Material covered

- BC Hall section 1 + normal subgroups, quotients, semidirect products, isomorphism theorem:

  A *normal subgroup* of a group G is a subgroup N such that $ghg^{-1} \in N$ for all $g \in G$ and $h \in N$. We can also write this $gNg^{-1} \subset N$ for all $g \in G$. To denote that N is a normal subgroup of G, we write

  $$N \triangleleft G.$$

  In general, to denote that H is a subgroup of G, we write $H < G$. Normal subgroups allow us to define *quotients*. Given two sets of group elements $S_1 \subset G$ and $S_2 \subset G$ (not necessarily subgroups), the multiplication is defined as the element-wise product, $S_1 S_2 = \{s_1 s_2 : s_1 \in S_1,\ s_2 \in S_2\}$. The g-left coset of a set S is the set of group elements $gS = \{gs : s \in S\}$. If N is a normal subgroup of G, then the set of left cosets $G/N = \{gN : g \in G\}$ forms a group. Indeed, the product of 2 left cosets is a left coset:

  $$g_1 N g_2 N = g_1 g_2 g_2^{-1} N g_2 N = g_1 g_2 N$$

  where in the last line we used the fact that $g_2^{-1} N g_2 \subset N$ so that $g_2^{-1} N g_2 N \subset N$, and that N contains the identity id, so that $g_2^{-1} N g_2$ also contains id, hence $g_2^{-1} N g_2 N$ contains N. Associativity of the product is obvious. The inverse also exists in G/N:

  $$(gN)^{-1} = g^{-1} N$$

  as well as the identity element $\mathrm{id}N = N$, which can be checked by similar calculations.

  The (first) isomorphism theorem tells us that, given a group homomorphism $\phi : G \to H$, we have:

  $$\begin{aligned} \ker\phi &\triangleleft G \\ \mathrm{Im}\phi &< G \\ \mathrm{Im}\phi &\cong G/\ker\phi \end{aligned} \tag{1}$$

  (where $\cong$ means isomorphic).

  Finally, a concept related to that of normal subgroup is that of semi-direct product. Consider two groups G and N, and let $\psi : G \to Aut(N)$ be a group homorphism from G to the group of automorphisms of N. Let us denote by ϕ_g the automorphism that is image of g, that is, $\psi(g) = \phi_g$. The semi-direct product $\ltimes_\psi$ of the two groups G and N, associated to the homomorphism ψ, is a group structure on the cartesian product $G \times N$ of elements of G and N. For $(g, h) \in G \ltimes_\psi N$ and $(g', h') \in G \ltimes_\psi M$, the product is defined by

$$(g, h)(g', h') = (gg', h\phi_g(h')).$$

To check that this defines a group, we must check associativity,

$$(g, h)((g', h')(g'', h'')) = (g, h)(g'g'', h'\phi_{g'}(h'')) = (gg'g'', h\phi_g(h'\phi_{g'}(h'')))$$

where the second member in the last term can be written $h\phi_g(h')\phi_g(\phi_{g'}(h'')) = h\phi_g(h')\phi_{gg'}(h'')$. On the other hand,

$$((g, h)(g', h'))(g'', h'') = (gg', h\phi_g(h'))(g'', h'') = (gg'g'', h\phi_g(h')\phi_{gg'}(h'')$$

which is in agreement with the previous result. We must also check the presence of an identity, $\mathrm{id} = (\mathrm{id}, \mathrm{id})$ (obvious using $\phi_{\mathrm{id}} = \mathrm{id}$ and $\phi_g(\mathrm{id}) = \mathrm{id}$, recall the general properties of homomorphisms). Finally, we must check that an inverse exists. It is given by

$$(g, h)^{-1} = (g^{-1}, \phi_{g^{-1}}(h^{-1}))$$

because we have

$$(g, h)^{-1}(g, h) = (\mathrm{id}, \phi_{g^{-1}}(h^{-1})\phi_{g^{-1}}(h)) = (\mathrm{id}, \mathrm{id})$$

and

$$(g, h)(g, h)^{-1} = (\mathrm{id}, h\phi_g(\phi_{g^{-1}}(h^{-1}))) = (\mathrm{id}, h\phi_{\mathrm{id}}(h^{-1})) = (\mathrm{id}, \mathrm{id}).$$

It turns out that $G \cong \{(g, \mathrm{id}) : g \in G\}$ is a subgroup, and that $N \cong \{(\mathrm{id}, h) : h \in N\}$ is a *normal* subgroup. Indeed,

$$(g, h)^{-1}(\mathrm{id}, h')(g, h) = (g^{-1}, \phi_{g^{-1}}(h^{-1}h'))(g, h) = (\mathrm{id}, \phi_{g^{-1}}(h^{-1}h'h)).$$

A special case of the semi-direct product is the direct product, where $\phi_g = \mathrm{id}$ for all $g \in G$. In this case, both G and N are normal subgroups.

An example of a semi-direct product occurs when looking at, for instance, $O(n)$ with n even. We know that matrices in $O(n)$ have determinant ± 1. Any $A \in O(n)$ can be written $A = Ua$ where $U \in SO(n)$ and $a = \mathrm{diag}(-1, 1, 1, \ldots, 1)$ or $a = \mathbf{1}$, and this in a unique way (the choice of a is uniquely determined by the determinant of A, and then $U = Aa^{-1}$ is unique). Clearly, the set of a as described is isomorphic to $\mathbb{Z}_2$ under multiplications; we will simply denote this set by $\mathbb{Z}_2$. Hence, we have a bijective map $O(n) \leftrightarrow \{(a, U) : a \in \mathbb{Z}_2,\ u \in SO(n)\}$ by $A = uA \leftrightarrow (a, U)$. We may give the pairs (a, U) a semi-direct product structure

$$(a, U)(a', U') = (aa', U\phi_a(U'))$$

with $\phi_a(U') = aU'a^{-1}$. Clearly, ϕ_a is an automorphism of $SO(n)$, and the map $\psi : a \mapsto \phi_a$ is a homomorphism from Z_2 to $\mathrm{Aut}(SO(n))$. But this semidirect product is in agreement with the product in $O(n)$, because we have

$$AA' = UaU'a' = UaU'a^{-1}aa' = U\phi_a(U')aa'.$$

Hence, we have found that

$$O(n) \cong \mathbb{Z}_2 \ltimes_\psi SO(n).$$

So in particular, $SO(n)$ is a normal subgroup of $O(n)$, but $\mathbb{Z}_2$ is not.

- BC Hall chapter 2 complete
- BC Hall chapter 3 complete, but section 7 and 9 not in details
- BC Hall chapter 4 sections 1 and 2 only (please understand the principle of why we want the BCH formula). In section 2, the proof of formula 4.10, essential in the proof of the BCH formula, was greatly simplified:

  A crucial step in proving the BCH formula is the following formula, for $X(t)$ a smooth matrix-valued function of t:

$$\frac{d}{dt}e^{X(t)} = e^{X(t)}\left[\frac{\mathbf{1}-e^{-\mathrm{ad}X(t)}}{\mathrm{ad}X(t)}\left(\frac{dX(t)}{dt}\right)\right].$$

  We now prove it. The exponential can always be expanded in a power series, and differentiated term by term. Then, we have

$$\frac{d}{dt}e^{X(t)} = \sum_{j=0}^{\infty}\frac{1}{j!}\frac{d}{dt}(X(t))^j = \sum_{j=0}^{\infty}\frac{1}{j!}\sum_{k=0}^{j-1}X(t)^k\left(\frac{d}{dt}X(t)\right)X(t)^{j-k-1}.$$

  Let us consider the formal element Y such that $[Y, X(t)] = \frac{d}{dt}X(t)$. Since $\mathrm{ad}Y$ is a derivation, i.e. $[Y, X_1X_2] = [Y, X_1]X_2 + X_1[Y, X_2]$, we have

$$\frac{d}{dt}e^{X(t)} = [Y, e^{X(t)}] = Ye^{X(t)} - e^{X(t)}Y = e^{X(t)}(-Y + e^{-X(t)}Ye^{X(t)}) = e^{X(t)}(-1+e^{-\mathrm{ad}X(t)})(Y).$$

  We note that the power series of $-1+e^{-\mathrm{ad}X(t)}$ has no constant term, so we can write it as

$$-1+e^{-\mathrm{ad}X(t)} = \frac{-1+e^{-\mathrm{ad}X(t)}}{\mathrm{ad}X(t)}\mathrm{ad}X(t).$$

  Since $\mathrm{ad}X(t)(Y) = [X(t), Y] = -\frac{d}{dt}X(t)$, we then find

$$\frac{d}{dt}e^{X(t)} = e^{X(t)}\frac{-1+e^{-\mathrm{ad}X(t)}}{\mathrm{ad}X(t)}\left(-\frac{d}{dt}X(t)\right)$$

  which is the desired result.

From this, we get the BCH formula. We consider $e^{X(t)} = e^X e^{tY}$, and $e^{-X(t)} de^{X(t)}/dt = e^{-tY} e^{-X} e^X e^{tY} Y = Y$. But also,

$$e^{-X(t)} \frac{d}{dt} e^{X(t)} = \frac{\mathbf{1} - e^{-\mathrm{ad}X(t)}}{\mathrm{ad}X(t)} \left(\frac{dX(t)}{dt} \right).$$

Equating and inverting, we find

$$\frac{dX(t)}{dt} = \frac{\mathrm{ad}X(t)}{\mathbf{1} - e^{-\mathrm{ad}X(t)}}(Y)$$

hence

$$\log(e^X e^Y) = X + \int_0^1 dt \, \frac{\log(e^{\mathrm{ad}X} e^{t\mathrm{ad}Y})}{\mathbf{1} - (e^{\mathrm{ad}X} e^{t\mathrm{ad}Y})^{-1}}(Y)$$

where we used

$$e^{\mathrm{ad}X(t)} = \mathrm{Ad} e^{X(t)} = \mathrm{Ad}(e^X)\mathrm{Ad}(e^{tY}) = e^{\mathrm{ad}X} e^{t\mathrm{ad}Y}.$$

- BC Hall, chapter 5, first page.
- JE Humphreys: we always restrict the field $\mathbb{F}$ to that of the complex numbers $\mathbb{C}$. Sections: 1 (but just first 2 paragraphs of 1.2), 2, 3.1, 5.1 (except restrictions of killing form), 5.2, 5.3, 6.1, 6.2, 6.3 (but don't worry about the proof), 7, 8.1, 8.2 (but again, don't worry about the proof).

  Here are some of the main definitions and concepts:

  A Lie algebra L is a vector space with a product $[\cdot\cdot]$, with the properties:

$$\begin{aligned} \text{bilinearity} &: \quad [(ax+by)z] = a[xz] + b[yz], \quad [x(ay+bz)] = a[xy] + b[xz] \\ \text{antisymmetry} &: \quad [xy] = -[yx] \\ \text{Jacobi identity} &: \quad [x[yz]] + [y[zx]] + [z[xy]] = 0. \end{aligned} \tag{2}$$

  The adjoint representation ad maps L to $gl(L)$, with the action given by

$$\mathrm{ad}x(y) = [xy].$$

  The adjoint map is a homomorphism of Lie algebras, with $gl(L)$ seen as a Lie algebra where the bracket is given by the commutator of matrices $[A, B] = AB - BA$. Indeed, the Jacobi identity gives

$$[\mathrm{ad}x, \mathrm{ad}y](z) = [x[yz]] - [y[xz]] = [[xy]z] = \mathrm{ad}[xy](z)$$

  hence $[\mathrm{ad}x, \mathrm{ad}y] = \mathrm{ad}[xy]$. This is why it is called the adjoint representation. Also, for any $x \in L$, the operator $\mathrm{ad}x$ is a derivation on L. In general, a derivation $D \in gl(L)$ has the property

$$D([xy]) = [D(x)y] + [xD(y)].$$

  We indeed have, again thanks to the Jacobi identity,

$$\mathrm{ad}x([yz]) = [x[yz]] = [[xy]z] + [y[xz]] = [\mathrm{ad}x(y)z] + [y\mathrm{ad}x(z)].$$

A *subalgebra* I of L is a subspace of L which is closed under brackets, $[xy] \in I \; \forall x \in I, \; y \in I$.

An *ideal* I of L is a subspace of L with the property $[xy] \in I \; \forall x \in I, \; y \in L$. In particular, an ideal is a subalgebra (but the opposite is not true in general).

A *proper ideal* of a Lie algebra L is an ideal other than $\{0\}$ and than L itself.

The *derived subalgebra* of a Lie algebra L is a subalgebra of L obtained by taking linear combinations of all possilble brackets of elements in L. That is, $[LL] = \text{span}\{[xy] : x, y \in L\}$.

A Lie algebra L is *abelian* if $[LL] = \{0\}$, that is, if all elements of L commute with each other. Otherwise, it is *nonabelian*.

A *simple* Lie algebra is a Lie algebra that does not have proper ideal, and that is nonabelian.

If L is simple, then $L = [LL]$ (easy to prove)

The *derived series* associated to a Lie algebra L is the series of Lie algebras $L^0, \; L^1, \; L^2, \ldots$ obtained by setting $L^0 = L$ and $L^{n+1} = [L^n L^n]$ for $n = 1, 2, 3, \ldots$.

A Lie algebra L is *solvable* if its derived series terminate, i.e. if $L^n = 0$ for some n.

A *semisimple* Lie algebra is a Lie algebra which does not have solvable ideals other than $\{0\}$.

The *Killing form* is a bilinear form on a Lie algebra L given by

$$\kappa(x, y) = \text{Tr}(\text{ad}x \, \text{ad}y).$$

The Killing form is symmetric, $\kappa(x, y) = \kappa(y, x)$, and *associative*, $\kappa([xz], y) = \kappa(x, [zy])$, thanks to the homomorphism property of ad and to cyclicity of the trace (easy to prove).

A bilinear form β on L is *degenerate* if $\exists \; x : \beta(x, y) = 0 \; \forall \; y \in L$. Otherwise, it is *nondegenerate*.

Equivalent characterisations of a semisimple Lie algebra L are:

- L does not have abelian ideal (easy to prove)
- The Killing form on L is *nondegenerate* (harder to prove)
- $L = \coprod_j I_j$ (direct sum of I_j) where I_j are simple ideals (medium-easy to prove once the second point is proven).

The theorem of Weyl (hard to prove) says that *any finite-dimensional representaion of a semisimple Lie algebra is completely reducible*, i.e. is a direct sum of irreducible representations.

If L is semisimple, then any derivation D on L is of the form $D = \text{ad}z$ for some $z \in L$ (hard to prove; but see question 6.4).

By *Shur's lemma*, if V is a vector space on which we have an irreducible representation ϕ of a Lie algebra L, then any endomorphism of V that commutes with $\phi(x)$ for all $x \in L$ must be proportional to the identity endomorphism, $\mathbf{1}$.

Given a representation ϕ, we may construct the trace form $\beta_\phi(x, y) = \mathrm{Tr}(\phi(x)\phi(y))$. (The Killing form is β_{ad}.) If L is semisimple and the representation is faithful, then the trace form is nondegenerate (hard to prove; similar to nondegeneracy of Killing form).

