Notes 5: The Baker-Campbell -Hausdorff formula.

Version 0.00 — with misprints,

A usefull power series.

There is a power series that plays a central role in the formulation of the Baker-Campbell-Hausdorff formula, and we start notes5 with introducing that series: Let $\psi(z)$ be the complex function

$$\psi(z) = \frac{z \log z}{z-1}.$$

It is analytic in the open disk $\{ z \mid |z-1| < 1 \}$ since $\log z$ is analytic there and has a zero at $z = 1$. The functiom ψ has a power series expansion in the disk looking like:

$$\psi(z) = \frac{z \log z}{z-1} = z \sum_{n=0}^{\infty} \frac{(-1)^n (z-1)^n}{n+1},$$

or if we substitute $z = 1 + u$, the development takes the form

$$\psi(1+u) = (1+u) \sum_{n=0}^{\infty} \frac{(-1)^n u^n}{n+1} = 1 + \frac{u}{2} - \frac{u^2}{6} + \dots . \qquad (\ast)$$

This series is absolutly convergent for $|u| < 1$. As usual, if a is any matrix we can form $\psi(a)$ just by plugging in a in the series above, and the result has a meaning whenever the series converges. The plugged in series is absolutely convergent as least if $\|a - I\| < 1$ which folloes from standard estimates using the norm $\|a\|$. Hence for such matrices:

$$\psi(a) = a \sum_{n=0}^{\infty} \frac{(-1)^n (a-I)^n}{n+1}.$$

The property of ψ that interests us the most is the following:

$$\psi(e^z) = \frac{z}{1-e^{-z}}, \qquad (\maltese)$$

or

$$\psi(e^z)^{-1} = \frac{1-e^{-z}}{z} = \sum_{n=1}^{0} \frac{(-1)^n z^n}{(n+1)!}$$

These equalities are purely formal, and hold therefore for any matrix, as long as the involved series converge.

Formulation of th BCH-formula

The formulation is a little cryptic, and there are several variations of it around. The background of the formula is the natural desire to understand the realtion between the product in the Lie group and the commutator product in the Lie algebra. More precisely, the question is: Can one express the product $\exp v \cdot \exp w$ as the exponential of some tractable function of v and w and theire nested brackets? The answere is basically yes at least for $\exp v$ and $\exp w$ close to the unit element; and there is in fact a *universal* function. One can always discuss to what extent it is tractible, but there are explicit, if not closed, expressions for it.

In any Lie group G there are open sets $U' \subseteq \operatorname{Lie} G$ and $U \subseteq G$ such the restriction of the exponential map is a diffeomeorphism between U' and U. The inverse of exp map which is defined in U, is naturally babtized the *logarithm*, and equally naturally, it is denoted by $\log\colon U \to \operatorname{Lie} G$.

Here comes BCH formula:

Theorem 1 *Let G be a Lie group, and let v, w be elements in the Lie algebra $\operatorname{Lie} G$. Then*

$$\log(\exp v \exp w) = v + \int_0^1 \psi(\exp \operatorname{ad}_v \exp t\operatorname{ad}_w) w \, dt,$$

whenever $\log(\exp v \exp w)$ is defined.

This fabulous formula merits several explanatory remarks.

i) First of all, the appearence of log in the statement is not only because of typographical reasons. The content of the formula is of course that the product of $\exp v$ and $\exp w$ is the exp of the right side in the formula, but there is the subtility that for this to be true, the product $\exp v \exp w$ has to be close enough to the unit e so that log is defined.

EXAMPLE— $\operatorname{Sl}(2, \mathbb{R})$.

Recall from example ?? that the elements in $\operatorname{Sl}(2, \mathbb{R})$ which are of the form $\exp v$ with $v \in \operatorname{sl}(2, \mathbb{R})$ are those whose eigenvalues are either real and positive, or purely imaginary.

Regard the following two elements in $\operatorname{Sl}(2, \mathbb{R})$:

$$a = \begin{pmatrix} -x & y \\ -y^{-1} & 0 \end{pmatrix} \text{ and } b = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix}$$

where $\lambda > 0$, $y \neq 0$, and $x > 0$, but $x^2 < 4$. Then b is in the image of the exponential, since the eigenvalues are positive. The discriminant of the characteristic polynomial of a being $(\operatorname{tr} a)^2 - 4 = x^2 - 4$, the eigenvalues of a are purely imaginary and a is in the image of the exponential map. But $\operatorname{tr} ba = -\lambda x < 0$ and by chosing λ big enough, $(\operatorname{tr} ba)^2 = \lambda^2 x^2 > 4$ and ba has real negative eigenvalues. Hence it is not in the image of the exponential. ❋

