Handout III for the course GROUP THEORY IN PHYSICS — Mic$\hbar$ael Flohr
Relation between a Lie group G and its Lie algebra $\mathfrak{g}$ — 9. and 16. November 2004

BIG ‘Ad’ AND LITTLE ‘ad’

The representation theory of Lie groups G amounts essentailly in finding and studying Lie group homomorphisms $\rho : G \to H$, where $H \subset GL(V)$ is a matrix subgroup of the general linear group to a vector space V. The problem is that Lie groups are also (complicated) manifolds, such that ρ not only has to respect the group multiplication law, but also the differentiable structure of the manifold. The first step to the solution of this problem is to reduce everything to solely local information about the Lie group in a small neighborhood of the identity element. This information is encoded in the Lie algebra, i.e. the tangent space T_eG of the Lie group at the identity element. The (complicated) topology is then simply “forgotten”, and everything is then done in vector spaces which are much easier to handle. Of course, it is by no means clear that almost everything one wants to know about a Lie group can be expressed in local information. The following principles, however, ensure precisely that.

Principle (I). Let G and H be two Lie groups, G connected. A map $\rho : G \to H$ is uniquely determined by its differential $\mathrm{d}\rho_e : T_eG \to T_eH$ at the point corresponding to the identity element.

Principle (II). Let G and H be two Lie groups G, connected and simply connected. A linear map $T_eG \to T_eH$ is the differential of a group homomorphism $\rho : G \to H$ if and only if it respects the Lie bracket, i.e. $\mathrm{d}\rho_e([X,Y]) = [\mathrm{d}\rho_e(X), \mathrm{d}\rho_e(Y)]$ for all $X, Y \in T_eG$.

These two principles can be explicitly realised by introducing a certain map, namely the *adjoint* representation of the group on its own tangent space. This representation has per definitionem the dimension $\dim G$. The adjoint representation is the analog of the regular representation for finite groups. Of course, we cannot represent the group on itself, since the group is not a vector space.

Ad. The group multiplication from left $m_g : G \to G$, $m_g(h) = g \circ h$, is not well suited to be reduced to local information, because it has in general no fixpoint. In fact, group multiplication with a generic element g might lead us far away from the point h. However, group conjugation $\psi_g : G \to G$, which maps each group element h to $g \circ h \circ g^{-1}$, is much better. It will in general map a point h to a point near h. Its differential $\mathrm{Ad}(g) = (\mathrm{d}\psi_g)_e : T_eG \to T_eG$ should already contain some information about the structure of the group. It is important to understand that Ad(g) defines for each $g \in G$ a map $T_eG \to T_eG$ of the tangent space onto itself. Therefore, Ad: $G \to \mathrm{Aut}(T_eG)$ is per constructionem a representation of the group G on the vector space T_eG. The reader should check that this is indeed true, i.e. that Ad(g)Ad(h) =Ad$(g \circ h)$. The most important properties of Ad are shown in the two following diagrams which commute whenever $\rho : G \to H$ is a group homomorphism:

$$
\begin{array}{ccc}
G & \xrightarrow{\rho} & H \\
{\scriptstyle \psi_g}\downarrow & & \downarrow{\scriptstyle \psi_{\rho(g)}} \\
G & \xrightarrow{\rho} & H
\end{array}
\quad \xrightarrow{\mathrm{Ad}(g) = (\mathrm{d}\psi_g)_e} \quad
\begin{array}{ccc}
T_eG & \xrightarrow{(\mathrm{d}\rho)_e} & T_eH \\
{\scriptstyle \mathrm{Ad}(g)}\downarrow & & \downarrow{\scriptstyle \mathrm{Ad}(\rho(g))} \\
T_eG & \xrightarrow{(\mathrm{d}\rho)_e} & T_eH
\end{array}
\;.
$$