Given a basis $\{x_i\}$ of a semisimple Lie algebra L, we may construct the dual basis $\{y_i\}$ under β_ϕ. It is the basis satisfying $\beta_\phi(x_i, y_j) = \delta_{ij}$. The *Casimir element* of a representation ϕ acting on the vector space V is the endomorphism of V given by

$$C_\phi = \sum_i \phi(x_i)\phi(y_i).$$

It has the property that it *commutes with all* $\phi(x)$ *for* $x \in L$. That is, $[C_\phi, \phi(x)] = 0$ for all $x \in L$. In particular, if ϕ is irreducible, then $C_\phi = \mathrm{dim}L/\mathrm{dim}V\,\mathbf{1}$.

A *toral subalgebra* of a Lie algebra L is a subalgebra I of L such that all elements $x \in I$ are ad-diagonalisable, i.e. $\mathrm{ad}x$ is a diagonalisable matrix.

One can prove that a toral subalgebra is *abelian* (medium-easy to prove).

A *Cartan subalgebra* is a maximal toral subalgebra.

A *root* is a non-zero linear functional $\alpha : H \to \mathbb{C}$ on a Cartan subalgebra H which reproduces eigenvalues of $\mathrm{ad}h$ for all $h \in H$. That is, if v is a common eigenvector of all $h \in H$, then $\mathrm{ad}h(v) = [hv] = \alpha(h)v$. The set of roots is denoted Φ.

The eigenspace in L corresponding to a linear functional α is denoted L_α. That is,

$$L_\alpha = \{x \in L : [hx] = \alpha(h)x \;\forall\; h \in H\}.$$

If α does not reproduce the eigenvalues of h, then $L_\alpha = \{0\}$. α is a root if and only if $L_\alpha \neq \{0\}$ and $\alpha \neq 0$. If $\alpha = 0$, then L_0 is the set of elements of L that commute with all of H (the *centraliser* of H in L). A theorem says that $L_0 = H$ (hard to prove).

The *root space decomposition* of a Lie algebra L is the decomposition into eigenspaces of H. That is,

$$L = H \oplus \coprod_{\alpha\in\Phi} L_\alpha.$$

(see additional exercise 1).

Solutions to exercises

(note: I tend to denote by id the abstract identity in a group, as well as identity maps, and by $\mathbf{1}$ the identity matrix in any dimension, when no confusion is possible).

BC Hall, Chapter 1

1. The center of a group G is $Z(G) = \{g \in G \,|\, gh = hg \,\forall\, h \in G\}$. Certainly, the identity $\mathrm{id} \in Z(G)$. Also, if $g \in Z(G)$, then $gh = hg \Rightarrow h = g^{-1}hg \Rightarrow hg^{-1} = g^{-1}h$ for all $h \in G$, hence also $g^{-1} \in Z(G)$. Finally, if $g_1, g_2 \in Z(G)$, then $g_1 h = hg_1 \Rightarrow g_1 h g_2 = hg_1 g_2 \Rightarrow g_1 g_2 h = hg_1 g_2$ for all $h \in G$, where in the last step we used $g_2 h = hg_2$. Hence, also $g_1 g_2 \in Z(G)$. These are the three conditions in order to have a subgroup. (Note: here and below we use without stating it the associativity of groups).

2. (a) H **is not** a subgroup, because it does not contain the identity element 0 (here, the group is under addition).

   (b) H **is** a subgroup: it contains the identity 0, every integer of the form $3n$ (for $n \in \mathbb{Z}$) has an inverse $-3n$, and $3n + 3m = 3(n + m)$.

   (c) H **is not** a subgroup, there are elements that do not have an inverse: any matrix of determinant, say, 2, is in H, but has an inverse of determinant $1/2$, outside of H.

   (d) H **is** a subgroup: it contains the identity $\mathbf{1}$, the product of any two matrices with integer entries is a matrix with integer entries, and the inverse of a matrix is obtained by evaluating the determinants of submatrices (which are always integer numbers), and dividing by the determinant of the matrix (which is 1 because this is a subgroup of $SL(n;\mathbb{R})$).

   (e) H **is** a subgroup: it contains the identity $\mathbf{1}$, the product of any two matrices with rational entries is a matrix with rational entries, and the inverse of a matrix is obtained via rational functions of the entries – determinant of submatrices divided by the determinant of the matrix – since determinants are polynomials in the matrix entries.

   (f) H **is not** a subgroup: no element in H, except for the identity 0, possess an inverse, since in $\mathbb{Z}_9$ inverses of even numbers are odd numbers (e.g. the inverse of 8 is 1 since $1 + 8 = 9 = 0 \bmod 9$, but 1 is not in H).

3. First, we know that $g^{-1}g = gg^{-1} = \mathrm{id}$. Hence, $(g^{-1})^{-1}g^{-1} = \mathrm{id} \Rightarrow (g^{-1})^{-1} = g$ by right-multiplication by g. Second, $g^{-1}g = \mathrm{id} \Rightarrow h^{-1}g^{-1}gh = \mathrm{id}$ (*) by left-multiplication by h^{-1} and right-multiplication by h and using $h^{-1}h = \mathrm{id}$. Then, right-multiplication of (*) by $(gh)^{-1}$ gives $h^{-1}g^{-1}gh(gh)^{-1} = (gh)^{-1}$, hence using $gh(gh)^{-1} = \mathrm{id}$, we find $h^{-1}g^{-1} = (gh)^{-1}$. Finally, since $\mathrm{id}\mathrm{id} = \mathrm{id}$, we immediately see that $\mathrm{id}^{-1} = \mathrm{id}$ (we use uniqueness of the inverse, which satisfies $g^{-1}g = gg^{-1} = \mathrm{id}$).

4. If there is an isomorphism ϕ from G to H, then we have a bijective map $\phi : G \to H$ that preserves the group action of G, $\phi(g_1 g_2) = \phi(g_1)\phi(g_2)$ (*) for all $g_1, g_2 \in G$. We need to show that there is a bijective map $\psi : H \to G$ that preserves the group action of H, $\psi(h_1 h_2) = \psi(h_1)\psi(h_2)$ for all $h_1, h_2 \in H$. Since ϕ is bijective, then to every element $h \in H$ corresponds at least one element g of G (surjectivity), and in fact exactly one element g (injectivity). Hence, to every $h \in H$ we can associated a unique $g \in G$: this defines the inverse map ϕ^{-1}. The map ϕ^{-1} is bijective: surjective, every $g \in G$ is in the image of

ϕ^{-1}, since the map ϕ associates a $h \in H$ for every $g \in G$; and injective, there is only one h mapping to any given g by ϕ^{-1}, since the map g associates only one $h \in H$ for every $g \in G$. We may then take $\psi = \phi^{-1}$, and all we need to show is that it preserves the group action of H. Note that $\psi(\phi(g)) = g$ by construction. Hence, from (*) we have $g_1 g_2 = \psi(\phi(g_1)\phi(g_2))$. Now for any $h_1, h_2 \in H$, consider $g_1 = \psi(h_1)$ and $g_2 = \psi(h_2)$. Then we find $\psi(h_1 h_2) = \psi(h_1)\psi(h_2)$. This completes the proof.

Note here that when we say that we have *map* from G to H, we already say many things: to every $g \in G$ there is exactly one $h \in H$. The condition of bijectivity only gives us the opposite: to every $h \in H$ there is exactly one $g \in H$.

5. The set of positive numbers $\mathbb{R}_+$ contains the identity 1, every $x \in \mathbb{R}_+$ has an inverse $1/x \in \mathbb{R}_+$, and for $x_1, x_2 \in \mathbb{R}_+$ we have $x_1 x_2 \in \mathbb{R}_+$: hence $\mathbb{R}_+$ is a subgroup of $\mathbb{R}^*$ (the group of the non-zero reals under multiplication). To show isomorphism to R (the group of the reals under addition), we contruct an explicit bijective map that preserves the group action: take ϕ to be $\mathbb{R}_+ \to \mathbb{R} : x \mapsto \phi(x) = \log(x)$. Clearly, $\log(x_1 x_2) = \log(x_1) + \log(x_2)$ for all $x_1, x_2 \in \mathbb{R}^+$, so this preserves group action. It is also bijective: the set $\{\log(x) : x \in \mathbb{R}_+\}$ is the set $\mathbb{R}$, and if $\log(x_1) = \log(x_2)$, then $x_1 = x_2$.

6. Consider the set of bijective maps $Aut(G) = \{\phi : G \to G \,|\, \phi(gg') = \phi(g)\phi(g')\,\forall\, g, g' \in G\}$ (this is the set of authomorphisms: isomorphisms from G to G). First, the composition of maps $\phi_1, \phi_2 \in Aut(G)$ gives a map $\phi_1 \circ \phi_2$ that is in $Aut(G)$: it is still bijective from G to G (there is a unique pre-image $g' \in G$ for any $g \in G$ under ϕ_1 by bijectivity, and then a unique pre-image $g'' \in G$ for that g' under ϕ_2), and $\phi_1(\phi_2(gg')) = \phi_1(\phi_2(g)\phi_2(g')) = \phi_1(\phi_2(g))\phi_1(\phi_2(g'))$. Second, by construction we have associativity: $((\phi_1 \circ \phi_2) \circ \phi_3)(g) = (\phi_1 \circ \phi_2)(\phi_3(g)) = \phi_1(\phi_2(\phi_3(g)))$ and $(\phi_1 \circ (\phi_2 \circ \phi_3))(g) = \phi_1((\phi_2 \circ \phi_3)(g)) = \phi_1(\phi_2(\phi_3(g)))$. Third, the identity map $\mathrm{id} : g \mapsto g$ is in $Aut(G)$, and has the property $\mathrm{id} \circ \phi = \phi \circ \mathrm{id} = \phi$ for any $\phi \in Aut(G)$, so it is a candidate for the identity of the group. Finally, since $\phi \in Aut(G)$ is an isomorphism, there is a unique inverse ϕ^{-1} such that $\phi \circ \phi^{-1} = \phi^{-1} \circ \phi = \mathrm{id}$ (see the answer to question 4).

7. Certainly, ϕ_g is bijective: for any $h \in G$, we have that $\phi_g(g^{-1}hg) = h$, so ϕ_g is surjective. Also, if $\phi_g(h) = \phi_g(h')$, then $ghg^{-1} = gh'g^{-1} \Rightarrow hg^{-1} = h'g^{-1} \Rightarrow h = h'$, hence ϕ_g is injective. Also, $\phi_g(hh') = ghh'g^{-1} = ghg^{-1}gh'g^{-1} = \phi_g(h)\phi_g(h')$. Hence, ϕ_g is an automorphism.

   In order to have a homomorphism, we only need to check that the group multiplication of G gives rise to the group multiplication of $Aut(G)$. We have $\phi_{gg'}(h) = gg'h(gg')^{-1} = gg'h(g')^{-1}g^{-1} = \phi_g(\phi_{g'}(h))$ for all $h \in G$, hence $\phi_{gg'} = \phi_g \circ \phi_{g'}$.

   Finally, the kernel is the set of $g \in G$ such that $\phi_g = \mathrm{id}$. This is the set $\{g \in G \,|\, ghg^{-1} = h \,\forall\, h \in G\}$, which is $\{g \in G \,|\, gh = hg \,\forall\, h \in G\} = Z(G)$.

8. Consider, for instance,
$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$
They give
$$AB = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \neq BA = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

9. Same question as 1 !!!

10. We have
$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$$
and
$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}.$$

11. Take for instance $g(n) = 2n$ and $f(n) = [n/2]$ (where $[x]$ is the integer part of x, which can be taken as the integer number $m \leq x$ that minimises $x - m$; we can generalise to $[x]_N = N[x/N]$ for any positive integer N, giving the number $m \leq x$ that is a multiple of N and that minimises $x - m$). We find $(f \circ g)(n) = [(2n)/2] = [n] = n$ for all $n \in \mathbb{Z}$, hence $f \circ g = \text{id}$. On the other hand, $(g \circ f)(n) = 2[n/2] = [n]_2$, with $[n]_2 = n$ for n even and $[n]_2 = n - 1$ for n odd. Hence, $g \circ f \neq \text{id}$. Note that g is not onto (surjective), and that f is not one-to-one (injective).

12. Consider $a_1 + a_2 \mod n$. Suppose we can write $a_j = k_j n + b_j$ ($j = 1, 2$) for k_j integers and $b_j \in \{0, \ldots, n-1\}$. Then, $a_1 + a_2 = (k_1 + k_2)n + b_1 + b_2$. If $b_1 + b_2 < n$, then we have found $a_1 + a_2 \mod n = b_1 + b_2$. On the other hand, if $b_1 + b_2 \geq n$, then we can write $a_1 + a_2 = (k_1 + k_2 + 1)n + (b_1 + b_2 - n)$, and we know that $b_1 + b_2 - n \in \{0, \ldots, n-1\}$. Hence, in this case $a_1 + a_2 \mod n = b_1 + b_2 - n$. But likewise, if we do $b_1 + b_2$ in the group $\mathbb{Z}_n$, we obtain $b_1 + b_2$ is $b_1 + b_2 < n$, and $b_1 + b_2 - n$ if $b_1 + b_2 \geq n$. Hence, since $b_j = a_j \mod n$, we have found that $a_1 + a_2 \mod n = (a_1 \mod n) + (a_2 \mod n)$ where on the left-hand side, addition is in $\mathbb{Z}$, and on the right-hand side, it is in $\mathbb{Z}_n$. Hence, we have a homomorphism from $\mathbb{Z}$ to $\mathbb{Z}_n$.

13. If G is commutative, then $g_1 g_2 = g_2 g_1$ for any $g_1, g_2 \in G$. If H is a subgroup of G, then $ghg^{-1} = hgg^{-1} = h$ for any $g \in G$, hence H is normal.

    Certainly, if $h, g \in G$, then $ghg^{-1} \in G$, so G is normal. Also, $g\text{id}g^{-1} = gg^{-1} = \text{id}$, so the subgroup $\{\text{id}\}$ is normal.

    If $h \in Z(G)$, then $ghg^{-1} = hgg^{-1} = h$ for any $g \in G$. Hence, the group $Z(G)$ (we already showed that it is a group in question 1) is normal.

    If $\phi : G \to H$, then $\ker \phi = \{g \in G \,|\, \phi(g) = \text{id}\}$. First, we show that $\ker \phi$ is a subgroup. Consider $g, g' \in \ker \phi$, then $\phi(gg') = \phi(g)\phi(g') = \text{id}$, hence $gg' \in \ker \phi$. Also, for any homomorphism we have $\phi(\mathrm{id}) = \mathrm{id}$, hence, $\mathrm{id} \in \ker\phi$. Finally, if $g \in \ker\phi$, then $\phi(g^{-1}) = \phi(g)^{-1} = \mathrm{id}$ so that $g^{-1} \in \ker\phi$. Hence, $\ker\phi$ is a subgroup. Then, for $h \in \ker\phi$, we have $\phi(ghg^{-1}) = \phi(g)\phi(h)\phi(g^{-1}) = \phi(g)\mathrm{id}\phi(g^{-1}) = \phi(gg^{-1}) = \phi(\mathrm{id}) = \mathrm{id}$ so that $ghg^{-1} \in \ker\phi$ for any $g \in G$. Hence, $\ker\phi$ is a normal subgroup.

