CHAPTER 5

Lie Groups and Lie Algebras

5.1. Definition of a Lie group and examples. A **Lie group** is a group G that is at the same time an analytic manifold such that the group multiplication $G \times G \to G$ defined by $(g_1, g_2) \to g_1 g_2$ is analytic. It can be shown that then the mapping $g \to g^{-1}$ of $G \to G$ is also analytic (Cohn, 1957, Theorem 2.6.1). Equivalently, the multiplication $(g_1, g_2) \to g_1 g_2^{-1}$ can be required to be analytic. (Some authors replace "analytic" by "C^∞" in the above definition.)

Left and right translation. For fixed $g \in G$, the transformation $h \to gh$, $h \in G$, is a 1-1 transformation of G onto itself which is an analytic diffeomorphism. It is called a **left translation** with g and denoted L_g. Similarly, **right translation** with g, denoted R_g, is the transformation $h \to hg$, $h \in G$. Any chart at the identity element e of G can serve also as a chart at an arbitrary element $g \in G$ by transporting it with help of the left translation L_g (or by the right translation R_g). It is therefore usually sufficient to consider only charts at e.

An obvious example of a Lie group is R^d with group multiplication defined as vector addition. Here the whole group can be covered by one chart (the usual coordinate system) and analyticity of the group multiplication is immediate. Other examples that are very important for statistical applications are provided by matrix groups, i.e., the **general linear group** $GL(n)$ of all $n \times n$ real nonsingular matrices, and its Lie subgroups. Among the latter especially important are **orthogonal**, **triangular**, **diagonal**, and **block diagonal** matrices.

Denote by $O(n)$ all $n \times n$ orthogonal matrices, by $UT(n)$ all $n \times n$ upper triangular matrices with positive diagonal elements, and similarly $LT(n)$ for lower triangular matrices. In every $n \times n$ real matrix group the identity element e of the group is the $n \times n$ identity matrix $I_n = \mathrm{diag}(1, \ldots, 1)$.

The above listed groups are not defined as analytic manifolds until an analytic structure on them has been defined. For $G = GL(n)$ this can be done simply by regarding it as an open submanifold of R^{n^2}. The group can be covered by a single chart with coordinates g_{ij}, $i, j = 1, \ldots, n$ where g_{ij} is the (i,j)-element of $g \in G$. It follows that $\dim GL(n) = n^2$. Group multiplication in this chart, defined by $(gh)_{ik} = \sum_j g_{ij} h_{jk}$, is obviously analytic. A basis for the tangent space G_g at $g \in G$ is $\{\partial/\partial g_{ij} : i, j = 1, \ldots, n\}$ so that an arbitrary analytic vector field on G has the form $\sum_{ij} \alpha_{ij}(g)\partial/\partial g_{ij}$, where the α_{ij} are analytic real valued functions on G.

The triangular groups $UT(n)$ and $LT(n)$ can be handled in the same way as $GL(n)$ above. A single chart covers each of them, with coordinates g_{ij} ($i \leq j$ for UT, $i \geq j$ for LT), where $g_{ii} > 0$. We have $\dim UT(n) = \dim LT(n) = \frac{1}{2}n(n+1)$.

The orthogonal group $O(n)$ cannot be covered by a single chart. A possible choice of local coordinates x_{ij} for a chart at e is $x_{ij} = g_{ij}$ for $i < j$, provided the $|x_{ij}|$ are sufficiently small, and all g_{ij} with $i \geq j$ are analytic functions of the x_{ij}. It follows that $\dim O(n) = \frac{1}{2}n(n-1)$. This chart can be used to show the analyticity of the group multiplication. In Section 5.7 we shall introduce a different kind of chart, called *canonical*. In that chart each element in a neighborhood of e in $O(n)$ corresponds to an $n \times n$ skew symmetric matrix.

The Lie groups that often occur in statistical applications are usually built up from a subgroup of an additive matrix group and the above mentioned subgroups of the general linear group.

5.2. Invariant vector fields. Lie algebras. All vector fields considered in this chapter will be understood to be analytic without special mention. If X is a vector field on G, then it is called **left invariant** if $dL_g X = X$ for every $g \in G$, and **right invariant** if

$dR_g X = X$ for every $g \in G$. We shall mostly consider left invariant vector fields and they will often be called simply **invariant**. We have seen in Section 3.6 that any invariant X has the form $X(g) = dL_g t$ for some $t \in G_e$. Conversely, for every $t \in G_e$, $X(g) = dL_g t$ is invariant. We can express this by saying that X is invariant if and only if $X(g) = dL_g X(e)$. Thus, there is a natural 1-1 correspondence between the invariant vector fields and the elements of the tangent space G_e of G at e. Furthermore, if X and Y are both invariant then so is $[X, Y]$, by (3.6.2). Therefore, the space of all invariant vector fields on G is a d-dimensional vector space closed under the bracket operation, where $d = \dim G$. This is called the **Lie algebra of** G and denoted $\mathfrak{g}$. It is a special case of an **abstract Lie algebra**, which is any finite dimensional vector space, say $\mathfrak{a}$ (the elements of which will be denoted x, y, etc.), on which there is a function $\mathfrak{a} \times \mathfrak{a} \to \mathfrak{a}$, denoted $[\cdot,\cdot]$ and called **bracket**, that is required to be bilinear and to satisfy, for all $x, y, z \in \mathfrak{a}$, (i) $[x,x] = 0$; (ii) the Jacobi identity $[[x,y],z] + [[y,z],x] + [[z,x],y] = 0$. By replacing x by $x + y$ in (i) and using bilinearity it follows immediately that $[x,y] + [y,x] = 0$. An example of a Lie algebra other than $\mathfrak{g}$ above is furnished by all $n \times n$ real matrices and bracket defined by $[x,y] = xy - yx$. Here (i) is of course obvious, and the verification of (ii) is straightforward. We shall return to this Lie algebra presently.

5.2.1. EXAMPLE. $G = GL(n)$. Let $t_{ij} = \partial/\partial g_{ij}$ at $g = e$ so that $\{t_{ij},\ i,j = 1,\dots,n\}$ is a basis of G_e. Define $X_{ij} \in \mathfrak{g}$ by $X_{ij}(g) = dL_g t_{ij}$, then the X_{ij} form a basis of $\mathfrak{g}$. An explicit expression for $X_{ij}(g)$ in the chart (g_{ij}) can easily be obtained as follows. For fixed $g \in G$ write the left translation L_g as $h \to gh = k$, say, where h ranges over G. Compute $dL_g \partial/\partial h_{ij}$ by expressing $\partial/\partial h_{ij}$ in terms of the $\partial/\partial k_{rs}$: $\partial/\partial h_{ij} = \sum_{rs}(\partial k_{rs}/\partial h_{ij})\partial/\partial k_{rs}$. The partial derivatives $\partial k_{rs}/\partial h_{ij}$ follow from $k_{rs} = \sum_m g_{rm} h_{ms}$. The result is $\partial/\partial h_{ij} = \sum_r g_{ri}\partial/\partial k_{rj}$. Now evaluate this at $h = e$ so that $\partial/\partial h_{ij}$ becomes t_{ij} and $\partial/\partial k_{rj}$ becomes $\partial/\partial g_{rj}$. Hence, $dL_g t_{ij} = \sum_r g_{ri}\partial/\partial g_{rj}$, i.e.,

$$X_{ij}(g) = \sum_{r=1}^{n} g_{ri}\frac{\partial}{\partial g_{rj}}, \quad i,j = 1,\dots,n, \tag{5.2.1}$$

is an explicit expression for the invariant vector field $dL_g(\partial/\partial g_{ij}\mid_e)$. Taking the bracket of X_{ij} of (5.2.1) with a similar expression for X_{rs} ($1 \le i, j, r, s \le n$) yields

$$[X_{ij}, X_{rs}] = \delta_{jr}X_{is} - \delta_{is}X_{rj} \tag{5.2.2}$$

in which $\delta_{ij} = 1$ or 0 according as $i =$ or $\neq j$ (Kronecker delta). Equation (5.2.2) shows once more that $\mathfrak{g}$ is closed under the formation of brackets. □

