The proof of this lemma is a bit hard to read. Here are some extra details.

Lemma 8.2. *Let* $X, Y \in \mathfrak{g}$ *then*

$$d(\exp)_X(Y) = d(\ell_{\exp X})_e \left(\int_0^1 \mathrm{Ad}(\exp(-sX))Y \, ds \right),$$

where the integral is taken of a $\mathfrak{g}$*-valued function.*

Proof. As in the notes we define $F(X, Y) = d(\ell_{\exp X})_e^{-1} \circ d(\exp)_X(Y) \in \mathfrak{g}$. The statement of the lemma is then

$$F(X, Y) = \int_0^1 \mathrm{Ad}(\exp(-sX))Y \, ds. \tag{1}$$

We will prove the identity (1) in $\mathfrak{g}$ by applying both sides to an arbitrary smooth function $\varphi \in C^\infty(G)$. More precisely we will show that

$$d\varphi_e\left(F(X, Y)\right) = d\varphi_e \left(\int_0^1 \mathrm{Ad}(\exp(-sX))Y \, ds \right) \tag{2}$$

for all such functions, which clearly then implies (1).

Note that by definition of vector valued integration the linear functional $d\varphi_e$ passes under the integral sign on the right. Hence

$$d\varphi_e\left(F(X, Y)\right) = \int_0^1 d\varphi_e\left(\mathrm{Ad}(\exp(-sX))Y\right) \, ds \tag{3}$$

is an equivalent form of (2).

As mentioned in the notes we find from the chain rule that

$$F(X, Y) = \frac{\partial}{\partial t}\Big|_{t=0} \exp(-X)\exp(X + tY) \in T_eG = \mathfrak{g}.$$

For the proof of (3) we let

$$g(s, t) = \exp(-sX)\exp(s(X + tY)) \in G$$

for $s, t \in \mathbb{R}$, and note that

$$F(sX, sY) = \frac{\partial}{\partial t}\Big|_{t=0} g(s, t).$$

Then by the chain rule

$$d\varphi_e(F(sX, sY)) = \frac{\partial}{\partial t}\Big|_{t=0} \varphi(g(s, t)).$$

The reason for introducing the extra variable s is that we want to exploit the fundamental theorem of calculus to write $F(X,Y)$ as an integral. Since $F(0,0)=0$ we find

$$d\varphi_e(F(X,Y)) = \int_0^1 \frac{\partial}{\partial s} d\varphi_e(F(sX,sY))\,ds. \tag{4}$$

Now

$$\frac{\partial}{\partial s} d\varphi_e(F(sX,sY)) = \frac{\partial}{\partial s}\frac{\partial}{\partial t}\Big|_{t=0}\varphi(g(s,t)) = \frac{\partial}{\partial t}\Big|_{t=0}\frac{\partial}{\partial s}\varphi(g(s,t)). \tag{5}$$

by interchanging partial derivatives of a smooth real valued function.

Note that

$$g(s+u,t) = \exp(-sX)g(u,t)\exp(s(X+tY))$$

for $s,t,u \in \mathbb{R}$ and hence

$$\varphi(g(s+u,t)) = (\varphi \circ l_{\exp(-sX)} \circ r_{\exp(s(X+tY))})(g(u,t)).$$

Since $\frac{\partial}{\partial s}\varphi(g(s,t)) = \frac{\partial}{\partial u}\Big|_{u=0}\varphi(g(s+u,t))$ it follows

$$\frac{\partial}{\partial s}\varphi(g(s,t)) = d\left(\varphi \circ l_{\exp(-sX)} \circ r_{\exp(s(X+tY))}\right)_e \left(\frac{\partial}{\partial u}\Big|_{u=0} g(u,t)\right).$$

Now

$$\frac{\partial}{\partial u}\Big|_{u=0} g(u,t) = -X + (X+tY) = tY$$

by Lemma 5.3, and hence

$$\begin{aligned}\frac{\partial}{\partial s}\varphi(g(s,t)) &= d\left(\varphi \circ l_{\exp(-sX)} \circ r_{\exp(s(X+tY))}\right)_e (tY)\\ &= td\left(\varphi \circ l_{\exp(-sX)} \circ r_{\exp(s(X+tY))}\right)_e (Y).\end{aligned}$$

by linearity. Let $\psi(t) = d\left(\varphi \circ l_{\exp(-sX)} \circ r_{\exp(s(X+tY))}\right)_e (Y) \in \mathbb{R}$, so that

$$\frac{\partial}{\partial s}\varphi(g(s,t)) = t\psi(t).$$

The expression by which we defined $\psi(t)$ is seen to depend smoothly on t, hence $t\psi(t)$ has derivative $\psi(0)$ at $t=0$ and thus

$$\frac{\partial}{\partial t}\Big|_{t=0}\frac{\partial}{\partial s}\varphi(g(s,t)) = \psi(0) \tag{6}$$

Finally, since $l_{\exp(-sX)} \circ r_{\exp(sX)}$ is conjugation by $\exp(-sX)$ we obtain for $t=0$

$$\psi(0) = d\left(\varphi \circ l_{\exp(-sX)} \circ r_{\exp(sX)}\right)_e (Y) = d\varphi_e\left(\mathrm{Ad}(\exp(-sX))Y\right) \tag{7}$$

by the chain rule together with the definition of $\mathrm{Ad}(x)$ as the differential of conjugation by x.

We finish the proof of (3) by inserting (7) into (6) into (5) into (4).

HS