ then

$$\lim_{t \to 0} \frac{f(x+tv) - f(x)}{t} = [\nabla_v f](x)$$

exists and defines a smooth function on U. This is a linear function of both f and v. It is a derivation in the sense that

(3.1) $$\nabla_v(fg) = (\nabla_v f) \cdot g + f \cdot (\nabla_v g)\,.$$

Because of this, ∇_v vanishes on every $\mathfrak{m}_x^2$, and hence defines a linear functional on the finite-dimensional vector space $\mathfrak{m}_x/\mathfrak{m}_x^2$. Likewise, f determines a linear functional df on $\boldsymbol{R}^n$—i.e. the image of $f - f(x)$ in $\mathfrak{m}_x/\mathfrak{m}_x^2$. The variables x_i determine a basis dx_i of $\mathfrak{m}_x/\mathfrak{m}_x^2$, with dual basis $\partial/\partial x_i$ associated to the standard basis vector ε_i. Explicitly, we can write

$$df = \sum \left(\frac{\partial f}{\partial x_i} \right) dx_i$$

and then

$$\langle v, df \rangle = \sum v_i \left(\frac{\partial f}{\partial x_i} \right)$$

if $v = \sum v_i(\partial/\partial x_i)$.

We now have the linear approximation

$$f(x+v) = f(x) + \langle v, df \rangle_x + O(\|v\|^2)$$

near any point x of U. This holds equally for the components of any smooth function of several variables. If $x(t)$ is a smooth function from $U \subseteq \mathbb{R}^n$ to $\mathbb{R}^m$ we have for small h in $\boldsymbol{R}^n$ the linear approximation

$$x(t+h) \sim x(t) + [\,h_1 \quad \ldots \quad h_n\,] \begin{bmatrix} \partial x_1/\partial t_1 & \ldots & \partial x_m/\partial t_1 \\ \ldots & \ldots & \ldots \\ \partial x_1/\partial t_n & \ldots & \partial x_m/\partial t_n \end{bmatrix} = x(t) + h \left[\frac{\partial x}{\partial t} \right].$$

This leads immediately to the chain rule. Suppose we are given two maps

$$\begin{aligned} t &\longrightarrow x(t), \quad \mathbb{R}^n \longrightarrow \mathbb{R}^n \\ x &\longrightarrow f(x), \quad \mathbb{R}^n \longrightarrow \mathbb{R} \end{aligned}$$

Define $F(t)$ to be the scalar function $f(x(t))$. Then the composition of functions amounts to matrix multiplication of the linear approximations. Hence

$$\left[\frac{\partial F}{\partial t}\right]_t = \left[\frac{\partial x}{\partial t}\right]_t \left[\frac{\partial f}{\partial x}\right]_{x(t)},$$

or

$$\left[\frac{\partial f}{\partial x}\right]_{x(t)} = \left[\frac{\partial x}{\partial t}\right]_t^{-1} \left[\frac{\partial F}{\partial t}\right]_t .$$

In these circumstances $F = x^*f$, so we can rephrase this:

$$x^* \left[\frac{\partial f}{\partial x}\right] = \left[\frac{\partial x}{\partial t}\right]_t^{-1} \left[\frac{\partial (x^*f)}{\partial t}\right]_t .$$

And in a commutative diagram:

$$\begin{array}{ccc} f(x) & \xrightarrow{\varphi^*} & F(t) \\ \downarrow & & \downarrow \\ \dfrac{\partial f}{\partial x} & \xrightarrow[\varphi^*]{} & \left[\dfrac{\partial x}{\partial t}\right]^{-1} \dfrac{\partial F}{\partial t} \end{array}$$

There is yet another formulation. Suppose first of all that

$$\varphi\colon\ t \longmapsto x(t), \quad U \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

to be any smooth map. If f is any smooth scalar function on a subset of $x(U)$ then x^*f is a smooth function on U, and if f is in $\mathfrak{m}^k_{x(t)}$ then x^*f is in $\mathfrak{m}^k_t$. The map x thus induces a linear map from $\mathfrak{m}_{x(t)}/\mathfrak{m}^2_{x(t)}$ to $\mathfrak{m}_t/\mathfrak{m}^2_t$, hence also a dual map x_* from tangent vectors at t to those at $x(t)$.

Now suppose x to be locally invertible. Then at $x(t)$ we have two bases of the tangent space—the $\partial/\partial x_i$ and the $x_*(\partial/\partial t_i)$. The first can therefore be expressed in terms of the second. The diagram above asserts that in fact

(3.2) $$\left[\frac{\partial}{\partial x}\right] = \left[\frac{\partial x}{\partial t}\right]^{-1} x_* \left[\frac{\partial}{\partial t}\right] .$$

Example. Let $f(x)$ be a smooth function of one variable, and set $F(t) = f(x(t))$. The basic formula now reads

$$\frac{df}{dx} = \frac{dF}{dt}\Big/\frac{dx}{dt}$$

To see how this goes, say $x = 1/t$, and $F(t) = f(1/t)$. Then $dx/dt = -1/t^2$ and

$$f'(x) = -t^2 F'(t) .$$

Example. Polar coordinates amount to a map from $\mathbb{R}^2$ to $\mathbb{R}^2$:

$$(r, \theta) \longmapsto (x, y) = (r\cos(\theta), r\sin(\theta)) .$$

Thus

$$\begin{bmatrix} \partial x/\partial r & \partial y/\partial r \\ \partial x/\partial\theta & \partial y/\partial\theta \end{bmatrix} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -r\sin(\theta) & r\cos(\theta) \end{bmatrix}$$

and then

$$\begin{aligned} \frac{\partial}{\partial x} &= \cos(\theta)\,\frac{\partial}{\partial r} - \frac{\sin(\theta)}{r}\,\frac{\partial}{\partial\theta} \\ \frac{\partial}{\partial y} &= \sin(\theta)\,\frac{\partial}{\partial r} + \frac{\cos(\theta)}{r}\,\frac{\partial}{\partial\theta}\,. \end{aligned}$$

∘———∘

CHANGE OF VARIABLES. The characterization of tangent vectors at a point x in $\mathbb{R}^n$ as linear functionals on $\mathfrak{m}_x/\mathfrak{m}_x^2$ can be transferred immediately as a definition to an arbitrary manifold.

3.3. Definition. *Suppose M to be a manifold. A tangent vector at a point x of M is a linear functional on the space $\mathfrak{m}_x/\mathfrak{m}_x^2$.*

Proposition 2.13 assures us that the dimension of $\mathfrak{m}_x/\mathfrak{m}_x^2$ is the same as that of M, say n. The set of all tangent vectors is that of all pairs (x, v) in which v is a tangent vector at x. If $M \subseteq \mathbb{R}^n$ we know that this can be identified with the direct product $M \times \boldsymbol{R}^n$.

But now suppose we are given two pairs (U, φ) and (V, ψ). Let x, y be the associated coordinate functions on U, V, and that v is a tangent vector at a point p in $U \cap V$. If

$$v = \sum a_i \cdot \partial/\partial x_i\,,$$

is a vector expressed in x-coordinates, what is its expression

$$v = \sum b_i \cdot \partial/\partial y_i$$

in y-coordinates? The following is an immediate consequence of (3.2) :

3.4. Proposition. *The vectors*

$$u = \sum a_i \cdot \partial/\partial x_i, \quad v = \sum b_i \cdot \partial/\partial y_i)$$

are equal if and only if

$$b = a\,(\partial x/\partial y)^{-1}\,.$$

This leads to:

3.5. Corollary. *The set of tangent vectors on a manifold is a vector bundle on it.*

Proof. Local isomorphisms on overlaps are specified by the Proposition. ▮

4. Vector fields and derivations

A **vector field** on M assigns a vector $v(p)$ to every point p of M. It is said to be smooth if in every coordinate patch it is of the form

$$\sum v_i(x)\,\frac{\partial}{\partial x_i}$$

for smooth functions $v_i(x)$.

There is an algebraic characterization of vector fields. A **derivation** of a ring R over $\mathbb{R}$ is an $\mathbb{R}$-linear map D from R to itself such that

$$D(rs) = D(r)s + rD(s)$$

for all r, s in R. The condition of $\mathbb{R}$-linearity means neither more nor less than that it annihilates the constants. It is clear that a vector field on M determines a derivation of $C^\infty(M)$, but even better a derivation of $C^\infty(U)$ for every open $U \subseteq M$. Conversely:

4.1. Theorem. *Every derivation of $C^\infty(M)$ is that associated to some smooth vector field on M.*

The proof begins with the case that M is an open subset of $\mathbb{R}^n$.

4.2. Proposition. *If U is an open subset of $\mathbb{R}^n$, then every $\mathbb{R}$-linear derivation of $C^\infty(U)$ is that determined by a smooth vector field.*

Proof. Suppose D to be a derivation of $C^\infty(U)$. Let $v_i = D(x_i)$ (a function of x). We want to show that

$$D(f) = \sum_i v_i\,\frac{\partial f_i}{\partial x_i}\,,$$

and it suffices to show this at any point $p = (p_i)$ of U. In some convex neighbourhood of p

$$\begin{aligned} f(x) - f(p) &= \sum (x_i - p_i) f_i(x) \\ Df &= \sum_i D(x_i) f_i + (x_i - p_i) Df_i \\ [Df](p) &= \sum_i v_i(p) f_i(p) \\ [Df](p) &= \sum_i v_i(p)\,\frac{\partial f}{\partial x_i}(p)\,. \end{aligned}$$

4.3. Lemma. *If D is a derivation of $C^\infty(M)$ then the support of Df is contained in the support of f.*

Proof. It has to be shown that if f vanishes on an open set U in M, then so does Df. Suppose p to be a point of U. We can find a smooth function φ which is 1 on the support of f and vanishes at p. Then $\varphi f = f$ and

$$\big[D(\varphi f)\big](p) = [Df](p)\varphi(p) + f(p)[D\varphi](p) = 0\,.$$

So $Df = D(\varphi f)$ also vanishes at every point of U.

As a consequence:

4.4. Lemma. *If $f = \sum f_i$ is a locally finite sum of functions in $C^\infty(M)$ and D is a derivation of $C^\infty(M)$, then the sum $\sum Df_i$ is also locally finite and equal to Df.*

4.5. Proposition. *If M is an arbitrary manifold, U an open subset of M, and D a vector field on M, there exists a unique derivation of $C^\infty(U)$ that agrees with the restriction of functions from M to U.*

Proof. Suppose f in $C^\infty(U)$. We must define Df as a smooth function on U. Choose a partition of unity $(\psi i, U_i)$ and compatible coordinate charts $\varphi_i)$ on U. if f lies in $C^\infty(U)$, we have

$$f = \sum f_i \quad (f_i = \psi_i f)\,.$$

Each f_i may be extended to a smooth function on all of M, equal to 0 outside the support of f_i Since the U_i make a locally finite covering, the sum $D_U f = \sum D f_i$ is also locally finite, and hence a smooth function on U. I claim that D_U is a derivation of $C^\infty(U)$. But if $f = \sum f_i$ and $g = \sum g_i$ then the sum

$$fg = \sum_{i,j} f_i g_j$$

is also locally finite. According to the previous Lemma, so are the sums

$$D(fg) = \sum D(f_i g_j) = \sum D(f_i) g_j + f_i D(g_j) = D_U(f) \cdot g + f \cdot D_U(g) .$$

MISCELLANEOUS. • If X and Y are two vector fields, their **Poisson bracket** $[X, Y]$ is the difference $XY - YX$. It is *a priori* a differential operator possibly of order 2, but in fact it is a vector field. This can be seen by a local computation, but also by checking directly that it is a derivation.

• The vector bundle dual to the tangent bundle is the **cotangent bundle** $T_M^\vee$, whose fibre at a point p is $\mathfrak{m}_p/\mathfrak{m}_p^2$. It is not quite obvious how to directly construct that bundle, nor how to construct more generally a bundle over M whose fibre at p is $\mathfrak{m}_p^{m-1}/\mathfrak{m}_p^m$, nor yet a bundle whose fibre at p is $\mathcal{C}_p/\mathfrak{m}_p^m$.

• Any vector field on a smooth manifold determines on it a **flow**, which determines a trajectory through each point tangential to the vector field at that point. In local coordinates, this comes down to solving an ordinary differential equation. More precisely, if the vector field is X then the trajectory of the flow starting at the point x of G is a smooth path $t \mapsto E(t) = E_X(t)$ satisfying the differential equation $E'(t) = X_{E(t)}$ and initial condition $E(0) = x$. In local coordinates, if $X = \sum X_i(x)(\partial/\partial x_i)$ and $E(t) = \big(x_i(t)\big)$ then d/dt maps to $\sum_i (dx_i/dt)(\partial/\partial x_i)$ and the differential equation is that determined by system of equations

$$dx_i/dt = X_i(x) .$$

Locally, it follows from well known results on about dependence on initial conditions that the flow determines a one-parameter family of diffeomorphisms.

Part II. Lie algebras and vector fields

5. The Lie algebra of a Lie group

Now suppose G to be a Lie group—that is to say, a smooth differentiable manifold with a smooth structure as topological group. This means that multiplication $G \times G \to G$ and the inverse $G \to G$ are smooth maps.

The **Lie algebra** $\mathrm{Lie}(G)$ of G is defined to be the space of left-invariant vector fields on G. If g is any element of G and f a smooth function defined in the neighbourhood U of g then $L_g^* f$ is the pullback of f to $g^{-1}U$, a neighbourhood of I, defined by the formula $L_g^* f(x) = f(gx)$. This induces a linear map $L_{g,*}$ from the tangent space at I to that at g:

$$\langle L_{g,*}X, f\rangle = \langle X, L_g^* f\rangle$$

That the function of g is smooth is a consequence of the following elementary remark: suppose M and N to be manifolds, n a point of N and V a vector at n. It determines at each point of $M \times \{n\}$ a 'vertical' vector field. For any smooth function on $M \times N$ the function $V \cdot f$ evaluated on M is smooth. This is to be applied to $G \times \{1\}$. Conversely, any element X of $\mathfrak{g}$ determines a smooth vector field, which I'll call R_X for reasons to be seen in a moment, on all of G. It is invariant with respect to all left G-translations. It is unique subject to the two conditions that it be left-invariant and agree with X at I. Therefore:

5.1. Proposition. *The map taking a left-invariant vector field to its value at* 1 *is an isomorphism of* $\mathrm{Lie}(G)$ *with the tangent space at* 1.

The flow associated to R_X is also left-invariant—if $E_X(t)$ is the trajectory starting at I, the trajectory at g is $gE_X(t)$. The flow on G is generated by the left translations of the trajectory $E_X(t)$ starting at I. The defining equation for E_X implies that $E_X(s)E_X(t) = E_X(s+t)$, which because of continuity implies that $E_X(t) = E_{tX}(1)$. This last is written as the Lie exponential map $\exp(tX)$. That is to say

$$R_X F(g) = \frac{d}{dt} F\big(g\, \exp(tX)\big)\Big|_{t=0} .$$

Let's look at some examples.

Example. *The additive group* $\mathbb{R}$ *of real numbers.* The space $\mathfrak{g}$ is $\mathbb{R}$. The invariant vector fields are the $a\,d/dx$ where a is a constant. The corresponding flow $x(t)$ satisfies the differential equation

$$x'(t) = a$$

and the flow starting at x is $x(t) = x + at$.

Example. *The multiplicative group of non-zero real numbers* $\mathbb{R}^\times$*.* The space $\mathfrak{g}$ may again be identified with $\mathbb{R}$, where a corresponds to the vector ad/dx at 1. Multiplicative translation by x takes a at 1 to $ax\,d/dx$. The invariant vector fields are therefore the $ax\,d/dx$, the differential equation of the flow $x'(t) = ax(t)$, and the corresponding flow at 1 takes t to $\exp(ta)$. The connected component of positive real numbers is isomorphic to the additive group through the exponential map, but both $\exp$ and its inverse $\log$ are analytic functions. In terms of algebraic structure the two groups are thus distinct.

Example. *The multiplicative group* $\mathbb{S}$ *of complex numbers of unit magnitude.* The space $\mathfrak{g}$ is again $\mathbb{R}$, with a corresponding to $a\,d/d\theta$ at 1, where θ is the argument. The invariant vector fields are the $a\,d/d\theta$. The flow starting at 1 is the complex exponential function $\exp(iat)$.

Example. *The multiplicative group of non-zero complex numbers* $\mathbb{C}^\times$*.* The space $\mathfrak{g}$ may be identified with $\mathbb{C}$. The flows at 1 are the complex exponential $\exp(tz)$, and the trajectories will usually be spirals. This group contains the previous two as subgroups.

Example. *The multiplicative group* $\mathrm{GL}_n(\mathbb{C})$ *of invertible real* $n \times n$ *matrices.* This group is an open subset of the matrices $M_n(\mathbb{C})$, and the space $\mathfrak{g}$ may be identified with it. The invariant vector fields are the $X\partial/\partial X$ for invertible X. The flow at I tangent to A is the matrix exponential $\exp(tA)$, where

$$e^X = \exp(X) = I + X + \frac{X^2}{2!} + \cdots$$

The series converges for all X, and by the implicit function theorem maps a neighbourhood of 0 in $\mathrm{M}_n(\mathbb{C})$ isomorphically onto a neighbourhood of I in $\mathrm{GL}_n(\mathbb{C})$. We can calculate explicitly some simple examples for $n = 2$:

$$\begin{aligned}
\exp\left(\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}\right) &= \begin{bmatrix} e^a & 0 \\ 0 & e^b \end{bmatrix} \\
\exp\left(\begin{bmatrix} 0 & x \\ 0 & 0 \end{bmatrix}\right) &= \begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix} \\
\exp\left(\begin{bmatrix} 0 & -t \\ t & 0 \end{bmatrix}\right) &= \begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix} .
\end{aligned}$$

For matrices near enough to I we can write out explicitly an inverse to the exponential map, by means of the series

$$\log(I + X) = X - \frac{X^2}{2} + \frac{X^3}{3} - \cdots ,$$

which converges for X small.

In general, the relationship between the group and the Lie algebra $\mathfrak{g}$ is intuitive. Elements in a group very close to I are of the form $I + X+$ *higher order terms*, where X is a very small element of the Lie algebra. Intuitively, the Lie algebra represents very, very small motions in the group, those not differing much from the identity. The exponential map makes this rough idea precise.