ii) Recall that ad_v and ad_w are endomorphisms of $\operatorname{Lie} G$, given as $\operatorname{ad}_v x = [v, x]$ and $\operatorname{ad}_w x = [w, x]$. The exponential in the expressions $\exp \operatorname{ad}_v$ and $\exp t\operatorname{ad}_w$ are the usual exponential of endomorphisms (or matrices if you want to work with a basis), given by the usual power series, so they are also endomorphisms of $\operatorname{Lie} G$, and they are given as the power series

$$\begin{aligned} \exp \operatorname{ad}_v &= I + \operatorname{ad}_v + \operatorname{ad}_v^2/2 + \operatorname{ad}_v^3/3! + \dots \\ \exp t\operatorname{ad}_w &= I + t\operatorname{ad}_w + t^2\operatorname{ad}_w^2/2 + t^3\operatorname{ad}_w^3/3! + \dots . \end{aligned}$$

In the formula in the theorem we are to subsitute their product as endomorphisms into the power series for ψ, apply the resulting endomorphism to w and then integrate with respect to t.

There is a simplifying aspect in this rather complex procedure. Resulting terms having an ad_w to uttermost the right, will disappear when applied to w, so there is no need to keep track of them in the calculations.

It is worthwhile to see what this procedure gives for the first few terms:

$$\begin{aligned} u = \exp \operatorname{ad}_v \exp t\operatorname{ad}_w - I &= \operatorname{ad}_v + \frac{1}{2}\operatorname{ad}_v^2 + t\operatorname{ad}_w + \frac{t^2}{2}\operatorname{ad}_w^2 + \dots \\ u^2 &= \operatorname{ad}_v^2 + t\operatorname{ad}_w\operatorname{ad}_v + \dots \end{aligned}$$

and then

$$\int_o^1 (1 - \frac{u}{2} - \frac{u^2}{6} \dots) w \, dt = w + \frac{1}{2}\operatorname{ad}_v w + \frac{1}{12}\operatorname{ad}_v^2 w - \frac{1}{12}\operatorname{ad}_w\operatorname{ad}_v w$$

and hence

$$\log(\exp v \exp w) = v + w + \frac{1}{2}[v, w] + \frac{1}{12}[v, [v, w]] + \frac{1}{12}[w, [w, v]] + \dots$$

iii) May be the most important feature of the Baker-Campbel-Hausdorff formula, is that the expression for $\log \exp v \exp w$ only involves iterated brackets og v and w, so $\log \exp v \exp w$ is completely determined by the subalgebra generated by v and w. A nice and important consequence of this is the following theorem that goes back to Sophus Lie:

Theorem 2 *Let G be connected Lie group and let $L \subseteq \operatorname{Lie} G$ be a subalgebra, then the subgroup generated by $\exp L$ is an immersed subgroup with Lie algebra L.*

PROOF: As above, the open sets $U' \subseteq \operatorname{Lie} G$ and $U \subseteq G$ are such the restriction of the exponential map is a diffeomeorphism between U' and U, and we let $W \subseteq G$ be an open set with $W \cdot W \subseteq U$

Let $V' = U' \cap L$, and let $V = \exp V'$, then $V \subseteq G$ is a submanifold. We put $V_0 = W \cap V$. We claim that then $V_0 \cdot V_0 \subseteq V$. Indeed, elements in $V_0 = W \cap V$ are of the form $\exp v$ with $v \in V'$, so the product of two of them $\exp v \exp w$ is in U where logarithm is defined, hence by the BCH-formula, is in V.

Thus V_0 defines a local Lie group with algebra L, and $h_I V_0$ where h_I runs through all finite products $h_1, \ldots, h_s$ of elements from V_0, is an atlas [1] for the group generated by V_0, namely: $\bigcup h_I V_0$. ❑

Corollary 1 *Let G and H be Lie groups and assume that G is simply connected (and connected). If $\Phi\colon \operatorname{Lie} G \to \operatorname{Lie} H$ is a Lie algebra homorphism, then there exists a unique map $\phi\colon G \to H$ of Lie groups with $\Phi = d_e\phi$.*

PROOF: Let $L \subseteq \operatorname{Lie} G \times \operatorname{Lie} H$ be the graph of Φ. Then in fact $L = \operatorname{Lie} G$, and about the two projections we can say that $\pi_{\operatorname{Lie} G}$ is the identity and $\pi_{\operatorname{Lie} H} = \Phi$. The algebra L is a subalgebra of $\operatorname{Lie} G \times \operatorname{Lie} H = \operatorname{Lie}(G \times H)$, and by the theorem it is the Lie algebra of a (weak) subgroup $\Gamma \subseteq G \times H$. The first projection π_G induces — by the standard translation argument and the inverse funcion theorem — a local diffeomorphism $\Gamma \to G$ since its derivative at (e, e) is the identity. Since G is simply connected $\pi_G|_\Gamma$ must be an isomorphism, and Γ is the graph of $\phi = \pi_H \circ \pi_G|_\Gamma^{-1}$, whose derivative is $d_e\phi = d_{e,e}\pi_H d_e|_\Gamma^{-1} = d\pi_{\operatorname{Lie} H} = \Phi$.