The second diagram reads as a formula as the condition $\mathrm{d}\rho(\mathrm{Ad}(g)(X)) = \mathrm{Ad}(\rho(g))(\mathrm{d}\rho(X))$ for all X, which are elements of the tangent space.

ad. The condition given above has one nasty drawback, namely that the map ρ still appears at one place explicitly. We can avoid this by considering the differential of Ad, ad: $T_eG \to \mathrm{End}(T_eG)$. It is crucial to understand that ad(X) defines for each $X \in T_eG$ a map $T_eG \to T_eG$ of the tangent space onto itself, i.e. ad$(X)(Y)$ is a linear map ad(X) of tangent vectors Y, and yields again a tangent vector in T_eG. Please note, however, that ad$(\cdot)$ might only be an endomorphisms, in contrast to Ad$(\cdot)$, i.e. it is not necessarily an automorphism. Considering ad and Ad (in some representation) as matrices, this means that ad(X) may have determinant zero, which cannot be the case for Ad(g), since this would contradict the group axioms. Now, ad$(X)(Y) : T_eG \times T_eG \to T_eG$ obviously is a bilinear map which motivates to introduce some bracket notation for it. The reader should check the anti-symmetry property! We define the *Lie bracket* as $[X,Y] \equiv \mathrm{ad}(X)(Y)$. A Lie group homomorphism $\rho : G \to H$ is characterised by the property that its differential respects the Lie bracket. This means that the diagram

$$
\begin{array}{ccc}
T_eG & \xrightarrow{(\mathrm{d}\rho)_e} & T_eH \\
{\scriptstyle \mathrm{ad}(X)}\downarrow & & \downarrow{\scriptstyle \mathrm{ad}(\mathrm{d}\rho(X))} \\
T_eG & \xrightarrow{(\mathrm{d}\rho)_e} & T_eH
\end{array}
$$

commutes. Expressed in formulæ, this amounts to the condition $\mathrm{d}\rho_e(\mathrm{ad}(X)(Y)) = \mathrm{ad}(\mathrm{d}\rho_e(X))(\mathrm{d}\rho_e(Y))$, or equivalently in the notaion with the Lie bracket, $\mathrm{d}\rho_e([X,Y]) = [\mathrm{d}\rho_e(X), \mathrm{d}\rho_e(Y)]$ for all $X, Y \in T_eG$.

In order to understand these rather abstract definitions of ‘Ad’ and ‘ad’, it helps a lot to view $g \in G$ and $X \in T_eG$ as matrices from $\mathrm{Aut}(V)$ or $\mathrm{End}(V)$, respectively, for an arbitrarily chosen V. This is what we have done in the lecture. For instance, one can put $G = \mathrm{GL}_n\mathbb{R}$. Then we would have $\mathrm{End}(\mathbb{R}^n) = M_n\mathbb{R}$, the set of $n \times n$ matrices with real entries. The operations ‘Ad’ and ‘ad’ can then be given explicitly. Chose an arbitrarily parametrized path $\gamma : I \to G$ in the manifold G with the properties $\gamma(0) = e$ und $\gamma'(0) = X$ for any fixed tangent vector $X \in T_eG$. Here and in the following, I is some finite interval. Without loss of generality, one can always assume $I = [-1, 1]$ Then we have $\mathrm{Ad}(\gamma(t))(Y) = \gamma(t) \circ Y \circ \gamma(t)^{-1}$ and the Lie bracket indeed takes the form of a commutator we all know:

$$[X, Y] = \mathrm{ad}(X)(Y) = \frac{\mathrm{d}}{\mathrm{d}t} \left(\mathrm{Ad}(\gamma(t))(Y)\right)\Big|_{t=0} = X \cdot Y - Y \cdot Y \, .$$