$SL(n;\mathbb{R})$ is the subgroup $\ker\det$, where det is the homomorphism $GL(n;\mathbb{R}) \to \mathbb{C}$ given by the determinant of the matrix (homomorphism since $\det(gg') = \det(g)\det(g')$). Hence, $SL(n;\mathbb{R})$ is a normal subgroup by the proof above.

BC Hall, Chapter 2

1. A subset of a topological space is dense in the space if its closure is the whole space. To find the closure of $K = \{e^{2\pi i n a} : n \in \mathbb{Z}\}$ for a irrational, we look at the complement in S^1, and take the union of all open sets included in the complement (the closure is the complement of that). Any open set of S^1 (with the standard topology) contains an open arc. But all open arcs contain at least one point of K (since given any two points in K, we find always more points in K by repeating the arc they form, and eventually some point falling in the original arc, so we will find that there are two points forming an arc less than half the length). Hence the complement of K does not contain any interval, so the closure of K is S^1. Now we can write $G = \left\{ \begin{pmatrix} e^{it} & 0 \\ 0 & e^{iatn} \end{pmatrix} \mid t \in [0, 2\pi],\, n \in \mathbb{Z} \right\}$, and by the previous argument we find this is dense in the space of diagonal matrices with elements in S^1 on the diagonal. Hence, $\overline{G}$ is as described in the question.

2. If A preserves the inner product, then $\langle Ax, Ay\rangle = \langle x, y\rangle$ for all x, y, so that $\sum_{i,j,k} A_{i,j} x_j A_{i,k} y_k = \sum_i x_i, y_i$. Making x and y running independently through basis elements, we find $A_{i,j}A_{i,k} = \delta_{j,k}$, which is the statement that the column vectors of A are orthonormal. The opposite direction of the argument works in a similar fashion, by writing x and y as arbitrary linear combinations of basis elements.

   We have $\langle Bx, y\rangle = \sum_{i,j} B_{i,j} x_j y_i = \sum_{i,j} x_j (B^T)_{j,i} y_i = \langle x, B^T y\rangle$. Hence, if $\langle Ax, Ay\rangle = \langle x, y\rangle$ for all x, y, then $\langle x, A^T A y\rangle = \langle x, y\rangle$ for all x, y. Since the inner product is non-degenerate, this implies that $A^T A = \mathbf{1}$.

3. Following the solution to 2, we see that $\langle Bx, y\rangle = \langle x, B^\dagger y\rangle$. Hence, $\langle Ax, Ay\rangle = \langle x, A^\dagger A y\rangle$. Suppose $A^\dagger A = \mathbf{1}$. We immediately conclude that $\langle Ax, Ay\rangle = \langle x, y\rangle$ for all x, y. On the other hand, suppose that $\langle Ax, Ay\rangle = \langle x, y\rangle$ for all x, y. Then we have $\langle x, A^\dagger A y\rangle = \langle x, y\rangle$, and since the inner product is non-degenerate, we deduce $A^\dagger A = \mathbf{1}$.

4. We immediately have $[x, y]_{n,k} = \sum_{i=1}^{n+k} g_{ii} x_i y_i = \sum_{i,j=1}^{n+k} x_i g_{i,j} y_j = \langle x, gy\rangle$. Hence, if $[Ax, Ay]_{n,k} = [x, y]_{n,k}$ for all x, y, then $\langle x, A^T g A y\rangle = \langle x, gy\rangle$ for all x, y so that $A^T g A = g$. The opposite direction is straightforward.

Clearly $O(n,k)$ and $SO(n,k)$ are subsets of $GL(n+k,\mathbb{R})$, since they are sets of matrices with non-zero determinant. Moreover, they are closed under matrix multiplication, they contain the identity, and they contain the inverse of every element. Hence they are subgroups. In order to show that they are matrix Lie groups, we only need to show that they are closed in $GL(n+k,\mathbb{R})$. But if a sequence of matrices $\{A_m : m = 0,1,2,3,\ldots\}$ in $O(n,k)$ or $SO(n,k)$ converges to a matrix $A = \lim_{m\to\infty} A_m$, then the sequence of determinants $\{\det(A_m) : m = 0,1,2,3,\ldots\}$ converges on the reals, since the determinant is a continuous function (from $\mathbb{R}^{n^2}$ to $\mathbb{R}$ with the usual topologies). But since $\det(A_m) \in \{-1,1\}$ for all m (in the $O(n,k)$ case) or $\det(A_m) = 1$ for all n (in the $SO(n,k)$ case), the fact that the sequence of determinant converges means that $\det(A) \in \{-1,1\}$ (in the $O(n,k)$ case) or $\det(A) = 1$ (in the $SO(n,k)$ case), so that the limit matrix A is in $O(n,k)$ or $SO(n,k)$. Hence, the subgroup is closed.

5. Similar to the above.

6. Using $\cos^2\theta + \sin^2\theta = 1$ we immediately find

$$A^T A = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

which shows that $A \in SO(2)$. Let us denote $A = A(\theta)$. Similarly, using $\cos\theta\sin\phi + \sin\theta\cos\phi = \sin(\theta+\phi)$ and $\cos\theta\cos\phi - \sin\theta\sin\phi = \cos(\theta+\phi)$ we find $A(\theta)A(\phi) = A(\theta+\phi)$. If some matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is in $O(2)$, then we have three conditions: $a^2 + c^2 = 1$, $b^2 + d^2 = 1$ and $ab + cd = 0$. The set of all solutions (over the reals) to the first equation is the set $(a,c) \in \{(\cos\theta, \sin\theta) : \theta \in [0, 2\pi)\}$ (because we know that $-1 \le a \le 1$, so we can write $a = \cos\theta$ for some $\theta \in [0,\pi)$, and we find $c = \pm\sin\theta$, where the sign can be taken care of by extending the range of θ to $[0, 2\pi)$); similarly for the second equation (where we will use the parameter ϕ instead of θ). Hence the third equation becomes $\cos\theta\cos\phi + \sin\theta\sin\phi = 0$, hence $\cos(\theta - \phi) = 0$. The set of all solutions is given by the set of solutions to the equations $\theta - \phi = \pm\pi/2, \pm 3\pi/2$. For any given θ, there are exactly two solutions for ϕ, which are $\theta - 3\pi/2$, $\theta - \pi/2$; or $\theta - \pi/2$, $\theta + \pi/2$; or $\theta + \pi/2$, $\theta + 3\pi/2$; whichever keep $\phi \in [0, 2\pi)$. In all cases, the two solutions for (b,d) are $(-\sin\theta, \cos\theta)$, $(\sin\theta, -\cos\theta)$.

7. We need to check that $A^T \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. We have:

$$\begin{pmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{pmatrix} =$$
$$= \begin{pmatrix} \cosh t & -\sinh t \\ \sinh t & -\cosh t \end{pmatrix} \begin{pmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

using $\cosh^2 t - \sinh^2 t = 1$. For the multiplication rule, the argument is like that of the previous question.

For some matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ to be in $O(1;1)$, there are three conditions (over the reals): $a^2 - c^2 = 1$, $b^2 - d^2 = -1$ and $ab - cd = 0$. Consider the first equation. We have $a \geq 1$ or $a \leq 1$ so we can always write $a = \eta \cosh t$ for $\eta \in \{+1, -1\}$ and $t \in [0, \infty)$. Then, we find $c = \pm \sinh t$, and the sign can be taken care of by extending the range of t to $\mathbb{R}$. Hence, the set of solutions to the first equation is $(a, c) = \{(\eta \cosh t, \sinh t) : t \in \mathbb{R}, \eta \in \{+1, -1\}\}$. likewise, for the set of solutions to the second equation we have $(b, d) = \{(\sinh r, \epsilon \cosh r) : r \in \mathbb{R}, \epsilon \in \{+1, -1\}\}$. The last equation is then $\eta \cosh t \sinh r - \epsilon \sinh t \cosh r = 0 \Rightarrow \sinh(r - \eta\epsilon t) = 0$. The unique solution is $r = \eta\epsilon t$, and we find $(b, d) = (\eta\epsilon \sinh t, \epsilon \cosh t)$. This gives the four possibilities displayed in the question.

8. For the group $SU(2)$, the three conditions on the matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ coming from $A^\dagger A = \mathbf{1}$ are $|a|^2 + |c|^2 = 1$, $|b|^2 + |d|^2 = 1$ and $a^* b + c^* d = 0$, and the condition saying that A has unit determinant is $ad - bc = 1$. Given a and c satisfying the first equation, we can solve the third for b, giving $b = -c^* d / a^*$. The unit determinant condition then is $ad + |c|^2 d / a^* = 1$ so that $d = a^*$, which implies $b = -c^*$.

9. A matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ in $Sp(1;\mathbb{R})$ (or $Sp(1;\mathbb{C})$) satisfies $A^T J A = J$ for $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. We have $A^T J A = \begin{pmatrix} 0 & ad - bc \\ bc - ad & 0 \end{pmatrix}$, so that we only have one condition, $ad - bc = \det(A) = 1$. Hence, $Sp(1;\mathbb{R}) = SL(2;\mathbb{R})$ and $Sp(1;\mathbb{C}) = SL(2;\mathbb{C})$. Since $Sp(1) = Sp(1;\mathbb{C}) \cap U(2)$, we have that $Sp(1)$ is the set of unitary 2-by-2 matrices ($U(2)$), with the only additional condition that they have determinant 1 (since $Sp(1;\mathbb{R}) = SL(2;\mathbb{R})$). Hence, this is $SU(2)$.

10. The center $Z(H)$ of a group H is the set of all group elements that commute with the whole group: $Z(H) = \{h \,|\, hg = gh \,\forall\, g \in H \ : \ h \in H\}$. This is certainly a subgroup, which is abelian and normal. With

$$h = \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix}$$

and h' of a similar form, we find

$$hh' = \begin{pmatrix} 1 & a' + a & b' + ac' + b \\ 0 & 1 & c' + c \\ 0 & 0 & 1 \end{pmatrix}.$$

For h to commute with everything, we need $ac' = a'c$ for all a' and c', hence $a = c = 0$. This is isomorphic to the abelian group $\mathbb{R}$. The quotient group $H/Z(H)$ is the group of left-cosets, $\{hZ(H) : h \in H\}$. This is the group of "matrices"

$$q = \begin{pmatrix} 1 & a & \mathbb{R} \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix}.$$

For $q, q' \in H/Z(H)$, the product is

$$qq' = \begin{pmatrix} 1 & a' + a & \mathbb{R} \\ 0 & 1 & c' + c \\ 0 & 0 & 1 \end{pmatrix}.$$

This is clearly abelian, $qq' = q'q$. In fact we see that $H/\mathbb{Z}(H) \cong \mathbb{R} \times \mathbb{R}$.

11. We essentially follow the suggested solution. First, connectedness is clear in the case $n = 2$ by question 6: the continous path from $A(0) = \mathbf{1}$ to $A(\theta)$ can be chosen to be simply $A(t/\theta) : t \in [0, 1]$. Second, we must realise that $SO(n)$ acts faithfully on $\mathbb{R}^n$: any non-identity element of $SO(n)$ changes some vector of $\mathbb{R}^n$ (by definition of the group itself). Similarly, it is clear that any two unit vectors in $\mathbb{R}^2$ can be connected by a $SO(2)$ transformation. Let us consider some non-identity element $A \in SO(n)$, and a (unit) vector w which is affected by it: $Aw = v \neq w$. Let us look at the subspace of $\mathbb{R}^n$ spanned by w and v. This a two-dimensional subspace with orthonormal basis w and $(v - (w \cdot v)w)/|v - (w \cdot v)w|$, for instance. Let us take $B(t)$ in the subgroup of $SO(n)$ isomorphic to $SO(2)$ and corresponding to rotations in the $w - v$ plane. By the previous argument on $SO(2)$, there is a continuous path $B(t) : t \in [0, 1]$ such that $B(0) = \mathbf{1}$ and $B(1)v = w$. Hence, let us consider $A(t) = B(t)A$; this is such that $A(0) = A$ and that $A(1)$ preserves w. Since $A(1)$ preserves w, it must be element of the subgroup of $SO(n)$ that preserves w (it is easy to check that this is a Lie subgroup); this subgroup acts faithfully on the hyperplane $\{x \,|\, x \cdot w = 0 \,:\, x \in \mathbb{R}^n\}$ and preserves lengths there, hence it is isomorphic to $SO(n-1)$. Therefore, we have shown that any element of $SO(n)$ can be connected continuously to an element of $SO(n-1)$, and by induction to an element of $SO(2)$, which we showed can be connected to the identity.

12. Here we must simply use the fact that any symmetric, positive-definite matrix has a unique symmetric, positive-definite square root. Certainly, for $A \in SL(n; \mathbb{R})$, the matrix $A^T A$ is symmetric and positive definite (the latter because $\langle x, A^T Ax\rangle = \langle Ax, Ax\rangle > 0$ by positive-definiteness of the inner-product). Hence, we may write $A^T A = H^2$ where H is symmetric and positive-definite (with determinant 1). Let us consider the matrix $R = AH^{-1}$. We have $R^T R = H^{-1}A^T AH^{-1} = H^{-1}H^2H^{-1} = \mathbf{1}$. Hence, this matrix is in $SO(n)$. This shows that it is possible to write $A = RH$. To show uniqueness, suppose we may write $A = R'H'$ with R' and H' different matrices (with the properties stated). Then, $A^T A = (H')^2$, and by uniqueness of the square root, we must have $H' = H$, which implies that we have $R' = R$.

13. Consider $A \in SL(n;\mathbb{R})$, and its polar decomposition $A = RH$. Since $SO(n)$ is connected, there is a path $R(t) : t \in [0,1]$ such that $R(0) = R$ and $R(1) = \mathbf{1}$, so we may form the path $A(t) = R(t)H$, that reaches $A(1) = H$. Now since H is real and symmetric, we can diagonalise with a real orthogonal matrix T, that is, $H = T^{-1}DT$ with D diagonal with determinant 1. It is then simple to make a path from D to $\mathbf{1}$ (while of course keeping the determinant 1): bring the elements of $D_{ii} : i \in \{0, \ldots, n-1\}$ to 1 in a continuous way without crossing 0 (this can be achieved, because $D_{ii} > 0 \,\forall\, i$ by positive-definiteness of D) and write $D_{nn} = 1/\prod_{i=1}^{n-1} D_{ii}$. This makes a path $D(t) : t \in [1,2]$ with $D(1) = D$ and $D(2) = \mathbf{1}$, and write $H(t) = T^{-1}D(t)T$ so that $A(t) : t \in [1,2]$ is continuous with $A(1) = H$ and $A(2) = \mathbf{1}$. Thus, we have a continuous path $A(t) : t \in [0,2]$ from $A(0) = A$ to $A(2) = \mathbf{1}$.