5.2. Proposition. *On any Lie group, the exponential map has these properties:*

(a) *if X and Y commute then $e^{X+Y} = e^X e^Y$;*
(b) *the inverse of e^X is e^{-X};*
(c) *if g is invertible then $e^{\mathrm{Ad}(g)\,Y\mathrm{Ad}(g)^{-1}} = ge^Y g^{-1}$;*

If $G = \mathrm{GL}_n(\mathbb{R})$:

(d) *the exponential of a diagonal matrix with entries d_i is the diagonal matrix with entries e^{d_i};*
(e) *the exponential of the transpose is the transpose of the exponential;*
(f) *the determinant of e^X is $e^{\mathrm{trace}(X)}$.*

Proof. All are straightforward except possibly (f), for which I'll give two proofs.

First proof. If X is an arbitrary complex matrix then some conjugate $X_* = YXY^{-1}$ will be in Jordan form $S + N$, where S is diagonal, N is nilpotent, and the two commute. The matrices X and X_* will have the same determinant and the same trace, so by (c) we may as well let $X = X_*$. But

$$e^X = e^S e^N$$

will again be in Jordan form, with diagonal equal to that of e^S. Hence the trace of Y is that of S and the determinant of e^X is that of e^S. But for S the claim is clear.

Second proof. It suffices to show that the function $D(t) = \det(e^{tX})$ satisfies the ordinary differential equation

$$D'(t) = \mathrm{trace}(X)D(t)\ ,$$

since it certainly satisfies the initial condition $D(0) = 1$. But

$$\begin{aligned} D'(t) &= \lim_{h\to 0} \frac{D(t+h) - D(t)}{h} \\ \frac{D(t+h) - D(t)}{h} &= \frac{\det\left(e^{(t+h)X}\right) - \det\left(e^{tX}\right)}{h} \\ &= \frac{\det\left(e^{tX}\right)\det\left(e^{hX}\right) - \det\left(e^{tX}\right)}{h} \\ &= \det\left(e^{tX}\right)\left[\frac{\det\left(e^{hX}\right) - 1}{h}\right] \end{aligned}$$

which has as limit $D(t)\,\mathrm{trace}(X)$ since $\det(I + hX + \cdots) = 1 + h\,\mathrm{trace}(X) + \cdots$

As one consequence, the Lie algebra of SL_n is the space of matrices with vanishing trace. This technique can be formalized. If G is a group defined by polynomials $P(x) = 0$ in $M_n(\mathbb{C})$, then the tangent space at I is defined by conditions $P(I + \varepsilon X) = 0$ modulo second order terms in ε. But $P(I + \varepsilon X) = P(I) + \varepsilon\langle\nabla P, X\rangle$ modulo ε^2, and this is just $\varepsilon\langle dP, X\rangle$ since I is in the group and $P(I) = 0$. Therefore the Lie algebra is defined by the equations

$$\langle dP(X), X\rangle = 0\,.$$

This agrees with what we just showed for SL_n, since the differential of $\det$ is trace. This will show, for example, that the Lie algebra of the symplectic group Sp is the space of matrices X such that ${}^tX\,J + JX = 0$ where

$$J = \begin{bmatrix} 0 & -I \\ I & 0 \end{bmatrix}$$

Of course the exponential maps real matrices onto real matrices, so the tangent space of $\mathrm{GL}_n(\mathbb{R})$ at I may be identified with $\mathrm{M}_n(\mathbb{R})$.

Any tangent vector at I determines by right as well as left translation vector fields Λ_X as well as R_X on all of G. Vector fields are differential operators, acting on scalar functions on a Lie group. The derivative of F at g with respect to R_X is

$$R_X F(g) = \frac{d}{dt}\, F\big(g\, \exp(tX)\big)\Big|_{t=0} \;,$$

and that with respect to Λ_X is

$$\Lambda_X F(g) = \frac{d}{dt}\, F\big(\exp(tX)\, g\big)\Big|_{t=0} \;,$$

The vector fields Λ_X are those which are right-invariant.

If U and V are left-invariant so is $UV - VU$, so we get from the Poisson bracket in this way the **Lie bracket** operation on a Lie algebra.

If $G = \mathrm{GL}_n$, we can ask for an explicit formula for $[R_X, R_Y]$, where X and Y are given matrices. Let M be the tautologous map from $\mathrm{GL}_n(\mathbb{R})$ to $M_n(\mathbb{R})$, taking g to its matrix $M(g)$. Thus

$$\begin{aligned} R_X M(g) &= \left[\frac{d}{dt}\, M(g\, \exp(tX))\right]_{t=0} \\ &= \left[\frac{d}{dt}\, M(g)\exp(tX))\right]_{t=0} \\ &= M(g)X \;. \end{aligned}$$

Then

$$R_X R_Y M - R_Y R_X M = R_X MY - R_Y MX = MXY - MYX = M(XY - YX) \;.$$

Therefore:

5.3. Proposition. *In $M_n(\mathbb{R})$, the Lie algebra of $\mathrm{GL}_n(\mathbb{R})$, the Lie bracket is the commutator:*

$$[X, Y] = XY - YX \;.$$

There is one peculiarity about the right-invariant vector fields. If Λ_X and Λ_Y are two of these, then their Poisson bracket $[\Lambda_X, \Lambda_Y]$ is equal to $\Lambda_{[Y,X]}$ rather than $\Lambda_{[X,Y]}$. This is related to the following point—the left-invariant vector fields R_X derive from the right-regular action of the group G on smooth functions on G, according to which

$$[R_g F](x) = F(xg) \;.$$

The right action on the space defines a left action of G on functions, since

$$]R_g R_h F](x) = [R_h F](xg) = F(xgh) = [R_{gh} F](x), \quad R_{gh} F = R_g R_h F \;.$$

The left-regular action is defined by

$$L_g F(x) = F(g^{-1}x) \;.$$

The associated left action of the Lie algebra is $L_X = \Lambda_{-X}$. Thus the bracket equation for right-invariant vector fields is, as it should be,

$$[L_X, L_Y] = L_{[X,Y]} \;.$$

The adjoint action has a simple expression in the case of GL_n—it turns out to be the obvious matrix calculation:

5.4. Proposition. *For any g in $\mathrm{GL}_n(\mathbb{R})$ and X in $M_n(\mathfrak{g})$*

$$\mathrm{Ad}(g)X = gXg^{-1} .$$

Proof. Very similar to that of Proposition 5.3.

Thus we can see also why the earlier claim about ad is true: X in $\mathfrak{g}$ gets mapped to the endomorphism

$$\mathrm{ad}_X\colon\ Y \longmapsto [X, Y]$$

since

$$\begin{aligned}(I + tA + \cdots)X(I + tA + \cdots)^{-1} &= (I + tA + \cdots)X(I - tA + \cdots)\\ &= X + t(AX - XA) + \cdots .\end{aligned}$$

6. Normal coordinates

The exponential map $\mathfrak{g} \to G$ is a local diffeomorphism, hence a coordinate system on $\mathfrak{g}$ induces a coordinate system in an open neighbourhood U of the identity on G. If we choose a basis (X_i) of $\mathfrak{g}$, then the coordinate system maps mapping $\exp(t_1X_1 + \cdots + t_nX_n)$ to $(t_1, \ldots, t_n)$.

Let $\mathfrak{m}$ be the maximal ideal of functions in $C^\infty(U)$ vanishing at 1. Its power $\mathfrak{m}^m$ is the subset of smooth functions on U such that $\partial^k f/\partial x^k = 0$ for all $|k| < m$, or equivalently those which can be represented as a sum $t^k fk$ with $|k| = m$ and each f_k in $C^\infty(U)$. The vector field X_i and $\partial/\partial t_i$ agree at 1, hence

$$X_i = \frac{\partial}{\partial t_i} \sum_{j \neq i} f_j \frac{\partial}{\partial x_j}$$

with each f_j in $\mathfrak{m}$. More generally, as a consequence, if $|k| \leq |\ell|$ then

$$X^k t^\ell = \begin{cases} \ell! & \text{if } k = \ell \\ 0 & \text{otherwise.} \end{cases}$$

Therefore:

6.1. Proposition. *The pairing*

$$\langle X, f \rangle = XF(1)$$

identifies $U_{n-1}(\mathfrak{g})$ with the annihilator of $\mathfrak{m}^n$, and $U_n(\mathfrak{g})/U_{n-1}(\mathfrak{g})$ is isomorphic to $S^n(\mathfrak{g})$.

6.2. Corollary. *The operators R_X for X in $U(\mathfrak{g})$ are exactly the left-invariant differential operators on G.*

Proof. Because $U(\mathfrak{g})$ exhausts the invariant symbols.

7. The Lie algebra of SL(2)

Now let G be $\mathrm{SL}_2(\mathbb{R})$ and $\mathfrak{g} = \mathfrak{sl}_2$ its Lie algebra, the space of 2×2 matrices of trace 0. Let A be the subgroup of diagonal matrices, N that of unipotent upper triangular matrices, K the special orthogonal group SO_2. Several distinct bases for $\mathfrak{g}$ are useful in different situations. One very useful basis is this:

$$e_+ := \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad h := \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \kappa := \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},$$

which are, respectively, elements in $\mathfrak{n}$, $\mathfrak{a}$, $\mathfrak{k}$ (the Lie algebras of N, A, and K). That $\mathfrak{g} = \mathfrak{n} + \mathfrak{a} + \mathfrak{k}$ is the infinitesimal version of the Iwasawa decomposition $G = NAK$.

Another useful basis is made up of e_+, h and

$$e_- := \begin{bmatrix} 0 & 0 \\ -1 & 0 \end{bmatrix}.$$

I depart here from the standard convention, which replaces e_- by its negative. I follow [Tits:1966] in this. The advantage of the notation he and I use is symmetry: the map $e_\pm \mapsto e_\mp$, $h \mapsto -h$ is the canonical involution of $\mathfrak{sl}_2$, an automorphism of order two.

This new element e_- spans the Lie algebra $\overline{\mathfrak{n}}$ of the lower triangular unipotent elements of SL_2, and the decomposition $\mathfrak{g} = \mathfrak{n} + \mathfrak{a} + \overline{\mathfrak{n}}$ is an infinitesimal version of the Bruhat decomposition $G = P\overline{N} \cup Pw$ where

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

This w is the same as κ, but it is important to distinguish them—one lies in SL_2 and the other in $\mathfrak{sl}_2$.

We have the Lie algebra bracket formulae

$$[h, e_\pm] = \pm 2 e_\pm, \quad [e_+, e_-] = -h .$$

Thus $e_\pm$ are eigenvectors of ad_h, called **root vectors** of $\mathfrak{g}$ with respect to $\mathfrak{a}$.

There is a third useful basis. The element κ also spans the Lie algebra of the torus K of G, but it is a compact one, made up of rotation matrices. Inside $G(\mathbb{C})$ the two tori A and K are conjugate, and more precisely the Cayley transform

$$\boldsymbol{C} = \begin{bmatrix} 1 & -i \\ 1 & i \end{bmatrix}$$

conjugates $K(\mathbb{C})$ to $A(\mathbb{C})$. This can be seen geometrically—K is the isotropy subgroup of $\mathrm{SL}_2(\mathbb{R})$ acting by Möbius transformations on the upper half plane, while A preserves the origin. Consistently with this, the Cayley transform takes i to 0. In the complexified Lie algebra

$$\mathrm{Ad}(\boldsymbol{C}^{-1})\, h = i\kappa .$$

Since $K_{\mathbb{C}}$ is a torus, the complexified Lie algebra decomposes into root spaces, the images under $\boldsymbol{C}$ of those for $A_{\mathbb{C}}$. Explicitly they are spanned by

$$x_\pm = \begin{bmatrix} 1 & \mp i \\ \mp i & -1 \end{bmatrix} .$$

We have now

$$[\kappa, x_\pm] = \pm 2i x_\pm$$

and

$$\mathrm{Ad}(\boldsymbol{C}^{-1})\, e_\pm = (1/2)\, x_\mp .$$

7.1. Lemma. *The center of $\mathfrak{sl}_2$ is $\{0\}$, and any element of $\mathfrak{sl}_2$ can be expressed as a linear combination of commutators $[X, Y]$.*

Proof. It suffices to verify it for elements of a basis, which I leave as an exercise.

Yet more strongly:

7.2. Proposition. *Any homomorphism from $\mathfrak{sl}_2$ to another Lie algebra is either trivial or injective.*

Proof. If it is not injective, the image must be a Lie algebra of dimension 1 or 2. If it is onto an abelian algebra, then $[X, Y]$ is mapped to 0 for all X and Y, and therefore by the Lemma so is all of $\mathfrak{sl}_2$. If the image is not abelian, then the image must be the unique non-abelian algebra of two dimensions described in Proposition 9.2, and in that case too, since that algebra has the trivial algebra as quotient, all of $\mathfrak{sl}_2$ is mapped to 0.

8. Vector fields associated to G-actions

The action of a Lie group G on a manifold M determines also vector fields corresponding to vectors in its Lie algebra, the flows along the orbits of one-parameter subgroups $\exp(tX)$. Thus the vector at a point m corresponding to X in $\mathfrak{g}$ is the image of d/dt at $t = 0$ under the map from $\mathbb{R}$ to M taking t to $\exp(tX)\cdot m$. This is a special case of the general problem of calculating the image of d/dt under a map onto a manifold. If we are given local coordinates around m and the map takes

$$t \longmapsto \big(x_i(t)\big)$$

then

$$\frac{d}{dt} \longmapsto \sum \frac{dx_i}{dt}\,\frac{\partial}{\partial x_i}\,.$$

This might seem sometimes to involve a formidable calculation, and it is often useful to use Taylor series to simplify it. The point is that it is essentially a first order computation in which terms of second order can be neglected. Roughly speaking, up to first order $\exp(\varepsilon X) = I + \varepsilon X$, so the element X in $\mathfrak{g}$ determines at m the vector

$$\frac{(I + \varepsilon X)\cdot m - m}{\varepsilon}$$

where we may assume $\varepsilon^2 = 0$. The coefficients of the $\partial/\partial x_i$ are read off as the coordinates of ε in the expression for $(I + \varepsilon X)\cdot m$.

Let's look at the example of $\mathrm{SL}_2(\mathbb{R})$ acting on the upper half plane $\mathcal{H}$ by Möbius transformations

$$z \longmapsto \frac{az + b}{cz + d}\,.$$

In the formulas for vector fields associated to elements of the Lie algebra $\mathfrak{sl}_2$, some simplification is possible because the vector fields are real and the group acts holomorphically. The natural result of these calculations will a complex-valued function. A complex analytic function $f(z) = p + iq$ is to be interpreted as as a real vector field according to the formula

$$\begin{aligned} p\,\frac{\partial}{\partial x} + q\,\frac{\partial}{\partial y} &= (p + iq)\left(\frac{1}{2}\right)\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right) \\ &\qquad + (p - iq)\left(\frac{1}{2}\right)\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right) \\ &= f(z)\,\frac{\partial}{\partial z} + \overline{f(z)}\,\frac{\partial}{\partial \overline{z}}\,. \end{aligned}$$

8.1. Proposition. *We have*

$$\begin{aligned}\Lambda_{e_+} &= \frac{\partial}{\partial x}\\ \Lambda_h &= 2x\,\frac{\partial}{\partial x} + 2y\,\frac{\partial}{\partial y}\\ \Lambda_{e_-} &= (x^2 - y^2)\frac{\partial}{\partial x} + 2xy\,\frac{\partial}{\partial y}\,.\end{aligned}$$

Proof. • The simplest is e_+. Here

$$I + \varepsilon e_+ = \begin{bmatrix} 1 & \varepsilon \\ 0 & 1 \end{bmatrix}$$

and this takes

$$z \longmapsto \frac{z+\varepsilon}{1} = z + \varepsilon, \quad (x, y) \longmapsto (x + \varepsilon, y)$$

Therefore

$$e_+ \rightsquigarrow \partial/\partial x\,.$$

• Now for h. Here

$$I + \varepsilon h = \begin{bmatrix} 1+\varepsilon & 0 \\ 0 & 1-\varepsilon \end{bmatrix}$$

and this takes

$$\begin{aligned} z \longmapsto & \frac{(1+\varepsilon)z}{(1-\varepsilon)}\\ &= z(1+\varepsilon)(1+\varepsilon+\varepsilon^2+\cdots)\\ &= z(1+2\varepsilon) = z + 2\varepsilon z\\ (x,y) \longmapsto & (x + 2\varepsilon x, y + 2\varepsilon y)\end{aligned}$$

so

$$h \rightsquigarrow 2x\,\frac{\partial}{\partial x} + 2y\,\frac{\partial}{\partial y}$$

• Finally

$$I + \varepsilon e_- = \begin{bmatrix} 1 & 0 \\ \varepsilon & -1 \end{bmatrix}$$

$$z \longmapsto \frac{z}{\varepsilon - z + 1} = z + \varepsilon z^2$$

$$e_- \rightsquigarrow (x^2 - y^2)\frac{\partial}{\partial x} + 2xy\,\frac{\partial}{\partial y}\,. \quad \blacksquare$$

Part III. Lie algebras on their own

9. Abstract Lie algebras

For X, Y, Z in the Lie algebra of a Lie group G we have an identity of differential operators

$$\begin{aligned}[X,[Y,Z]] &= X(YZ-ZY)-(YZ-ZY)X \\ &= XYZ-XZY-YZX+ZYX \\ [Y,[Z,X]] &= Y(ZX-XZ)-(ZX-XZ)Y \\ &= YZX-YXZ-ZXY+XZY \\ [Z,[X,Y]] &= Z(XY-YX)-(XY-YX)Z \\ &= ZXY-ZYX-XYZ+YXZ\end{aligned}$$

and summing we get

$$[X,[Y,Z]]+[Y,[Z,X]]+[Z,[X,Y]]=0\,.$$

This last equation is called the **Jacobi identity**. It is easy to recall to mind, since it is the sum of three terms obtained by applying a cyclic permutation of X, Y, Z to $[X,[Y,Z]]$.

There are several ways to interpret it, aside from the formal calculation. Even the formal computation can be analyzed intelligently. It can be seen immediately that the expression

$$[X,[Y,Z]]+[Y,[Z,X]]+[Z,[X,Y]]$$

changes sign if any two of the variable are swapped. Swaps generate all of $\mathfrak{S}_3$, and one can conclude that the sum must be a scalar multiple of the alternating sum

$$A(X,Y,Z)=\sum_{\sigma\in\mathfrak{S}_3}\operatorname{sgn}(\sigma)\,\sigma(X,Y,Z)=XYZ-XZY+YZX-YXZ+ZXY-ZYX\,.$$

But a simple look at the coefficients of $[X,[Y,Z]]$ in the Jacobi sum makes it clear that its twelve terms amount to $A(X,Y,Z)-A(X,Y,Z)=0$.