Uniqueness was allready proved in Notes1 .

❑

This shows that the functor [2] Lie: $\mathsf{Liegrps}_0 \to \mathsf{Liealg}$ is fully faithfull; *i.e.*,

$$d_e(\star)\colon \operatorname{Hom}_{\mathsf{Liegrps}_0}(G, H) \simeq \operatorname{Hom}_{\mathsf{Liealg}}(\operatorname{Lie} G, \operatorname{Lie} H).$$

That Lie is essentially surjective follows from the theorem and the following result due to Ado:

[1] Admittedly, this is not very clear, one has to do somelthing like constrducting a manifold as the disjoint union $\dot{\bigcup} h_I V_0$ mod the equivalence relation generated by relations like $h_I v = h_J v'$ and see that that manifold immerses into G. See chapter 8 in van den Bans notes for details.

[2] Where as usual $\mathsf{Liegrps}_0$ denotes the category of simply-connected Lie groups.

Theorem 3 *If L is a finite dimensional real Lie algebra, then L has a faithfull representation, or stated a bit differently, for some natural number n, there is an embedding $L \subseteq \mathrm{gl}(n, \mathbb{R})$.*

A formula for the derivative of $\exp v + tw$

The core of the proof of BCH-formula we shall present, is a formula for the derivative of the exponential map not only at origin, but at any point $v \in \operatorname{Lie} G$. To simplyfy the notaion, and thus making the proof more transperant and easier to absorbe, we shall give the proof of the formula only in the matrix case. The formulation in that case reads as follows in proposition 1. In the general case, stated below in proposition 2, the formulation is completely analogue, just a factor $e^\star$ being replaces by the derivative of a translation $d_e\lambda_{\exp w}$.

Proposition 1 *Let v and w be matrices in $\mathrm{M_n}(\mathbb{K})$. Then*

$$\frac{d}{dt}|_{t=0} e^{v+tw} = e^v \int_0^1 e^{-s\,\mathrm{ad}_v w}\, ds = e^v \frac{I - e^{-\mathrm{ad}_v}}{\mathrm{ad}_v} w$$

Or in the general formulation:

Proposition 2 *Let G be a Lie group and let v and w be elements in the Lie algebra* $\operatorname{Lie} G$. *Then*

$$d_v \exp w = d_e\lambda_{\exp v} \int_0^1 \exp(-s\,\mathrm{ad}_v w)\, ds = d_e\lambda_{\exp v} \frac{I - \exp(-\mathrm{ad}_v)}{\mathrm{ad}_v} w.$$

The expression $\frac{I-\exp(-\mathrm{ad}_v)}{\mathrm{ad}_v} w$ merits an explanation; the function $(1 - e^{-z}/z)$ has a power series expansion

$$(1 - e^{-z}/z) = \sum_{n=1}^{\infty}) \frac{(-1)^{n+1} z^{n-1}}{n!} = \int_0^1 \sum_{n=0}^{\infty} \frac{(-1)^n s^n z^n}{n!}\, ds,$$

converging for any z. The meaning of the expression is to plugg in ad_v in that power series and apply the resulting operator to w, that is

$$\frac{I - \exp(-\mathrm{ad}_v)}{\mathrm{ad}_v} \cdot w = \sum_{n=1}^{\infty} \frac{(-1)^{n+1} \mathrm{ad}_v^{n-1}}{n!} \cdot w = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n!} \underbrace{[\dots[v,[v,[v,w]]]\dots]}$$

where in the last sum there are nested $n-1$ brackets.

PROOF: The trick is to study the function

$$e^{-sv}e^{s(v+tw)}$$

and give a formula for the double derivative of. First taking the derivative with respect to t and putting $t = 0$, and subsequently taking the derivative with resprect to s, we obtain a function whose integral with respect to s is the desired expression; that is

$$\frac{\partial}{\partial t}\Big|_{t=0} e^{v+tw} = e^v \int_0^1 \frac{\partial}{\partial s}\frac{\partial}{\partial t}\Big|_{t=0} e^{-sv}e^{s(v+tw)}\, ds$$