14. Consider $A \in GL(n;\mathbb{R})^+$. We may consider the continuous path $A(t) = f(t)A : t \in [0,1]$ with $f(t) \in \mathbb{R}^+$ a continuous function, and $f(0) = 1$, $f(1) = 1/(\det(A))^{1/n}$. Then, $A(0) = A$ and $A(1) \in SL(n;\mathbb{R})$, and we can use the results of the previous question to make a continuous path from $A(1)$ to $\mathbf{1}$.

15. Done in class.

16. Consider a Lie group homomorphism ϕ from $\mathbb{R}$ to S^1. We have $\phi(a+b) = \phi(a)\phi(b)$ by the homomorphism property. In particular $\phi(ka) = \phi(a)^k$ and $\phi(1) = \phi(1/n)^n$ for all $k = 1,2,3,\ldots$ and $n = 1,2,3,\ldots$. This gives $\phi(k/n) = \phi(1/n)^k = \phi(1)^{k/n}$, hence $\phi(r) = \phi(1)^r$ for all rational numbers r. For any real number x, there is a sequence of rational numbers $\{r_j : j = 1,2,3,\ldots\}$ so that $\lim_{j\to\infty} r_j = x$. Then, by continuity of the homomorphism ϕ (since it is a Lie group homomorphism, it is continuous), we have that $\lim_{j\to\infty} \phi(r_j) = \phi(x)$, from which we find $\phi(x) = \phi(1)^x$ (using, of course, continuity of the exponentiation). Since $\phi(1) \in S^1$, we may write $\phi(1) = e^{ia}$ for some $a \in \mathbb{R}$, and we obtain $\phi(x) = e^{iax}$.

BC Hall, Chapter 3

Most but not all questions of this chapter are relevant to / were covered in the course. The questions which are *not* so relevant are: 1,3,14,16,19 (and for question 4 only the general idea is of interest).

1. –

2. The Jordan canonical form of a M by M matrix A is a matrix J that can be obtained form A by a similarity transformation, $J = CAC^{-1}$, and which has a block-diagonal form $J = J_1 \oplus J_2 \oplus \cdots \oplus J_n$, with each block associated to an eigenvalue $\lambda_1, \ldots, \lambda_n$ and of

dimensions, say, M_i by M_i (of course, $M = M_1 + \ldots + M_n$). The blocks are given by

$$J_i = \begin{pmatrix} \lambda_i & 1 & 0 & 0 & \cdots \\ 0 & \lambda_i & 1 & 0 & \cdots \\ 0 & 0 & \lambda_i & 1 & \cdots \\ \vdots & \vdots & \vdots & \ddots & \cdots \\ 0 & 0 & \cdots & 0 & \lambda_i \end{pmatrix}.$$

The eigenvalues $\lambda_1, \ldots, \lambda_n$ are the roots of the characteristic polynomial $\det(A - \lambda \mathbf{1}) = 0$. The multiplicity of a given eigenvalue λ_i corresponds to the total number of times it appears in all blocks of J. Hence, if all roots of the characteristic polynomial are distincts (we often say simply that "all eigenvalues are distincts"), then all "blocks" are just 1-by-1 matrices, so the block diagonal form is just exactly diagonal. Hence, if all eigenvalues are distincts, the matric A can be diagonalised. On the other hand, if they are not all distincts, then the matrix A may or may not be diagonalisable (e.g. we may have many 1 by 1 blocks with the same eigenvalue - hence A may still be diagonalisable) Otherwise, we see that J is of the form $J = D + T$, where D is diagonal with the eigenvalues on the diagonal, and T is upper-triangular (with in fact just 1's or 0's on the line just to the right of the diagonal, and 0's everywhere else). Of course, each block is of the form $J_i = \lambda_i \mathbf{1} + T_i$ with T_i upper triangular. Consider the big M by M matrix $\hat{J}_i$ which is zero everywhere except for the M_i by M_i block J_i (and similarly for $\hat{T}_i$). Clearly we have $J = \hat{J}_1 + \ldots + \hat{J}_n$, and all these matrices commute with each other, $[\hat{J}_i, \hat{J}_j] = 0$; likewise, $T = \hat{T}_1 + \ldots + \hat{T}_n$ and $[\hat{T}_i, \hat{T}_j] = [\hat{T}_i, \hat{J}_j] = 0$. Hence, we see that $[T, D] = 0$. Inverting the transformation, we have

$$A = S + N$$

where $S = C^{-1}DC$ and $N = C^{-1}TC$, with $[S, N] = C^{-1}[D, T]C = 0$. Also, since T is upper triangular, we know that $T^M = 0$, hence that $N^M = C^{-1}T^M C = 0$. Hence, we find that S is diagonalisable (this is the *semi-simple part* of A) and that N is nilpotent (this is the *nilpotent part* of A).

Note that for every Jordan block J_i, there is only *one* corresponding eigenvector of A with eigenvalue λ_i (easy to construct from the explicit form of J). Hence if A is not diagonalisable, then the number of eigenvectors is in general less then M; but at worst there will be one big Jordan block, and one eigenvector - so there is always at least one eigenvector (all this is over $\mathbb{C}$).

3. –

4. We only give the general idea. Consider some matrix A. If it's diagonalisable, we're done. If not, then we know from above that there must be eigenvalues that have > 1 multiplicity – some roots of the characteristic polynomial are multiple. But we know that when this happens for a polynomial, we can modify the coefficients of the polynomial

by an infinitesimal amount to break the multiplicity and make all roots distinct (i.e. the polynomial is $\prod_i(\lambda-\lambda_i)$, and we may change the coefficients by making a change $\lambda_i \mapsto \lambda_i+\delta\lambda$). This corresponds to an infinitesimal change of matrix elements of A, and the non-diagonalisable case can be obtained by taking this infinitesimal changes to zero. Check for instance in the 2 by 2, real traceless case: $A=\begin{pmatrix} a & b \\ c & -a \end{pmatrix}$, $P(\lambda)=\lambda^2-a^2-bc$ so the roots are $\lambda=\pm\sqrt{a^2+bc}$ which are not distinct iff $a^2+bc=0$. So we may make them distinct simply by $b \mapsto \mp deltab$ (if $c \neq 0$) so that $a^2+bc \mapsto 0+\delta bc$, or $b \mapsto b+\delta b$ and $c \mapsto c+\delta c$ (if $c=0$) so that $a^2+bc \mapsto 0+\delta b\delta c$.

5. Take $SL(2;\mathbb{R})$ and $sl(2;\mathbb{R})$. We may take $X=i\pi\mathbf{1}$ which has trace $2i\pi \neq 0$, so $X \notin sl(2;\mathbb{R})$, yet we have $\det(e^X)=e^{\mathrm{Tr}(X)}=e^{2i\pi}=1$ so $e^X \in SL(2;\mathbb{R})$.

6. Suppose $G \cong H$, and consider the corresponding Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$. We know that there is an isomorphism $\phi: G \to H$, and we may construct the corresponding Lie algebra homomorphism $\tilde{\phi}: \mathfrak{g} \to \mathfrak{h}$ by
$$t\phi(X)=\left.\frac{d}{dt}\phi(e^{tX})\right|_{t=0}.$$
Since ϕ is an isomorphism, it maps homeomorphically (i.e. continuously and bijectively) neighbourhoods of $\mathbf{1} \in G$ to neighbourhoods of $\mathbf{1} \in H$. We also know that the exponential mapping is a homeomorphism from Lie algebra neighbourhoods around 0 to Lie group neighbourhoods around $\mathbf{1}$, for small enough neighbourhoods. Hence, not only for any $X \in \mathfrak{g}$ there is a unique $Y \in \mathfrak{h}$ such that $\phi(e^{tX})=e^{tY}$ for all t (theorem 3.18), but also, for any $Y \in \mathfrak{h}$ there is a unique $X \in \mathfrak{g}$ such that $\phi(e^{tX})=e^{tY}$ for all t. The reason for the last assertion is: for any $Y \in \mathfrak{h}$ and any given $t \in \mathbb{R}$ there is a unique $e^{tY} \in H$, hence there is a unique group element $g_t \in G$ such that $\phi(g_t)=e^{tY}$ (since ϕ is an isomorphism). But looking at t small enough, g_t is near to the dentity (since ϕ is continuous), so there is a unique X such that $e^{tX}=g_t$ for all such small t (since exp is locally a homeomorphism), hence $e^{tX}=g_t$ for all t (since g_t form a one-parameter subgroup and by theorem 3.12). Hence, for any $Y \in \mathfrak{h}$ there is a unique $X \in \mathfrak{g}$ such that $\tilde{\phi}(X)=Y$. Hence $\tilde{\phi}$ is bijective, so it is an isomorphism.

7. We have the requirement $A^TJA=J$ with $J=\mathrm{diag}(1,1,1,-1)$. Hence, $JA^TJ=A^{-1}$ using $J^2=\mathbf{1}$. Let us write $A=e^{tX}$. We find $e^{tJX^TJ}=e^{-tX}$ for all t, which implies and is implied by $JX^TJ=-X$. These are the matrices in $so(3,1)$. For the general form, consider first the diagonal: it must be 0 because J acts twice on it, hence it must equal minus itself. Clearly, we may choose the 6 upper triangular elements at will, and the 6 lower triangular are determined. This gives
$$\begin{pmatrix} 0 & a & b & c \\ -a & 0 & d & e \\ -b & -d & 0 & f \\ c & e & f & 0 \end{pmatrix}.$$

8. The Lie algebra $su(n)$ of $SU(n)$ is the set of anti-hermitian matrices of trace 0 (as a vector space over the reals). Consider a unitary, determinant-1 matrix $A \in SU(n)$, and consider $X \in su(n)$. Then, AXA^{-1} still has trace 0 by cyclicity of the trace, and is still anti-hermitian since $(AXA^{-1})^\dagger = AX^\dagger A^{-1} = -AXA^{-1}$. It is also obviously true that sX as trace zero and is anti-hermitian for s real, and that for $X, Y \in su(n)$ the sum $X + Y$ has trace zero and is anti-hermitian. Also, we have $Tr([X, Y]) = \mathrm{Tr}(XY - YX) = 0$ by cyclicity, and $[X, Y]^\dagger = (XY - YX)^\dagger = (Y^\dagger X^\dagger - X^\dagger Y^\dagger) = YX - XY = -[X, Y]$, so $[X, Y] \in su(n)$.

9. The Lie algebra $su(2)$ is the real linear space of all matrices that are anti-hermitian and that have trace zero. With
$$X = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$
this gives the conditions $a^* = -a$, $d^* = -d$, $b^* = -c$ as well as $a + d = 0$. Hence, there are three real parameters: say A, B, C with $a = iA/2$, $d = -iA/2$, $b = (B + iC)/2$, $c = (-B + iC)/2$; all these parameters get added up under addition of matrices, and get multiplied by a scalar under scalar multiplication of matrices, so we have linearly mapped the linear space $su(2)$ to the linear space $\mathbb{R}^3$ (the space of triplets (A, B, C)). We can denote this linear map, which is invertible, by $\phi : su(2) \to \mathbb{R}^3$. Obviously, a basis in $\mathbb{R}^3$ is $(1, 0, 0)$, $(0, 1, 0)$ and $(0, 0, 1)$, so a corresponding basis in $su(2)$ as obtained by ϕ^{-1} is that given by E_1, E_2 and E_3. Direct calculations give
$$[E_1, E_2] = E_3, \quad [E_2, E_3] = E_1, \quad [E_3, E_1] = E_2.$$
This can be written as
$$[E_i, E_j] = \sum_k \epsilon_{ijk} E_k. \tag{3}$$
Hence, we see that
$$\phi([E_i, E_j]) = \sum_k \epsilon_{ijk} \phi(E_k) = \phi(E_i) \times \phi(E_j)$$
by the standard definition of the vector product in $\mathbb{R}^3$.

10. $su(2)$ was described above. On the other hand, $so(3)$ is the real linear space of anti-symmetric 3 by 3 real matrices:
$$so(3) = \left\{ X = \begin{pmatrix} 0 & a & -b \\ -a & 0 & c \\ b & -c & 0 \end{pmatrix} : (a, b, c) \in \mathbb{R}^3 \right\}.$$
To show isomorphism, we may simply choose a basis. We may use similar ideas as above: consider a map $\psi : so(3) \to \mathbb{R}^3$ given by $X \mapsto \psi(X) = (a, b, c)$ in the notation above. This

is an invertible linear map, and the usual basis in $\mathbb{R}^3$ gives rise to the corresponding basis in $so(3)$:

$$F_1 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad F_2 = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad F_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}$$

(that is, $\psi(F_1) = (1,0,0)$, $\psi(F_2) = (0,1,0)$, $\psi(F_3) = (0,0,1)$). We see that the commutation relations give

$$[F_i, F_j] = \sum_k \epsilon_{ijk} F_k.$$

Hence, we may consider an invertible linear map $\Phi : su(2) \to so(3)$ defined by $\Phi(E_i) = F_i,\ i = 1,2,3$. By linearity, the maps is defined for all elements of $su(2)$, and it is clearly invertible, with $\Phi^{-1}(F_i) = E_i$ (hence it is bijective). We immediately find that

$$\Phi([E_i, E_j]) = \Phi(\sum_k \epsilon_{ijk} E_k) = \sum_k \epsilon_{ijk} F_k = [F_i, F_j] = [\Phi(E_i), \Phi(E_j)].$$

Hence, by using bilinearity we can extend this relation to all elements of $su(2)$:

$$\Phi([X, X'])[\Phi(X), \Phi(X')].$$

which shows that Φ is also a homomorphism. Hence it is an isomorphism.

11. $sl(2;\mathbb{R})$ is the real linear space of real traceless 2 by 2 matrices,

$$sl(2;\mathbb{R}) = \left\{ X = \begin{pmatrix} a & b \\ c & -a \end{pmatrix} : (a,b,c) \in \mathbb{R}^3 \right\}.$$

We may proceed by constructing a basis in a similar fashion as above, since we have an invertible linear map $\zeta : sl(2;\mathbb{R}) \to \mathbb{R}^3$ given by $\zeta(X) = (a,b,c)$. The basis is

$$G_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad G_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad G_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

The commutation relations are

$$[G_1, G_2] = 2G_2, \quad [G_1, G_3] = -2G_3, \quad [G_2, G_3] = G_1.$$

Following the hint, we see that $\mathbb{C}G_1 + \mathbb{C}G_2$, for instance, is a two-dimensional Lie subalgebra. On the other hand, the result (3) of question 9 can be written more generally, using bilinearity and anti-symmmetry of both sides:

$$\phi([X, X']) = \phi(X) \times \phi(X')$$

for all $X, X' \in su(2)$. Suppose there were a two-dimensional subalgebra in $su(2)$. Since it has to be a subspace, we can map it to a 2-dimensional subspace of $\mathbb{R}^3$ via ϕ: a plane in $\mathbb{R}^3$.