An arbitrary element $X \in \mathfrak{g}$ can be expressed in terms of the basis elements $X_{ij} : X = \sum_{ij} a_{ij}X_{ij}$ with any matrix $A = ((a_{ij}))$ of coefficients. Since A depends on X we shall sometimes write $A(X)$. In particular, $A(X_{ij}) = E_{ij}$, where

$$E_{ij} = \text{ matrix with 1 in position } (i,j), \text{ 0 elsewhere.} \tag{5.2.3}$$

For two such matrices, say E_{ij} and E_{rs}, one easily computes their bracket $E_{ij}E_{rs} - E_{rs}E_{ij}$ (matrix multiplication):

$$[E_{ij}, E_{rs}] = \delta_{jr}E_{is} - \delta_{is}E_{rj}. \tag{5.2.4}$$

The left-hand side is the bracket of the coefficient matrices of X_{ij} and X_{rs}, while the right-hand side is the coefficient matrix of the bracket of X_{ij} and X_{rs}, by (5.2.2). Hence, for the basis elements X_{ij}, A preserves the bracket operation (A of bracket equals bracket of A's). Using the bilinearity of the bracket, it is easily seen that the same holds true for any elements of $\mathfrak{g}$:

$$A([X,Y]) = [A(X), A(Y)], \quad X, Y \in \mathfrak{g}. \tag{5.2.5}$$

Furthermore, A is obviously linear in its argument. Thus, there is a 1-1 correspondence between elements of $\mathfrak{g}$ and $n \times n$ real matrices that preserves linear operations and brackets. The Lie algebra of $n \times n$ real matrices is denoted $\mathfrak{gl}(n)$. Thus, there is an isomorphism between $\mathfrak{g}$ and $\mathfrak{gl}(n)$ as Lie algebras.

With help of the coefficient matrix $A(X)$ of $X \in \mathfrak{g}$ it is also possible to get a simple explicit expression for the value of X at any $g \in G$. Let $A(X;g)$ with elements $a_{ij}(X;g)$ be defined by

$$X(g) = \sum_{ij} a_{ij}(X;g)\frac{\partial}{\partial g_{ij}}, \tag{5.2.6}$$

and call this the coefficient matrix of $X(g)$. (Thus, $A(X;e) = A(X)$ defined earlier.) From (5.2.1) the coefficient matrix of $X_{ij}(g)$ is seen to be $A(X_{ij};g) = gE_{ij}$, with E_{ij} defined in (5.2.3). Furthermore, $A(X_{ij}) = E_{ij}$, so that $A(X_{ij};g) = gA(X_{ij})$. By the linearity of $A(\cdot)$ and of $A(\cdot;g)$, this equation holds for every $X \in \mathfrak{g}$:

$$A(X;g) = gA(X), \quad X \in \mathfrak{g}. \tag{5.2.7}$$

Consider now the elements g_{ij} of g as n^2 real valued functions, and differentiation of g is to be carried out elementwise on each of these g_{ij}. From (5.2.6) and (5.2.7) we compute

$$X(g)g = gA(X) \tag{5.2.8}$$

and repeating this process with g replaced by $X(g)g = gA(X)$, etc., we have

$$X^k(g)g = g(A(X))^k, \quad k = 0, 1, \ldots . \tag{5.2.9}$$

In particular, at $g = e = I_n$ this becomes

$$X^k(e)g = (A(X))^k. \tag{5.2.10}$$

This may be expressed in words: the kth order derivative of g at $g = e$ with the vector field X may be obtained by raising the coefficient matrix of $X(e)$ to the kth power.

5.3. Invariant differential forms of maximum degree. Together with left and right invariance of vector fields we consider the analagous invariance of differential forms. For notational convenience we shall often write $\omega(g)$ instead of ω_g for the differential form ω at $g \in G$. The differentials dL_g and dR_g of the left and right translations L_g and R_g have adjoints δL_g and δR_g, respectively (Section 4.5). A differential form ω on G will be called **left invariant** if $\delta L_g \omega = \omega$ for every $g \in G$, and **right invariant** if $\delta R_g \omega = \omega$. More explicitly, for δL_g we have

$$\delta L_g \omega(g_1) = \omega(g^{-1} g_1), \quad g, g_1 \in G, \tag{5.3.1}$$

and a similar equation for δR_g (remember that in general for any $f : M \to N$, δf "pulls back" from N to M). Apply (5.3.1) with $g_1 = e$ and g replaced by g^{-1}:

$$\omega(g) = \delta L_{g^{-1}} \omega(e) \tag{5.3.2}$$

for a left invariant differential form ω, which shows that such an ω is defined on the whole of G by its value at e. Similarly, a right invariant ω satisfies (5.3.2) with $\delta L_{g^{-1}}$ replaced by $\delta R_{g^{-1}}$. It follows immediately from (4.5.4) (applied to $M = N = G$ and $f = L_{g^{-1}}$) that the wedge product of two differential forms of the form (5.3.2) is again left invariant. This extends to any number of such differential forms and is especially useful for building a left invariant k-form from the wedge product of k linearly independent left invariant 1-forms.

We are interested mainly in invariant d-forms ω, where $d = \dim G$. Since V_d at e has dimension 1 (Section 4.1), $\omega(e)$ is unique except for a multiplicative constant, and the same is then true for a left invariant d-form ω, by (5.3.2). Similarly for a right invariant d-form. In order to obtain a more explicit expression for a left invariant d-form let $x = (x_1, \dots, x_d)$ be local coordinates at g, and $y = (y_1, \dots, y_d)$ at e, then the function $L_{g^{-1}}$ (that maps g to e) causes y locally to be an analytic function of x. Take $\omega(e) = dy_1 \wedge \cdots \wedge dy_d$ at e and apply (5.3.2) and (4.5.3) with $f = L_{g^{-1}}$, then

$$\omega(g) = \frac{\partial(y)}{\partial(x)} dx_1 \wedge \cdots \wedge dx_d \tag{5.3.3}$$

in which $y(x)$ is the function $x \to y$ that parametrically represents the transformation $h \to g^{-1}h$, $h \in G$, and the right-hand side of (5.3.3) is to be evaluated at x corresponding to g. An arbitrary left invariant d-form is any constant multiple of (5.3.3). For a right invariant d-form the only change in the above is that $x \to y$ corresponds to $h \to hg^{-1}$, $h \in G$.