> **Definition.** *Suppose $\mathfrak{g}$ to be an arbitrary finite-dimensional vector space over an arbitrary field F, and suppose that it is given a bilinear map $(x,y)\mapsto[x,y]$ from $\mathfrak{g}\times\mathfrak{g}$ to $\mathfrak{g}$ itself. This is said to define $\mathfrak{g}$ as a Lie algebra if this bilinear map is anti-symmetric and satisfies Jacobi's identity.*

It is not immediately apparent that this definition captures completely what one wishes. One justification is the 'Third Theorem of Lie' (apparently first proved by Eli Cartan): *If $F=\mathbb{R}$ or $\mathbb{C}$ then every Lie algebra is the Lie algebra of an analytic group over F.* A clear exposition can be found in LG §5.8 of [Serre:1965]. A related result, with a more direct argument, is Serre's construction of 'group chunks' in LG §5.4.

Here is another way to understand the Jacobi identity. Let $\mathfrak{g}$ be an arbitrary Lie algebra defined over $\mathbb{R}$, and let $A=\operatorname{Aut}(\mathfrak{g})$ be the group of automorphisms of $\mathfrak{g}$. This is an algebraic group defined over $\mathbb{R}$. It is by definition embedded in $\mathrm{GL}_{\mathbb{R}}(\mathfrak{g})$, and this representation gives rise to its differential, an embedding of its Lie algebra $\mathfrak{a}$ in the space of linear endomorphisms of $\mathfrak{g}$. For a in $\mathfrak{a}$, t in $\mathbb{R}$ let α_t be the automorphism $\exp(ta)$ of $\mathfrak{g}$. Then

$$d\alpha(x)=\frac{d}{dt}\,\alpha_t(x)\Big|_{t=0}$$

Since each α_t is an automorphism of $\mathfrak{g}$, the product rule for derivatives implies that $d\alpha$ is a derivation:

$$d\alpha([x,y])=[d\alpha(x),y]+[x,d\alpha(y)]\,.$$

9.1. Proposition. *The space of derivations of $\mathfrak{g}$ is the Lie algebra of* $\operatorname{Aut}(\mathfrak{g})$.

Proof. The group of derivations of $\mathfrak{g}$ is an algebraic subgroup of $\mathrm{GL}(\mathfrak{g})$. We know from a remark in Part I that an element X in $\mathfrak{gl}(\mathfrak{g})$ is in its Lie algebra if and only if $I + \varepsilon X$ lies in the group, where $\varepsilon^2 = 0$. This happens if and only if X is a derivation.

One can also verify directly that if X is a derivation then $\exp(X)$ is an automorphism.

Among the automorphisms of the Lie algebra of a group G are its inner automorphisms, given by the adjoint action. For $G = \mathrm{GL}_n(\mathbb{R})$ this is matrix conjugation:

$$\mathrm{Ad}(g)X = gXg^{-1}\,.$$

The map $\mathrm{Ad}(g)$ is an automorphism of the Lie algebra structure. The differential of Ad is

$$\mathrm{ad}\colon\ X \longmapsto \text{ the map taking } Y \text{ to } [X, Y]$$

since

$$\begin{aligned}(I + tX + \cdots)Y(I + tX + \cdots)^{-1} &= (I + tX + \cdots)Y(I - tX + \cdots) \\ &= Y + t(XY - YX) + \cdots\end{aligned}$$

Because each $\mathrm{Ad}(g)$ is an automorphism of $\mathfrak{g}$, each ad_X is a derivation:

$$\mathrm{ad}_X[Y, Z] = [\mathrm{ad}_X Y, Z] + [Y, \mathrm{ad}_X Z]$$

which is precisely the Jacobi identity.

> *From now on, unless I specify otherwise, Lie algebras will be assumed to be defined over an arbitrary coefficient field F of characteristic 0.*

One Lie algebra is the matrix algebra $\mathrm{End}_F(V)$ with $[X, Y] = XY - YX$. To distinguish this as a Lie algebra from this as a ring, I'll write the Lie algebra as $\mathfrak{gl}(V)$. If $V = F^n$ the endomorphism ring will be $M_n(F)$ and the Lie algebra will be $\mathfrak{gl}_n(F)$.

One mildly useful exercise is to classify all Lie algebras of dimension 1 or 2. For dimension 1, the only possibility is F with the trivial bracket.

9.2. Proposition. *There are, up to isomorphism, exactly two Lie algebras of dimension 2, the trivial one where all brackets are 0, and that with basis X, Y where $[X, Y] = Y$.*

The first is the Lie algebra of the additive group F^2, or of the group of diagonal matrices in $\mathrm{GL}_2(F)$, while the second is that of the group of matrices

$$\begin{bmatrix} a & x \\ 0 & 1 \end{bmatrix}.$$

This second algebra $\mathfrak{g}$ fits into an exact sequence of Lie algebras

$$0 \longrightarrow F \longrightarrow \mathfrak{g} \longrightarrow F \longrightarrow 0\,.$$

Proof. Let X, Y be a basis, and suppose $[X, Y] = aX + bY$. Suppose one of the coefficients, say b, is not 0. Then

$$[cX, aX + bY] = cb[X, Y] = cb(aX + bY)$$

which is equal to $aX + bY$ if $c = b^{-1}$. So if $Y_* = aX + bY$ and $X_* = cX$ we then have

$$[X_*, Y_*] = Y_*\,.$$

10. Representations

A **Lie algebra homomorphism** is a linear map $\varphi\colon \mathfrak{g} \to \mathfrak{h}$ such that

$$[\varphi(X), \varphi(Y)] = \varphi([X, Y]) .$$

If $g = \exp(tX)$, $h = \exp(tY)$ then

$$ghg^{-1}h^{-1} = I + t^2[X, Y] + \text{ higher order terms.}$$

and therefore a homomorphism φ of Lie groups induces a homomorphism $d\varphi$ of Lie algebras, its differential.

An **ideal** in $\mathfrak{g}$ is a linear subspace $\mathfrak{h}$ such that $[X, Y]$ lies in $\mathfrak{h}$ for all X in $\mathfrak{g}$ and Y in $\mathfrak{h}$. The Lie bracket on $\mathfrak{g}$ thus induces the structure of a Lie algebra on the quotient space $\mathfrak{g}/\mathfrak{h}$. Conversely, the kernel of any homomorphism of Lie algebras is an ideal.

A **representation** of a Lie algebra on a vector space V is a Lie algebra homomorphism φ into $\mathfrak{gl}(V)$. This is a linear map into $\mathrm{End}(V)$(such that $\varphi(X)\varphi(Y) - \varphi(Y)\varphi(X) = \varphi([X, Y])$. In this way, V becomes a module over $\mathfrak{g}$.

If U and V are $\mathfrak{g}$-modules then so is their tensor product:

$$X\,(u \otimes v) = Xu \otimes v + u \otimes Xv$$

and the space $\mathrm{Hom}_F(U, V)$:

$$[Xf](u) = X\bigl(f(u)\bigr) - f(Xv) .$$

In particular the linear dual $\widehat{U}$ becomes the **dual** representation. So does the tensor algebra $\bigotimes^\bullet V \otimes \bigotimes^\bullet \widehat{V}$, and the exterior algebra $\bigwedge^\bullet V$ embedded in it.

These definitions are compatible with, and indeed motivated by, the corresponding definitions of representations of a group, since

$$\begin{aligned}(I + tX + \cdots)(u \otimes v) &= (I + tX + \cdots)u \otimes (I + tX + \cdots)v \\ &= u \otimes v + tXu \otimes v + u \otimes tXv + \cdots .\end{aligned}$$

Every Lie algebra has at least one representation, the adjoint representation by the linear maps ad_X (in which it is Jacobi's identity that verifies it is a representation).

A one-dimensional representation of a Lie algebra is called a **character**. If φ is a character of $\mathfrak{g}$ then $\varphi([x, y]) = 0$. Define $D\mathfrak{g}$ to be the span of all comutators $[x, y]$.

10.1. Proposition. *The subspace $\mathcal{D}\mathfrak{g}$ is an ideal of $\mathfrak{g}$.*

Proof. Immedite from the Jacobi identity.

Any charcacter φ vanishes on $\mathcal{D}\mathfrak{g}$. In fact:

10.2. Proposition. *The characters of a Lie algebra may be identified with the linear dual of $\mathfrak{g}/D\mathfrak{g}$.*

In particular, if $\mathfrak{g} = D\mathfrak{g}$ then the only character of $\mathfrak{g}$ is the trivial one. This is true, for example, of $\mathfrak{sl}_2$, according to Lemma 7.1.

11. Nilpotent Lie algebras

The simplest Lie algebras are the abelian ones, for which $[X, Y]$ identically vanishes. Next simplest are the **nilpotent** ones.

The **upper central series** of the Lie algebra $\mathfrak{g}$ is the succession of ideals $\mathcal{C}^i(\mathfrak{g})$ in $\mathfrak{g}$ defined recursively:

$$\begin{aligned}\mathcal{C}^0 &= \mathfrak{g} \\ \mathcal{C}^{n+1} &= [\mathfrak{g}, \mathcal{C}^n]\,.\end{aligned}$$

This is a weakly decreasing sequence. Each $\mathcal{C}^n$ is an ideal of $\mathfrak{g}$, and each quotient $\mathcal{C}^n/\mathcal{C}^{n+1}$ is in the centre of $\mathfrak{g}/\mathcal{C}^{n+1}$.

If $\mathfrak{g}$ is any Lie algebra, let $\mathcal{C}(\mathfrak{g})$ be its center. The **lower central series** of $\mathfrak{g}$ is the succession of ideals $\mathcal{C}_i(\mathfrak{g})$ in $\mathfrak{g}$ defined recursively:

$$\begin{aligned}\mathcal{C}_0 &= \{0\} \\ \mathcal{C}_{n+1} &= \text{ the inverse image in } \mathfrak{g} \text{ of } \mathcal{C}(\mathfrak{g}/\mathcal{C}_n)\,.\end{aligned}$$

Thus in particular $\mathcal{C}_1 = \mathcal{C}(\mathfrak{g})$. This is a weakly increasing sequence.

It can happen that both series are trivial, with $\mathcal{C}^0 = \mathfrak{g}$ and $C_0 = 0$. This happens for $\mathfrak{sl}_2$. The following result is about the other extreme:

11.1. Proposition. *Suppose $\mathfrak{g}$ to be a Lie algebra. The following are equivalent:*

(a) some $\mathcal{C}^n = \{0\}$;
(b) some $\mathcal{C}_n = \mathfrak{g}$;
(c) the Lie algebra $\mathfrak{g}$ possesses a strictly increasing filtration by ideals $\mathfrak{g}_n$ such that $[\mathfrak{g}, \mathfrak{g}_{n+1}] \subset \mathfrak{g}_n$;
(d) the Lie algebra $\mathfrak{g}$ possesses a strictly increasing filtration by ideals $\mathfrak{g}_n$ such that $[\mathfrak{g}, \mathfrak{g}_{n+1}] \subset \mathfrak{g}_n$ with each $\mathfrak{g}_{n+1}/\mathfrak{g}_n$ of dimension one.

I leave this as an exercise.

The Lie algebra $\mathfrak{g}$ is defined to be nilpotent if one of these conditions holds. Any quotient or subalgebra of a nilpotent Lie algebra is nilpotent.

One nilpotent Lie algebra is $\mathfrak{n}_n$, the subalgebra of $\mathfrak{gl}_n$ of matrices whose entries are 0 on and below the diagonal. It has as basis the matrices $e_{i,j}$ with $i < j$, with a single entry 1 at position (i, j). We have

$$e_{i,j}e_{k,\ell} = \begin{cases} e_{i,\ell} & \text{if } j = k \\ -e_{k,j} & \text{if } \ell = i \\ 0 & \text{otherwise.}\end{cases}$$

Hence $\mathcal{C}^d$ is the subspace spanned by the $e_{i,j}$ with $j - i > d$. For example, when $n = 3$ we have the basis

$$e_{1,2} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad e_{1,3} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad e_{2,3} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

with

$$[e_{1,2}, e_{2,3}] = e_{1,3}, \quad [e_{1,2}, e_{1,3}] = 0, \quad [e_{2,3}, e_{1,3}] = 0\,.$$

The next theorem (which was apparently first proven by Killing, in spite of the name by which it is frequently called) asserts that the $\mathfrak{n}_n$ are in a very strong sense universal. It is an immediate consequence of the definition that a Lie algebra $\mathfrak{g}$ is nilpotent if and only if its image in $\mathfrak{gl}(\mathfrak{g})$ with respect to the adjoint map is contained in a conjugate of the nilpotent upper triangular matrices. A much weaker criterion is in fact valid.

An element of a Lie algebra is said to be nilpotent if the endomorphism $\mathrm{ad}_{\mathfrak{n}}(X)$ is nilpotent or, equivalently, conjugate to an upper triangular nilpotent matrix. Every element X of a nilpotent Lie algebra is nilpotent in this sense. Conversely:

11.2. Theorem. (Engel's Criterion) *A Lie algebra is nilpotent if and only if every one of its elements is nilpotent.*

In other words, it is globally nilpotent if and only if it is locally nilpotent. This is immediate from:

11.3. Proposition. *Suppose $\mathfrak{g}$ to be a Lie subalgebra of $\mathfrak{gl}_n$, and suppose that every element of $\mathfrak{g}$ is nilpotent in M_n. Then $\mathfrak{g}$ is contained in a conjugate of $\mathfrak{n}_n$.*

Here is a modest reformulation, apparently first found in [Merrin:1994], that makes possible a constructive proof on which one can base a practical algorithm.

11.4. Lemma. *Suppose $\mathfrak{g}$ to be a Lie subalgebra of $\mathfrak{gl}_n$. Either it is contained in a conjugate of $\mathfrak{n}_n$, or there exists X in $\mathfrak{g}$ that is non-nilpotent in M_n.*

By an easy induction argument, Lemma 11.4 follows from:

11.5. Lemma. *Suppose $\mathfrak{g}$ to be a Lie subalgebra of $\mathfrak{gl}_n$. Either there exists $v \neq 0$ in V with $Xv = 0$ for all X in $\mathfrak{g}$, or there exists X in $\mathfrak{g}$ which is not nilpotent in M_n.*

The proof will in fact be constructive, and it would be easy to design a practical algorithm based on it. In this algorithm, we would start with a basis of matrices X_i in $\mathfrak{g}$, and the output would be either (1) vector $v \neq 0$ with $X_i v = 0$ for all i, or (2) some linear combination of the X_i that is certifiably not nilpotent.

Proof. I begin the proof of the Proposition by looking in detail at the case in which $\mathfrak{g}$ has dimension one—i.e. at a single linear transformation of a finite-dimensional vector space V.

11.6. Lemma. (Fitting's Lemma) *If T is a linear transformation of the finite-dimensional vector space V, then V has a unique T-stable decomposition*

$$V = V(T^n) \oplus T^n V$$

for $n \geq \dim V$. The operator T is nilpotent on $V(T^n)$ and invertible on $T^n V$.

Here $V(A)$ means the kernel of A.

Proof. I recall first of all a variant of Gauss-Jordan elimination. Given any finite set of vectors in a vector space V, let M be a matrix with these as columns. Applying elementary column operations will reduce M to a matrix in echelon form whose columns are a basis for the subspace of V they span. This reduced matrix is in fact unique, and there is a bijection between subspaces and matrices in column echelon form. If M is the matrix of a linear transformation, this allows us to compute its image. Working with row operations instead, one can compute its kernel.

Now suppose T to be a linear transformation of V and compute successively the images of the powers T^n. If $T^{n+1}V = T^n V$ then $T^m V = T^n V$ for all $m \geq n$. Therefore the weakly decreasing sequence $T^n V$ is eventually stable, and for some $n \leq \dim V$. Similarly, the kernels $V(T^n)$ of T^n in V are a weakly increasing sequence, also eventually stable. I write $(\!(T)\!)V$ for $T^n V$ and $V(\!(T)\!)$ for $\ker(T^n)$ with $n \gg 0$. If $T \cdot T^n V = T^n V$, the transformation T is invertible on $T^n V$, and hence $(\!(T)\!)V$ is complementary to $V(\!(T)\!)$. The operator T is nilpotent if and only if $V(\!(T)\!) = V$ and $(\!(T)\!)V = 0$.

Fitting's Lemma asserts that every v can be expressed as $v_0 + v_1$ with $v_0 \in V(T^n)$, $v_1 \in T^n V$. We can make this decomposition more explicit. Suppose the characteristic polynomial of T factors as $Q(x)x^n$ with $Q(x)$ not divisible by x. Then $Q(T) = 0$ on $T^n V$. The Euclidean algorithm gives us $A(x)$, $B(x)$ such that

$$1 = A(x)x^n + B(x)Q(x)\,.$$

But then $A(T)T^n$ is the projection of V onto $T^n V$, and $B(T)Q(T)$ is the projection onto $V(T^n)$.

The subspaces $V(\!(T)\!)$ and $(\!(T)\!)V$ are called the **Fitting components** of T.

Now back to the proof of Lemma 11.5. Suppose $\mathfrak{g}$ to be a Lie subalgebra of $\mathfrak{gl}(V)$.

First an elementary observation: *If X is a nilpotent operator on V, then so is ad_X nilpotent acting on* $\mathrm{End}(V)$. To see this, suppose X in $\mathfrak{g}$ is nilpotent, say $X^m = 0$. Therefore the operators

$$\Lambda_X\colon A \longmapsto XA$$
$$R_X\colon A \longmapsto AX$$

acting on matrices A satisfy say $\Lambda_X^n = R_X^n = 0$.

But these commute with each other. Since $\mathrm{ad}_X = \Lambda_X - R_X$, we have therefore

$$\mathrm{ad}_X^{2n} = (\Lambda_X - R_X)^{2n} = \sum \binom{2n}{k} (-1)^k L_X^k R_X^{2n-k} = 0\,,$$

since either $k\mathfrak{e}n$ or $2n - k\mathfrak{e}n$.

We are given $\mathfrak{g} \subset \mathfrak{gl}_n$, and want to (a) show $\mathfrak{g}^n V = 0$ for $n \gg 0$ or (b) find X in $\mathfrak{g}$ which is not nilpotent. Choose $X = X_1$ in $\mathfrak{g}$ (say, the first element of a given basis). Either X is nilpotent, or it is not. If not, we are through. If it is then (a) the subspace U_1 of vectors annihilated by X is not 0, and (b) by the remark just above, ad_X is nilpotent. It takes the one-dimensional Lie algebra $\mathfrak{h} = \mathfrak{h}_1 \subseteq \mathfrak{g}$ spanned by X into itself, so also acts linearly on the quotient $\mathfrak{g}/\mathfrak{h}$, and we can find $X_2 \notin \mathfrak{h}$ in $\mathfrak{g}$ such that $\mathrm{ad}_X(X_2)$ lies in $\mathfrak{h}$. Thus the space $\mathfrak{h}_2$ spanned by X_1 and X_2 is a Lie subalgebra of $\mathfrak{g}$.