But the vector product of any two vectors is a vector that is perpendicular to both, hence a vector that is outside of the plane. With $[X, X'] = \phi^{-1}(\phi(X) \times \phi(X'))$, we would find that $[X, X']$ is not in the original 2-dimensional subsapce of $su(2)$, a contradiction. Hence, there cannot be a 2-dimensional subalgebra in $su(2)$. But there were an isomorphism between $su(2)$ and $sl(2; \mathbb{R})$, we could map the 2-dimensional subalgebra $\mathbb{C}G_1 + \mathbb{C}G_2$ to a 2-dimensional subalgebra of $su(2)$. Hence, there are no isomorphism.

If we complexify the algebras, however, things are different. Complexifying means considering the space of matrices over the complex instead of the reals (everything still works). Of course, with complex vectors, the vector product does not necessarily gives something perpendicular, so the argument above breaks down. We can write down an explicit isomorphism between $su(2)_{\mathbb{C}}$ and $sl(2; \mathbb{R})_{\mathbb{C}}$. Indeed, simply write

$$\Psi(G_1) = -2iE_1, \quad \Psi(G_2) = E_2 - iE_3, \quad \Psi(G_3) = -E_2 - iE_3$$

which is possible because we are allowed to have complex coefficients. We immediately see that $\Psi(G_i)$, as a matrix, is nothing else than the standard matrix form of G_i for all i, hence it is immediate that this is an isomorphism. But it is easy to work out more explicitly the commutation relations and see that $\Psi([G_i, G_j]) = [\Psi(G_1), \Psi(G_j)]$, and to see that Ψ is invertible,

$$\Psi^{-1}(E_1) = \frac{i}{2}G_1, \quad \Psi^{-1}(E_2) = \frac{G_2 - G_3}{2}, \quad \Psi^{-1}(E_3) = i\frac{G_2 + G_3}{2}.$$

Hence, Ψ is an isomorphism from $sl(2; \mathbb{R})_{\mathbb{C}}$ to $su(2)_{\mathbb{C}}$. Note finally that $sl(2; \mathbb{R})_{\mathbb{C}}$ is real-isomorphic to $sl(2; \mathbb{C})$, which is the real vector space of all traceless 2 by 2 complex matrices (i.e. there is a real invertible linear map that preserves the bracket relations).

12. We have, for $X, Y \in \mathfrak{g}$, the homomorphism property

$$\begin{aligned}
[\mathrm{Ad}A(X), \mathrm{Ad}A(Y)] &= [AXA^{-1}, AYA^{-1}] \\
&= AXA^{-1}AYA^{-1} - AYA^{-1}AXA^{-1} \\
&= A(XY - YX)A^{-1} = \\
&= \mathrm{Ad}A([X, Y]),
\end{aligned} \tag{4}$$

and also $\mathrm{Ad}A$ is linear invertible (hence bijective) $\mathfrak{g} \to \mathfrak{g}$, because $\mathrm{Ad}A(X) = Y \Rightarrow \mathrm{Ad}A^{-1}(Y) = X$.

13. I wish to present here a different "direct calculation", slightly more conceptual. We know that we can see $\mathrm{ad}X$ as an operator on the linear space of matrices, given by

$$\mathrm{ad}X(Y) = [X, Y].$$

There are naturally two other operators that we can think of, on the space of matrices: $\mathrm{L}X$ and $\mathrm{R}X$, which are left- and right-multiplication,

$$\mathrm{L}X(Y) = XY, \quad \mathrm{R}X(Y) = YX.$$

Hence, on the space of matrices, we can write the operator $\mathrm{ad}X$ as

$$\mathrm{ad}X = \mathrm{L}X - \mathrm{R}X.$$

Now, notice that the operators $\mathrm{L}X$ and $\mathrm{R}X$ commute with each other:

$$\mathrm{L}X\,\mathrm{R}X(Y) - \mathrm{R}X\,\mathrm{L}X(Y) = \mathrm{L}X(YX) - \mathrm{R}X(XY) = XYX - XYX = 0.$$

But we know that the exponential of the sum of commuting operators is the product of the exponentials:

$$e^{\mathrm{ad}X} = e^{\mathrm{L}X - \mathrm{R}X} = e^{\mathrm{L}X} e^{-\mathrm{R}X}.$$

Applying this to a matrix Y, we get

$$e^{\mathrm{ad}X}(Y) = e^{\mathrm{L}X} e^{-\mathrm{R}X}(Y) = e^{\mathrm{L}X}(Ye^{-X}) = e^X Y e^{-X} = \mathrm{Ad}e^X(Y).$$

14. –

15. Consider a matrix $X \in sl(2;\mathbb{R})$,

$$X = \begin{pmatrix} a & b \\ c & -a \end{pmatrix}.$$

The eigenvalues are given by $\lambda_\pm = \pm\sqrt{a^2+bc}$. Assuming that X can be diagonalised,

$$X = U^{-1}\begin{pmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{pmatrix} U \Rightarrow e^X = U^{-1}\begin{pmatrix} e^{\lambda_+} & 0 \\ 0 & e^{\lambda_-} \end{pmatrix} U.$$

Then, $\mathrm{Tr}(e^X) = e^{\lambda_+} + e^{\lambda_-}$. If $a^2+bc \geq 0$, then this is clearly ≥ -2. If $a^2+bc < 0$, then this is $2\cos\sqrt{-a^2-bc}$ which is ≥ -2. This shows that the exponential map maps $sl(2;\mathbb{R})$ so a subset of $SL(2;\mathbb{R})$ containing only matrices A with $\mathrm{Tr}(A) \geq -2$. The case where the trace is exactly -2 is the case $a^2+bc = -(2n+1)^2\pi^2$ with n an integer. But then, $e^{\lambda_\pm} = -1$ so that $X = -\mathbf{1}$, hence the only case with trace -2 is $e^X = -\mathbf{1}$. If X cannot be diagonalised, then the eigenvalues are the same, so that $a^2+bc = 0$. But then, we can just modify $X \mapsto X'$ by $b \mapsto b' = b + \delta b$ and, if necessary, $c \mapsto c' = c + \delta c$, so that X' can be diagonalised. The argument above then shows that $\mathrm{Tr}(X')$ is near to 2, and as $\delta b, \delta c \to 0$, the trace tends to 2. Since the trace is a continuous function of the matrix elements, when X cannot be diagonalised, then $\mathrm{Tr}(X) = 2 \geq -2$.

To go in the other direction, take a matrix A such that $\mathrm{Tr}(A) > -2$. We want to find a X such that $e^X = A$. Let us write again

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

The conditions are $\det(A) = 1 \Rightarrow ad - bc = 1$ and $\mathrm{Tr}(A) > -2 \Rightarrow a + d > -2$. Again, we consider the eigenvalues, and assume that the matrix can be diagonalised. The eigenvalues satisfy $\lambda_1\lambda_2 = 1$ and $\lambda_1 + \lambda_2 > -2$. Certainly, since A is a real matrix, we also have $\lambda_1 + \lambda_2 \in \mathbb{R}$. Hence if λ_1 is complex $\lambda_1 = re^{i\theta}$ ($r > 0$, $\theta \in [0, 2\pi)$), then $re^{i\theta} + r^{-1}e^{-i\theta} \in \mathbb{R}$, hence $r = r^{-1} = 1$ or $\theta = 0$ or $\theta = \pi$. In the case $\theta = 0$, we always have $\lambda_1 + \lambda_2 > -2$. For $\theta = \pi$, this never occurs. Finally, for $r = 1$, then we must have $\theta \neq \pi$. In all these cases, with

$$A = U^{-1} \begin{pmatrix} re^{i\theta} & 0 \\ 0 & r^{-1}e^{-i\theta} \end{pmatrix} U,$$

we see that we can write $A = e^X$ with

$$X = U^{-1} \begin{pmatrix} \log r & 0 \\ 0 & -\log r \end{pmatrix} U \text{ or } X = U^{-1} \begin{pmatrix} i\theta & 0 \\ 0 & -i\theta \end{pmatrix} U$$

where the first is for the case $\theta = 0$ and the second for the case $r = 1$. Recall that A is real. We must check that X is real as well. In the first case looking at the Taylor series expansion for r around 1, we see that X is a Taylor series in powers of A with real coefficients, hence X is real for all r near enough to 1. By analytic continuation, it is then real for all $r > 0$. In the second case, we can likewise write $i\theta = \log(e^{i\theta})$ and make a Taylor series expansion for $e^{i\theta}$ near to 1 – this is in fact convergent for any $\theta \neq \pi$. We then have X as a convergent power series in A with real coefficients, hence X is real as well. Hence, we see that X is real and that it has trace 0. So, for any $A \in SL(2; \mathbb{R})$ with $\mathrm{Tr}(A) > -2$, we have a $X \in sl(2; \mathbb{R})$ such that $e^X = A$.

16. –

17. The Lie algebra of the Heisenberg group is the set of matrices of the form

$$\left\{ X = \begin{pmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{pmatrix} : (a, b, c) \in \mathbb{R}^3 \right\}.$$

Explicitly exponentiating gives

$$e^X = \mathbf{1} + \begin{pmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 0 & 0 & ac \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & a & b + ac/2 \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix}.$$

On the other hand, the matrices in the Heisenberg group are

$$\left\{ M = \begin{pmatrix} 1 & A & B \\ 0 & 1 & C \\ 0 & 0 & 1 \end{pmatrix} : (A, B, C) \in \mathbb{R}^3 \right\}.$$

Hence, we find that the exponential map is

$$(a, b, c) \mapsto (A, B, C) = (a, b + ac/2, c).$$

The inverse map can easily be obtained:

$$(A, B, C) \mapsto (a, b, c) = (A, B - AC/2, C).$$

This makes it clear that the exponential map is onto (surjective: every (A, B, C) has a pre-image) and 1-1 (injective: if $(A, B, C) \neq (A', B', C')$, then $(A, B - AC/2, C) \neq (A', B' - A'C'/2, C')$ - clear if $A \neq A'$ or $C \neq C'$, and if $A = A'$ and $C = C'$, it becomes clear for $B' \neq B'$ as well).

18. The algebra $u(n)$ is the algebra of anti-hermitian n by n matrices. Hermitian matrices can always be diagonalised by a unitary transformation (hence also anti-hermitian matrices), so for $X \in su(n)$ we can always write

$$X = UDU^\dagger$$

where U is unitary and $D = i\,\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ with $\lambda_i \in \mathbb{R}\ \forall i$. Then, we find

$$e^X = Ue^D U^\dagger.$$

We see that with $\lambda_i \mapsto \lambda_i + 2\pi$, the matrix X changes, by e^X stays the same. Hence, the exponential mapping is not injective (1-1). However, any matrix M in the group $U(n)$ of unitary matrices can be diagonalised (since it has a basis of orthonormal eigenvectors), so we can write

$$M = VQV^\dagger$$

where V is unitary and $Q = \mathrm{diag}(e^{i\theta_1}, \ldots, e^{i\theta_n})$ with $\theta_i \in \mathbb{R}\ \forall i$. Clearly, then, for any M we can form a X by choosing $U = V$ and $\lambda_i = \theta_i$, so the exponential map is surjective (onto).

19. –

20. If X is diagonalisable, then $X = MDM^{-1} = \mathrm{Ad}M(D)$ for some invertible matrix M and some diagonal matrix D. Then, we have

$$\begin{aligned}
\mathrm{ad}X(Y) &= [X, Y] \\
&= [MDM^{-1}, Y] \\
&= M[D, M^{-1}YM]M^{-1} \\
&= M[D, \mathrm{Ad}M^{-1}(Y)]M^{-1} \\
&= M\mathrm{ad}D\,\mathrm{Ad}M^{-1}(Y)M^{-1} \\
&= \mathrm{Ad}M\,\mathrm{ad}D\,\mathrm{Ad}M^{-1}(Y) \\
&= \mathrm{Ad}M\,\mathrm{ad}D\,(\mathrm{Ad}M)^{-1}(Y).
\end{aligned}$$

Now let us analyse $\mathrm{ad}D$. Taking the basis in $gl(n;\mathbb{C})$ given by the matrices e_{ij} with a 1 at the position i, j and zero everywhere else, $(e_{ij})_{kl} = \delta_{ik}\delta_{jl}$, we see that

$$\mathrm{ad}D(e_{ij}) = [D, e_{ij}] = (D_i - D_j)e_{ij}.$$

That is, ad$D$ acts diagonally in that basis. Hence, we have found a basis (the basis of matrices e_{ij}), an invertible linear map AdM on the space of matrices, and a map ad$D$ that is diagonal in that basis, such that

$$\mathrm{ad}X = \mathrm{Ad}M\,\mathrm{ad}D\,(\mathrm{Ad}M)^{-1}.$$

That is, we have shown that ad$X$ is diagonalisable.

BC Hall, Chapter 4

Problems 1-3 are relevant; problems 6, 7 are about connected Lie subgroup, not covered in class; problems 4,5 are doable with the material taught in class, but slightly tedious, and they have to do with the explicit series expansion of the BCH formula, which we didn't really cover in class.

1. The center $Z(\mathfrak{h})$ is the set of matrices $\begin{pmatrix} 0 & u & v \\ 0 & 0 & w \\ 0 & 0 & 0 \end{pmatrix}$ that commute with all of $\mathfrak{h}$. That is, we want to find all $u, v, w \in \mathbb{R}$ such that

   $$\left[\begin{pmatrix} 0 & u & v \\ 0 & 0 & w \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & \alpha & \beta \\ 0 & 0 & \gamma \\ 0 & 0 & 0 \end{pmatrix}\right] = 0 \quad \forall \quad \alpha, \beta, \gamma \in \mathbb{R}.$$

   Direct calculation of the commutator gives

   $$\begin{pmatrix} 0 & 0 & u\gamma - \alpha w \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = 0$$

   hence $u = 0$, $w = 0$. So the center is

   $$Z(\mathfrak{h}) = \left\{ \begin{pmatrix} 0 & 0 & v \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} : v \in \mathbb{R} \right\}.$$

   Clearly, the commutator computer above just has an element in the upper right corner, so it is in the center. Note that in general, a *Heisenberg Lie algebra* is an algebra such that the center is one-dimensional, and such that the commutator of any two algebra element is in the center.