5.3.1. EXAMPLE. $G = GL(n)$ with elements $g = ((g_{ij})) : n \times n$. The g_{ij} serve as coordinates on all of G, but it is sometimes convenient to replace them temporarily by x_{ij} and/or y_{ij}, as in (5.3.3), and consider $x = ((x_{ij}))$ and $y = ((y_{ij}))$ also as $n \times n$ matrices. Then the function $y(x)$ in (5.3.3) is $y = g^{-1}x$. In order to obtain its Jacobian note that this function can be written $y_{(j)} = g^{-1}x_{(j)}$, $j = 1, \ldots, n$, where $x_{(j)}$ is the jth column of x, and similarly $y_{(j)}$. Since for each j the Jacobian is $(\det g)^{-1}$, the Jacobian of $y = g^{-1}x$ is $\partial(y)/\partial(x) = (\det g)^{-n}$. Substitute this into (5.3.3):

$$\omega(g) = (\det g)^{-n} \bigwedge_{i,j=1}^{n} dg_{ij}, \tag{5.3.4}$$

then an arbitrary left invariant d-form ($d = n^2$) is a constant multiple of (5.3.4). Note that the order of writing down the factors in the wedge product on the right-hand side of (5.3.4) has been left unspecified, in spite of the fact that changing the order may result in changing the sign. However, this does not matter for two reasons. First, a change in sign is absorbed by the undetermined constant factor. Second, we shall use a left invariant d-form ω to determine a left Haar measure (Section 7.7) and that only uses the absolute value of ω.

A right invariant d-form ω on $GL(n)$ is derived in the same way and involves the Jacobian of $y = xg^{-1}$. This also equals $(\det g)^{-n}$ so that the same formula (5.3.4) is obtained. □

5.3.2 EXAMPLE. $G = LT(n)$ with elements $g = ((g_{ij}))$, $1 \le j \le i \le n$. Again, the Jacobian of $y = g^{-1}x$ has to be evaluated. Temporarily denote g^{-1} by $A = ((a_{ij}))$, $1 \le j \le i \le n$. Put $A_1 = A$, and for $j = 2, \dots, n$, $A_j = A$ with its first $(j-1)$ columns deleted. Then because of the lower triangular nature of the matrices, $y_{(j)} = Ax_{(j)}$ is the same as $y_{(j)} = A_j x_{(j)}$. This has Jacobian $\det A_j = a_{jj} \cdots a_{nn}$. Taking the product over $j = 1, \dots, n$ and replacing a_{jj} by $(g^{-1})_{jj} = g_{jj}^{-1}$ yields $\partial(y)/\partial(x) = \prod_1^n g_{jj}^{-j}$. Thus, an arbitrary left invariant d-form ($d = n(n+1)/2$) on $LT(n)$ is a constant multiple of

$$\omega(g) = g_{11}^{-1} g_{22}^{-2} \cdots g_{nn}^{-n} \bigwedge_{1 \le j \le i \le n} dg_{ij}. \tag{5.3.5}$$

A similar computation shows that an arbitrary right invariant d-form on $LT(n)$ is a constant multiple of

$$\omega(g) = g_{11}^{-n} g_{22}^{-n+1} \cdots g_{nn}^{-1} \bigwedge_{1 \le j \le i \le n} dg_{ij}. \quad \square \tag{5.3.6}$$

5.3.3 EXAMPLE. $G = UT(n)$ with elements $g = ((g_{ij}))$, $1 \le i \le j \le n$. The left invariant d-form ω is now given by (5.3.6), the right invariant ω by (5.3.5), except that the wedge product is now over $1 \le i \le j \le n$. (The Jacobian of $y = g^{-1}x$ in $UT(n)$ is of course the same as that of $y = xg^{-1}$ in $LT(n)$.) $\square$

We shall exhibit now a different method of obtaining an explicit expression for an invariant d-form on a subgroup of $GL(n)$, which is more convenient in the case of the orthogonal group $O(n)$. Take first $G = GL(n)$ itself. Let $g \in G$ and together with the tangent space G_g at g consider its dual G_g^* whose elements are the 1-forms at g (Sections 3.3, 4.1). A basis of G_g^* is the set of all dg_{ij}, $1 \le i, j \le n$. In particular, take $g = e$, but it is now notationally more convenient to denote the matrix in a neighborhood of e by $y = ((y_{ij}))$. For fixed $g \in G$ let $y = g^{-1}x$, so that x denotes a matrix in a neighborhood of g. In (5.3.2), for fixed i, j take $\omega_{ij}(e) = dy_{ij}$ at $y = e$, and denote the left-hand side by $\omega_{ij}(g)$. In order to compute the right-hand side we have to express y_{ij} as a function of x, take the differential, and evaluate at $x = g$. Let $a_i = (a_{i1}, \dots, a_{in})$ be the ith row of g^{-1}, then $y_{ij} = \sum_{k=1}^n a_{ik} x_{kj}$, so $\omega_{ij}(g) = \delta L_{g^{-1}} dy_{ij} \mid_e = \sum_{k=1}^n a_{ik} dx_{kj} \mid_g = \sum_{k=1}^n a_{ik} dg_{kj}$. The result can be put conveniently in the form

$$\underset{\sim}{\omega}(g) = g^{-1} dg \tag{5.3.7}$$

in which $\underset{\sim}{\omega}(g)$ is an $n \times n$ matrix whose (i, j) element is the transform under $\delta L_{g^{-1}}$ of $dg_{ij}\,|_{g=e}$, and dg is the $n \times n$ matrix $((dg_{ij}))$. Since the $\omega_{ij}(e)$ on the right-hand side of (5.3.2) form a basis of G_e^*, the $\omega_{ij}(g)$ on the left-hand side of (5.3.2) form a basis of G_g^* and are therefore linearly independent. Since all these 1-forms $\omega_{ij}(g)$ are left invariant, so is their wedge product, by an earlier remark in connection with (5.3.2). Hence, the wedge product of the $d = n^2$ elements of the matrix on the right-hand side of (5.3.7) is a left invariant d-form on G. Similarly, the wedge product of the elements of the matrix $dg\, g^{-1}$ is a right invariant d-form. If G is a Lie subgroup of $GL(n)$ with dimensions $d < n^2$, then the n^2 elements of the matrix (5.3.7) are of course no longer linearly independent. But the dg_{ij} at $g = e$ still span G_e^* so that the elements of the matrix (5.3.7) span G_g^*. Then choose from these n^2 elements d linearly independent ones, and their wedge product is a left invariant d-form on G. A right invariant d-form is similarly chosen from the elements of $dg\, g^{-1}$. It is also of interest to see what happens to the matrix (5.3.7) under a right translation $R_{g_1} : h \to hg_1$, $h \in G$, with any fixed $g_1 \in G$. Put $g = hg_1$ and express (5.3.7) as a function of h, then we get

$$\delta R_{g_1} \underset{\sim}{\omega}(g) = g_1^{-1}(h^{-1}dh)g_1, \quad g = hg_1. \tag{5.3.8}$$

5.3.4. Example. $G = GL(n)$ revisited. Fix $1 \leq j \leq n$ and use (5.3.7) and Lemma 4.3.1 to obtain $\bigwedge_{i=1}^{n} \omega_{ij}(g) = (\det g)^{-1} \bigwedge_{i=1}^{n} dg_{ij}$. Then take wedge product over $j = 1, \dots, n$ to obtain the previous result (5.3.4). □

5.3.5. Example. $G = LT(n)$ revisited. Now the matrix (5.3.7) is also lower triangular and we have to take the wedge product over $\omega_{ij}(g)$ with $1 \leq j \leq i \leq n$. For the jth column on both sides of (5.3.7) we have $\omega_{(j)}(g) = A_j dg_{(j)}$ with A_j of Example 5.3.2 and $dg_{(j)}$ having elements dg_{ij}, $i = j, \dots, d$. Then by Lemma 4.3.1, $\bigwedge_{i=j}^{n} \omega_{ij}(g) = (\det A_j) \bigwedge_{i=j}^{n} dg_{ij}$, and taking wedge product over $j = 1, \dots, n$ yields the previous result (5.3.5). □