Lie algebra. A *Lie algebra* $\mathfrak{g}$ is a vector space together with a bilinear skew-symmetric map $[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, which satisfies the Jacobi identity. This definition implies implicitly a statement with far reaching consequences, which we can extract from the operation ‘ad’. A vector space together with a bilinear operation is a tangent space at the identity element of a Lie group if and only if this bilinear operation is skew-symmetric and fulfills the Jacobi identity. Our definiton of ad does start from a given Lie group G, and therefore yields per constructionem a skew-symmetric Lie bracket, $[X, Y] = -[Y, X]$, which automatically satisfies the Jacobi identity, since it is realised as commutation in every representation. However, the converse is also true. If we have a Lie algebra $\mathfrak{g}$, then we can (re)construct from its Lie bracket a group multiplication law and thus a Lie group.

A representation of a Lie algebra $\mathfrak{g}$ on a vector space V is simply a map between Lie algebras $\rho : \mathfrak{g} \to \mathfrak{gl}(V) = \mathrm{End}(V)$, i.e. a map which respects the Lie bracket such that we have for all $v \in V$ an operation of $\mathfrak{g}$ on V given by $[X, Y](v) = X(Y(v)) - Y(X(v))$.

Lie group versus Lie algebra. Let us summarize what we got so far: The tangent space $\mathfrak{g}$ at the identity element of a Lie group G is equipped in a natural way with the structure of a Lie algebra. Furthermore, if G and H are two Lie groups with G conencted and simply connected, then the maps $\rho : G \to H$ are in one-to-one correspondence to maps between associated Lie algebras by associating to each ρ the differential $(\mathrm{d}\rho)_e : \mathfrak{g} \to \mathfrak{h}$.

The Exponential Map

We (hopefully) have now a feeling, how we can reduce a Lie group to local information which is encoded in its Lie algebra. Now, we would like to see how we can get back from the Lie algebra to the group. Indeed, it is possible to reconstruct (almost all of) the group structure. To do so, we consider a fixed given tangent vector $X \in \mathfrak{g} = T_eG$. Further, let $\Gamma_X = \{\gamma : I \to G : \gamma(0) = e, \gamma'(0) = X\}$ be the set of all parametrized paths in G, which start at the identity element and leading there in the direction of X.

Vector fields. Givem a manifold M, we can define the ring of the differentiable functions $C^\infty(M)$. Each function $f \in C^\infty(M)$ assigns to each point $p \in M$ a (real) value, i.e. a point $x = f(p) \in \mathbb{R}$. Now, a *vector field* v assigns for a given f to each point $p \in M$ the tangent vector of f at the point p, i.e. $(v(f))(p) = v_p(f) \in T_pM$.

An elementary but fundamental theorem in differential geometry tells us that vector fields v on M can be integrated to functions $\phi : I \to M$ with boundary conditions $\phi(0) = p$ for some $p \in M$, and $\phi'(t) = v_{\phi(t)}$. The function ϕ is uniquely characterised by the choice of boundary (or initial) conditions. We remark here that the existence of ϕ is ensured by the fact that for each path $\gamma : t \in I \mapsto \gamma(t) \in M$ with I an open set of $\mathbb{R}$ containing zero, and for each function $q : t \in I \mapsto \mathbb{R}$, one can find a function $f \in C^\infty(M)$ such that $f(\gamma(t)) = g(t)$.

Left-translations. One of the implications that the manifold carries a group structure is that there are families of differentiable mappings of the Lie group manifold into itself, which act *transitively*. This means that for any two group elements g and g', there is a member of the family which maps g to g'. These mappins are given by the so-called *left-translations* $m_g : h \mapsto g \circ h$, which are nothing else than the group multiplications with g from left. Of course, one could define right-translations in the same fashion. The translations which map a given group element $g \in G$ to a prescribed $g' \in G$ are $m_{g' \circ g^{-1}}$. We note that the translations can in particular be used to map any point of G to the unit element of the group multiplication, e. It is easy to see that left-translations transport the basis of any tangent space T_gG to any other point of G in an invertible way.