   Finally, let us compute the BCH formula in the Heisenber algebra case. First, specialising the formula itself (theorem 4.3, p. 57): Since $[X, Y]$ commutes with both X and Y, this means that in the BCH formula, we only need to keep the terms containing only one commutator. Hence, we expand the exponentials in ad$X$ and ad$Y$, and keep only the terms that lead to first order in these matrices. It is convenient to first expand the function:

$$g(1+x) = \frac{\log(1+x)}{1-(1+x)^{-1}} = \frac{x - x^2/2 + \dots}{1-(1-x+x^2)+\dots} = \frac{1-x/2+\dots}{1-x+\dots} = 1 + x/2 + \dots$$

With $x = e^{\mathrm{ad}X}e^{t\mathrm{ad}Y} - \mathbf{1}$, we only need to keep the first order of the exponentials, $x = \mathrm{ad}X + t\mathrm{ad}Y + \dots$. This gives

$$\log(e^X e^Y) = X + \int_0^1 dt\, \left(\mathbf{1} + \frac{1}{2}\left(\mathrm{ad}X + t\mathrm{ad}Y\right)\right)(Y) = X + Y + \frac{1}{2}[X,Y].$$

Second, by direct computation: take any two elements X, Y in $\mathfrak{h}$, in matrix form $X = \begin{pmatrix} 0 & u & v \\ 0 & 0 & w \\ 0 & 0 & 0 \end{pmatrix}$ and $Y = \begin{pmatrix} 0 & \alpha & \beta \\ 0 & 0 & \gamma \\ 0 & 0 & 0 \end{pmatrix}$, and evaluate their exponential:

$$\exp\begin{pmatrix} 0 & u & v \\ 0 & 0 & w \\ 0 & 0 & 0 \end{pmatrix} = \mathbf{1} + \begin{pmatrix} 0 & u & v \\ 0 & 0 & w \\ 0 & 0 & 0 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 0 & u & v \\ 0 & 0 & w \\ 0 & 0 & 0 \end{pmatrix}^2 + \dots$$

where we have

$$\begin{pmatrix} 0 & u & v \\ 0 & 0 & w \\ 0 & 0 & 0 \end{pmatrix}^2 = \begin{pmatrix} 0 & 0 & uw \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

and $\begin{pmatrix} 0 & u & v \\ 0 & 0 & w \\ 0 & 0 & 0 \end{pmatrix}^n = 0$ for all $n \geq 3$. Hence,

$$\exp\begin{pmatrix} 0 & u & v \\ 0 & 0 & w \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & u & v + uw/2 \\ 0 & 1 & w \\ 0 & 0 & 1 \end{pmatrix}. \tag{5}$$

Then,

$$\begin{aligned} \exp\begin{pmatrix} 0 & u & v \\ 0 & 0 & w \\ 0 & 0 & 0 \end{pmatrix} \exp\begin{pmatrix} 0 & \alpha & \beta \\ 0 & 0 & \gamma \\ 0 & 0 & 0 \end{pmatrix} &= \begin{pmatrix} 1 & u & v + uw/2 \\ 0 & 1 & w \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & \alpha & \beta + \alpha\gamma/2 \\ 0 & 1 & \gamma \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & \alpha + u & \beta + v + u\gamma + \alpha\gamma/2 + uw/2 \\ 0 & 1 & \gamma + w \\ 0 & 0 & 1 \end{pmatrix} \end{aligned} \tag{6}$$

On the other hand,

$$\exp\left(X + Y + \frac{1}{2}[X,Y]\right) = \exp\begin{pmatrix} 0 & u+\alpha & v + \beta + (u\gamma - \alpha w)/2 \\ 0 & 0 & w + \gamma \\ 0 & 0 & 0 \end{pmatrix}$$

and, using the previous calculation (5) for the exponential, we have

$$\exp\begin{pmatrix} 0 & u+\alpha & v+\beta+(u\gamma-\alpha w)/2 \\ 0 & 0 & w+\gamma \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & u+\alpha & v+\beta+(u\gamma-\alpha w)/2+(u+\alpha)(w+\gamma)/2 \\ 0 & 1 & w+\gamma \\ 0 & 0 & 1 \end{pmatrix}$$

which agrees with (6)

2. In order for the matrix to be invertible, its determinant must be non-zero. If X is diagonalisable, $X = UDU^{-1}$, then

$$\frac{\mathbf{1}-e^{-X}}{X} = U\begin{pmatrix} f(\lambda_1) & 0 & \cdots \\ 0 & f(\lambda_2) & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}U^{-1}$$

where λ_j are the eigenvalues, and where

$$f(z) = \frac{1-e^{-z}}{z}.$$

Then, the determinant is $\prod_j f(\lambda_j)$, so we require $f(\lambda_j) \neq 0$ for all j. The zeroes of f are $z = 2\pi i n$ with n any non-zero integer (at $z = 0$ there is no problem, because there $f(0) = 1$). This gives the required condition. If X is not diagonalisable, then we can bring it in a Jordan normal form, and we have

$$\frac{\mathbf{1}-e^{-X}}{X} = U\begin{pmatrix} f(\lambda_1) & * & \cdots \\ 0 & f(\lambda_2) & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}U^{-1}.$$

Again, the determinant is $\prod_j f(\lambda_j)$, so we have the same condition.

3. If X and Y commute, then we have both $\mathrm{ad}Y = 0$ and $\mathrm{ad}X = 0$ as matrices on the vector space $\mathbb{R}X + \mathbb{R}Y$, so that $g(e^{\mathrm{ad}X}e^{t\mathrm{ad}Y}) = g(\mathbf{1}) = \mathbf{1}$ (see problem 1 for g). This matrix acting on vector Y gives $g(e^{\mathrm{ad}X}e^{t\mathrm{ad}Y})(Y) = Y$, which we have to put in the BCH formula p. 57 (Hall). Hence, we have $\log(e^X e^Y) = X + \int_0^1 Y\,dt = X + Y$ as it should.

4. –

5. –

6. –

7. –

JE Humphreys, Section 1

Here all questions are written and self-contained (no need to refer to Humphreys' book). I'm keeping the field of numbers to be $\mathbb{F} = \mathbb{C}$, so I avoid all questions having to do with the particularly of the field over which we work.

Chapter 1 is essentially a warm up, to make sure basic principles of Lie algebras are understood. I didn't cover the $A_\ell, B_\ell, C_\ell, D_\ell$ description of Lie algebras, but looking at the definitions in Humphreys one sees that they correspond to the Lie algebras that we saw, in class. The questions related to these are interesting, but for now I do not cover them. I give solutions to the questions of more general character and more directly related to concepts discussed in class.

1. Let L be the real vector space $\mathbb{R}^3$. Define $[xy] = x \times y$ (cross product of vectors) for $x, y \in L$, and verify that L is a Lie algebra. Write down the structure constants relative to the usual basis of $\mathbb{R}^3$.

   Solution

   In order to have a Lie algebra, we need a vector space (ok), a bilinear form (ok) which is anti-symmetric (ok), and which satisfies the Jacobi identity. To verify the later, we evaluate:

   $$[x[yz]] + [y[zx]] + [z[xy]] = x \times (y \times z) + y \times (z \times x) + z \times (x \times y) \tag{7}$$

   where we can use the "bac-cab" formula:

   $$a \times (b \times c) = b(a \cdot c) - c(a \cdot b)$$

   which gives, on the r.h.s. of (7)

   $$y(x \cdot z) - z(x \cdot y) + z(y \cdot x) - x(y \cdot z) + x(z \cdot y) - y(z \cdot x) = 0.$$

   The structure constants are the numbers c_{ijk} that arise when we evaluate the bracket using an explicit basis $e_i \in L$:

   $$[e_i e_j] = \sum_k c_{ijk} e_k.$$

   If e_i represent the usual orthonormal basis in $\mathbb{R}^3$, then we have $e_i \times e_j = \sum_k \epsilon_{ijk} e_k$ where ϵ_{ijk} is the completely anti-symmetric symbol. Hence, we find the structure constants

   $$c_{ijk} = \epsilon_{ijk}.$$

2. Verify that the following equations, along with bilinearity and anti-symmetry, define a Lie algebra structure on a 3-dimensional vector space with basis (x, y, z):

   $$[xy] = z, [xz] = y, [yz] = 0$$

   Solution

Since the product is bilinear and anti-symmetric by assumption, we only need to show the Jacobi identity, $[u[vw]]+[v[wu]]+[w[uv]]=0$, for the triplet u,v,w being any triplet formed with x,y,z (i.e. (x,x,x), (x,x,y), (x,z,y), (y,z,x), etc.). Indeed, the validity of the Jacobi identity for any basis element implies its validity for the whole vector space, by bilinearity. Note that the Jacobi identity is automatic by bilinearity and anti-symmetry whenever any two of u,v,w are equal. Hence, we need to calculate only 6 terms, corresponding to the 6 triplets (x,y,z), (y,z,x), (z,x,y) (and opposite orders). Then, we calculate $[u[vw]]$ for these triplets: $[x[yz]]=[x0]=0$, $[y[xz]]=[yy]=0$ and $[z[xy]=[zz]=0$ and the opposite orders of the elements in the inner bracket (giving the negative of these answers). Hence, the Jacobi identity immediately holds.

3. Let $x=\begin{pmatrix}0&1\\0&0\end{pmatrix}$, $h=\begin{pmatrix}1&0\\0&-1\end{pmatrix}$, $y=\begin{pmatrix}0&0\\1&0\end{pmatrix}$ be an ordered basis for $sl(2;\mathbb{C})$. Compute the matrices $\mathrm{ad}x$, $\mathrm{ad}h$ and $\mathrm{ad}y$ relative to this basis.

   Solution

   The basis is taken in the order x,h,y, so we may make the association

$$x\mapsto\begin{pmatrix}1\\0\\0\end{pmatrix},\quad h\mapsto\begin{pmatrix}0\\1\\0\end{pmatrix},\quad y\mapsto\begin{pmatrix}0\\0\\1\end{pmatrix}.$$

   Then, we evaluate:

$$[xh]=-2x,\ [xy]=h,\ [yh]=2y.$$

   This gives $\mathrm{ad}x(h)=-2x$ and $\mathrm{ad}x(y)=h$, which, along with $\mathrm{ad}x(x)=0$, tells us that

$$\mathrm{ad}x=\begin{pmatrix}0&-2&0\\0&0&1\\0&0&0\end{pmatrix}.$$

   Similarly, we find

$$\mathrm{ad}h=\begin{pmatrix}2&0&0\\0&0&0\\0&0&-2\end{pmatrix}$$

   and

$$\mathrm{ad}y=\begin{pmatrix}0&0&0\\-1&0&0\\0&2&0\end{pmatrix}.$$

4. Find a Lie algebra of matrices that is isomorphic to the 2-dimensional Lie algebra L (with basis x,y) determined by the bilinear anti-symmetric Lie bracket $[xy]=x$.

   Solution

   What we have here is an abstract specification of a Lie algebra: we are given the Lie bracket only (we can easily verify that it satisfies the Jacobi identity, hence this indeed defines a Lie algebra). We want to see this as an algebra of matrices, with bracket being the usual matrix commutator. A way of doing this is by considering the adjoint representaion, which, as we know, gives us a homomorphism from L to some Lie algebra of matrices. At the end, we just have to check that ad actually gives an isomorphism in the present case. Evaluating the matrices ad$x$ and ad$y$ in the same way as in the previous question, with $x \mapsto \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $y \mapsto \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, we find

$$\mathrm{ad}x = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \mathrm{ad}y = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

One can explicitly check that $[\mathrm{ad}x, \mathrm{ad}y] = \mathrm{ad}x$ with matrix commutator. Hence, the Lie algebra of matrices that we find is

$$\left\{ \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix} : a, b \in \mathbb{C} \right\}.$$

Since this is 2-dimensional, the ad homomorphism is actually an isomorphism.

5. –

6. Let $x \in gl(n;\mathbb{C})$ have n distinct eigenvalues $a_1, \dots, a_n$. Prove that the eigenvalues of ad$x$ are precisely the n^2 scalars $a_i - a_j$, $(1 \le i, j \le n)$, which of course need not be distinct.

   Solution

   If the eigenvalues of x are distinct, then for sure the matrix x can be diagonalised. Hence, we can write

$$x = UDU^{-1} = U \begin{pmatrix} a_1 & 0 & \cdots \\ 0 & a_2 & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} U^{-1}$$

   for some invertible matrix U. Then, let us consider a basis of matrices in the linear space $gl(n;\mathbb{C})$, given by

$$f_{ij} = Ue_{ij}U^{-1}$$

   where e_{ij} are the n by n matrices with 0 everywhere, except for a 1 at the position i, j; that is, $(e_{ij})_{kl} = \delta_{ik}\delta_{jl}$. The f_{ij} form a basis, because we can always write any matrix A as $A = \sum_{ij}(A)_{ij}e_{ij}$, so we can always write any matrix B as $B = UU^{-1}BUU^{-1} = U\sum_{ij}(U^{-1}BU)_{ij}e_{ij}U^{-1} = \sum_{ij}(U^{-1}BU)_{ij}f_{ij}$. Also, it is clear that the f_{ij} are independent. Then, we may evaluate the adjoint action of x on the basis elements f_{ij}:

$$\mathrm{ad}x(f_{ij}) = [x, f_{ij}] = U[D, e_{ij}]U^{-1} = (a_i - a_j)Ue_{ij}U^{-1} = (a_i - a_j)f_{ij}.$$

   That is, we have found a basis of matrices in $gl(n;\mathbb{C})$ which diagonalises the action of ad$x$. Hence, the numbers $a_i - a_j$ are the n^2 eigenvalues of ad$x$.

7. Let $s(n;\mathbb{C})$ denote the one-dimensional space of the *scalar matrices* (i.e. the scalar multiples of the identity matrix) in $gl(n;\mathbb{C})$. Prove that $gl(n;\mathbb{C}) = sl(n;\mathbb{C}) \oplus s(n;\mathbb{C})$ as vector spaces, and that both $s(n;\mathbb{C})$ and $sl(n;\mathbb{C})$ are ideals.

   Solution

   Take any matrix x in $gl(n;\mathbb{C})$. We may always construct the new matrix $x' = x - \mathrm{Tr}(x)\mathbf{1}$, which is traceless, $\mathrm{Tr}(x') = 0$. Hence, $x' \in sl(n;\mathbb{C})$. Moreover, there is a unique element y of $s(n;\mathbb{C})$ such that $x = x' + y$ with $x' \in sl(n;\mathbb{C})$; it is the element $y = \mathrm{Tr}(x)\mathbf{1}$. Hence, we have a 1-1 correspondence between $sl(n;\mathbb{C}) \oplus s(n;\mathbb{C})$ and $gl(n;\mathbb{C})$ as vector space, which proves the first assertion. Finally, it is clear that $s(n;\mathbb{C})$ commutes with everything: in fact this is the center $Z(gl(n;\mathbb{C}))$, because the only matrices that commute with all n by n matrices are the scalar n by n matrices. Hence, it is an ideal (because the center is always an ideal). Also, $sl(n;\mathbb{C})$ is certainly a subalgebra (because commutator of traceless matrices are always traceless): it is the commutator subalgebra of $gl(n;\mathbb{C})$, i.e. $sl(n;\mathbb{C}) = [gl(n;\mathbb{C}), gl(n;\mathbb{C})]$. But the commutator subalgebra is also an ideal, so $sl(n;\mathbb{C})$ is an ideal. In fact, here it's quite easy to see that it is, because it clearly commutes with all of $s(n;\mathbb{C})$, i.e. $[sl(n;\mathbb{C}), s(n;\mathbb{C})] = 0$.

   Note that a Lie algebra that can be written as a direct sum of *simple ideals* plus *abelian ideals* is called **reductive**. Hence, although $gl(n;\mathbb{C})$ is neither simple nor semisimple, it is in fact reductive, because $sl(n;\mathbb{C})$ is simple and $s(n;\mathbb{C})$ is abelian.

8. –

9. –

10. –

11. Verify that the commutator of two derivations of a Lie algebra is again a derivation, whereas the ordinary product need not be.