5.3.6. Example. $G = O(n)$. Since there is no single chart that covers all of G, it is not possible to get as explicit an expression for an invariant d-form as in the case of $GL(n)$, $LT(n)$, and $UT(n)$. Since at $g = e$ a chart may be formed by taking as coordinates the elements of g below (or above) the diagonal, a left invariant d-form ($d = n(n-1)/2$) is obtained by taking the wedge product of the elements of the matrix (5.3.7) below (or above) the diagonal. We shall generically write Γ for $g \in O(n)$ and use $'$ to denote transpose of a matrix. Let γ_j be the jth column of Γ and observe that $\Gamma^{-1} = \Gamma'$ with rows γ_i', $i = 1, \ldots, n$. If we choose the lower triangular part of (5.3.7), which is $\Gamma' d\Gamma$, then a left invariant d-form on G is

$$\omega(\Gamma) = \bigwedge_{1 \le j < i \le n} \gamma_i' d\gamma_j. \tag{5.3.9}$$

Note that at $g = e$, i.e., $\Gamma = I_n$, the right-hand side reduces to $\bigwedge_{1 \le j < i \le n} dg_{ij} \,|_{g=e}$. From $\Gamma'\Gamma = I_n$, after taking the differential and using (3.3.9) one obtains $\Gamma' d\Gamma + (\Gamma' d\Gamma)' = 0$. This shows that $\Gamma' d\Gamma$ is skew symmetric, so that taking the wedge product of the elements of $\Gamma' d\Gamma$ above the diagonal would have produced the same result (5.3.9) except, possibly, for a sign change. The right-hand side of (5.3.9) may be further expressed in terms of local coordinates if desired. As a simple example where $\omega(\Gamma)$ can be expressed explicitly take $n = 2$, so $d = 1$, and $\Gamma = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$, then the only element on the right-hand side of (5.3.9) is $d\theta$, the usual element of length on the unit circle.

A right invariant d-form on $O(n)$ could be constructed in an analogous way from the matrix $(d\Gamma)\Gamma'$ but in fact the left invariant d-form (5.3.9) is also right invariant, at least on the component of the identity. In order to show this consider first in general for any $n \times n$ skew symmetric matrix A the $\frac{1}{2}n(n-1) \times 1$ vector $x = x(A)$ consisting of the a_{ij} with $1 \le j < i \le n$, in an arbitrary but fixed order. Let B be another skew symmetric matrix, then the inner product of the two vectors $x(A)$ and $x(B)$ can be written as $\frac{1}{2} \operatorname{tr} A'B$. This is invariant under $A \to \Gamma' A \Gamma$, $B \to \Gamma' B \Gamma$, for any $\Gamma \in O(n)$ which shows that to $A \to \Gamma' A \Gamma$ corresponds an orthogonal transformation of $x : x(A) \to \Delta x(A)$, with $\Delta \in O(\frac{1}{2}n(n-1))$. If Γ is restricted to the component of the identity, then by a continuity argument $\det \Delta = 1$. Then apply this to the skew symmetric matrix $h^{-1}dh = h'dh$ ($h \in O(n)$) on the right-hand side of (5.3.8) and apply Lemma 4.3.1.

Invariant differential forms on a homogeneous space. If X is a C^∞ manifold on which a C^∞-group G acts differentiably and transitively to the left, then we also have the concept of (left) invariant differential form on X: if L_g is the diffeomorphism $x \to gx$, $x \in X$, $g \in G$, then a d-form on X is invariant if $\delta L_g \omega = \omega$ for every $g \in G$. Below is an example where construction of such an invariant ω is easy.

5.3.7. EXAMPLE. **Invariant differential form on the space of positive definite matrices.** Let $X = PD(n)$ as in Example 2.1.7 and $G = GL(n)$ with action

$$S \to CSC', \quad S \in X, \ C \in G, \tag{5.3.10}$$

where the multiplication on the right-hand side of (5.3.10) is matrix multiplication. Here S may be considered the coset space G/H, where $H = O(n)$ is the isotropy subgroup of G at $S = I_n$, and therefore S is an analytic manifold (see Section 5.8). A chart may be put on all of S by taking the coordinates as the $d = \frac{1}{2}n(n+1)$ elements s_{ij} with (say) $i \geq j$. We are interested in an invariant d-form ω on X, which has to be of the form

$$\omega(S) = \alpha(S) \bigwedge_{1 \leq j \leq i \leq n} ds_{ij}. \tag{5.3.11}$$

In order to find the function α consider the left translation (5.3.10): $S \to \phi(S)$ and rewrite this as

$$R = CSC' \equiv \phi(S), \quad S \in X, \ C \in G. \tag{5.3.12}$$

On order to obtain $\delta\phi(\omega)$ at S we have to take $\omega(R) = \alpha(R) \wedge_{ij} dr_{ij}$ and express this in terms of S and the ds_{ij}. For this we need the Jacobian of the transformation $S \to R$ of (5.3.12), which equals

$$\frac{\partial(R)}{\partial(S)} = |C|^{n+1} \tag{5.3.13}$$

in which $|C| = \text{abs}\det(C)$. Thus, we get

$$(\delta\phi)\omega(S) = \alpha(CSC')|C|^{n+1} \bigwedge_{ij} ds_{ij}. \tag{5.3.14}$$

In order that $(\delta\phi)\omega = \omega$ we have to have $\alpha(CSC')|C|^{n+1} = \alpha(S)$ for every $S \in X$, $C \in G$. Setting first $S = I_n$ and then $CC' = S$ we find $\alpha(S) = \alpha(I_n)|S|^{-\frac{1}{2}(n+1)}$. Omitting the arbitrary constant multiplier $\alpha(I_n)$ we have then

$$\omega(S) = |S|^{-\frac{1}{2}(n+1)} \bigwedge_{1\le j\le i\le 1} ds_{ij} \tag{5.3.15}$$

as an invariant d-form on X under the transformations (5.3.10).

We could equally well have taken G to be $LT(n)$ or $UT(n)$ instead of $GL(n)$, with the same action (5.3.10), for then G is still transitive over X. For such choice of G the isotropy subgroup at $S = I_n$ is $\{e\}$ so that there is a 1-1 correspondence between X and G, given by the equation $S = TT'$ (**Cholesky decomposition**). The d-form (5.3.15) is of course still invariant under $LT(n)$ or $UT(n)$. Since there is a diffeomorphism between T and S, a left invariant d-form on one space provides one on the other. Thus, the forms (5.3.5) and (5.3.15) can be identified, except for a positive factor. It is easy to determine this factor. Take $S = TT'$ and differentiate at $T = I_n : dS = dT + dT'$. Then $ds_{ii} = 2dt_{ii}$ and for $i > j$, $ds_{ij} = dt_{ij}$. Take the wedge product over all $ds_{ij} : \bigwedge_{i\ge j} ds_{ij} = 2^n \bigwedge_{i\ge j} dt_{ij}$. Then comparing (5.3.5) and (5.3.15) (after replacing g_{ij} by t_{ij} in (5.3.5)) we find

$$|S|^{-\frac{1}{2}(n+1)} \bigwedge_{i\ge j} ds_{ij} = 2^n \prod_{1}^{n} t_{ii}^{-i} \bigwedge_{i\ge j} dt_{ij}. \quad \square \tag{5.3.16}$$

5.4. Subgroups and subalgebras. Let G be a Lie group and H a subgroup of G that is both a Lie group and an analytic submanifold of G. We shall then simply say that H is a Lie subgroup of G. Let $\mathfrak{g}$ and $\mathfrak{h}$ be the Lie algebras of G and H, respectively. The inclusion map (Section 3.4) $i : H \to G$ has the property that at any element $h \in H$, $di : H_h \to G_h$ is 1-1, by definition of submanifold. Thus, there is a linear isometry between H_h and $di(H_h)$, the latter being a linear subspace of G_h. Furthermore, the 1-1 correspondence preserves bracket formation, by (3.6.2) applied to $f = i$. (One can characterize the tangent vectors of G_h that correspond to tangent vectors of H_h as any $X \in G_h$ such that $X(h)f = 0$ for any analytic f that vanishes on H in a neighborhood of h; see Cohn, 1957, Sect. 3.4.) It is convenient to identify elements of H_h and G_h that correspond under the mapping di, so that we may regard H_h as a linear subspace of G_h. Since also dL_h, for arbitrary $h \in H$, preserves linear and bracket operations, the above identification can be extended to the invariant vector fields. That is, we may regard $\mathfrak{h}$ as a subset of $\mathfrak{g}$. Since $\mathfrak{h}$ is closed under linear and bracket operations it is called a **Lie subalgebra** of $\mathfrak{g}$.