Left-invariant vector fields. The existence of the left-translations admits to construct very special global vector fields with the property that they vanish nowhere on G. We can associate to each tangent vector $X \in \mathfrak{g}$ exactly one so-called *left-invariant vector field* $\tilde{X}$, such that $\tilde{X}_e = X$. First of all, $\tilde{X}$ is a vector field that assigns to a function $f \in C^\infty(G)$ and for each element $g \in G$ a tangent vector $\tilde{X}_g(f) \in T_gG$. The fact that it is left-invariant means that this assignment is compatible with the group multiplication from left, $m_g : G \to G$. A vector field $\tilde{X}$ with this property is easy to find:

$$\tilde{X}_g(g) = \mathrm{d}m_g X(f) = X(f \circ m_g) = \frac{\mathrm{d}}{\mathrm{d}t} \, f(g \circ \gamma(t))\big|_{t=0} \, ,$$

where γ is an arbitrary element of Γ_X. In fact, we obviously then have

$$\mathrm{d}m_h\tilde{X}_g(f) = \mathrm{d}m_h\mathrm{d}m_gX(f) = X(f\circ m_g\circ m_h) = \frac{\mathrm{d}}{\mathrm{d}t}\, f(h\circ g\circ\gamma(t))|_{t=0} = \tilde{X}_{hg}(f)\,.$$

The left-invariant vector field $\tilde{X}$ does nothing else than to transport the tangent vector $X\in T_eG$ in a way compatible with the group multiplication law m_g to a tanget vector in T_gG. Thinking of a basis in T_eG, we see that we get a moving frame which moves in accordance with the group multiplication law to a basis in T_gG. This means that the vector bundle over G where we assign to each g the tangent space T_gG is trivial. Such manifolds are called *parallelizable*.

Of course, left-invariant vector fields can be integrated as well, and our boundary conditions are now that $\phi(0) = e$ and $\phi'(t) = \tilde{X}_{\phi(t)}$. The left-invariance of the special vector field $\tilde{X}$ together with the uniqueness of the integration curve has the consequence that ϕ, whereever it is defined, is a group homomorphism, i.e. $\phi(s+t) = \phi(s)\circ\phi(t)$ for $s,t\in I$. Let us write this down in the following way

$$\tilde{X}_{\phi(s)}(f) = \frac{\mathrm{d}}{\mathrm{d}t}\, f(\phi(s)\circ\phi(t))|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t}\, f(\phi(s+t))|_{t=0}\ ,$$

since it is clear from the boundary conditions for ϕ that $\phi\in\Gamma_X$.

One-parameter subgroups. The existence of left-invariant vector fields to given tangent vectors $X\in\mathfrak{g}$ gives us integral curves, which simultaneously are group homomorphisms. These are therefore called *one-parameter subgroups*. Due to the group structure, these one-parameter subgroups $\phi_X(t)$, which we only defined locally for $t\in I$ and thus for a small neighborhood around e, are automatically well defined for $t\in\mathbb{R}$. Another way to say this is the statement that for each $X\in\mathfrak{g}$ there exists exactly one path in the family Γ_X, which is a group homomorphism. This path is precisely the integral curve of the left-invariant vector field $\tilde{X}$. Since this works for all $X\in T_eG$, the set of all one-parameter subgroups will completely cover a neighborhood of the identity element. Since, by use of the group law, any arbitrarily small neighborhood of the identity generates the full (connected component of the identity element of the) group G, we finally get the desired result that the information encoded in the Lie algbera $\mathfrak{g}$ suffices to reconstruct the group (to a large extent).