    Solution

    Given a Lie algebra L with Lie bracket $[\cdot\cdot]$, a derivation $D \in gl(L)$ is a linear map from L to L with the property that $D([xy]) = [D(x)y] + [xD(y)]$ for all $x, y \in L$. Suppose we have two derivations, D_1 and D_2. Consider the commutator $[D_1, D_2] = D_1D_2 - D_2D_1$. We need to check that this is still a derivation, so we need to apply this to $[xy]$. We have

$$\begin{aligned}
[D_1, D_2]([xy]) &= D_1([D_2(x)y] + [xD_2(y)]) - D_2([D_1(x)y] + [xD_1(y)]) \\
&= [D_1(D_2(x))y] + [D_2(x)D_1(y)] + [D_1(x)D_2(y)] + [xD_1(D_2(y))] - (D_1 \leftrightarrow D_2) \\
&= [[D_1, D_2](x)y] + [x[D_1, D_2](y)]
\end{aligned}$$

    hence indeed $[D_1, D_2]$ is a derivation. Note that for the ordinary product, we get

$$D_1(D_2([xy])) = [D_1(D_2(x))y] + [D_2(x)D_1(y)] + [D_1(x)D_2(y)] + [xD_1(D_2(y))]$$

where we have the two spurious terms $[D_2(x)D_1(y)]+[D_1(x)D_2(y)]$ which break the derivation property in general.

12. Let L be a Lie algebra and let $x \in L$. Prove that the subspace of L spanned by the eigenvectors of ad$x$ is a Lie subalgebra.

    Solution: The eigenvectors of ad$x$ are the elements in $V = \{y \in L \,|\, [xy] \propto y\}$. We want to look at the linear span of V (i.e. all linear combinations of elements of V). This is a subalgebra, because 1) it is a subspace, since we take the linear span, and 2) given any $y, y' \in V$, we have $[x[yy']] = [y[xy']] + [[xy]y']$ by the Jacobi identity, which is proportional to $[yy']$, hence $[yy'] \in V$; since the bracket is bilinear, the bracket also preserves the linear span.

JE Humphreys, selected problems from sections 2-8

2.1 Prove that the set of all *inner derivations* of a Lie algebra L, i.e. derivations D of the form $D = \mathrm{ad}x$ for some $x \in L$, is an ideal of the Lie algebra of derivations Der L (see section 1, question 11 for the fact that Der L is a Lie algebra, with Lie bracket the usual commutator).

Solution

What we have to prove is that given any derivation D, and any inner derivation ad$x$, the commutator $[D, \mathrm{ad}x]$ is of the form ad$y$ for some $y \in L$, i.e. is still an inner derivation. In order to do so, we act on an arbitrary Lie algebra element z to see what happens. We have

$$\begin{aligned} [D, \mathrm{ad}x](z) &= D(\mathrm{ad}x(z)) - \mathrm{ad}x(D(z)) \\ &= D([xz]) - [xD(z)] \\ &= [D(x)z] + [xD(z)] - [xD(z)] \\ &= [D(x)z] \\ &= \mathrm{ad}(D(x))(z) \end{aligned}$$

so that $[D, \mathrm{ad}x] = \mathrm{ad}y$ for $y = D(x)$.

2.2 Show that $sl(n; \mathbb{C})$ is precisely the derived algebra of $gl(n; \mathbb{C})$; the derived algebra of a Lie algebra L is $[LL] = \mathrm{span}\{[xy] : x, y \in L\}$, the set of all linear combinations of elements of the form $[xy]$.

Solution

Since the commutator of any two matrices has zero trace, $\mathrm{Tr}([A, B]) = \mathrm{Tr}(AB - BA) = \mathrm{Tr}(AB - AB) = 0$, then certainly $[gl(n; \mathbb{C}), gl(n; \mathbb{C})] \subset sl(n; \mathbb{C})$. To prove equality, let us consider the basis of matrices in $gl(n;\mathbb{C})$ given by the matrices e_{ij} for $i,j=1,\ldots,n$ (see section 1, question 6). We want to see what we can get by taking all possible commutators $[e_{ij},e_{i'j'}]$. We have

$$[e_{ij},e_{i'j'}]=\delta_{i'j}e_{ij'}-\delta_{ij'}e_{i'j}.$$

Hence, with $i'=j$ and $i>j'$ we get a basis for the upper-triangular matrices (i.e. we get all matrices with only a 1 somewhere in the upper triangular region), and with $i'=j$ and $i<j'$ we get a basis for the lower-triangular matrices. In order to get all possible matrices in $sl(n;\mathbb{C})$, we now need to get all diagonal matrices with trace 0. Choosing $i'=j$ and $i=j'$, we find that the commutator is $e_{ij'}-e_{i'j}=e_{ii}-e_{i'i'}$. Hence, we get all diagonal matrices with a 1 at (i,i) and a -1 at (i',i'), and 0 everywhere else. Keeping always the -1, say, at the position $(1,1)$ in the matrix, we see that by taking linear combinations we can obtain a matrix with any chosen number we want at the position $(2,2)$ up to (n,n), and that at the position $(1,1)$ there is always the unique number necessary to make it traceless. Hence, we do get all matrices in $sl(n;\mathbb{C})$.

2.3 Prove that the center of $gl(n;\mathbb{C})$ equals $s(n;\mathbb{C})$ (the scalar multiples of $\mathbf{1}$). Prove that $sl(n;\mathbb{C})$ has center $\{0\}$.

Solution

We consider again the basis of matrices $e_{ij}:i,j=1,\ldots n$, and look for matrices x that commute with all of them. We find

$$0=([e_{ij},x])_{kl}=\sum_m((e_{ij})_{km}x_{ml}-x_{km}(e_{ij})_{ml})=\delta_{ik}x_{jl}-\delta_{jl}x_{ki}$$

so, taking $i=k$, we see that we need $x_{jl}=0$ for all $j\neq l$, and taking $i=k,\ j=l$, that we need $x_{jj}-x_{kk}=0$ for all j,k. Hence, all elements on the diagonal have to be equal, which shows that we need an element of $s(n;\mathbb{C})$.

For the case $sl(n;\mathbb{C})$, the only difference is that the basis of matrices that are diagonal is not e_{ii} but $e_{ii}-e_{jj}$ (see the previous question). This does not affect the arguments above, because we can always choose $i\neq j$. But the only matrix in $s(n;\mathbb{C})$ that is in $sl(n;\mathbb{C})$ is the zero matrix.

2.4 Show that (up to isomorphism) there is a unique Lie algebra L of dimension 3 whose derived algebra $[LL]$ has dimension 1 and lies in its center $Z(L)$.

Solution

Consider a basis for L given by x,y,z. The derived subalgebra is composed of all elements of the form $a[xy]+b[zx]+c[yz]$ for any $a,b,c\in\mathbb{C}$. We want this to be of dimension 1, so $[xy]\propto[zx]\propto[yz]$. That is, there is a triplet $(u,v,w)\in\mathbb{C}^3$, and complex numbers α,β, such that $[xy]=ux+vy+wz$, $[zx]=\alpha(ux+vy+wz)$, $[yz]=\beta(ux+vy+wz)$ (here, we chose to make sure that at least $[xy]$ is non-zero; one can always choose a basis in this way). But also, the element $ux + vy + wz$ must lie in the center, so $[x\,(ux + vy + wz)] = [y\,(ux + vy + wz)] = [z\,(ux + vy + wz)] = 0$. This tells us that

$$v[xy] + w[xz] = u[yx] + w[yz] = u[zx] + v[zy] = 0$$

hence

$$v[xy] - w\alpha[xy] = -u[xy] + w\beta[xy] = -u\alpha[xy] - v\beta[xy] = 0.$$

This gives us the equations:

$$v - w\alpha = 0,\ -u + w\beta = 0,\ -u\alpha - v\beta = 0.$$

So we have $v = w\alpha$, $u = w\beta$ and $w\alpha\beta = 0$. If $w = 0$, then also $u = 0$ and $v = 0$, so that $[LL]$ is in fact zero-dimensional, which we do not want. Hence, $w \neq 0$, and $\alpha = 0$ or $\beta = 0$. By exchanging $x \leftrightarrow -y$, we see that the commutators are not changed except for $\alpha \leftrightarrow \beta$, so we can choose $\beta = 0$ and $\alpha \neq 0$ by an appropriate choice of basis. Hence, we find

$$v = w\alpha,\ u = 0,\ \beta = 0$$

with w, α arbitrary but non-zero. This gives us

$$[xy] = w(\alpha y + z), \quad [zx] = w\alpha(\alpha y + z), \quad [yz] = 0.$$

Since w, α are nonzero, we can always write $x = w\alpha x'$ and $y = y'/\alpha$, and we find

$$[x'y'] = (y' + z), \quad [zx'] = (y + z), \quad [y'z] = 0.$$

This defines the unique (up to isomorphism) algebra that we wanted.

2.5 Suppose that the dimension of a Lie algebra L is 3, and that $L = [LL]$. Prove that L must be simple. Recover the simplicity of $sl(2;\mathbb{C})$.

Solution

Let us look for a proper ideal I in L. It can only have dimension 1 or 2. Suppose I has dimension 1, $I = \mathbb{C}x$ for some $x \in L$. Then, taking two other basis elements in L to be y, z, we must have $[xy] \propto x$ and $[xz] \propto x$ because I is an ideal. But then, $[LL]$ is only at most two-dimensional, because $[LL] = \mathbb{C}[xy] + \mathbb{C}[yz] + \mathbb{C}[zx] = \mathbb{C}x + \mathbb{C}[yz]$, so we cannot have $L = [LL]$, a contradiction. Hence, the proper ideal I, if it exists, must be of dimension 2. Suppose it is spanned by basis elements x, y, i.e. $I = \mathbb{C}x + \mathbb{C}y$. With z a basis element outside I, we have $[xy] \in I$, $[xz] \in I$ and $[zy] \in I$, so that $[LL] = \mathbb{C}[xy] + \mathbb{C}[yz] + \mathbb{C}[zx] = I$. Again, we see that $[LL]$ then has dimension 2, so we cannot have $L = [LL]$, a contradiction. Hence, there is no proper ideal. Additionally, since $[LL] = L$, then L is not abelian. Hence, L is simple. An analysis similar to that of question 2.2 above shows that $sl(2,\mathbb{C}) = [sl(2,\mathbb{C}), sl(2,\mathbb{C})]$, and we know that $sl(2,\mathbb{C})$ has dimension 3. Hence, $sl(2,\mathbb{C})$ is simple.

3.1 (part of it) In general, the *derived series* of a Lie algebra L is $(L^0, L^1, L^2, \ldots)$ with

$$L^0 = L, \quad L^1 = [L^0 L^0], \quad L^2 = [L^1 L^1], \ldots, L^{n+1} = [L^n L^n], \ldots$$

Let I be an ideal of a Lie algebra L. Prove that each member I^n of the derived series of I is also an ideal of L.

Solution

By assumption, $[IL] \subset I$. This shows the statement for the member I^0. Let us proceed by induction: assume that I^n is an ideal of L, i.e. that $[I^n L] \subset I^n$. Consider any $x, y \in I^n$, so that $[xy] \in I^{n+1}$. Consider any $z \in L$. Then, by the Jacobi identity,

$$[[xy]z] = [[xz]y] + [[zy]x].$$

Since $[xz] \in I^n$ and $[zy] \in I^n$ by the induction assumption, we have that $[[xy]z] \in [I^n I^n] = I^{n+1}$. Since any element in I^{n+1} is a linear combination of elements of the form $[xy]$ for $x, y \in I^n$, this shows that $[I^{n+1} L] \subset I^{n+1}$. Hence, I^{n+1} is an ideal of L. This completes the induction.

3.4 (part of it) Prove that L is solvable if and only if $\mathrm{ad}L$ is solvable.

Solution

Consider the derived series of L (see question 3.1). By the homomorphism property of ad, we have $[\mathrm{ad}L^n, \mathrm{ad}L^n] = \mathrm{ad}[L^n L^n] = \mathrm{ad}L^{n+1}$. Hence, by an easy induction, the derived series of $\mathrm{ad}L$ is given by the members $(\mathrm{ad}L)^n = \mathrm{ad}L^n$ for $n = 0, 1, 2, \ldots$. If L is solvable, then $L^m = 0$ for some m, hence $(\mathrm{ad}L)^m = \mathrm{ad}L^m = 0$, so that $\mathrm{ad}L$ is solvable as well. On the other hand, is $\mathrm{ad}L$ is solvable, then $(\mathrm{ad}L)^m = 0$ for some m, hence $\mathrm{ad}L^m = 0$, so that L^m is in the center of L (it commutes with everything). Then, clearly $L^{m+1} = [L^m L^m] = 0$, hence L is solvable as well.

3.5 (part of it) Prove that the nonabelian 2-dimensional Lie algebra of section 1, question 4 is solvable. Do the same for the Lie algebra of section 1, question 2.

Solution

First, we consider the 2-dimensional Lie algebra L determined by $[xy] = x$. Clearly, the derived subalgebra $L^1 = [LL]$ is $\mathbb{C}x$, which is one-dimensional hence abelian, so that $L^2 = [L^1 L^1] = 0$. Hence, L is solvable.

Second, we consider the 3-dimensional Lie algebra L determined by $[xy] = z$, $[xz] = y$, $[yz] = 0$. Clearly, the derived subalgenra $L^1 = [LL]$ is $\mathbb{C}y + \mathbb{C}z$, which is abelian thanks to $[yz] = 0$, so that $L^2 = [L^1 L^1] = 0$. Hence, L is solvable.

5.3 Let L be the 2-dimensional Lie algebra of section 1, question 4, which is solvable (see 3.5). Prove that L has non-trivial (i.e. not identically zero) Killing form.

Solution

We may just calculate the Killing form explicitly. We have $[xy] = x$, so, with $x \mapsto \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $y \mapsto \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, we find

$$\mathrm{ad}x = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \mathrm{ad}y = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Hence, the Killing form is determined by

$$\kappa(x,x) = \mathrm{Tr}\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0, \quad \kappa(x,y) = \mathrm{Tr}\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0, \quad \kappa(y,y) = \mathrm{Tr}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = 1.$$

Hence, there is one element that is non-zero.

5.4 Let L be the 3-dimensional Lie algebra of section 1, question 2. Compute the radical of its Killing form.

Solution

The Lie algebra is determined by $[xy] = z$, $[xz] = y$, $[yz] = 0$. We first compute the adjoint representation, with $x \mapsto \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $y \mapsto \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, and $z \mapsto \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. We find

$$\mathrm{ad}x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \mathrm{ad}y = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad \mathrm{ad}z = \begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Hence, with $u = ax + by + cz$ for some complex numbers a, b, c, and similarly $u' = a'x + b'y + c'z$, we ask for the set of a, b, c such that $\kappa(u, u') = 0$ for all a', b', c'. We find

$$\kappa(u,u') = \mathrm{Tr}\begin{pmatrix} 0 & 0 & 0 \\ -c & 0 & a \\ -b & a & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 & 0 \\ -c' & 0 & a' \\ -b' & a' & 0 \end{pmatrix} = \mathrm{Tr}\begin{pmatrix} 0 & 0 & 0 \\ -ab' & aa' & 0 \\ -ac' & 0 & aa' \end{pmatrix} = 2aa'.$$

Then, we need $a = 0$, so that the radical is the subspace $\mathbb{C}y + \mathbb{C}z$ (which is an ideal, as it should).