We can continue. At each stage we have a Lie subalgebra $\mathfrak{h}_i$ of $\mathfrak{g}$ together with a non-zero subspace U_i of V annihilated by it, and we also have a subalgebra $\mathfrak{h}_{i+1}$ spanned by X_{i+1} and $\mathfrak{h}_i$ with $[X_{i+1}, \mathfrak{h}_i] \subseteq \mathfrak{h}_i$. The space U_i is stable under $\mathfrak{h}_{i+1}$ since for X in $\mathfrak{h}_i$

$$XX_{i+1}u = X_{i+1}Xu + [X, X_{i+1}]u = 0\,.$$

We now set U_{i+1} equal to the space of all vectors annihilated by X_{i+1}. Either it is non-trivial and we continue on, or X_{i+1} acting on U_i is non-nilpotent. There are in the end one of two outcomes: either we find some element of $\mathfrak{g}$ that is not nilpotent on V, or $\mathfrak{h}_n = \mathfrak{g}$ and we have found a non-zero subspace of V annihilated by $\mathfrak{g}$.

12. Representations of a nilpotent Lie algebra

If $\mathfrak{g}$ is an abelian Lie algebra and F is algebraically closed, any finite-dimensional vector space on which $\mathfrak{g}$ acts is the sum of certain primary components associated to maximal ideals of the polynomial algebra $F[\mathfrak{g}]$. The component associated to the maximal ideal $\mathfrak{m}$ is the subspace of $F[\mathfrak{g}]$ annihilated by some power of $\mathfrak{m}$. This can be proven by a simple induction argument on the dimension of $\mathfrak{g}$, starting with Fitting's Lemma applied to operators $T - \lambda I$ for eigenvalues λ of T. This works because if $(X - \lambda)^m v = 0$ and Y commutes with X then

$$(X - \lambda)^m Yv = Y(X - \lambda)^m v = 0\,.$$

In this section I'll show that something analogous is valid for any nilpotent Lie algebra, although the argument is a bit less simple.

Suppose $\mathfrak{g}$ to be a nilpotent Lie algebra and V to be a $\mathfrak{g}$-module. This means that we are given a map φ from $\mathfrak{g}$ to the Lie algebra $\mathfrak{gl}(V)$, but for the moment I'll assume $\mathfrak{g}$ to be embedded into $\mathrm{End}(V)$. With this assumption, I can ignore φ in notation.

For every n there exists a linear map

$$\Phi_n\colon \bigotimes^n \mathfrak{g} \otimes V \longrightarrow V, \quad X_1 \otimes \ldots \otimes X_n \otimes v \longmapsto X_1 \ldots X_n v\,.$$

Define $\mathfrak{g}^n V$ to be its image. There also exists a canonical map

$$V \longrightarrow \mathrm{Hom}(\otimes^n \mathfrak{g}, V)$$

that takes v to the map taking T in $\bigotimes^n \mathfrak{g}$ to $\Phi_n(T \otimes v)$. Let $V(\mathfrak{g}^n)$ be its kernel. The first sequence is weakly decreasing, the second weakly increasing, and they both eventually stabilize. Let $(\!(\mathfrak{g})\!)V$ be the intersection of all the $\mathfrak{g}^n V$, and $V(\!(\mathfrak{g})\!)$ be the union of the $V(\mathfrak{g}^n)$.

12.1. Lemma. *For* $n \mathfrak{e} \dim V$

$$V(\!(\mathfrak{g})\!) = \bigcap\nolimits_{X \in \mathfrak{g}} V(\!(X)\!) \,.$$

Proof. I'll show in a moment that each space $V(\!(X)\!)$ is stable under $\mathfrak{g}$. Their intersection is then also $\mathfrak{g}$-stable. It certainly contains $V(\!(\mathfrak{g})\!)$, and is equal to it by Engel's Criterion.

12.2. Lemma. *Suppose* X, Y *to be two endomorphisms of the finite-dimensional vector space* V*. If* $\mathrm{ad}_X^n Y = 0$ *for some* n*, then the Fitting components of* X *are stable under* Y*.*

Proof. Suppose $X^\ell v = 0$. We want to know that for $m \gg 0$ we also have $X^m Y v = 0$. We start out:

$$\begin{aligned}
XY &= YX + \mathrm{ad}_X Y \\
X^2 Y &= X(XY) \\
&= X(YX + \mathrm{ad}_X Y) \\
&= (XY)X + X(\mathrm{ad}_X Y) \\
&= (YX + \mathrm{ad}_X Y)X + (\mathrm{ad}_X Y)X + \mathrm{ad}_X^2 Y \\
&= YX^2 + 2(\mathrm{ad}_X Y)X + \mathrm{ad}_X^2 Y \,,
\end{aligned}$$

which leads us to try proving by induction that

$$\begin{aligned}
X^m Y &= YX^m + m(\mathrm{ad}_X Y)X^{m-1} + \cdots + m(\mathrm{ad}_X^{m-1} Y)X + \mathrm{ad}_X^m Y \\
&= \sum_{k=0}^{m} \binom{m}{k} (\mathrm{ad}_X^k Y) X^{m-k} \,.
\end{aligned}$$

We can make the inductive transition:

$$\begin{aligned}
X^m Y &= \sum_{k=0}^{m} \binom{m}{k} (\mathrm{ad}_X^{m-k} Y) X^k \\
X^{m+1} Y &= \sum_{k=0}^{m} \binom{m}{k} X(\mathrm{ad}_X^{m-k} Y) X^k \\
X^{m+1} Y &= \sum_{k=0}^{m} \binom{m}{k} \big((\mathrm{ad}_X^{m-k} Y) X^{k+1} + (\mathrm{ad}_X^{m+1-k} Y) X^k\big) \\
&= \sum_{k=0}^{m+1} \binom{m+1}{k} (\mathrm{ad}_X^{m+1-k} Y) X^k \,.
\end{aligned}$$

Now if $X^\ell v = 0$ and $\mathrm{ad}_X^n Y = 0$ then

$$X^{\ell+n} Y v = \sum_{0}^{\ell+n} \binom{\ell+n}{k} (\mathrm{ad}_X^k Y) X^{\ell+n-k} v = 0$$

since in the sum either $k \mathfrak{e} n$ or $\ell + n - k > \ell$. A similar argument will show that $(\!(X)\!)V$ is $\mathfrak{g}$-stable, once we prove

$$YX^m = X^m Y - mX^{m-1}(\mathrm{ad}_X Y) + \cdots \pm mX(\mathrm{ad}_X^{m-1} Y) \mp \mathrm{ad}_X^m Y \,.$$

Here is a nilpotent extension of Fitting's Lemma:

12.3. Proposition. *Any $\mathfrak{g}$-module V is the direct sum of $V(\!(\mathfrak{g})\!)$ and $(\!(\mathfrak{g})\!)V$.*

Proof. We proceed by induction on $\dim V$. If every X in $\mathfrak{g}$ acts nilpotently then by Engel's Criterion $V = V(\!(\mathfrak{g})\!)$ and $(\!(\mathfrak{g})\!)V = 0$. Otherwise, there exists X in $\mathfrak{g}$ with $(\!(X)\!)V \neq 0$. If $(\!(X)\!)V = V$ then necessarily $V(\!(\mathfrak{g})\!) = 0$ and we are through. Otherwise we have the decomposition

$$V = V(\!(X)\!)V \oplus (\!(X)\!)V$$

into two proper subspaces. By Lemma 12.2 each of these is stable under $\mathfrak{g}$, and the second is contained in $(\!(\mathfrak{g})\!)V$. Let the first be U. We may apply the induction hypothesis to decompose it into a sum $U(\!(\mathfrak{g})\!) \oplus (\!(\mathfrak{g})\!)U$. This gives us in turn the decomposition

$$V = U(\!(\mathfrak{g})\!) \oplus (\!(\mathfrak{g})\!)U \oplus (\!(X)\!)V$$

But the first is the same as $V(\!(\mathfrak{g})\!)$ and the second is contained in $(\!(\mathfrak{g})\!)V$. On the other hand, if $U = V(\!(\mathfrak{g})\!)$ and $\mathfrak{g}^n U = 0$ then

$$\mathfrak{g}^n V \subseteq (\!(\mathfrak{g})\!)U \oplus (\!(X)\!)V$$

so in fact

$$(\!(\mathfrak{g})\!)V = (\!(\mathfrak{g})\!)U \oplus (\!(X)\!)V\,.$$

This might prove useful sometime:

12.4. Corollary. *The functors*

$$V \rightsquigarrow V(\!(\mathfrak{g})\!)$$
$$V \rightsquigarrow (\!(\mathfrak{g})\!)V$$

are both exact.

Proof. From the Proposition, since the first is clearly left exact and the second right exact.

I must now take into account the Lie homomorphism φ from $\mathfrak{g}$ to $\mathrm{End}(V)$. If λ is a character of $\mathfrak{g}$, we can use it to define a 'twisted' representation of $\mathfrak{g}$ on V according to the formula

$$\varphi - \lambda I\colon\ X \longmapsto \varphi(X) - \lambda(X)I\,.$$

This is a representation because

$$[\varphi(X) - \lambda(X)I, \varphi(Y) - \lambda(Y)I] = [X, Y] = [X, Y] - \lambda([X, Y])$$

since $\lambda([X, Y]) = 0$ by definition of character. In this new representation, the vectors annihilated by $\mathfrak{g}^n$ are those annihilated by all products

$$\prod_1^n \big(\varphi(X_i) - \lambda(X_i)I\big)$$

in the original.

Assume for the rest of this section that F is algebraically closed.

If we apply the previous Proposition successively to the $\mathfrak{g}$-module $\mathfrak{g}^n V$, we get a decomposition of V into a direct sum of components $V(\!(\mathfrak{g}, \lambda)\!)$ annihilated by such products, given this:

12.5. Lemma. *If V is a finite-dimensional module over the nilpotent Lie algebra $\mathfrak{g}$, there exists at least one eigenvector for it.*

Proof. I am going to prove something a bit more general that will be useful later on.

12.6. Lemma. *Suppose $\mathfrak{g}$ to be a Lie algebra possessing an increasing sequence of Lie subalgebras*

$$\mathfrak{g}_0 = 0 \subset \mathfrak{g}_1 \subset \ldots \subset \mathfrak{g}_n = \mathfrak{g}$$

with the property that $\mathcal{D}\mathfrak{g}_{n+1} \subseteq \mathfrak{g}_n$. Any finite-dimensional module over $\mathfrak{g}$ possesses an eigenvector for it.

The proof of this is by induction on the dimension m of $\mathfrak{g}$. The case that $m = 1$ is trivial, so suppose $m > 1$. The Lemma will follow by induction from this new Lemma:

12.7. Lemma. *Suppose $\mathfrak{h} \subset \mathfrak{g}$ of codimension one, with $\mathfrak{h} \subseteq \mathcal{D}\mathfrak{g}$. Any eigenspace for $\mathfrak{h}$ is stable under $\mathfrak{g}$.*

An eigenvector of all of $\mathfrak{g}$ is one for $\mathfrak{h}$, so this is a very optimistic hope about a possible converse.

Proof. Let $X \neq 0$ be an element of $\mathfrak{g}$ not in $\mathfrak{h}$. Let $v \neq 0$ be an eigenvector of $\mathfrak{h}$, so that

$$H \cdot v = \lambda(H) v$$

for all H in $\mathfrak{h}$, where λ is a character of $\mathfrak{h}$. We would like to know that all $v_k = X^k v$ are also eigenvectors for $\mathfrak{h}$. At any rate, let V_m be the space spanned by the vectors v_k for $k \le m$. Since $XV_m \subseteq V_{m+1}$, the union of these is at least a finite-dimensional space V_* taken into itself by X. I shall show that each space V_m is stable under $\mathfrak{h}$, and that the representation of $\mathfrak{h}$ on V_m/V_{m-1} is by λ. I do this by induction. It is true for $m = 0$ by assumption.

For H in $\mathfrak{h}$ and $m\mathfrak{e}1$ we have

$$\begin{aligned} \big(H - \lambda(H)\big) v_m &= \big(H - \lambda(H)\big) X v_{m-1} \\ &= HXv_{m-1} - \lambda(H) X v_{m-1} \\ &= XHv_{m-1} + [H, X] v_{m-1} - X\lambda(H) v_{m-1} \\ &= X\big(H - \lambda(H)\big) v_{m-1} + [H, X] v_{m-1} \,. \end{aligned}$$

But by induction $\big(H - \lambda(H)I\big) v_{m-1}$ lies in V_{m-2}, so the first term lies in V_{m-1}. But, again by induction, since $[H, X]$ lies in $\mathfrak{h}$ the third term also lies in V_{m-1}.

We now therefore know that V_* is stable with respect to both X and $\mathfrak{h}$, hence all of $\mathfrak{g}$, and we also know that $H - \lambda(H)$ is nilpotent on it for every H in $\mathfrak{h}$.

I now claim that in fact every vector in V_* is an eigenvector for $\mathfrak{h}$. First of all, the trace of any H in $\mathfrak{h}$ is equal to $d_* \cdot \lambda(H)$, where d_* is the dimension of V_*. On the other hand, the trace of $[H, X] = HX - XH$ is 0. Therefore $\lambda([H, X]) = 0$ for every H in $\mathfrak{h}$. But if we assume that $Hv_{m-1} = \lambda v_{m-1}$ for all H in $\mathfrak{g}$ then

$$Hv_m = HX \cdot v_{m-1} = XH \cdot v_{m-1} + [H, X] v_{m-1} = \lambda v_m + \lambda([H, X]) v_{m-1} = \lambda v_m \,.$$

Any eigenvector for X in V_* will be an eigenvector for all of $\mathfrak{g}$.

In summary:

12.8. Theorem. *If V is a module over the nilpotent Lie algebra $\mathfrak{g}$, there exists a finite set Λ of characters λ of $\mathfrak{g}$ and a primary decomposition*

$$V = \bigoplus\nolimits_\Lambda V(\!(\mathfrak{g}, \lambda)\!) \,.$$

These are called the **primary components** of the $\mathfrak{g}$-module V.

We'll see later that $\mathfrak{g}$ is embedded into its universal enveloping algebra $U(\mathfrak{g})$, an associative algebra which it generates. Every $\mathfrak{g}$-module is automatically a module over $U(\mathfrak{g})$. If $\mathfrak{g}$ is abelian, its universal enveloping algebra is a polynomial algebra. This Proposition can be seen as an indication that basic results about modules over polynomial algebras extend to the universal enveloping algebras of nilpotent Lie algebras. For more along these lines see [McConnell:1967] and [Gabriel-Nouazé:1967].

13. Cartan subalgebras

If $\mathfrak{g}$ is a Lie algebra, the **normalizer** of any subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ is the space of all X such that $\mathrm{ad}_X \mathfrak{h} \subseteq \mathfrak{h}$. It is a Lie algebra that contains $\mathfrak{h}$ itself as an ideal.

A **Cartan subalgebra** is a nilpotent subalgebra whose normalizer is itself. It is not completely obvious that Cartan subalgebras exist, but in fact:

13.1. Theorem. *Every Lie algebra $\mathfrak{g}$ possesses at least one Cartan subalgebra.*

Proof. As [Merrin:1994] points out, a constructive proof is an easy consequence of Engel's Criterion. Recall that for every X in $\mathfrak{g}$

$$\mathfrak{g}(\!(\mathrm{ad}_X)\!) = \{Y \in \mathfrak{g} \mid \mathrm{ad}_X^n Y = 0 \text{ for some } n > 0\}\,.$$

13.2. Lemma. (Leibniz' rule) *If D is a derivation of a Lie algebra then*

$$D^n[A,B] = \sum_{k=0}^{n} \binom{n}{k} [D^k A, D^{n-k} B]\,.$$

Proof. By a straightforward induction.

As a consequence of Leibniz' rule, $\mathfrak{g}(\!(\mathrm{ad}_X)\!)$ *is a Lie subalgebra of* $\mathfrak{g}$.

Suppose Y to lie in the normalizer of $\mathfrak{g}(\!(\mathrm{ad}_X)\!)$. Since X itself lies in $\mathfrak{g}(\!(\mathrm{ad}_X)\!)$ we have $[X,Y] = -[Y,X] \in \mathfrak{g}(\!(\mathrm{ad}_X)\!)$, and hence Y itself will lie in $\mathfrak{g}(\!(\mathrm{ad}_X)\!)$. So *the normalizer of* $\mathfrak{g}(\!(\mathrm{ad}_X)\!)$ *is itself*. It will be a Cartan subalgebra if and only if it is nilpotent. In this case, for the moment, X will be called **regular**.

Therefore in order to prove Theorem 13.1 it suffices to prove that there exists a regular X in $\mathfrak{g}$.

We start off with an arbitrary element X in $\mathfrak{g}$. I shall show that if $\mathfrak{g}(\!(\mathrm{ad}_X)\!)$ is not nilpotent, there exists Z in $\mathfrak{g}(\!(\mathrm{ad}_X)\!)$ such that $\mathfrak{g}(\!(\mathrm{ad}_Z)\!)$ is a proper subalgebra of it. The decreasing sequence of subalgebras must eventually stop with a Cartan subalgebra.