EXAMPLE. Lie subgroups of $G = GL(n)$. Let $H = UT(n)$, then the coefficient matrix $A(X)$ of an arbitrary element $\sum_{i\le j} a_{ij}\partial/\partial g_{ij}$ of $\mathfrak{h}$ is an arbitrary $n \times n$ upper triangular matrix. The set of all such matrices is clearly a Lie algebra under scalar multiplication, addition, and matrix multiplication and is isomorphic to $\mathfrak{h}$. Similarly, if $H = LT(n)$, the set of all $n\times n$ lower triangular matrices is isomorphic to $\mathfrak{h}$. Now let $H = O(n)$, then it turns out that $\mathfrak{h}$ is isomorphic to the set of all $n \times n$ skew symmetric matrices, which is easily verified to be a Lie algebra. We can derive this result as follows. Choose a chart at e with local coordinates x_{ij}, $1 \le i < j \le n$, where $x_{ij} = g_{ij}$ and the $|x_{ij}|$ are sufficiently small. Then for $r \ge s$, the g_{rs} are analytic functions of the x_{ij}. We need the derivatives $\partial g_{rs}/\partial x_{ij}$ evaluated at 0. This follows from differentiating the defining equations for an orthogonal matrix: $\sum_k g_{rk}g_{sk} = \delta_{rs}$ with respect to x_{ij} and then setting $g_{rs} = \delta_{rs}$. The result is

$$(5.4.1) \qquad \partial g_{rs}/\partial x_{ij} + \partial g_{sr}/\partial x_{ij} = 0 \quad \text{at } g = e.$$

In (5.4.1) first take $r = s$, then $\partial g_{rr}/\partial x_{ij} = 0$ at $g = e$. Next, take $r > s$. If $(s,r) \ne (i,j)$, then $\partial g_{sr}/\partial x_{ij} = 0$ so that $\partial g_{rs}/\partial x_{ij} = 0$ at $g = e$. There remains the case $(s,r) = (i,j)$, and then (5.4.1) yields

$\partial g_{ji}/\partial x_{ij} = -1$ at $g = e$. In summary,

(5.4.2) $$\partial g_{ij}/\partial x_{ij} = 1, \quad \partial g_{ji}/\partial x_{ij} = -1, \quad (i < j) \text{ at } g = e,$$

and all other derivatives are 0. Now write the set of coordinates x_{ij} as x and the dependence of g on x as $g = f(x)$. In terms of the chart (x) at e a basis of G_e is formed by the $\partial/\partial x_{ij}$, $i < j$. In terms of the $\partial/\partial g_{rs}$ the tangent vector $\partial/\partial x_{ij}$ becomes

(5.4.3) $$df\left(\frac{\partial}{\partial x_{ij}}\right) = \sum_{rs} \frac{\partial g_{rs}}{\partial x_{ij}} \frac{\partial}{\partial g_{rs}}$$

where the differentiations have to be performed at $g = e$. Due to (5.4.2) only two terms on the right-hand side of (5.4.3) survive:

(5.4.4) $$df\left(\frac{\partial}{\partial x_{ij}}\right) = \frac{\partial}{\partial g_{ij}} - \frac{\partial}{\partial g_{ji}} \quad \text{at } g = e.$$

The right-hand side of (5.4.4) shows that the coefficient matrix $((a_{ij}))$ in $\sum_{ij} a_{ij}\partial/\partial g_{ij}$ corresponding to $\partial/\partial x_{ij}$ is the skew symmetric matrix $\delta_{ij} - \delta_{ji}$. These matrices, for $1 \le i < j \le n$, form a basis of the set of all $n \times n$ skew symmetric matrices, hence the claim has been proved. A shorter way of establishing this result will appear after the exponential map has been introduced in Section 5.6. □

5.5. One-dimensional subgroups. Let G be a Lie group and $X \in \mathfrak{g}$ a given and fixed nonzero invariant vector field. Let $\gamma(u)$, $-a < u < b$, be an integral curve of X starting at e (Section 3.5), with $\gamma(0) = e$ and $0 < a, b \le \infty$. That is, γ is an analytic curve in G and

(5.5.1) $$d\gamma\left(\frac{d}{du}\right) = X(\gamma(u)), \quad -a < u < b.$$

We shall assume that the integral curve is maximal, i.e., neither a nor b can be increased. It will now be shown that we must have $a = b = \infty$

and that the points $\gamma(u)$ form a one-dimensional Lie subgroup H of G. In fact, it will be shown that for any $u_1, u_2 \in R$,

$$\gamma(u_1 + u_2) = \gamma(u_1)\gamma(u_2), \tag{5.5.2}$$

so that as a group, H is a homomorphic image of the additive group R. Let u_1 be arbitrary, with $-a < u_1 < b$. Since $X \in \mathfrak{g}$, $dL_g X = X$ for any $g \in G$; i.e., the transformations L_g leave X unchanged. In particular, take $g = \gamma(u_1)$ then $L_{g^{-1}}$ maps $\gamma(u_1)$ on e and $dL_{g^{-1}}$ maps $X(\gamma(u_1))$ on $X(e) = t$. Now consider an integral curve of X starting at $\gamma(u_1)$, i.e., a curve $\gamma(u_1 + u)$ with u in an interval about 0 such that $d\gamma \frac{d}{du} |_{u_1+u} = X(\gamma(u + u_1))$. Due to the invariance of the vector field X under the transformation $L_{g^{-1}}$ we can solve this problem after the transformation where it becomes an integral curve of X starting at e, and then transform back. In the transformed form the solution is as before $\gamma(u)$, with $-a < u < b$. Transform back, and the solution is $L_g\gamma(u) = \gamma(u_1)\gamma(u)$. Therefore

$$\gamma(u_1 + u) = \gamma(u_1)\gamma(u), \quad -a < u_1, u < b. \tag{5.5.3}$$

This shows that $-a < u_1 + u < b$. But this can be true for all $-a < u_1, u < b$ only if $a = b = \infty$. Therefore, $\gamma(u)$ is defined for all $-\infty < u < \infty$ and (5.5.3) holds for all $u_1, u \in R$. Moreover, since at every $u \in R$ the right-hand side of (5.5.1) is $\neq 0$ (remember that X was assumed to be a nonzero invariant vector field), $d\gamma$ is 1-1 at every point. Therefore, at every $\gamma(u_1) \in H$ there is a chart by giving $\gamma(u)$ the coordinate u, for u in some interval about u_1. (It may not be possible to cover H by a single chart, for instance if γ maps R onto the circle group.) Therefore, H is a 1-dimensional Lie group. Moreover, in this parametrization, at any $\gamma(u) \in H$ the tangent space of H is spanned by $\frac{d}{du}$ and if $i : H \to G$ is the inclusion map, then $di\left(\frac{d}{du}\right) = X(\gamma(u)) \neq 0$ so that di is 1-1. It follows that H is a submanifold of G so that H is a 1-dimensional Lie subgroup of G. (These results can be established more elegantly by first studying the relation between an analytic homomorphism of one Lie group into another one and the induced homomorphism of the Lie algebras. See Chevalley, 1946, Chap. IV, §VI.)