5.5 Compute the basis of $sl(2;\mathbb{C})$ that is dual to the standard basis (see section 1, question 3) relative to the Killing form.

Solution

Using the expressions for $\mathrm{ad}x$, $\mathrm{ad}y$ and $\mathrm{ad}h$ obtained in section 1, question 3, we can compute the Killing form. It is simplest to express it at a matrix: we say $\kappa(u, v) = u^T K v$ where $u, v \in L$ are seen as column vectors as usual. We find

$$K = \begin{pmatrix} 0 & 0 & 4 \\ 0 & 8 & 0 \\ 4 & 0 & 0 \end{pmatrix}.$$

In order to find the dual basis, which we will denote by $\tilde{x}$, $\tilde{y}$ and $\tilde{z}$, we want to find a matrix M such that $KM = \mathbf{1}$. Then, we will simply define $\tilde{x} = Mx$, $\tilde{z} = Mz$ and $\tilde{z} = Mz$, and we will immediately see that $x^T K\tilde{x} = x^T KMx = x^t x = 1$, $x^T K\tilde{y} = x^T KMy = x^T y = 0$, etc. The matrix M is simply $M = K^{-1}$, and since K has non-zero determinant (i.e. the form κ is non-degenerate), we can compute it:

$$M = \frac{1}{8}\begin{pmatrix} 0 & 0 & 2 \\ 0 & 1 & 0 \\ 2 & 0 & 0 \end{pmatrix}.$$

Hence, the dual basis is

$$\tilde{x} = y/4, \quad \tilde{h} = h/8, \quad \tilde{y} = x/4.$$

6.1 (simplified) Using the standard basis for $sl(2;\mathbb{C})$ (see section 1, question 3), write down the Casimir element of the adjoint representation (see question 5.5). So the same for the usual 2-dimensional representation of $sl(2;\mathbb{C})$.

Solution

Given a basis x_i and its corresponding dual basis y_i according to the trace form $beta(x,y) = \mathrm{Tr}(\phi(x)\phi(y))$ of a representation ϕ, the Casimir element is $C_\phi = \sum_i \phi(x_i)\phi(y_i)$. Here, we consider the adjoint representation $\phi = \mathrm{ad}$, so the trace form is the Killin form. We computed the dual basis in 5.5. Hence, we can write

$$C_\kappa = \mathrm{ad}x\,\mathrm{ad}\tilde{x} + \mathrm{ad}y\,\mathrm{ad}\tilde{y} + \mathrm{ad}z\,\mathrm{ad}\tilde{z} = \frac{1}{4}\mathrm{ad}x\,\mathrm{ad}y + \frac{1}{4}\mathrm{ad}y\,\mathrm{ad}x + \frac{1}{8}(\mathrm{ad}h)^2.$$

Explicitly, as a matrix, this is simply

$$C_\kappa = \mathbf{1}$$

which is in agreement with the general formula $C_\phi = \dim(L)/\dim(V)\mathbf{1}$ for an irreducible representation space V. Note that here, the adjoint representation, which always has the same dimension as L itself, is irreducible because L is simple.

For the 2-dimensional representation ϕ, we have that $\phi(x)$, $\phi(y)$ and $\phi(z)$ are directly the matrices displayed in section 1, question 3 (first line). We must compute the dual basis according to the trace form. First, the form is $\beta(x,y) = x^T By$ with

$$B = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

so that the dual basis is

$$\tilde{x} = y, \quad \tilde{h} = h/2, \quad \tilde{y} = x.$$

Working out again the Casimir, we find

$$C_\phi = \phi(x)\phi(y) + \phi(y)\phi(x) + \frac{1}{2}\phi(h)\phi(h) = \frac{3}{2}\mathbf{1}$$

which is again in agreement with $C_\phi = \dim(L)/\dim(V)\mathbf{1}$ where now V has dimension 2.

6.4 Use Weyl's theorem to give a proof that for L semisimple, $\mathrm{ad}L = \mathrm{Der}(L)$, that is, all derivations of L are of the form $\mathrm{ad}y_0$ for some $y_0 \in L$. **Hint:** for any $D \in \mathrm{Der}(L)$, make the direct sum $\mathbb{C} \oplus L$ into a representation space for L (i.e an L-module) via the rule $\phi_D(x)(a, y) = (0, aD(x) + [xy])$.

Solution

First, let us check that $\mathbb{C} \oplus L$ is indeed a representation space. We have

$$\begin{aligned}[\phi_D(x), \phi_D(y)](a, z) &= \phi_D(x)(0, aD(y) + [yz]) - \phi_D(y)(0, aD(x) + [xz]) \\ &= (0, a[xD(y)] + [x[yz]]) - (0, a[yD(x)] + [y[xz]]).\end{aligned}$$

But also

$$\phi_D([xy])(a, z) = (0, aD([xy]) + [[xy]z]).$$

The first member is 0, so agree. The last terms in the second member also agree by the Jacobi identity, $[x[yz]] - [y[xz]] = [[xy]z]$. The first terms in the second member also agree for any a by the derivation property, $D([xy]) = [xD(y)] + [D(x)y]$. Hence, we indeed have $[\phi_D(x), \phi_D(y)] = \phi_D([xy])$. According to Weyl's theorem, any finite-dimensional representation of a semisimple Lie algebra is completely reducible. Here we do have a finite-dimensional representation. We also clearly have a sub-representation, if we restrict to the space $\{0\} \oplus L = \{(0, z) : z \in L\}$ (just set $a = 0$ above): this is certainly an invariant subspace. Since the representation must be completely reducible, then there must be a complement of this subspace that is also an invariant subspace. The complement is certainly one-dimensional, so there must exist a $y_0 \in L$ (that depends on the representation ϕ_D, hence on D) such that $S := \{(a, ay_0) : a \in \mathbb{C}\}$ is invariant (this is a complement, because the only element in common with $\{0\} \oplus L$ is $(0, 0)$). But for S to be invariant, i.e. $\phi_D(x)S \in S \;\forall\, x \in L$, we need

$$\phi_D(x)(a, ay_0) = (0, aD(x) + a[xy_0]) \in S \;\Rightarrow\; D(x) + [xy_0] = 0.$$

Hence, we find that $D(x) = [y_0 x] = \mathrm{ad}y_0(x)$, which shows the statement.

6.6 Let L be a simple Lie algebra. Let $\beta(x, y)$ and $\gamma(x, y)$ be two symmetric associative bilinear forms on L. If β and γ are nondegenerate, prove that β and γ are proportional.

Solution

First, the properties of β and γ are symmetry ($\beta(x, y) = \beta(y, x)$), associativity ($\beta([xy], z) = \beta(x, [yz])$) and bilinearity (and the same for γ). Let us use the notation $\beta(x, y) = x^T By$ and $\gamma(x, y) = x^T Cy$, with the column-vector notation for elements of L (see, e.g., question 5.5). Symmetry means $B^T = B$ and $C^T = C$, and associativity means $-(\mathrm{ad}x)^T B = B\mathrm{ad}x$ and $-(\mathrm{ad}x)^T C = C\mathrm{ad}x$ for all $x \in L$. Nondegeneracy means that both B and C are invertible. Hence, we have

$$B\,\mathrm{ad}x\,B^{-1} = C\,\mathrm{ad}x\,C^{-1} \;\forall\, x \in L.$$

Then, let us consider the adjoint representation. Since L has no proper ideals, then there is no invariant subspace of L under ad$L$, hence the adjoint representation is irreducible. Now consider the equation above. It gives $C^{-1}B\,\mathrm{ad}x = \mathrm{ad}x\,C^{-1}B$ for all $x \in L$. Using Shur's lemma, this means that $C^{-1}B = \mathbf{1}$. This proves the statement.

7.3 Verify that the formula

$$\begin{aligned}\phi(h)v_i &= (\lambda - 2i)v_i \\ \phi(y)v_i &= (i+1)v_{i+1} \\ \phi(x)v_i &= (\lambda - i + 1)v_{i-1}\end{aligned}$$

define a representation of $sl(2;\mathbb{C})$.

7.4 (have a look – but not essential)

8.4 Prove that each Cartan subalgebra of $sl(2;\mathbb{C})$ is one dimensional.

Solution

All we have to show is that there is no subalgebra of $sl(2;\mathbb{C})$ of dimension higher than 1 that is abelian. Since $sl(2;\mathbb{C})$ has dimension 3, and is nonabelian, we only need to look for a subalgebra of dimension 2. Suppose the subalgebra contains both x and y. Then it must contain $[xy] = h$, so it must have dimension 3, a contradiction. Hence, the subalgebra must contain only a one-dimensional subspace of the 2-dimensional space spanned by x and y. Suppose this is spanned by $ax + by$. The subalgebra is then spanned by $ax + by$ and h. The commutator is $[h(ax + by)] = 2ax - 2by$. This must be proportional to $ax + by$, hence we must have $a = 0$ or $b = 0$. But the subalgebras $\mathbb{C}h + \mathbb{C}x$ and $\mathbb{C}h + \mathbb{C}y$ are not abelian, because $[hx] = 2x$ and $[hy] = -2y$. Hence, there is no 2-dimensional abelian subalgebra.

Some additional Lie algebra exercises

1. Consider $sl(2;\mathbb{C})$ and the Cartan subalgebra

$$H = \left\{ \begin{pmatrix} 0 & a \\ a & 0 \end{pmatrix} : a \in \mathbb{C} \right\}.$$

Find the roots and the corresponding root space decomposition.

Solution

As an extra, let us first check that this is indeed a Cartan subalgebra. It is maximal because any bigger subalgebra will be the whole of $sl(2;\mathbb{C})$, hence will not be abelian. So we just have to check that it is ad-diagonalisable. We consider as usual the basis x, y and

h (see section 1, question 3). Note that $H = \mathbb{C}(x+y)$. In the adjoint representation, we have

$$\mathrm{ad}(x+y) = \begin{pmatrix} 0 & -2 & 0 \\ -1 & 0 & 1 \\ 0 & 2 & 0 \end{pmatrix}.$$

We have to check that this is diagonalisable. The eigenvalues are the roots of the characteristic polynomial,

$$-\lambda(\lambda^2 - 2) + 2\lambda = \lambda(4 - \lambda^2)$$

so that the roots of the polynomial are $\lambda = 0$ and $\lambda = \pm 2$. Since they are distinct, the matrix is diagonalisable. By the usual linear algebra techniques, we can calculate the corresponding eigenvectors:

$$x_2 = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} = y - x + h, \quad x_{-2} = \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} = y - x - h, \quad x_0 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = x + y.$$

Hence, we find two roots, α_2 and α_{-2}, with

$$\alpha_2(x+y) = 2, \quad \alpha_{-2}(x+y) = -2.$$

So, the root space decomposition relative to $H = \mathbb{C}(x+y)$ is simply

$$L = \mathbb{C}(x+y) \oplus \mathbb{C}(y - x + h) \oplus \mathbb{C}(y - x - h).$$

2. Prove that the *derived subalgebra* $[LL]$ of a Lie algebra L is an ideal. Recall that the derived subalgebra of a Lie algebra L is $[LL] = \mathrm{span}\{[xy] : x, y \in L\}$, the set of all linear combinations of elements of the form $[xy]$.

   Solution

   Consider any $x \in [LL]$ and any $z \in L$. We have $x = \sum_i [x_i y_i]$ by the definition of the derived subalgebra. Certainly, $[z[x_i y_i]] \in [LL]$, hence $[zx] \in [LL]$. Hence, $[LL]$ is an ideal.

3. Prove that the *center* $Z(L) = \{x \in L : [xy] = 0 \; \forall \; y \in L\}$ is an ideal.

   Solution

   Let $x \in Z(L)$ and $z \in L$. We have $[zx] = 0 \in Z(L)$, hence $Z(L)$ is an ideal.

4. Prove that the *radical* $\mathrm{rad}\beta = \{x \in L : \beta(x, y) = 0 \; \forall \; y \in L\}$ of any bilinear symmetric associative form $\beta(x, y)$ is an ideal of L.

   Solution—

   Properties: symmetric means $\beta(x, y) = \beta(y, x)$ and associative means $\beta([xz], y) = \beta(x, [zy])$. Then, let $x \in \mathrm{rad}\beta$ and $y, z \in L$. We have $\beta([xz], y) = \beta(x, [zy]) = 0$ by definition of $\mathrm{rad}\beta$. Hence, $[xz] \in \mathrm{rad}\beta$, so $\mathrm{rad}\beta$ is an ideal.

5. Let I be an ideal of L. Define $I^{\perp} = \{x \in L : \kappa(x, y) = 0 \; \forall \; y \in I\}$ where κ is the Killing form, $\kappa(x, y) = \mathrm{Tr}(\mathrm{ad}x \, \mathrm{ad}y)$. Show that $I^{\perp}$ is also an ideal of L.

   Solution

   Let $x \in I^{\perp}$, $y \in I$ and $z \in L$. Then $\kappa([xz], y) = \kappa(x, [zy]) = 0$ because $[zy] \in I$. Hence $[xz] \in I^{\perp}$, so that $I^{\perp}$ is an ideal.

6. Given a derivation D on L, prove that the bilinear form $\alpha(x, y) = \kappa(x, D(y))$, where κ is the Killing form $\kappa(x, y) = \mathrm{Tr}(\mathrm{ad}x \, \mathrm{ad}y)$, is *anti-symmetric*, $\alpha(x, y) = -\alpha(y, x)$.

   Solution

   The Killing form is $\kappa(x, y) = \mathrm{Tr}(\mathrm{ad}x\mathrm{ad}y)$. Consider $\kappa(x, D(y))$. Since D is a derivation, it satisfies $D([xy]) = [D(x)y] + [xD(y)]$. Note that

   $$\mathrm{ad}(D(y))(z) = [D(y)z] = D([yz]) - [yD(z)] = D(\mathrm{ad}y(z)) - \mathrm{ad}y(D(z)) = [D, \mathrm{ad}y](z).$$

   As operators on L, this equation means $\mathrm{ad}(D(y)) = [D, \mathrm{ad}y]$. Hence, $\kappa(x, D(y)) = \mathrm{Tr}(\mathrm{ad}x \, [D, \mathrm{ad}y]) = \mathrm{Tr}([\mathrm{ad}x, D] \, \mathrm{ad}y)$ using cyclicity of the trace. But then,

   $$[\mathrm{ad}x, D] = -[D, \mathrm{ad}x] = -\mathrm{ad}(D(x))$$

   hence we find $\kappa(x, D(y)) = -\kappa(D(x), y)$. Using symmetry of the Killing form, this is $\kappa(x, D(y)) = -\kappa(y, D(x))$, which is anti-symmmetry of $\alpha(x, y) = \kappa(x, D(y))$.