If $\mathfrak{g}(\!(\mathrm{ad}_X)\!)$ is not nilpotent, then by Engel's Criterion there exists Y in it with ad_Y not nilpotent. By Fitting's Lemma we can write $\mathfrak{g} = \mathfrak{g}(\!(\mathrm{ad}_X)\!) \oplus (\!(\mathrm{ad}_X)\!)\mathfrak{g}$, with ad_X invertible on the second term. We can therefore write the matrix for ad_X as

$$M_X = \begin{bmatrix} A_X & \cdot \\ \cdot & C_X \end{bmatrix}\,.$$

and that for ad_Y as

$$M_Y = \begin{bmatrix} A_Y & B_Y \\ \cdot & C_Y \end{bmatrix}\,.$$

Here A_X is nilpotent, C_X is invertible, and A_Y is *not* nilpotent. I claim now that *some linear combination $Z = \alpha X + \beta Y$ satisfies the condition that $\mathfrak{g}(\!(\mathrm{ad}_Z)\!)$ is a proper subspace of $\mathfrak{g}(\!(\mathrm{ad}_X)\!)$, or in other words that (a) $\alpha A_X + \beta A_Y$ is not nilpotent but (b) $\alpha C_X + \beta C_Y$ is still invertible.* Let α, β be variables. Consider the characteristic polynomial

$$\begin{aligned} D_M(t,\alpha,\beta) &= \det(tI - \alpha M_X - \beta M_Y) \\ &= \det(tI - \alpha A_X - \beta A_Y)\det(tI - \alpha C_X - \beta C_Y) \\ &= D_A(t,\alpha,\beta) D_C(t,\alpha,\beta)\,. \end{aligned}$$

If $k = \dim \mathfrak{g}(\!(\mathrm{ad}_X)\!)$, then t^k cannot divide $D_A(t,\alpha,\beta)$ because this would contradict the fact that A_Y is not nilpotent. And t cannot divide $D_C(t,\alpha,\beta)$ since C_X is invertible. Thus if

$$D_M(t,\alpha,\beta) = \sum t^\ell c_\ell(\alpha,\beta)$$

then some c_ℓ with $\ell < k$ must not vanish identically as a polynomial in α, β. Choose numbers α, β so it doesn't vanish. In these circumstances, $\mathfrak{g}(\!(\mathrm{ad}_Z)\!)$ is a proper subspace of $\mathfrak{g}(\!(\mathrm{ad}_X)\!)$. Induction allows us to conclude. ∎

If $\mathfrak{h}$ is a Cartan subalgebra of $\mathfrak{g}$ then according to Theorem 12.8 $\mathfrak{g}$ decomposes into a direct sum of components $\mathfrak{g}(\!(\mathrm{ad}_\mathfrak{h}, \lambda)\!)$. The characters λ that occur are called the **roots** of $\mathfrak{g}$ with respect to $\mathfrak{h}$. The algebra itself is certainly contained in the root space $\mathfrak{g}(\!(\mathrm{ad}_\mathfrak{h})\!)$.

13.3. Proposition. *If $\mathfrak{h}$ is a nilpotent Lie subalgebra of $\mathfrak{g}$, the following are equivalent:*

(a) it is a Cartan subalgebra;
(b) it is the same as $\mathfrak{g}(\!(\mathrm{ad}_\mathfrak{h})\!)$.

Proof. Suppose $\mathfrak{h}$ nilpotent, let $\mathfrak{n}$ be its normalizer in $\mathfrak{g}$, and let $\mathfrak{r} = \mathfrak{g}(\!(\mathrm{ad}_\mathfrak{h})\!)$. If $Y \in \mathfrak{n}$ then for any $X \in \mathfrak{h}$

$$[X, Y] = -[Y, X] \in \mathfrak{h}$$

and $\mathrm{ad}_X^n Y = 0$ for some n, so

$$\mathfrak{h} \subseteq \mathfrak{n} \subseteq \mathfrak{r}\,.$$

Thus if $\mathfrak{h} = \mathfrak{r}$ then $\mathfrak{h} = \mathfrak{n}$ and $\mathfrak{h}$ is a Cartan subalgebra.

Conversely, suppose $\mathfrak{h} = \mathfrak{n}$. Then $\mathfrak{h}$ acts nilpotently on $\overline{\mathfrak{r}} = \mathfrak{r}/\mathfrak{h}$, so by Engel's Theorem either $\overline{\mathfrak{r}} = \{0\}$ or there exists $\overline{Y} \neq 0$ such that $\mathrm{ad}_X \overline{Y} = 0$ for all X in $\mathfrak{h}$. But if Y in $\mathfrak{r}$ has image $\overline{Y}$ then Y must lie in $\mathfrak{n}$, a contradiction. ∎

13.4. Proposition. *For characters λ, μ of a Cartan subalgebra*

$$[\mathfrak{g}_\lambda, \mathfrak{g}_\mu] \subseteq \mathfrak{g}_{\lambda+\mu}\,.$$

Proof. Suppose X in $\mathfrak{a}$, U in $\mathfrak{g}_\lambda$, Y in $\mathfrak{g}_\mu$. Let $D = \mathrm{ad}_X$. The case most commonly seen is that in which the Cartan subalgebra is abelian and the root spaces are eigenspaces. Since D is a derivation, we have

$$(D - \lambda - \mu)[X, Y] = [(D - \lambda)X, Y] + [X, (D - \mu)Y] = 0\,,$$

which proves the proposition in this simplest case. But the formula

$$(D - \lambda - \mu)[X, Y] = [(D - \lambda)X, Y] + [X, (D - \mu)Y]$$

is valid for any X, Y at all. We may apply it repeatedly to get

$$(D - \lambda - \mu)[X, Y] = \sum_{p=0}^{n} \binom{n}{p} [(D - \lambda)^p\, X, (D - \mu)^{n-p}\, Y]\,,$$

from which the Proposition follows if we just take n large enough. ∎

14. Conjugacy of Cartan subalgebras

In this section, let $\mathfrak{g}$ be a Lie algebra defined over the algebraically closed field F (still of characteristic 0). Recall that $\mathfrak{g}$ decomposes into root spaces with respect to any Cartan subalgebra.

If x is a nilpotent element of $\mathfrak{g}$, and in particular if it is in some non-trivial root space, the exponential $\exp(\mathrm{ad}_x)$ is an automorphism of $\mathfrak{g}$. Let $\mathrm{Int}(\mathfrak{g})$ be the group generated by these. In this section I'll prove, along with other related results:

14.1. Theorem. *Any two Cartan subalgebras of $\mathfrak{g}$ are conjugate by an element of $\mathrm{Int}(\mathfrak{g})$.*

I follow the treatment attributed in [Serre:1966] to Chevalley and presented in [Cartier:1955].

The proof starts with a simple, useful (and perhaps well known) result from algebraic geometry. It is a partial substitute for the implicit function theorem, which does not remain valid.

14.2. Lemma. *Suppose $\varphi: U \to V$ to be a polynomial map from one vector space over F to another. If $d\varphi$ is surjective at some point of U, then φ is generically surjective.*

The precise conclusion is that

> *if we are given a polynomial P on U, we can find a polynomial Q on V with the property that every point v of V with $Q(v) \neq 0$ is equal to $\varphi(u)$ for some u in U with $P(u) \neq 0$.*

Intuitively, this is quite plausible. Algebraic closure is certainly necessary, as the map $x \mapsto x^2$ from $\mathbb{R}$ to itself shows.

Proof. Say U has dimension m, V has dimension n. Let (x_i) be the coordinate system on U, (y_j) that on V. Under the hypothesis of the Lemma, and by a change of coordinates if necessary, we may assume that φ is given by the formula

$$y_j = x_j + \text{ second order terms for } 1 \leq j \leq m\,.$$

Step 1. I first show that composition with f, which maps $F[V]$ to $F[U]$, is an injection. Intuitively, this means that the image of U in V is not a proper subvariety of V, since if I is the ideal of $F[V]$ vanishing on the image, the map from $F[V]$ to $F[U]$ is filtered through $F[V]/I$. Given $f(y) \neq 0$, suppose it has order N at 0, so

$$f = \sum_{|k|=N} f_k y^k + \text{ terms of order } > N$$

with some $f_k \neq 0$. But then

$$f(\varphi(x)) = \sum_{|k|=N} f(k) x^k + \text{ terms of order } > N\,.$$

Hence the image of f in $F[U]$ does not vanish.

Step 2. From this point on, we do not need the assumption about $d\varphi$ any more—we just need the ring $F[V]$ to be embedded in $F[U]$. A point of U may be identified with a ring homomorphism π_u from $F[U]$ to F, and similarly for a point of V. If P is a polynomial in $F[U]$ and a u point in U then $P(u) = 0$ if and only if $\pi_u(P) = 0$. Thus what must be proven is this:

> *Given P in $F[U]$, we can find Q in $F[V]$ such that any homomorphism $\pi: F[V] \to F$ with $\pi(Q) \neq 0$ may be extended to a ring homomorphism $\Pi: F[U] \to F$ with $\Pi(P) \neq 0$.*

Define by recursion

$$R_0 = F[V], \quad R_k = R_{k-1}[x_k]\,.$$

Thus $R_m = F[U]$. An easy induction reduces the proof of the claim to this:

14.3. Lemma. *Suppose R to be an integral domain containing F, $S = R[s]$ an integral domain containing R and generated over R by a single element s. If $\sigma \neq 0$ is an element of S, there exists an element ρ in R*

with the following property: every homomorphism $\pi_R\colon R \to F$ *with* $\pi_R(\rho) \neq 0$ *lifts to a homomorphism* $\pi_S\colon S \to F$ *with* $\pi_S(\sigma) \neq 0$*.*

Proof of the Lemma. Represent S as $R[x]/I$, where x is a variable and I an ideal. Because S contains R, we must have $R \cap I = (0)$. There are two cases:

(a) $I = (0)$ and $S = R[x]$. Suppose $\sigma = \sum r_i x^i$. If $\pi : r \mapsto \overline{r}$ is a ring homomorphim, the extensions to S correspond to a choice of image $\overline{x}$ of x, which is arbitrary. The image of σ is $\sum \overline{r}_i \overline{x}^i$. Since $\sigma \neq 0$ we must have some $r_k \neq 0$ in R. Set $\rho = r_k$. If $\overline{\rho} \neq 0$ the polynomial $\sum \overline{r}_i x^i$ does not vanish identically, and only has a finite number of roots. Since F is algebraically closed it is infinite, and we may choose $\overline{x}$ in F such that $\sum \overline{r}_i \overline{x}^i \neq 0$.

(b) $I \neq (0)$. If $\sum q_i x^i$ lies in I, then some $q_i \neq 0$ with $i > 0$ since $I \cap R = (0)$. Let $P(x) = \sum p_i x^i$ be of minimal degree d in I. If $Q(x)$ is any other polynomial in I then the division algorithm gives

$$p_d^n Q(x) = U(x)P(x) + V(x)$$

with the degree of V less than d. But then $V(x)$ must also lie in I, hence $V(x) = 0$. Because S is an integral domain, I is a prime ideal, and since $I \cap R = (0)$, we conclude:

A polynomial $Q(x)$ *lies in* I *if and only if* $p_d^n Q(x)$ *is a multiple of* $P(x)$ *in* $R[x]$ *for some* n*.*

Suppose $r \mapsto \overline{r}$ is a ring homomorphims from R to F, and let $\mathfrak{m}$ be its kernel, a maximal ideal of R. Extensions to S amount to homomorphisms from $\overline{S} = S/\mathfrak{m}S$ to F, or in other words to maximal ideals of $\overline{S}$ to F. If $\overline{p}_d \neq 0$, the quotient $\overline{S}$ is the same as $F[x]/(\overline{p})$. Suppose P is the image in S of the polynomial $\Pi = \sum \pi_i x^i$. Apply the division algorithm to it to get

$$p_d^n \Pi(x) = U(x)P(x) + V(x)$$

with the degree of V less than d. Because $P \neq 0$ in S, $\Pi(x)$ is not in I and $V(x) \neq 0$. Let v be one of its coefficients in R. Now set $Q = p_d v$. If $\overline{Q} \neq 0$ then the image $\overline{P}$ of Π in $\overline{S}$ is not 0. Since F is algebraically closed, we may find a homomorphism from $\overline{S}$ to F that does not annihilate $\overline{P}$.

I now turn to the original question of conjugacy of Cartan subalgebras in a Lie algebra. This turns out to be closely related to other important properties of Cartan subalgebras that I shall prove at the same time.

If A is any linear transformation of a finite-dimensional vector space of dimension n over an algebraically closed field, its characteristic polynomial $\det(T - A)$ will be of the form $\delta_\ell T^\ell + \cdots + T^n$ where ℓ is least such that $a_\ell \neq 0$. The integer ℓ is called the **nilpotent rank** of A. The **rank** of a Lie algebra is the smallest nilpotent rank of all linear transformations ad_X for X in $\mathfrak{g}$. It is at most n since the coefficient of T^n is 1, and it is at least 1 since $[X, X] = 0$. An element X is called **regular** if its nilpotent rank is the same as that of $\mathfrak{g}$. If $\mathfrak{g}$ is assigned a coordinate system the coefficients of the characteristic polynomial of ad_X will be a polynomial in the coefficients of X, so the regular elements are the complement in $\mathfrak{g}$ of a proper algebraic subvariety.

It will follow from Theorem 14.1 that all Cartan algebras have the same dimension. More precisely:

14.4. Proposition. *All Cartan subalgebras have dimension equal to the rank of* $\mathfrak{g}$*.*

14.5. Proposition. *Suppose* $\mathfrak{a}$ *to be a Cartan subalgebra of* $\mathfrak{g}$*, and suppose* x *to lie in* $\mathfrak{a}$*. Then* $\langle \lambda, x \rangle \neq 0$ *for all roots* λ *if and only if* x *is regular in* $\mathfrak{g}$*.*

I'll prove these and Theorem 14.1 all at the same time. Let ℓ be the rank of $\mathfrak{g}$, δ_ℓ the corresponding coefficient of the characteristic polynomials, so that $\delta_\ell(X) \neq 0$ if and only if X is regular. Fix for the moment the Cartan subalgebra $\mathfrak{a}$ and the associated root decomposition of $\mathfrak{g}$. Let $\mathfrak{g}_{\neq 0}$ be the direct sum of the $\mathfrak{g}_\lambda$ with $\lambda \neq 0$. Choose a basis (X_i) (say for $1 \leq i \leq n$) of $\mathfrak{g}_{\neq 0}$ with X_i in $\mathfrak{g}_{\lambda_i}$, and define the map

$$\varphi\colon \left(\bigoplus_{\lambda \neq 0} \mathfrak{g}_\lambda\right) \oplus \mathfrak{a} \longrightarrow \mathfrak{g}$$

taking

$$(X_i) \times H \longmapsto \exp(\mathrm{ad}_{X_1}) \dots \exp(\mathrm{ad}_{X_n}) \cdot H .$$

Since

$$\exp(\mathrm{ad}_{tX})Y = I + \mathrm{ad}_{tX}Y + \cdots = I + t[X, Y] + \cdots$$

the Jacobian map $d\varphi$ at h in $\mathfrak{a}$ takes $(X_i) \oplus H$ to $(\langle \lambda_i, h \rangle X_i) \oplus H$. It is thus an isomorphism of $U = V = \mathfrak{g}$ with itself as long as $\langle \lambda_i, h \rangle \neq 0$ for all i. Let P be the product $\prod \lambda(H)$ on $\big((t_i X_i), H\big)$, and let $\mathfrak{a}'$ be the set of H in $\mathfrak{a}$ with $P(H) \neq 0$. By the Proposition, there exists a polynomial $Q(x)$ on $\mathfrak{g}$ with the property that if $Q(X) \neq 0$ then X is conjugate by $\mathrm{Int}(\mathfrak{g})$ to some H in $\mathfrak{a}'$. The set of regular elements of $\mathfrak{g}$ is Zariski-open in $\mathfrak{g}$ and invariant under $\mathrm{Int}(\mathfrak{g})$, so any X in $\mathfrak{g}$ with $Q(X)P(X) \neq 0$ is $\int(\mathfrak{g})$-conjugate to some regular H in $\mathfrak{a}$. But for such an H, the space on which ad_H is nilpotent is all of $\mathfrak{a}$ on the one hand, and of dimension ℓ on the other. Hence $\mathfrak{a}$ has dimension ℓ.

If $\mathfrak{b}$ is another Cartan subalgebra, the same holds for it, and there must exist some x in $\mathfrak{g}$ which is at once conjugate to a regular element of $\mathfrak{a}$ as well as one of $\mathfrak{b}$, which implies that the nilpotent component of x is conjugate to $\mathfrak{a}$ as well as $\mathfrak{b}$, so they must also be conjugate to each other.

15. Killing forms

If ρ is a (finite-dimensional) representation of $\mathfrak{g}$, we can define a symmetric inner product

$$K_\rho\colon\ \mathfrak{g} \otimes \mathfrak{g} \to F, \quad X \otimes Y \mapsto \mathrm{trace}\big(\rho(X)\rho(Y)\big) .$$

It is called the **Killing form** associated to ρ. If ρ is not specified, it is assumed to be the adjoint representation ad.

15.1. Proposition. *For any representation ρ the bilinear form K_ρ is $\mathfrak{g}$-invariant. If $\mathfrak{g}$ is the Lie algebra of G and ρ is the differential of a smooth representation of G then it is also G-invariant.*

The first assertion means that the adjoint representation of $\mathfrak{g}$ on itself takes $\mathfrak{g}$ into the Lie algebra of the orthogonal group preserving this bilinear form.

Proof. First I'll show invariance under $\mathfrak{g}$. We must show that

$$K_\rho(\mathrm{ad}_X Y, Z) + K_\rho(Y, \mathrm{ad}_X Z) = 0 .$$

This translates to the condition

$$\mathrm{trace}\big((\rho(\mathrm{ad}_X Y)\rho(Z)\big) + \mathrm{trace}\big(\rho(Y)\rho(\mathrm{ad}_X Z)\big) = 0$$

and then to

$$\begin{aligned} &\mathrm{trace}\big((\rho([X, Y]\rho(Z)\big) + \mathrm{trace}\big(\rho(Y)\rho([X, Z])\big) \\ &\qquad = \mathrm{trace}\big(\rho(X)\rho(Y)\rho(Z) - \rho(Y)\rho(X)\rho(Z) + \rho(Y)\rho(X)\rho(Z) - \rho(Y)\rho(Z)\rho(X)\big) \\ &\qquad = \mathrm{trace}\big(\rho(X)\rho(Y)\rho(Z) - \rho(Y)\rho(Z)\rho(X)\big) \\ &\qquad = 0 . \end{aligned}$$

Now the matter of G-invariance. For ease of reading I'll write $\mathrm{Ad}(g)X$ as gXg^{-1}. Then

$$\begin{aligned} \mathrm{trace}\big(\rho(gXg^{-1})\rho(gYg^{-1})\big) &= \mathrm{trace}\big(\rho(g)\rho(X)\rho(g^{-1})\rho(g)\rho(Y)\rho(g^{-1})\big) \\ &= \mathrm{trace}\big(\rho(g)\rho(X)\rho(Y)\rho(g^{-1})\big) \\ &= \mathrm{trace}\big(\rho(X)\rho(Y)\big) . \end{aligned}$$

The Killing form on $\mathfrak{g}$ itself is *characteristic*:

15.2. Proposition. *The Killing form of $\mathfrak{g}$ is invariant under any automorphism of $\mathfrak{g}$.*

Proof. If α is an automorphism of $\mathfrak{g}$ then by definition $[\alpha(X), \alpha(Y)] = \alpha([X, Y])$. In other words

$$\mathrm{ad}_{\alpha(X)} = \alpha\, \mathrm{ad}_X\, \alpha^{-1} .$$

But then

$$\begin{aligned} \mathrm{trace}\big(\mathrm{ad}_{\alpha(X)}\mathrm{ad}_{\alpha(Y)}\big) &= \mathrm{trace}(\alpha\, \mathrm{ad}_X\, \alpha^{-1} \cdot \alpha\, \mathrm{ad}_Y\, \alpha^{-1}) \\ &= \mathrm{trace}(\alpha\, \mathrm{ad}_X \mathrm{ad}_y\, \alpha^{-1}) \\ &= \mathrm{trace}(\mathrm{ad}_X \mathrm{ad}_y) . \end{aligned}$$

If K is a $\mathfrak{g}$-invariant bilinear form on $\mathfrak{g}$, its **radical** is the subspace of X in $\mathfrak{g}$ such that $K(X, \mathfrak{g}) = 0$.