EXAMPLES. Let $G = GL(n)$ and let $X \in \mathfrak{g}$ have coefficient matrix $A = A(X)$ (Section 5.2). We seek a solution of the equation (5.5.1) for $\gamma(u) \in G$. Let $f : G \to R$ be analytic, then (5.5.1) can be rewritten (see (3.5.1))

$$\frac{d}{du} f(\gamma(u)) = X(\gamma(u))f. \tag{5.5.4}$$

Take in particular $f(g) = g_{ij}$ for any fixed (i, j), then (5.5.4) can be written $\frac{d}{du} g_{ij} = X(g)g_{ij}$ in which $g = g(u) = \gamma(u)$. In matrix form

$$\frac{d}{du} g = X(g)g, \quad g = g(u). \tag{5.5.5}$$

Write $X(g)$ in its form (5.2.6) and observe that for any (i, j), $(\partial/\partial g_{ij})g = E_{ij}$ (defined in (5.2.3)) so that

$$X(g)g = A(X; g), \tag{5.5.6}$$

and use (5.2.7), then (5.5.5) becomes

$$\frac{d}{du} g(u) = gA, \quad A = A(X) \tag{5.5.7}$$

with initial condition $g(0) = I_n$. The solution can be put in the simple form

$$g(u) = e^{uA}, \quad -\infty < u < \infty \tag{5.5.8}$$

where for any square matrix B, e^B is defined by

$$e^B = \sum_{k=0}^{\infty} \frac{1}{k!} B^k, \tag{5.5.9}$$

which converges absolutely. Any arbitrary $n \times n$ matrix A generates via (5.5.8) a 1-dimensional subgroup of $GL(n)$. In some cases a closed expression for the power series can be found easily. For instance, if $A = E_{11}$ (see (5.2.3)), which is the coefficient matrix of $\partial/\partial g_{11}$ at e,

$g(u) = \operatorname{diag}(e^u, 1, \ldots, 1)$, and if $A = E_{12}$ (corresponding to $\partial/\partial g_{12}$), $g(u) = I_n + uE_{12}$, which is $\in UT(n)$. Other E_{ij} follow from the above two by interchanging rows and columns. Now take A in the Lie subalgebra of $O(n)$. For simplicity let $n = 2$ and take A to have elements $a_{11} = a_{22} = 0$, $a_{12} = -a_{21} = 1$. Then (5.5.8) yields $g_{11}(u) = g_{22}(u) = \cos u$, $g_{12}(u) = -g_{21}(u) = \sin u$, i.e., $g(u)$ is a rotation matrix through angle u. In this last example the homomorphic image of R under γ is the unit circle.

5.6. The exponential map. In Section 5.5 the definition of a 1-dimensional subgroup involved a fixed vector field X and its corresponding integral curve γ. If X is allowed to be an arbitrary nonzero element of $\mathfrak{g}$ write γ_X for its integral curve starting at e. When $X = 0$ define $\gamma_X(u) = e$ for all u. By definition (5.5.1) of γ_X:

$$d\gamma_X\left(\frac{d}{du}\right) = X(\gamma_X(u)). \tag{5.6.1}$$

We shall now establish

$$\gamma_{tX}(u) = \gamma_X(tu), \quad t, u \in R. \tag{5.6.2}$$

To see this, write for simplicity γ for γ_X and define $\tau : R \to G$ by $\tau(u) = \gamma(tu)$ (when $t \neq 0$, the point set $\{\tau(u) : u \in R\}$ is the same as the point set $\{\gamma(u) : u \in R\}$, but the parametrization is different). If f is a real valued differentiable function on G then by differentiating $f(\tau(u)) = f(\gamma(tu))$ with respect to u one finds $d\tau\left(\frac{d}{du}\right) = t d\gamma\left(\frac{d}{dv}\right)_{v=tu} =$ (by (5.6.1)) $tX(\gamma(tu)) = tX(\tau(u))$ so that (again by (5.6.1) applied now to τ rather than γ) τ is the integral curve of tX, establishing (5.6.2).

The **exponential map**, written **exp**, is the function $\mathfrak{g} \to G$ defined by

$$\exp(X) = \gamma_X(1), \quad X \in \mathfrak{g}. \tag{5.6.3}$$

It follows then from (5.6.2) that

$$\exp(tX) = \gamma_X(t), \quad X \in \mathfrak{g},\ t \in R. \tag{5.6.4}$$

The 1-dimensional subgroup of Section 5.5 determined by X can therefore also be written $\{\exp(tX) : -\infty < t < \infty\}$. By (5.5.2) we have

$$\exp(t_1X)\exp(t_2X) = \exp((t_1+t_2)X), \quad X \in \mathfrak{g},\ t_1, t_2 \in R. \tag{5.6.5}$$

However, it is not true in general (unless G is abelian) that for any $X, Y \in \mathfrak{g}$, $\exp(X)\exp(Y) = \exp(X+Y)$. By taking $t_1 = -t_2 = 1$ in (5.6.5) one finds

$$(\exp(X))^{-1} = \exp(-X). \tag{5.6.6}$$

It is of interest to take the left-hand side of (5.6.4) and write formally

$$\exp(tX) = \sum_{k=0}^{\infty} \frac{1}{k!} t^k X^k, \quad t \in R. \tag{5.6.7}$$

This can be interpreted as follows. Let $f : G \to R$ be analytic, then $h(t) \equiv f(\gamma_X(t))$ is an analytic function of t and has therefore an infinite Taylor series expansion about $t = 0$. By definition of the integral curve γ_X, $h'(0) = Xf(e)$. Repeating the argument for the analytic function $Xf : G \to R$, etc., we see that $h^{(k)}(0) = X^k f(e)$, $k = 0, 1, \ldots$. Thus, for $|t|$ small enough,

$$f(\exp(tX)) = \sum_{k=0}^{\infty} \frac{1}{k!} t^k X^k f(e). \tag{5.6.8}$$

In particular, let $G = GL(n)$. Take in (5.6.8) $t = 1$ and take f to be the (i,j) element g_{ij} of $g \in G$, for $i, j = 1, \ldots, n$; then express the result in matrix form. On the left-hand side of (5.6.8) we have the matrix g corresponding to $\exp(X)$, where X is given by (5.2.6); on the right-hand side we get $e^{A(X)}$ by (5.2.10) and (5.5.9). This yields

$$\exp\left(\sum a_{ij} X_{ij}\right) = e^A \tag{5.6.9}$$

where $X_{ij} = \partial/\partial g_{ij}$ and $A = ((a_{ij}))$. Thus, for $GL(n)$ the function exp reduces to ordinary exponentiation. Note that by (5.6.6)

$$(e^A)^{-1} = e^{-A} \tag{5.6.10}$$

and by (5.5.9), written both for $B = A$ and for $B = A'$,

$$(e^A)' = e^{A'}. \tag{5.6.11}$$

Let H be a Lie subgroup of G and $\mathfrak{h}$ its Lie algebra, which is a Lie subalgebra of $\mathfrak{g}$ (Sec. 5.4). If in (5.6.3) X is restricted to $\mathfrak{h}$, then the resulting $\exp(X)$ is an element of H. If $G = GL(n)$, then $X \in \mathfrak{h}$ can be represented by a matrix $A = A(X)$ that is the member of the Lie subalgebra of $\mathfrak{gl}(n)$ corresponding to $\mathfrak{h}$. For instance, if A is skew symmetric ($A' = -A$), then by (5.6.10) and (5.6.11) e^A is orthogonal. This shows once more that the Lie subalgebra of all $n \times n$ skew symmetric matrices corresponds to the subgroup $H = O(n)$ of all $n \times n$ orthogonal matrices (see Section 5.4).