Exponential map. The integral curve satisfies the functional equation $\phi(s+t) = \phi(s)\circ\phi(t)$. This is exactly the functional equation of the exponential function. One therefore defines

$$\begin{array}{rcl}\exp:\mathfrak{g} & \to & G\\ \exp(X) & = & \phi_X(1)\end{array}\ .$$

Since ϕ is unique, we obviously have $\phi_{\lambda X}(t) = \phi_X(\lambda t)$. The exponential map, restricted to lines through the origin of $T_eG = \mathfrak{g}$, exactly yields the one-parameter subgroups. More precisely, the exponential map is the *unique* map $\mathfrak{g}\to G$, which sends the origin to the identity element, $0\mapsto e$, whose differential at the origin is the identity, i.e.

$$(\mathrm{d}\exp)_0 : T_0\mathfrak{g} = \mathfrak{g}\to T_eG = \mathfrak{g}\,,$$

and whose restriction to lines through the origin yields the one-parameter subgroups. This map is natural in the sense that for arbitrary Lie group mappings $\rho: G\to H$ we have that the diagram

$$\begin{array}{ccc}\mathfrak{g} & \xrightarrow{(\mathrm{d}\rho)} & \mathfrak{h}\\ {\scriptstyle\exp}\downarrow & & \downarrow{\scriptstyle\exp}\\ G & \xrightarrow{\rho} & H\end{array}$$

commutes. This statement allows us to study the theory of representations of a Lie group by looking at the representations of its Lie algebra!

Since $(\mathrm{d}\exp)_0$ in $\mathfrak{g}$ is an isomorphism, the image $\mathrm{Im}(\exp)\supset U$ must contain a neighborhood of the identity element e in G. If G is connected, then U generates the whole group G which puts Principle (I) on firm grounds. Moreover, we get the following simple relation between 'Ad' and 'ad': $\mathrm{Ad}(\exp(X)) = \exp(\mathrm{ad}(X))$, which one deduces with the help of the so-called *Baker-Campbell-Hausdorff formelæ*. Thus, we also obtain principle (II).

Baker-Campbell-Hausdorff. We can use the exponential map to assign to elements of the Lie algebra $\mathfrak{g}$ elements of the corresponding Lie group G. But how is the group multiplication law implemented, i.e. how do we find the element $Z\in\mathfrak{g}$, such that $\exp(X)\circ\exp(Y) = \exp(Z)$ is satisfied? One can do this explicitly, when (in a given representation) one realizes the Lie group and its algebra by matrices. Then, the exponential map is nothing else than

$$\exp(X) = \sum_n \frac{1}{n!} X^n \,,$$

which converges and is invertible with the inverse $\exp(-X)$. Obviously, we have $(\mathrm{d}\exp)_0 = 1$. We immediately obtain for the one-parameter subgroups

$$\exp(\lambda X)\exp(\mu X) = \sum_n \sum_m \frac{1}{n!}\frac{1}{m!}\lambda^n \mu^m X^{n+m} = \sum_N \sum_{k=0}^{N} \frac{1}{N!}\binom{N}{k}\lambda^k \mu^{N-k} X^N = \sum_N (\lambda{+}\mu)^N X^N = \exp((\lambda{+}\mu)X)\,.$$

But the whole group structure is already hidden in its Lie algebra. To see this, let X, Y be chosen from a sufficiently small neighborhood of the origin $0 \in \mathfrak{g}$. Furthermore, consider for $g \in G \subset \mathrm{GL}_n\mathbb{R}$ the map

$$\log(g) = -\sum_n \frac{(-)^n}{n}(g-e)^n \in \mathfrak{gl}_n\mathbb{R}\,,$$

which is, of course, only valid for such g which lie in a sufficiently small neighborhood of the identity element. But wherever this map is defined, it is obviously the inverse of the exponential map. With all this we are ready to define the *Baker-Campbell-Hausdorff product*

$$X * Y = \log(\exp(X) \circ \exp(Y))\,.$$

The crucial point is not the explicit form of $X * Y$, but the fact that the result depends only on X, Y and the operations ad(X) and ad(Y). The first few terms read as follows:

$$\begin{aligned} X * Y &= (X+Y) + \tfrac{1}{2}[X,Y] + \tfrac{1}{12}\left([X,[X,Y]] + [Y,[Y,X]]\right) + \ldots \\ &= (X+Y) + \tfrac{1}{2}\mathrm{ad}(X)(Y) + \tfrac{1}{12}\left(\mathrm{ad}(\mathrm{ad}(X))(Y) + \mathrm{ad}(\mathrm{ad}(Y))(X)\right)\ldots \\ &= \left(1 + \tfrac{1}{4}(\mathrm{ad}X - \mathrm{ad}Y) + \tfrac{1}{12}(\mathrm{ad}^2 X + \mathrm{ad}^2 Y) + \ldots\right)(X+Y)\,. \end{aligned}$$

Note that there are in particular no terms such as X^n, but all terms can be collected in such a way that they can be expressed solely in terms of the Lie algebra and their Lie brackets. This implies that we do not need matrix multiplication (which we need to explicitly compute commutators) but only linearity on $\mathfrak{g}$ and the Lie bracket. The proof of such formulæ is not easy, but Dynkin managed to find a closed form of the B-C-H product. This handout closes with the uncommented display of an intergral representation of the B-C-H product,

$$X * Y = X + \int_0^1 g(\exp(\mathrm{ad}X) \circ \exp(t\,\mathrm{ad}Y))(Y)\mathrm{d}t\,, \quad g(z) = \frac{\log(z)}{1-\frac{1}{z}} = 1 - \sum_{n=1}^{\infty} \frac{(-)^n}{n(n+1)}(z-1)^n\,.$$

The key point is that again X and Y themselves appear only linearly, all other terms involve only the commutator operation (the Lie bracket). Furthermore, one sees that the expansion in a series makes sense, since the term with the identity element just cancels,

$$X * Y = X + Y + \int_0^1 \mathrm{d}t \sum_{n=1}^{\infty} \frac{(-)^{n+1}}{n(n+1)}\left(\exp(\mathrm{ad}X) \circ \exp(t\,\mathrm{ad}Y) - e\right)^n(Y)\,.$$

A useful formula. The problem of computing exponentials of non-commuting objects is well-known from quantum mechanics. In particular, one often has to transform a given observable A by a unitary transformation, i.e. $A \mapsto UAU^\dagger$. Thus, as we now can appretiate a bit better, the unitary transformation will typically be given as the exponential of the generator of the corresponding infinitesimal transformation, namely $U = U_\lambda(X) = \exp(\mathrm{i}\lambda X)$. Thus, one needs an explicit expression for $\exp(\mathrm{i}\lambda X)A\exp(-\mathrm{i}\lambda X)$. Of course, X is a Hermitean operator and λ a real parameter for the one-parameter subgroup, which is generated by X. Note that we now switched back to the convention most often used in physics, where the generatros of transformations are chosen Hermitean. In mathematics, one typically uses anti-Hermitean generators, but real instead of purley imaginary coefficients. With the B-C-H product one finds

$$\begin{aligned} \exp(\mathrm{i}\lambda X)A\exp(-\mathrm{i}\lambda X) &= A + \mathrm{i}\lambda[X,A] + \frac{(\mathrm{i}\lambda)^2}{2!}[X,[X,A]] + \ldots + \frac{(\mathrm{i}\lambda)^n}{n!}[X,[X,[X,\ldots[X,A]]]\ldots] + \ldots \\ &= \left(\sum_{n=0}^{\infty} \frac{(\mathrm{i}\lambda)^n}{n!}\mathrm{ad}^n X\right)(A) \\ &= \exp(\mathrm{i}\lambda\,\mathrm{ad}X)A\,, \end{aligned}$$

which also demonstrates once more, how useful the notation ad$(X)(\cdot) = \mathrm{ad}X(\cdot)$ is, even if we mean nothing else than the commutator $[X, \cdot\,]$.