15.3. Proposition. *The radical of any $\mathfrak{g}$-invariant bilinear form is an ideal of $\mathfrak{g}$.*

Proof. Immediate.

15.4. Proposition. *If $\mathfrak{g} = \mathfrak{sl}_n$ then $K_{\mathrm{ad}}(X, Y) = 2n\, \mathrm{trace}(XY)$.*

In particular, it is non-degenerate.

Proof. The Lie algebra $\mathfrak{sl}_n$ is the space of $n \times n$ matrices of trace 0, with

$$\mathrm{ad}_X(Y) = XY - YX .$$

This action extends to the space of all matrices, and the action on the complement of $\mathfrak{sl}_n$ is trivial. Therefore the Killing form K_{ad} is also that associated to the adjoint action on $\mathfrak{gl}_n = M_n$. So for any X and Y in M_n the Killing form is $\mathrm{trace}(\mathrm{ad}_X \mathrm{ad}_Y)$, where $\mathrm{ad}_X \mathrm{ad}_Y$ takes

$$Z \longmapsto YZ - ZY \longmapsto X(YZ - ZY) - (YZ - ZY)X = XYZ - XZY - YZX + ZYX .$$

If A is any matrix, the trace of each of the maps

$$Z \longmapsto AZ, \quad Z \longmapsto ZA$$

is $n\, \mathrm{trace}(A)$, since as a left or right module over M_n the space M_n is a sum of n copies of the module F. Therefore the trace of the first and last maps is $2n\, \mathrm{trace}(XY)$. The trace of each of the middle two $Z \mapsto XZY$, $Z \mapsto YZX$ is the product $\mathrm{trace}(X)\mathrm{trace}(Y)$, hence 0 on $\mathfrak{sl}_n$.

16. Solvable Lie algebras

The **derived idea** $\mathcal{D}\mathfrak{g}$ of $\mathfrak{g}$ is that spanned by all commutators $[X, Y]$ (see Proposition 10.1). It is an ideal in $\mathfrak{g}$, and the quotient $\mathfrak{g}/D\mathfrak{g}$ is abelian. The **derived series** is that of the $\mathcal{D}^i\mathfrak{g}$ where

$$\mathcal{D}^0 = \mathfrak{g}, \quad \mathcal{D}^{i+1} = [\mathcal{D}^i, \mathcal{D}^i] .$$

Each $\mathcal{D}^i$ is an ideal in $\mathcal{D}^{i-1}$, but not necessarily an ideal of $\mathfrak{g}$. A Lie algebra is said to be **solvable** if some $\mathcal{D}^i = 0$. The ultimate origin of this term is presumably Galois' criterion for solvability of equations by radicals in terms of what we now call solvable Galois groups.

16.1. Lemma. *Suppose E a field extension of F. The Lie algebra $\mathfrak{g}$ over F is solvable if and only $\mathfrak{g} \otimes_F E$ is.*

Proof. Since

$$\mathcal{D}\mathfrak{g} \otimes_F E = \mathcal{D}\big(\mathfrak{g} \otimes_F E\big) .$$

A Lie algebra is solvable if and only if it possesses a filtration by subalgebras $\mathfrak{g}_i$ with each $\mathfrak{g}_i$ an ideal in $\mathfrak{g}_{i-1}$ and $\mathfrak{g}_{i-1}/\mathfrak{g}_i$ abelian. Any quotient or subalgebra of a solvable Lie algebra is solvable. Conversely, if a Lie algebra is an extension of solvable Lie algebras, it is solvable (this is not true for nilpotent Lie algebras). Every nilpotent Lie algebra is solvable, but not conversely. The prototypical solvable algebra is the Lie algebra $\mathfrak{b}_n$ of all upper triangular $n \times n$ matrices. In this case $\mathcal{D}\mathfrak{b}_n$ is $\mathfrak{n}_n$.

16.2. Theorem. (Lie's Criterion) *Assume F algebraically closed. A Lie subalgebra $\mathfrak{g}$ of $\mathfrak{gl}_n$ is solvable if and only if some conjugate lies in $\mathfrak{b}_n$.*

Algebraic closure of F is necessary. For example, if $F = \mathbb{Q}$ and $\mathfrak{g}$ is the Lie algebra of the multiplicative group of an algebraic extension of F of degree n, it is commutative, hence solvable, but not diagonalizable in $\mathfrak{gl}_n(F)$.

Proof. An induction argument reduces the proof to showing that there exists a common eigenvector for all X in $\mathfrak{g}$. But this we have seen already in Lemma 12.6.

16.3. Corollary. *A Lie algebra is solvable if and only if $\mathcal{D}\mathfrak{g}$ is nilpotent.*

Proof. One way is trivial. For the other, apply the Proposition to the image of $\mathfrak{g}$ in $\mathfrak{gl}(\mathfrak{g})$, after extending the base field to an algebraic closure of F.

16.4. Theorem. (Cartan's Criterion) *If the Lie algebra $\mathfrak{g}$ is solvable, then $\mathrm{trace}(\mathrm{ad}_X \mathrm{ad}_Y) = 0$ for all X in $\mathfrak{g}$, Y in $\mathcal{D}\mathfrak{g}$. Conversely, if $\mathrm{trace}(\mathrm{ad}_X \mathrm{ad}_Y) = 0$ for all X, Y in $\mathcal{D}\mathfrak{g}$ then $\mathfrak{g}$ is solvable.*

I recall that the inner product $\mathrm{trace}(\mathrm{ad}_X \mathrm{ad}_Y)$ is the Killing form of $\mathfrak{g}$.

Proof. Because of Theorem 16.2 the first claim is easy, since if b is an upper triangular matrix and n a nilpotent upper triangular matrix then nb is nilpotent.

As for the second, I follow the argument of [Jacobson:1962]. It was primarily for this purpose that I also followed him in speaking of the primary decomposition of modules over nilpotent Lie algebras.

It suffices to show that $\mathcal{D}\mathfrak{g}$ is nilpotent or, by Engel's Criterion, that each X in $\mathcal{D}\mathfrak{g}$ is nilpotent. This is clearly true if X lies in $\mathfrak{g}_\lambda$ with $\lambda \neq 0$, since then $\mathrm{ad}_X \mathfrak{g}_\mu \subseteq \mathfrak{g}_{\lambda+\mu}$ and some $\mathfrak{g}_{\lambda+n\mu} = 0$, so we may assume X to lie in $\mathfrak{g}_0 \cap \mathcal{D}\mathfrak{g}$, hence that $X = [X_\lambda, X_{-\lambda}]$ with $X_{\pm\lambda} \in \mathfrak{g}_{\pm\lambda}$.

16.5. Lemma. *If $X = [X_\lambda, X_{-\lambda}]$ then for every weight μ of X on any $\mathfrak{g}$-module V, $\mu(X)$ is a rational multiple of $\lambda(X)$.*

Proof. Given μ, the sum of weight spaces $\mathfrak{g}_{\mu+n\lambda}$ is stable under the operators $X_{\pm\lambda}$, and therefore the trace of $X = X_\lambda X_{-\lambda} - X_{-\lambda} X_\lambda$ is 0 on this space. But if n_ρ is the dimension of $\mathfrak{g}_\rho$, we therefore have

$$\sum_k n_{\mu+k\lambda}\big(\mu(X) + k\lambda(X)\big) = 0, \quad \Big(\sum_k n_{\mu+k\lambda}\Big)\mu(X) = -\lambda(X)\Big(\sum k n_{\mu+k\lambda}\Big)\,.$$

To finish the proof of Cartan's criterion, say $\mu(X) = c_{\mu,\lambda}\alpha(X)$ for all weights μ. Then

$$0 = \mathrm{trace}\, X^2 = \sum_\mu n_\mu \mu^2(X) = \lambda^2(X)\Big(\sum n_\mu c_{\mu,\lambda}^2\Big)$$

from which we conclude $\lambda(X) = 0$, and this then implies that ad_X is nilpotent.

17. The universal enveloping algebra

Roughly speaking, the **universal enveloping algebra** $U(\mathfrak{g})$ of a Lie algebra $\mathfrak{g}$ is the associative algebra generated by the space $\mathfrak{g}$ subject to the relations $XY - YX = [X, Y]$ for X, Y in $\mathfrak{g}$. More formally, it is the quotient of the tensor algebra $\bigotimes^\bullet \mathfrak{g}$ modulo these relations, in other words by the ideal $I = I_\mathfrak{g}$ of the tensor algebra spanned by all

$$A \otimes Y \otimes Z \otimes B - A \otimes Z \otimes Y \otimes B - A \otimes [Y, Z] \otimes B\,.$$

The following is immediate from the definition:

17.1. Theorem. *The universal enveloping algebra is universal in the sense that any linear map φ from $\mathfrak{g}$ to an associative algebra A such that*

$$\varphi([X, Y]) = \varphi(X)\varphi(Y) - \varphi(Y)\varphi(X)$$

determines a unique ring homomorphism from $U(\mathfrak{g})$ extending φ.

Choose a basis $\Xi = \{X_i\}$ of $\mathfrak{g}$, ordered by index. The monomials

$$X_{i_1} \otimes \ldots \otimes X_{i_n}$$

form a basis of $\bigotimes^\bullet \mathfrak{g}$. I call a monomial **reduced** if $i_k \le i_\ell$ whenever $k \le \ell$.

17.2. Lemma. *(*Poincaré-Birkhoff-Witt*) The image in $U(\mathfrak{g})$ of the ordered monomials in $\bigotimes^\bullet \mathfrak{g}$ are a basis.*

Proof. It will be long. I'll call a tensor **reduced** if all of the monomials occurring in it with non-zero coefficients are reduced. We want to show that for every tensor x there exists a unique tensor $\mathrm{red}(x)$ such that $x - \mathrm{red}(x)$ lies in I.

It is relatively easy to see that ordered monomials span the enveloping algebra. I need for this a measure of how non-reduced a tensor is. Define the degree of a monomial in $\bigotimes^\bullet \mathfrak{g}$:

$$\deg(Y_1 \otimes \ldots \otimes Y_n) = n\,.$$

Also for each monomial S define $|S|$ to be the number of its inversions, the pairs $k < \ell$ with $i_k > i_\ell$. To each non-zero tensor $T = \sum c_S S$ assign a vector $\rho(T) = (\rho_n)$ where

$$\rho_n = \sum_{\substack{\deg S = n \\ c_S \ne 0}} |S|$$

for $n > 0$. For convenience, set $\rho_n(0) = 0$ for all n. Thus ρ is a vector with an infinite number of coordinates in $\mathbb{N}$, all but a finite number equal to 0. Define the **reduced degree** of T to be largest n with $\rho_n(T) \ne 0$. A tensor T is reduced if and only if its reduced degree is 0. Order such vectors:

> $\nu < \rho$ if and only if either (a) the reduced degree of ν is less than that of ρ or (b) they have the same reduced degree d but $\rho_d(\nu) < \rho_d(\rho)$.

Any descending chain of such vectors must be finite. Write $S \prec T$ if $\rho(S) < \rho(T)$.

If a tensor T has a non-reduced term $A \otimes Y \otimes Z \otimes B$ with $Y > Z$, it is equivalent modulo I to

$$S = A \otimes Z \otimes Y \otimes B + A \otimes [Y, Z] \otimes B$$

with $S \prec T$. This can be continued only for a finite number of steps, and at the end we wind up with a reduced tensor. Thus the reduced tensors span the universal enveloping algebra. The hard (and interesting) part is to show that the images of reduced monomials in the enveloping algebra are linearly independent. This can't be quite trivial, because in the enveloping algebra the Jacobi identity holds automatically, which means that we should expect to use it somewhere.

I follow the proof to be found in [Bergman:1978]. Bergman's argument is not all that different from, but somewhat clearer than, the original one of [Birkhoff:1937].

The proof will specify the linear map red from $\bigotimes^\bullet \mathfrak{g}$ to the reduced tensors with these properties:

(a) $T - \mathrm{red}(T)$ lies in I;
(b) $\mathrm{red}(T) = T$ if all the monomials occurring in T are ordered;
(c) $\mathrm{red}(T) = 0$ if T lies in I.

This will suffice to prove the Theorem. First of all, since $T - \mathrm{red}(T)$ lies in I the image of T and $\mathrm{red}(T)$ in the enveloping algebra are the same. Second, if T is equivalent to two reduced tensors T_1 and T_2 then their difference lies in I. Hence $\mathrm{red}(T_1 - T_2) = 0$ but then

$$\begin{aligned} \mathrm{red}(T_1 - T_2) &= 0 \\ &= \mathrm{red}(T_1) - \mathrm{red}(T_2) \\ &= T_2 - T_2\,. \end{aligned}$$

Define $\sigma_{k,\ell}$ to be the linear operator defined on basis elements of the tensor algebra by the formulas

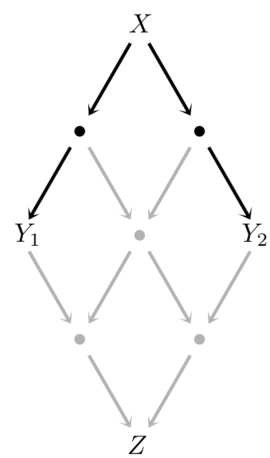
$$\begin{aligned}\sigma_{k,\ell}(T) &= T \\ &\quad \text{if } \deg T \neq k+2+\ell \\ \sigma_{k,\ell}(A\otimes X\otimes Y\otimes B) & \\ &= A\otimes X\otimes Y\otimes B \\ &\quad \text{if } \deg A = k,\ \deg B = \ell,\ X \le Y \\ &= A\otimes Y\otimes X\otimes B + A\otimes[X,Y]\otimes B \\ &\quad \text{if } \deg A = k,\ \deg B = \ell,\ X > Y\,.\end{aligned}$$

Thus $\sigma_{k,\ell}T - T$ always lies in I, and $\sigma_{k,\ell}T \prec T$. Write $S \to T$ if T is obtained from S by a single $\sigma_{k,\ell}$, and $S \xrightarrow{*} T$ if it is obtained from S by zero or more such operations.

All reductions σ leave irreducible expressions unchanged. If one takes B_1 and B_2 in the following Lemma to be irreducible, its conclusion is that $B_1 = B_2$.

17.3. Lemma. (PBW Confluence) *If A is a monomial with two reductions $A \xrightarrow{*} B_1$ and $A \xrightarrow{*} B_2$ then there exist reductions $B_1 \xrightarrow{*} C$ and $B_2 \xrightarrow{*} C$.*

Proof. As the following 'diamond' diagram suggests, the case of a simple reduction applied several times will prove the general case.



So we must show that if A is a monomial with two one-step reductions $\sigma: A \to B_1$ and $\tau: A \to B_2$ then there exist reductions $\tau': B_1 \to C$ and $\sigma': B_2 \to C$.

If the reductions are applied to non-overlapping pairs there is no problem. An overlap occurs for a term $A\otimes X\otimes Y\otimes Z\otimes B$ with $X > Y > Z$, say with $|A| = k$, $|B| = \ell$. It give rise to a choice of reductions:

$$\begin{aligned}X\otimes Y\otimes Z &\xrightarrow{\sigma_{k,\ell+1}} Y\otimes X\otimes Z + [X,Y]\otimes Z \\ X\otimes Y\otimes Z &\xrightarrow{\sigma_{k+1,\ell}} X\otimes Z\otimes Y + X\otimes[Y,Z]\,.\end{aligned}$$

But then

$$\begin{aligned}Y\otimes X\otimes Z + [X,Y]\otimes Z &\xrightarrow{\sigma_{k+1,\ell}} Y\otimes Z\otimes X + Y\otimes[X,Z] + [X,Y]\otimes Z \\ &\xrightarrow{\sigma_{k,\ell+1}} Z\otimes Y\otimes X + [Y,Z]\otimes X + Y\otimes[X,Z] + [X,Y]\otimes Z \\ X\otimes Z\otimes Y + X\otimes[Y,Z] &\xrightarrow{\sigma_{k,\ell+1}} Z\otimes X\otimes Y + [X,Z]\otimes Y + X\otimes[Y,Z] \\ &\xrightarrow{\sigma_{k+1,\ell}} Z\otimes Y\otimes X + Z\otimes[X,Y] + [X,Z]\otimes Y + X\otimes[Y,Z]\end{aligned}$$

and the difference

$$[[Y,Z],X]+[Y,[X,Z]]+[[X,Y],Z]$$

between the right hand sides, because of Jacobi's identity, lies in I.

I call a tensor **uniquely reducible** if there exists exactly one reduced tensor it can be reduced to. We now know that monomials are uniquely reducible. If S is any uniquely reducible element of $\bigotimes^\bullet \mathfrak{g}$, let $\mathrm{red}(S)$ the unique irreducible element it reduces to.

17.4. Lemma. *If S and T are uniquely reducible, so is $S+T$, and $\mathrm{red}(S+T)=\mathrm{red}(S)+\mathrm{red}(T)$.*

Proof. Suppose σ is a reduction taking $S+T$ to an irreducible expression W. According to the previous lemma, we can find a reduction σ' such that $\sigma'(\sigma(S))=\mathrm{red}(S)$. Since W is irreducible, on the one hand we have

$$\sigma'(\sigma(S+T))=\sigma'(W)=W$$

but on the other it is

$$\sigma'(\sigma(S))+\sigma'(\sigma(T))=\mathrm{red}(S)+\sigma'(\sigma(T))$$

Again according to the Lemma, we can find σ'' such that $\sigma''\sigma'\sigma(T)=\mathrm{red}(T)$. Then

$$w=\sigma''(W)=\sigma''(\mathrm{red}(S))+\sigma''(\sigma'(\sigma(T)))=\mathrm{red}(S)+\mathrm{red}(T)\,.$$

An induction argument now implies that every T in $\bigotimes^\bullet \mathfrak{g}$ is uniquely reducible. Define $\mathrm{red}(T)$ to be what it reduces to.