5.7. Canonical charts.

5.7.1. LEMMA. *The exponential map $X \to \exp(X)$ of $\mathfrak{g} \to G$ is an analytic diffeomorphism in a neighborhood of $X = 0$.*

PROOF. It will be shown first that exp is analytic in a neighborhood of 0. Let $x = (x_1, \dots, x_d)$ be a chart at $e \in G$ and let X_i be the element of $\mathfrak{g}$ for which $X_i = \partial/\partial x_i$ at $g = e$, $i = 1, \dots, d$. Then $X_1, \dots, X_d$ is a basis of $\mathfrak{g}$ so that any $X \in \mathfrak{g}$ is expressible as $X = \sum u_i X_i$. Regard $\mathfrak{g}$ as a linear manifold with $u = (u_1, \dots, u_d)$ a chart covering $\mathfrak{g}$. In (5.6.8) replace tX by $X = \sum u_i X_i$ and take $f = x_j$, with $1 \le j \le d$ fixed:

$$x_j\left(\exp\left(\sum_1^d u_i X_i\right)\right) = \sum_{k=0}^{\infty} \frac{1}{k!}\left(\sum_1^d u_i X_i\right)^k x_j(e). \tag{5.7.1}$$

On the right-hand side, in a neighborhood of e each X_i is a linear combination of $\partial/\partial x_1, \dots, \partial/\partial x_d$ with coefficients that are functions of x. Then after differentiation of x_j by the various powers of the X_i has been performed, the kth term is a polynomial of degree k in the u_i. Thus, the right-hand side of (5.7.1) is a power series in the u_i, and the only question is about its convergence. (The ideas of the following arguments are taken from the proof of Lemma (9.1.2) in Dieudonné, 1960.) The individual terms in the infinite series on the right-hand side of (5.7.1) are of the form

$$c_{i_1,\dots,i_d} u_1^{i_1} \dots u_d^{i_d} \tag{5.7.2}$$

in which the coefficients $c_{i_1,\dots,i_d}$ are real numbers, and $i_1,\dots,i_d$ take values $0, 1, \dots$. In (5.7.1) first take $u_1 = \dots = u_d = t$ with $t > 0$ to be chosen shortly, and put $X = \sum X_i$, then the right-hand side of (5.7.1) becomes the right-hand side of (5.6.8) (with $f = x_j$) and we know that there is a value of $t > 0$ for which the latter converges, say, $t = t_j > 0$. That implies that all terms (5.7.2) are bounded in absolute value by, say, $0 < A < \infty$ if $u_1 = \dots = u_d = t_j$. Now choose any $0 < \rho < 1$ and put $r_j = \rho t_j$, then if $|u_i| \leq r_j$, $i = 1, \dots, d$, the terms (5.7.2) are bounded in absolute value by $A\rho^{i_1+\dots+i_d}$. Then every partial sum of absolute values is bounded by $A \sum \rho^{i_1+\dots+i_d}$ where the summation is over $i_1, \dots, i_d$ from 0 to ∞. The latter sum equals $(1-\rho)^{-d} < \infty$. Hence, the infinite series on the right-hand side of (5.7.1) is absolutely convergent provided $|u_i| \leq r_j$, $i = 1, \dots, d$. Put $r = \min\{r_j : j = 1, \dots, d\}$, then we have proved that exp has a power series expansion that converges absolutely in the region $|u_i| \leq r$, $i = 1, \dots, d$. Hence, exp is analytic in a neighborhood of 0.

From (5.7.1) it follows that $\partial x_j / \partial u_i \mid_{u=0} = X_i x_j(e) = (\partial/\partial x_i)\, x_j(e) = \delta_{ij}$ so that the Jacobian matrix of the mapping $X \to \exp(X)$ at $X = 0$ is the $d \times d$ identity matrix. Hence, the Jacobian is positive at 0 so that by Theorem 3.1.1 there is a neighborhood of 0 on which exp is an analytic diffeomorphism. □

By Lemma 5.7.1 there is a neighborhood U of 0 in $\mathfrak{g}$ such that there is a 1-1 bi-analytic correspondence between points $X \in U$ and $\exp(X) \in \exp(U)$. This can be used to transfer a chart on U to a chart on $\exp(U)$. For any choice of basis $X_1, \dots, X_d$ of $\mathfrak{g}$ and corresponding coordinates $u = (u_1, \dots, u_d)$ as in the proof of Lemma 5.7.1 the chart on $\exp(U)$ that assigns coordinates $u = (u_1, \dots, u_d)$ to $\exp(\sum u_i X_i) \in \exp(U)$ (provided $\sum u_i X_i \in U$) is called a **canonical** chart. For instance (taking into account Section 5.6), a chart on a sufficiently small neighborhood of e in $O(n)$ can be chosen by assigning to e^A the elements of (say) the upper triangular part of the skew symmetric matrix A.

There are several other slightly different ways of defining a canonical chart. For instance, for integer $1 \leq m < d$ define the function $\phi : X = \sum_1^d u_i X_i \to \exp(\sum_1^m u_i X_i) \exp(\sum_{m+1}^d u_i X_i)$. For u in a sufficiently small neighborhood of $0 \in R^d$ this function is analytic since the product of two absolutely convergent power series is an absolutely convergent power series. The invertibility at $X = 0 \in \mathfrak{g}$ follows again from the Jacobian. Then in a sufficiently small neighborhood of $e \in G$ assign the coordinates $u_1, \dots, u_d$ to $\phi(\sum u_i X_i)$. This method can be extended by breaking up $\sum_1^d u_i X_i$ in more than two sums.

Another consequence of Lemma 5.7.1 is that if U is any neighborhood of $0 \in \mathfrak{g}$ on which exp is an analytic diffeomorphism, then $\{\exp(X) : X \in U\}$ is a neighborhood of $e \in G$, i.e., a nucleus (Section 2.3). If U is connected, then so is the above nucleus, and the latter generates therefore the identity component of G (Section 2.3). A fortiori, since $\mathfrak{g}$ is connected, $\{\exp(X) : X \in \mathfrak{g}\}$ is a connected nucleus of $e \in G$ and generates therefore the identity component. Hence all important properties of G follow from the elements $\exp(X)$, $X \in \mathfrak{g}$. The same is of course true for a Lie subgroup H of G and its Lie subalgebra $\mathfrak{h}$. For instance, the identity component of $O(n)$ is generated by all matrices e^A with A $n \times n$ skew symmetric.

The fact that a Lie group has a connected nucleus has the following consequence.

5.7.2. THEOREM. *In a Lie group G the identity component G_0 is both closed and open.*

PROOF. Closedness was stated in Section 2.2 (for any topological space). In order to show that G_0 is open let g_0 be an arbitrary point of G_0 and let V be a connected neighborhood of e. Then $g_0 V$

is a connected neighborhood of g_0. If g_0V would not entirely be contained in G_0, then $G_0 \cup g_0V$ would be a connected set that properly contains G_0. This contradicts the fact that G_0 is the component of g_0. Therefore, $g_0V \subset G_0$ so that G_0 is open. □

Theorem 5.7.2 and Proposition 2.3.18 together imply that if G is a Lie group, then G/G_0 (which is actually a group) is discrete. An example of a topological group that is not a Lie group and for which the conclusion of Theorem 5.7.2 fails is the set of rational numbers under addition. Then $G_0 = \{0\}$ which is closed but not open (Bourbaki, 1966b, I §11.5, "dangerous curve" remark).