Now, the order of the two differential operations can swapped, so we first differentiat wit respect to s, to get

$$\frac{\partial}{\partial s} e^{-sv}e^{s(v+tw)} = -ve^{-sv}e^{s(v+tw)} + e^{-sv}(v+tw)e^{s(v+tw)} = e^{-sv}twe^{s(v+tw)},$$

where we used the product rule, the differential equation satisfied by the exponential, and finally that v and e^{-sv} commute. Taking the derivative with respect to t and setting $t = 0$, we see that

$$\frac{\partial}{\partial s}\frac{\partial}{\partial t}|_{t=0} e^{-sv}e^{s(v+tw)} = \frac{\partial}{\partial t}|_{t=0}\frac{\partial}{\partial s} e^{-sv}e^{s(v+tw)} = e^{-sv}we^{sv} = e^{-s\mathrm{ad}_v w},$$

and we are through. In the general case, the proof is word by word the same, only that the factors $e^{\pm v}$ have to replaced by $d_e\lambda_{\exp \pm v}$. ❑

The following corollary follows directly from 2 by use of the chain rule:

Corollary 2 *If $v(t)$ is smooth function defined on an interval I taking values in* Lie G, *then*

$$d_{v(t)} \exp v'(t) = d_0\lambda_{\exp(v(t))} \frac{I - \exp(-\mathrm{ad}_{v(t)})}{\mathrm{ad}_{v(t)}} v'(t).$$

Proof of BCH-formula

Life should not be too confortable, so we give that proof in the general case — it is also a point that the reader should be reminded of the general formalism of Lie groups.

Let $z(t)$ be the curve $z(t) = \log(\exp v \exp tw)$, that is $\exp z(t) = \exp v \exp tw$. Computing the derivative of zt we get on one hand

$$d_{z(t)} \exp z'(t) = d_{\exp tw}\lambda_{\exp v} d_t \exp tw = d_{\exp tw}\lambda_{\exp v} d_0\lambda_{\exp tw} w = \qquad (\star)$$
$$= d_0\lambda_{\exp v \exp tw} w = d_0\lambda_{\exp z(t)} w.$$

On the other hand, proposition 2 gives

$$d_{z(t)} \exp z'(t) = d_0\lambda_{\exp(z(t))} \frac{I - \exp(-\mathrm{ad}_{z(t)})}{\mathrm{ad}_{v(t)}} z'(t).$$

Combining this with equation ☆ above and using the usefull property ✦ of ψ on page 1 , we simply get

$$w = \psi(\exp(\mathrm{ad}_{z(t)}))^{-1} z'(t)$$

wich gives on integration

$$z(1) = \log \exp v \exp w = \int_0^1 \psi(\exp(\mathrm{ad}_{z(t)})) w \, dt.$$

The last step in the proof is a closer look at the term $\exp(-\mathrm{ad}_{z(t)})$.

Recalling the proposition 6 on page 9 in Notes4 saying that $Ad_* \circ \exp = \exp \circ \mathrm{ad}_*$, we see that

$$\mathrm{Ad}_{\exp zt} = \mathrm{Ad}_{\exp v \exp tw} = \mathrm{Ad}_{\exp v} \mathrm{Ad}_{exptw} = \exp \mathrm{ad}_v \exp t\mathrm{ad}_w$$

and

$$\mathrm{Ad}_{\exp z(t)} = \exp \mathrm{ad}_{z(t)}$$

from which we get $\exp(\mathrm{ad}_{z(t)}) = \exp(\mathrm{ad}_v) \exp(t\mathrm{ad}_w)$, and we are through.

Two important theorems we need,

but unfortunately do not have time to prove.

The first one goes back to Elie Cartan. It is kind of strong, and is an example of the extensive "bootstrapping" one sees in the theory of Lie groups: Structures that *a priori* are only topological turn often out to be differentiable, or even analytic.

Theorem 4 *Let $H \subseteq G$ be a closed Lie group. Then H is a submanifold, i.e., it is a Lie subgroup in the strong sense*

Corollary 3 *If $\phi\colon G \to H$ is a continuous homorphism between two Lie groups, then ϕ is differentiable.*

PROOF: The graph $\Gamma \subseteq G \times H$ is a closed since ϕ is continuous and a subgroup since ϕ is a group homomorphism. Hence it is a submanifold by the theorem, and therefore ϕ is differentiable. ❑

Corollary 4 *If $\phi\colon G \to H$ is a group homorphism being a homeomorphism, then it is a diffeomorphism.*

Corollary 5 *A Lie group has only one differentiable structure.*

I have some notes about this, but only in Norwegian.

The other theorem we shall need, but do not prove is the following:

Theorem 5 *Let $H \subseteq G$ be a strong subgroup. Then the set G/H of cosets has a unique structure as a smooth manifold such that the canonical map $\pi\colon G \to G/H$ sending g to gH is differentiable.*

There is an exact sequence of tangent spaces

$$0 \longrightarrow \operatorname{Lie} H \longrightarrow \operatorname{Lie} G \xrightarrow{d_e\pi} T_{\pi e} G/H \longrightarrow 0$$

hence $\dim G/H = \dim G - \dim H$.

If H is a normal subgroup, G/H is a Lie group with π being a Lie group map.