Implicit here is what is called 'confluence' in the literature, a tool of great power in finding normal forms for algebraic expressions. It is part of the theory of **term rewriting** and **critical pairs**, and although it has been used informally for a very long time, the complete theory seems to have originated in [Knuth-Bendix:1965]. It has been rediscovered independently a number of times. A fairly complete bibliography as well as some discussion of the history can be found in [Bergman:1978].

The ring $U(\mathfrak{g})$ is filtered by **order**. Let $U_n(\mathfrak{g})$ be the linear combinations of at most n products of elements of $\mathfrak{g}$. This filtration is compatible with products, hence determines the **graded ring**

$$\mathrm{Gr}^\bullet\, U(\mathfrak{g})=\bigoplus_n U_n(\mathfrak{g})/U_{n-1}(\mathfrak{g})\ .$$

The **symbol** of a product of n elements in $\bigotimes^\bullet \mathfrak{g}$ the corresponding product in the degree n component $S^n(\mathfrak{g})$ of the symmetric algebra of $\mathfrak{g}$. The kernel contains $U_{n-1}(\mathfrak{g})$, and therefore there exists a canonical map

$$\mathrm{Gr}^\bullet\, U(\mathfrak{g})\longrightarrow S^\bullet(\mathfrak{g})\ .$$

17.5. Proposition. *This canonical map is an isomorphism of graded rings.*

17.6. Corollary. *If $\mathfrak{g}$ is the Lie algebra of the Lie group G, then the map from $U(\mathfrak{g})$ to the ring of all left-invariant differential operators on G generated by the vector fields in $\mathfrak{g}$ mapping*

$$X_1\otimes X_2\otimes\cdots\otimes X_n\longmapsto R_{X_1}R_{X_2}\ldots R_{X_n}$$

is an isomorphism.

In other words, the ring $U(\mathfrak{g})$ may be described concretely in this case.

Part IV. Semi-simple Lie algebras

18. Casimir elements

Suppose now K to be any invariant and non-degenerate inner product on the Lie algebra $\mathfrak{g}$, for example the Killing form K_ρ associated to a suitable representation ρ.

In the following Proposition, let (X_i) be a basis of $\mathfrak{g}$ and $(X_i^\vee)$ the basis dual with respect to K.

18.1. Proposition. *The element*

$$C_K := \sum X_i X_i^\vee$$

lies in the centre of $U(\mathfrak{g})$.

It is called the **Casimir element** associated to K. In the literature one can often find a factor of 2 in this definition, placed there in order to match the Casimir to the Laplacian on the symmetric space attached to the group. If $K = K_{\mathrm{ad}}$, this element is just called the Casimir operator C without reference to K.

Proof. A straightforward calculation would do, but there is a better way. The bilinear form K gives rise to a covariant linear map τ_K from $\mathfrak{g}$ to its linear dual $\mathfrak{g}^\vee$, and τ_K is an isomorphism because of non-degeneracy. This in turn gives rise to a $\mathfrak{g}$-covariant isomorphism of $\mathfrak{g} \otimes \mathfrak{g}$ with $\mathfrak{g}^\vee \otimes \mathfrak{g} = \mathrm{End}(\mathfrak{g})$.

The linear transformation of $\mathfrak{g}$

$$Y \longmapsto \sum_i \langle X_i^\vee, Y\rangle\, X_i$$

is the identity transformation, since it takes each X_i to itself. We have a sequence of maps

$$\mathrm{Hom}_{\mathbb{C}}(\mathfrak{g}, \mathfrak{g}) \cong \mathfrak{g}^\vee \otimes \mathfrak{g} \xrightarrow{\text{Killing}} \mathfrak{g} \otimes \mathfrak{g} \longrightarrow U(\mathfrak{g})$$

which are all $\mathfrak{g}$-covariant. The Casimir element C_K is thus intrinsically characterized as the image of the identity transformation I on the left. Since I commutes with $\mathfrak{g}$ the Casimir does too, and since $U(\mathfrak{g})$ is generated by $\mathfrak{g}$ it lies in the center of $U(\mathfrak{g})$. ▯

In the rest of this section, we'll look at the case $\mathfrak{g} = \mathfrak{sl}_2$.

The Lie algebra $\mathfrak{g} = \mathfrak{sl}_2(F)$ has

$$\begin{aligned} h &= \begin{bmatrix} 1 & \cdot \\ \cdot & -1 \end{bmatrix} \\ e_+ &= \begin{bmatrix} \cdot & 1 \\ \cdot & \cdot \end{bmatrix} \\ e_- &= \begin{bmatrix} \cdot & \cdot \\ -1 & \cdot \end{bmatrix} \end{aligned}$$

as a basis. The $e_\pm$ are eigenvectors of ad_h. The complete specification of the Lie bracket is:

$$\begin{aligned} [h, e_\pm] &= \pm 2 e_\pm \\ [e_+, e_-] &= -h\,. \end{aligned}$$

Proposition 15.4 tells us:

18.2. Corollary. *For the basis h, $e_\pm$ of $\mathfrak{sl}_2$ the matrix of the Killing form is*

$$\begin{bmatrix} 8 & 0 & 0 \\ 0 & 0 & -4 \\ 0 & -4 & 0 \end{bmatrix}.$$

Given the explicit calculation of the form, the Casimir element can be seen to be

$$C = (1/8)\, h^2 - (1/4)\, e_+e_- - (1/4)\, e_-e_+ \,,$$

with alternate expressions

$$C = h^2/8 - h/4 - e_+e_-/2 = h^2/8 + h/4 - e_-e_+/2 \,.$$

Now take F to be $\mathbb{R}$, $G = \mathrm{SL}_2(\mathbb{R})$, $K = \mathrm{SO}(2) \subset G$. One reason that the Casimir element is important is because it is related to the Laplacian operator on the Riemannian space $\mathcal{H} = G/K$. Explicitly:

18.3. Lemma. *Acting on functions on $\mathcal{H}$, the Casimir is the same as half the Laplacian.*

Proof. We know that twice the Casimir is

$$2C = h^2/4 - h/2 - e_-e_+ \,,$$

so its action on functions on $\mathcal{H}$ is as

$$\begin{aligned}
L_{2C} &= (\Lambda_h)^2/4 - \Lambda_h/2 + L_{e_-}L_{e_+} \\
&= \Big(x\,\frac{\partial}{\partial x} + y\,\frac{\partial}{\partial y}\Big)^2 - \Big(x\,\frac{\partial}{\partial x} + y\,\frac{\partial}{\partial y}\Big) - (x^2-y^2)\,\frac{\partial^2}{\partial x^2} - 2xy\,\frac{\partial^2}{\partial x \partial y} \\
&= x^2\,\frac{\partial^2}{\partial x^2} + 2xy\,\frac{\partial^2}{\partial x \partial y} + y^2\,\frac{\partial^2}{\partial y^2} - (x^2-y^2)\,\frac{\partial^2}{\partial x^2} - 2xy\,\frac{\partial^2}{\partial x \partial y} \\
&= y^2\,\frac{\partial^2}{\partial x^2} + y^2\,\frac{\partial^2}{\partial y^2} \\
&= \Delta_{\mathcal{H}} \,.
\end{aligned}$$

19. Semi-simple Lie algebras

If $\mathfrak{h}$ and $\mathfrak{k}$ are solvable ideals of $\mathfrak{g}$ then so is $\mathfrak{h} + \mathfrak{k}$. Therefore there exists a maximal solvable ideal, which I shall call the **Lie radical** of $\mathfrak{g}$. A Lie algebra is called **semi-simple** if its Lie radical is 0. The **Killing radical** of $\mathfrak{g}$ is the radical of the Killing form K_{ad}—i.e. the subspace of all X in $\mathfrak{g}$ such that $K(X, \mathfrak{g}) = 0$.

19.1. Lemma. *The Killing radical is a solvable ideal of $\mathfrak{g}$.*

Proof. Let $\mathfrak{r}$ be the **Killing radical** of $\mathfrak{g}$. It is immediate from the invariance of the Killing form that $\mathfrak{r}$ is an ideal of $\mathfrak{g}$. If x lies in $\mathfrak{r}$ then $K(x, y) = 0$ for all y in $\mathfrak{g}$, hence in particular for y in $\mathcal{D}\mathfrak{r}$. Hence by Cartan's criterion $\mathrm{ad}_{\mathfrak{g}}\mathfrak{r}$ is solvable. But $\mathfrak{r}$ is an extension of this by some subspace of the centre of $\mathfrak{g}$, so that $\mathfrak{r}$ itself must be solvable.

19.2. Theorem. *A Lie algebra is semi-simple if and only if its Killing form K_{ad} is non-degenerate.*

Proof. If the Killing form is degenerate then the Killing radical is non-trivial, and by the Lemma $\mathfrak{g}$ contains a non-trivial solvable ideal, hence cannot be semi-simple.

For the other half, suppose that the Lie radical $\mathfrak{r}$ is not 0. Then either $\mathfrak{r}$ is abelian, or $\mathcal{D}\mathfrak{r}$ is nilpotent and not 0. In the latter case, the centre of $\mathcal{D}\mathfrak{r}$ is an abelian ideal of $\mathfrak{g}$. In either case, we may assume that $\mathfrak{g}$ contains a non-trivial abelian ideal $\mathfrak{a}$. Then $K(a, x) = 0$ for all a in $\mathfrak{a}$, x in $\mathfrak{g}$, so $\mathfrak{a}$ is contained in the Killing radical of $\mathfrak{g}$.

19.3. Theorem. *If $\mathfrak{g}$ is semi-simple and $\mathfrak{h}$ is an ideal of $\mathfrak{g}$, then $\mathfrak{h}^\perp$ is a complementary ideal.*

Proof. Because, as the argument just finished shows, $\mathfrak{h} \cap \mathfrak{h}^\perp$ is solvable.

This can be improved somewhat:

19.4. Proposition. *If $\mathfrak{g}$ is a semi-simple ideal in a Lie algebra $\mathfrak{h}$, there exists a complementary ideal in $\mathfrak{h}$.*

Proof. Consider $\mathfrak{h}$ as a module over $\mathfrak{g}$, via the adjoint representation. Let $\mathfrak{a}$ be a $\mathfrak{g}$-stable complemnt to $\mathfrak{g}$ itself. It is easy to see that $[\mathfrak{g}, \mathfrak{a}] = 0$, and in fact $\mathfrak{a}$ is the space of all y such that $[\mathfrak{g}, y] = 0$. Therefore $\mathfrak{a}$ is uniquely characterized, and it is an ideal in $\mathfrak{h}$ because it is the annihilator of the $\mathfrak{h}$-ideal $\mathfrak{g}$.

A **simple** Lie algebra is one with no non-trivial ideals. If it has dimension greater than one, it will be semi-simple, in which case I'll call it a **proper** simple Lie algebra.

19.5. Corollary. *Any semi-simple Lie algebra is the direct sum of simple Lie algebras.*

19.6. Corollary. *If $\mathfrak{g}$ is semi-simple then $\mathfrak{g} = D\mathfrak{g}$.*

Proof. Because it is clearly true of a simple algebra.

The following is the principal result of this section, and motivates the term "semi-simple."

19.7. Theorem. *Any finite-dimensional representation of a semi-simple Lie algebra decomposes into a direct sum of irreducible representations.*

Proof. I follow LA §6.3 of [Serre:1965], which was perhaps the first clear account of the purely algebraic proof (as opposed to the one of Hermann Weyl for complex groups, which utilized the relationship between complex semi-simple groups and their maximal compact subgroups).

The proof is by induction on dimension. If $V = \{0\}$ there is no problem. Otherwise we can find a proper $\mathfrak{g}$-stable subspace U and a short exact sequence of $\mathfrak{g}$-representations

$$0 \longrightarrow U \longrightarrow V \longrightarrow W \longrightarrow 0$$

where W is irreducible. By induction, U decomposes into a direct sum of irreducibles, so the Theorem will follow from this:

19.8. Lemma. *An exact sequence of $\mathfrak{g}$-modules as above, with W irreducible, splits.*

Proof of the Lemma. We assume at first W to be the trivial representation $\mathbb{C}$, so our exact sequence is

$$0 \longrightarrow U \longrightarrow V \longrightarrow \mathbb{C} \longrightarrow 0$$

We proceed now by induction on the dimension of U. If it is 0, there is nothing to prove. Otherwise, U will contain an irreducible $\mathfrak{g}$ subspace U'.

Case 1. The subspace U' is not all of U. We have an exact sequence

$$0 \longrightarrow U/U' \longrightarrow V/U' \longrightarrow \mathbb{C} \longrightarrow 0$$

which splits by the induction assumption. Therefore V/U' contains a one dimensional $\mathfrak{g}$-stable subspace, so we can write

$$V/U' = U/U' \oplus \mathbb{C}\ .$$

If V' is the inverse image in V of $\mathbb{C}$, we have an exact sequence

$$0 \longrightarrow U' \longrightarrow V' \longrightarrow \mathbb{C} \longrightarrow 0$$

which again splits by induction, giving a one-dimensional subspace of the original V.

Case 2. The subspace U in the sequence is irreducible.

There is a further subdivision. (a) Suppose U is trivial. We are looking at this situation:

$$0 \longrightarrow F \longrightarrow V \longrightarrow F \longrightarrow 0\ .$$

If X and Y are elements of $\mathfrak{g}$ then XY and YX are both 0 on V, hence $[X, Y] = 0$ as well. But by Corollary 19.6 $\mathfrak{g} = \mathcal{D}\mathfrak{g}$ so all of $\mathfrak{g}$ acts trivially on V, which must be $F \oplus F$.

(b) The representation ρ on U is not trivial. The image $\mathfrak{g}_\rho$ of $\mathfrak{g}$ in $\mathrm{End}(U)$ factors through a direct sum of simple algebras, and the Killing form K_ρ also factors through $\mathfrak{g}_\rho$.

The Killing for K_ρ is non-degenerate on $\mathfrak{g}_\rho$.

Similar things have been proved before, I leave this claim as an exercise.

Because ρ is irreducible, the Casimir C_ρ acts as a scalar, say γ_ρ. Its trace is on the one hand $\gamma_\rho \dim V$, and on the other

$$\mathrm{trace} C_\rho = (1/2) \sum_i \mathrm{trace}\big(\rho(X_i)\rho(X_i^\vee)\big) = \frac{\dim \mathfrak{g}}{2},$$

so $\gamma_\rho \neq 0$. But C_ρ lies in the universal enveloping algebra of $\mathfrak{g}$, and acts as 0 on F. So V decomposes into a direct sum of eigenspaces with respect to C_ρ.

At this point, we know that any sequence

$$0 \longrightarrow U \longrightarrow V \longrightarrow F \longrightarrow 0$$

splits.

Suppose now that we have an arbitrary exact sequence of $\mathfrak{g}$-spaces

$$0 \longrightarrow U \longrightarrow V \longrightarrow W \longrightarrow 0 .$$

with W irreducible, and consider the exact sequence of $\mathfrak{g}$-modules

$$0 \longrightarrow \mathrm{Hom}_F(W, U) \longrightarrow \mathrm{Hom}_F(W, V) \longrightarrow \mathrm{Hom}_F(W, W) \longrightarrow 0 .$$

Let V' be the subspace of maps in $\mathrm{Hom}_F(W, V)$ mapping onto scalar multiplications. Then we have a sequence

$$0 \longrightarrow \mathrm{Hom}_F(W, U) \longrightarrow V' \longrightarrow F \longrightarrow 0$$

But since this splits as a $\mathfrak{g}$-module, there exists a $\mathfrak{g}$-invariant element in $V' \subseteq \mathrm{Hom}_F(W, V)$ mapping onto the identity map from W to itself. This amounts to a splitting.

20. Representations of SL(2)

The Lie algebra $\mathfrak{g} = \mathfrak{sl}_2(F)$ has

$$\begin{aligned} h &= \begin{bmatrix} 1 & \cdot \\ \cdot & -1 \end{bmatrix} \\ e_+ &= \begin{bmatrix} \cdot & 1 \\ \cdot & \cdot \end{bmatrix} \\ e_- &= \begin{bmatrix} \cdot & \cdot \\ -1 & \cdot \end{bmatrix} \end{aligned}$$

as a basis. The $e_\pm$ are eigenvectors of ad_h. The complete specification of the Lie bracket is:

$$\begin{aligned} [h, e_\pm] &= \pm 2 e_\pm \\ [e_+, e_-] &= -h . \end{aligned}$$

Suppose V to be an irreducible finite-dimensional module over $\mathfrak{g}$. If $\overline{V}$ is $V \otimes_F \overline{F}$ (where $\overline{F}$ is an algebraic closure of F) there exists at least one eigenvalue of h:

$$h{\cdot}v = \lambda v .$$

Then $e_\pm{\cdot}v$ is also an eigenvector with eigenvalue $\lambda \pm 2$, and likewise $e_\pm^k{\cdot}v$ is of eigenvalue $\lambda \pm 2k$. Since these eigenvalues are distinct, the non-zero vectors among these must be linearly independent. Since V has finite dimension, there exists k with $v_0 = e_+^k{\cdot}v \neq 0$ but $e_+{\cdot}v_0 = 0$.

Let $\mu = \lambda + 2k$, so $h{\cdot}v_0 = \mu v_0$. The space spanned by the vectors $v_\ell = e_-^\ell\, v_0$ is $\mathfrak{g}$-stable. It must therefore be all of $\overline{V}$. The vector v_ℓ is an eigenvector of h with eigenvalue $\mu - 2\ell$. Finite-dimensionality implies that $e_-{\cdot}u_n = 0$ for some n.

The Casimir element C of $U(\mathfrak{g})$ acts as a scalar on $\overline{V}$. Since

$$C = h^2/4 + h/2 + e_-e_+ = h^2/4 - h/2 + e_+e_-\,,$$

we have

$$\begin{aligned}
Cv_0 &= (\mu^2/4 + \mu/2)\, v_0 \\
Cu_0 &= \big((\mu - 2n)^2/4 - (\mu - 2n)/2\big)u_0 \\
\frac{\mu^2}{4} + \frac{\mu}{2} &= \frac{\mu^2 - 4\mu n + 4n^2}{4} - \frac{\mu - 2n}{2} \\
0 &= -\mu n + n^2 - \mu + n \\
(n+1)(\mu - n) &= 0 \\
n &= \mu\,.
\end{aligned}$$

Consequently, μ must be a non-negative integer. Hence there exist eigenvectors already in the original F-space V, and one may (and I shall) assume all the v_ℓ to be in V itself.

20.1. Proposition. *Suppose V to be a vector space over F, and an irreducible module over $\mathfrak{sl}_2(F)$. There exists a vector v_0 with $e_+{\cdot}v_0 = 0$ and $h{\cdot}v_0 = nv_0$ for some integer $n\mathfrak{e}0$. The space V is then spanned by the vectors $v_\ell = e_-^\ell{\cdot}v_0$ for $0 \le \ell \le n$ with*

$$h{\cdot}v_\ell = (n - 2\ell)v_\ell, \quad e_-{\cdot}v_n = 0\,.$$

Taking into account the following, we deduce the explicit representation:

20.2. Lemma. *If $e_+{\cdot}v_0 = 0$ and $h{\cdot}v_0 = \lambda v_0$ then*

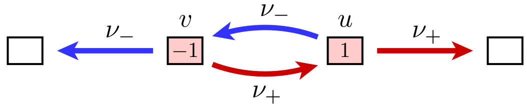
$$e_+e_-^k{\cdot}v_0 = k(k - 1 - \lambda)e_-^{k-1}{\cdot}v_0\,.$$

Proof. Say $e_+e_-^k{\cdot}v_0 = \lambda_k e_-^{k-1}{\cdot}v_0$. Then $\lambda_{k+1} = \lambda_k - \lambda + 2k$, and $\lambda_0 = 0$. The formula above follows by induction.

It can be verified directly by algebraic calculation that the formulas above define an irreducible representation of dimension $n+1$ of $\mathfrak{sl}_2$. But these representations can be constructed explicitly. Any finite-dimensional representation of $\mathrm{SL}_2(F)$ gives rise to a representation of its Lie algebra $\mathfrak{sl}_2(F)$. There are two obvious representations—the trivial one and the tautological representation π_1 on F^2. The associated representation of $\mathfrak{sl}_2$ acts like this on the standard basis $u = [1, 0]$, $v = [0, 1]$:

$$\begin{aligned}
h\ :\ u &\longmapsto u \\
v &\longmapsto -v \\
e_+\!:\ v &\longmapsto u \\
u &\longmapsto 0 \\
e_-\!:\ u &\longmapsto -v \\
v &\longmapsto 0
\end{aligned}$$

which can be illustrated (but not indicating a necessary $-$-sign):



This representation of $\mathrm{SL}_2(F)$ on $V = F^2$ gives rise to the representation π_n on the symmetric product S^nV with basis $u^k v^{n-k}$ for $0 \le k \le n$:

$$\pi_n(g)\colon\ u^k v^{n-k} \longmapsto (gu)^k (gv)^{n-k} .$$

Thus

$$\pi_n\left(\begin{bmatrix} 1 & x \\ \cdot & 1 \end{bmatrix}\right) :\ u^k v^{n-k} \longmapsto u^k (v + xu)^{n-k}$$

which implies

$$\pi_n(e_+)\colon\ u^k v^{n-k} \longmapsto (n-k) u^{k+1} v^{n-k-1} .$$

Similarly

$$\pi_n(e_-)\colon\ u^k v^{n-k} \longmapsto (-1)^k k u^{k-1} v^{n-k+1} .$$

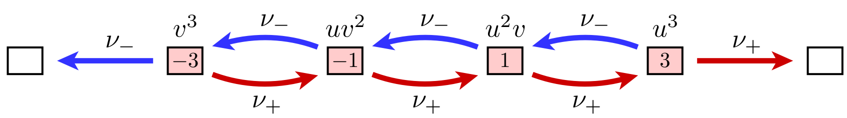
If $F = \mathbb{C}$ we can calculate $\pi_n(h)$ by applying the exponential and differentiating. In general, we can proceed formally by using the nil-ring $F[\varepsilon]$ with $\varepsilon^2 = 0$ or by using the formula $[e_+, e_-] = -h$ to deduce

$$\pi_n(h)\colon\ u^k v^{n-k} \longmapsto -(n-2k) u^k v^{n-k} .$$

Also

$$\begin{aligned} e_-e_+ &\colon u^k v^{n-k} \longmapsto (-1)^{k+1}(k+1)(n-k) u^k v^{n-k} \\ e_+e_- &\colon u^k v^{n-k} \longmapsto (-1)^k k(n-k+1) u^k v^{n-k} . \end{aligned}$$

For example, when $n = 3$:



There is one more simple formula that is useful. If

$$w = \begin{bmatrix} \cdot & 1 \\ -1 & \cdot \end{bmatrix}$$

then

$$\pi_n(w)\colon\ u^\ell v^{n-\ell} \longmapsto (1)^\ell u^{n-\ell} v^\ell ,$$

and if $n = 2\ell + 1$

$$\begin{aligned} \pi_n(w)&\colon\ u^\ell v^{\ell+1} \longmapsto (-1)^\ell u^{\ell+1} v^\ell \\ \pi_n(e_+)&\colon\ u^\ell v^{\ell+1} \longmapsto (-1)^\ell (\ell+1) \pi_n(w) u^\ell v^{\ell+1} . \end{aligned}$$

Every finite-dimensional representation of $\mathfrak{sl}_2$ is a direct sum of irreducible ones, and the irreducible ones are representations of $\mathrm{SL}_2(F)$. Hence:

20.3. Proposition. *Every finite-dimensional representation of $\mathfrak{sl}_2$ is derived from a representation of the group SL_2.*

This is the principal consequence of this section in what follows. But another consequence is this:

20.4. Proposition. *If V is a finite-dimensional module over $\mathfrak{sl}_2$, and v in V is an eigenvector for h annihilated by e_+, its eigenvalue with respect to h is non-negative.*

21. Tensor invariants

I learned the following result from LA §6.5 of [Serre:1966].

If we are given an embedding of $\mathfrak{g}$ into $\mathfrak{gl}(V)$, then there are associated representations of $\mathfrak{g}$ on the dual $\widehat{V}$ and on the tensor products

$$\bigotimes^{p,q} V = \bigotimes^p V \otimes \bigotimes^q \widehat{V} .$$

A **tensor invariant** for $\mathfrak{g}$ is any tensor annihilated by it. One example is the identity map in $\mathrm{End}(V)$, which may be identified with $\widehat{V} \otimes V$.

21.1. Theorem. *If $\mathfrak{g} \subseteq \mathfrak{gl}(V)$ is semi-simple, it is same as the subalgebra of $\mathfrak{gl}(V)$ leaving invariant all the tensor invariants of $\mathfrak{g}$.*

Proof. Let $\mathfrak{h}$ be the Lie algebra annihilating all the tensor invariants of $\mathfrak{g}$. We know that $\mathfrak{g} \subseteq \mathfrak{h}$, and we must show equality. This comes from a series of elementary steps interpreting tensors and tensor invariants suitably. For details, look at LA §6.5 of Serre's book.

(a) Any $\mathfrak{g}$-homomorphism from $\bigotimes^{p,q} V$ to $\bigotimes^{r,s} V$ is also an $\mathfrak{h}$-homomorphism.

(b) Any $\mathfrak{g}$-stable subspace of $\bigotimes^{p,q} V$ is also $\mathfrak{h}$-stable.

(c) By Proposition 19.4, we may write $\mathfrak{h} = \mathfrak{g} \oplus \mathfrak{c}$, where $\mathfrak{c}$ is an $\mathfrak{h}$-ideal commuting with $\mathfrak{g}$. But then by (a) it also commutes with $\mathfrak{h}$. This means that $\mathfrak{c}$ is the centre of $\mathfrak{h}$.

(d) If U is an irreducible $\mathfrak{g}$-submodule of V, then $\mathfrak{c}$ acts by scalars on it. In fact, it acts by 0. Why? The exterior products $\bigwedge^* W$ are contained in the tensor algebra. The semi-simple algebra $\mathfrak{g}$ acts trivially on its highest exterior product, and therefore so does $\mathfrak{h}$, hence $\mathfrak{c}$. But the action of any element X of $\mathfrak{gl}(V)$ on this power is by the scalar $\mathrm{trace}(X)$, which vanishes if X lies in $\mathfrak{c}$. ▮

If T is any linear transformation in $\mathrm{M}_n(F)$, where F is a field of characteristic 0, it may be written uniquely as $T_s + T_n$, where $T_s \in \mathrm{M}_n(F)$ is diagonalizable over an algebraic closure, $T_n \in |rmM_n(F)$ is nilpotent, and $T_s T_n = T_n T_s$. There exist polynopmials $P(x)$, $Q(x)$ with $P(0) = Q(0) = 0$ such that $P(T) = T$, $Q(T) = T_n$. This is called the **Jordan decomposition** of T. The endomorphism T_s is called the semi-simple component of T, T_n it nilpotent component.

21.2. Corollary. *If $\mathfrak{g}$ is semi-simple then any X in $\mathfrak{g}$ can be expressed as $X = X_s + X_n$, such that in any finite-dimensional representation (π, V) of $\mathfrak{g}$, $\pi(X_s)$ is the semi-simple component of $\pi(X)$ and $\pi(X_n)$ is its nilpotent component.*

Proof. Because $X_s = P(X)$ and $X_n = Q(X)$, both leave invariant anything that X does, hence any of the tensor invariants of $\mathfrak{g}$. ▮

22. The structure of semi-simple Lie algebras

In this section, let $\mathfrak{g}$ be a semi-simple Lie algebra. and $\mathfrak{a}$ a Cartan subalgebra, which exists by Theorem 13.1.

I recall that this means that $\mathfrak{a}$ is a nilpotent algebra that is its own normalizer in $\mathfrak{g}$. By Theorem 12.8, the adjoint action of $\mathfrak{a}$ on $\mathfrak{g}$ decomposes into a direct sum of primary modules $\mathfrak{g}_\lambda$. These are called **root spaces** when $\lambda \neq 0$.

Two spaces $\mathfrak{g}_\lambda$ and $\mathfrak{g}_\mu$ are orthogonal with respect to the Killing form $K = K_{\mathrm{ad}}$, so the decomposition

$$\mathfrak{g} = \mathfrak{a} + \bigoplus_{\pm\lambda \neq 0} \left(\mathfrak{g}_\lambda \oplus \mathfrak{g}_{-\lambda}\right)$$

is orthogonal. The restriction of K to any summand must be non-degenerate.

22.1. Proposition. *Every Cartan subalgebra of $\mathfrak{g}$ is abelian and is its own centralizer.*

Proof. According to Cartan's criterion, $\mathrm{trace}(\mathrm{ad}_X, \mathrm{ad}_Y) = 0$ for every X in $\mathfrak{a}$, Y in $\mathcal{D}\mathfrak{a}$. Since K is non-degenerate on $\mathfrak{a}$, we must have $\mathcal{D}(\mathfrak{a}) = 0$. Thus the centralizer of $\mathfrak{a}$ contains $\mathfrak{a}$ and is contained in its normalizer, hence must be $\mathfrak{a}$. ▮

22.2. Proposition. *Every element of $\mathfrak{a}$ is semi-simple.*

Proof. The nilpotent component in the Jordan decomposition, which lies in $\mathfrak{g}$ according to Corollary 21.2 is orthogonal to $\mathfrak{a}$.

For λ a root of $\mathfrak{g}$ with respect to $\mathfrak{a}$, let h_λ be the inverse image of λ under the isomorphism $\mathfrak{a} \to \widehat{\mathfrak{a}}$ induced by K.

22.3. Proposition. *Suppose X in $\mathfrak{g}_\lambda$, Y in $\mathfrak{g}_\mu$. Then $K(X,Y) = 0$ unless $\mu = -\lambda$, in which case $[X,Y] = K(X,Y)\,h_\lambda$.*

Proof. Only the last assertion is troublesome. For any H in $\mathfrak{a}$

$$K(H,[X,Y]) = K([H,X],Y) = \lambda(H){\cdot}K(X,Y) = K(H,h_\lambda)K(X,Y)\,,$$

since K is invariant.

22.4. Proposition. *There exists some multiple H_λ of h_λ such that $\langle \lambda, H_\lambda\rangle = 2$.*

Proof. It must be shown that $\langle \lambda, h_\lambda\rangle \neq 0$. Suppose the contrary. Since the restriction of K_{ad} is non-degenerate, by the previous result we may choose e in $\mathfrak{g}_\lambda$ and f in $\mathfrak{g}_{-\lambda}$ with $h = [e,f] \neq 0$. By assumption $\langle \lambda, h\rangle = 0$, so

$$[h,e] = 0, \quad [h,f] = 0, \quad [x,y] = h\,.$$

So the Lie algebra generated by them is solvable. Lie's criterion tells us that ad_h is nilpotent. But it is also semi-simple, so it must be 0. Contradiction.

If we choose $e \neq 0$ in $\mathfrak{g}_\lambda$, we may find f in $\mathfrak{g}_{-\lambda}$ such that $[e,f] = -h = -H_\lambda$. We then have

$$\begin{aligned} [h,e] &= 2e \\ [e,f] &= -h \\ [h,f] &= -2f \end{aligned}$$

which tells us that the subspace of $\mathfrak{g}$ spanned by e, h, and f is a Lie algebra isomorphic to $\mathfrak{sl}_2$. Hence we get the adjoint representation of this copy of $\mathfrak{sl}_2$ on $\mathfrak{g}$. It splits into a sum of irreducible representations, and this lifts to a representation of the group SL_2. These copies of $\mathfrak{sl}_2$ and the associated representations of SL_2 are part of the basic structure of a semi-simple Lie algebra.

22.5. Proposition. *The eigenspace $\mathfrak{g}_\lambda$ of the root λ has dimension one.*

Proof. It suffices to prove that every $\mathfrak{g}_{-\lambda}$ has dimension one. The subspace $e^\perp$ in $\mathfrak{g}_{-\lambda}$ has codimension one. Say $f_* \neq 0$ lies in $e^\perp$. By Proposition 22.3, $[e,f_*] = 0$. But then we have a vector in a finite-dimensional module over $\mathfrak{sl}_2$ annihilated by e of weight -2 with respect to h, which contradicts Proposition 20.4.

This means that a choice of $e_\lambda \neq 0$ in $\mathfrak{g}_\lambda$ is unique up to scalar, and then we get $e_{-\lambda}$, H_λ generating a unique copy of $\mathfrak{sl}_2$ in $\mathfrak{g}$ that I'll call $\mathfrak{sl}_{2,\lambda}$. The adjoint representation of this subalgebra on $\mathfrak{g}$ splits up into a direct sum of irreducible representations, and is derived from a direct sum of representations of the associated group $\mathrm{SL}_{2,\lambda}$. Let s_λ be the involutory automorphism of $\mathfrak{g}$ corresponding to the image of

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

in $\mathfrak{g}$. It commutes with the kernel of λ in $\mathfrak{a}$, which is a complement to H_λ. This involution take H_λ to $-H_\lambda$, so normalizes $\mathfrak{a}$. Since it acts as an automorphims on all of $\mathfrak{g}$, this proves:

22.6. Proposition. *The reflection s_α takes Σ to itself.*

22.7. Proposition. *The set Σ spans $\widehat{\mathfrak{h}}$.*

Proof. An element h in the complement acts trivially on $\mathfrak{g}$, hence lies in the centre of $\mathfrak{g}$.

The involutions s_λ also preserve the Killing form.

According to [Bourbaki:1972] a root system is a subset Σ of a real vector space V satisfying these conditions:

(a) Σ is finite, does not contain 0, and generates V;
(b) for each λ in Σ there exists $\lambda^\vee$ in the dual of V such that $\langle \lambda, \lambda^\vee \rangle = 2$ and the reflection

$$s_\lambda : v \longmapsto v - \langle v, \lambda^\vee \rangle \lambda$$

takes Σ to itself;
(c) for all λ in Σ, $\langle \lambda, \mu \rangle \in \mathbb{Z}$.

I have verified all of these for our Σ except (c). This follows from facts about finite-dimensional representations of SL_2. Therefore:

22.8. Proposition. *The set Σ is a root system in $\widehat{\mathfrak{h}}$.*

But what this really means, and what its consequences are, I do not consider here.

Part V. References

1. George Bergman,'The Diamond Lemma for ring theory', *Advances in Mathematics* **29** (1978), 178–218.

2. Nicolas Bourbaki, **Groupes et algèbres de Lie, Chapitres 4, 5, et 6**, Hermann, Paris, 1968.

3. Pierre Cartier, 'Théorème de conjugaison des algèbre de Cartan', exposé 15 (5 pages) in volume I of *Séminaire Sophus Lie* at lÉcole Normale Supérieure (1954–1955). This is available from NUMDAM.

4. Richard Courant, **Differential and integral calculus**, volume I (second edition), London and Glasgow: Blackie and Son Limited, 1937.

5. ——, **Differential and integral calculus**, volume II, London and Glasgow: Blackie and Son Limited, 1936.

6. Dan Freed, 'Homotopy invariance', course lecture available at

`https://web.ma.utexas.edu/users/dafr/M392C-2015/Notes/lecture2.pdf`

7. Peter Gabriel and Y. Nouazé, 'Idéaux premiers de l'algèbre enveloppante d'une algèbre de Lie nilpotente', *Jour. Alg.* **6** (1967), pp. 77–99.

8. Willem de Graaf, **Lie algebras: theory and algorithms**, Elsevier, 2000

9. Marco Gualtieri, course lecture notes at the University of Toronto, dated 2018:

`https://www.math.toronto.edu/mgualt/courses/18-1300/`

10. S. Helgason, **Differential geometry and symmetric spaces**, Academic Press, 1965.

11. Nathan Jacobson, **Lie algebras**, Wiley, 1962.

12. John McConnell, 'The intersection theorem for a class of non-commutative rings', *Proceedings of the London Mathematical Society* **17** (1967), 487–498.

13. Stephen Merrin, 'A strong constructive version of Engel's Theorem', *Communications in Algebra* **22** (1994), 1115–1125.

14. L. M. Milne-Thompson, **The calculus of finite differences**, A.M.S. Chelsea, originally published 1933.

15. John Milnor, **Topology from the differentiable veiwpoint**, University Press of Virginia, 1965.

16. Jet Nestruev, **Smooth manifolds and observables**, *Graduate Texts in Mathematics* **220**, Springer, 2003.

17. Isaac Newton, 'Methodus differentialis', in **The mathematical works of Isaac Newton**, volume 2, Johnson Reprint, 1967.

18. Georges de Rham, **Variétés différentiables**, Hermann, 1969.

19. Jean-Pierre Serre, **Lie algebras and Lie groups**, Benjamin, 1965.

20. ——, **Algèbres de Lie semi-simples complexes**, Benjamin, 1966.

21. Jacques Tits, 'Sur les constants de structure et the théorème d'existence des algèbres de Lie semi-simple', *Publications de l'I. H. E. S.* **31** (1966), 21–58.

22. Wikipedia:

`https://en.m.wikipedia.org/wiki/Exotic_sphere`