5.8. Coset space as an analytical manifold. Local cross section. Let H be a Lie subgroup of the Lie group G, and G/H the space of left cosets (Section 2.1). We would like to be able to put an analytic structure on G/H in order to make it into an analytic manifold. A necessary condition for this is that H be closed, by Proposition 2.3.2 since a manifold is Hausdorff. It turns out that closedness of H is also sufficient, as will be shown now.

5.8.1. THEOREM. *If H is a closed Lie subgroup of the Lie group G, then G/H is an analytic manifold.*

PROOF. Let $\mathfrak{h}$ be the Lie algebra of H. Suppose first that $\mathfrak{h} = \mathfrak{g}$. By Section 5.7, H contains the group generated by all $\exp(X)$, $X \in \mathfrak{h} = \mathfrak{g}$, which is the identity component G_0 of G. Since G_0 is open, by Theorem 5.7.2, gG_0 is open for each $g \in G$ since $x \to gx$ is a homeomorphism. Therefore, $H = \{hG_0 : h \in H\}$ is the union of open sets, hence H is open. By Proposition 2.3.18 G/H is discrete and is therefore a 0-dimensional manifold.

Now suppose $\mathfrak{h} \neq \mathfrak{g}$. Since $\mathfrak{h}$ as a linear space is a subspace of $\mathfrak{g}$ we must have $\dim \mathfrak{h} < \dim \mathfrak{g}$, say $\dim \mathfrak{h} = d - m$, $0 < m \leq d$. Let $X_1, \ldots, X_d$ be a basis of $\mathfrak{g}$, with $X_{m+1}, \ldots, X_d$ a basis of $\mathfrak{h}$ (the latter X's are absent if $m = d$). By Section 5.7 there is a canonical chart at

e with local coordinates $u_1, \dots, u_d$, defined by the function

$$(5.8.1) \qquad (u_1, \dots, u_d) \to \exp\left(\sum_1^m u_i X_i\right) \exp\left(\sum_{m+1}^d u_i X_i\right)$$

(the second factor on the right-hand side is absent if $m = d$), provided the u_i are small enough, say $|u_i| < u_0$ for all i and some $u_0 > 0$. Let U be a neighborhood of e in which those inequalities are satisfied. If $u_1 = \cdots = u_m = 0$, then $X = \sum_{m+1}^d u_i X_i$ is an element of $\mathfrak{h}$ and therefore $\exp(X) \in H$. That is, if $g \in U$ and g has coordinates u with $u_1 = \cdots = u_m = 0$, then $g \in H$. Conversely, it can be shown, by using the closedness of H, that $g \in H \cap U$ implies that the coordinates of g satisfy $u_1 = \cdots = u_m = 0$ (Cohn, 1957, proof of Theorem 6.5.1). Choose $V \subset U$, V a neighborhood of e such that $V^{-1}V \subset U$ (use the continuity of $(g_1, g_2) \to g_1^{-1} g_2$) and let $W = \{g \in V : g$ has coordinates $u_{m+1} = \cdots = u_d = 0\}$. The coset projection $\pi : G \to G/H$ is defined in (2.1.2). We shall show first that π is 1-1 on W. Suppose $g_1, g_2 \in W$ and $\pi(g_1) = \pi(g_2)$. Then $g_2 = g_1 h$ for some $h \in H$ so that $h = g_1^{-1} g_2 \in V^{-1}V \subset U$. Since g_1 and g_2 are in W, they are of the form $\exp(\sum_1^m u_i X_i)$ and $\exp(\sum_1^m v_i X_i)$ respectively. Since $h \in H \cap U$ it is of the form $\exp(\sum_{m+1}^d u_i X_i)$. The equation $g_2 = g_1 h$ then reads

$$\exp\left(\sum_1^m v_i X_i\right) = \exp\left(\sum_1^m u_i X_i\right) \exp\left(\sum_{m+1}^d u_i X_i\right)$$

and since to each point of U corresponds a unique set of coordinates we have $(v_1, \dots, v_m, 0, \dots, 0) = (u_1, \dots, u_m, u_{m+1}, \dots, u_d)$ from which follows $v_i = u_i$, $i = 1, \dots, m$, i.e., $g_2 = g_1$. Modifying the previous argument slightly, we see that if $g_1 \in W$, $g_2 \in V$ (not necessarily $\in W$) and $\pi(g_1) = \pi(g_2)$, then g_2 has the same first m coordinates as g_1. Therefore, $\pi(W) = \pi(V)$ and the latter is a neighborhood of $[e] \in G/H$ since π is open (Section 2.3). Let $\pi^{-1} : \pi(W) \to W$ be the inverse of π restricted to W. Put a chart on $\pi(W)$ by assigning to $x \in \pi(W)$ the first m coordinates of $\pi^{-1}(x)$.

The transitive action of G on G/H (see (2.1.3)) can be used to transfer the above chart on $\pi(W)$ to any point of G/H. That is, if $[g] \in G/H$ (where g is not unique) then to $x \in g\pi(W)$ we can assign the coordinates of $g^{-1}x \in \pi(W)$. However, it remains to be shown that in the intersection of any two charts the coordinates are analytic functions of each other. Suppose $x \in g_1\pi(W) \cap g_2\pi(W)$ with $g_1, g_2 \in G$. Put $x_i = g_i^{-1}x$, $i = 1, 2$, so that $x_i \in \pi(W)$ and $x_2 = g_0 x_1$ where $g_0 = g_2^{-1}g_1$. It remains to be shown that in the transformation $x_2 = g_0 x_1$ with $g_0 \in G$ fixed and both $x_i \in \pi(W)$ the coordinates of x_2 are analytic functions of the coordinates of x_1. Let $g_i = \pi^{-1}x_i \in W$, $i = 1, 2$, then $\pi(g_0 g_1) = g_0\pi(g_1) = g_0 x_1 = x_2 = \pi(g_2)$ so that there exists $h \in H$ such that

$$g_2 = g_0 g_1 h. \tag{5.8.2}$$

Now keep g_0 and h fixed but allow g_1 to vary in a neighborhood (restricted to W) and define g_2 by (5.8.2) rather than by $\pi^{-1}x_2$ (then g_2 is no longer constrained to W). We still have that the coordinates of x_i are the first m coordinates of g_i, $i = 1, 2$. Since $g_1 \to g_0 g_1 h$ is analytic, the coordinates of g_2 given by (5.8.2) are analytic functions of those of g_1, i.e., of the first m coordinates of g_1 since the last d-m coordinates of g_1 are 0. In particular, the first m coordinates of g_2 are analytic functions of the first m coordinates of g_1. □

The set W together with the function $\pi^{-1} : \pi(W) \to W$ in the proof of Theorem 5.8.1 is called a **local cross section** of G/H. The choice of a local cross section permits one to choose in each coset x in a neighborhood of $[e]$ a unique representative $\pi^{-1}(x)$ in an analytic way. Furthermore, relative to the analytic structure on G/H constructed with help of Theorem 5.8.1 we have that π is analytic, and on $\pi(W)$ π^{-1} is analytic and $\pi \circ \pi^{-1}$ is the identity.

5.8.2. EXAMPLE. Let $G = O(n)$ and H the subgroup consisting of all matrices $\operatorname{diag}(I_r, \Gamma_{22})$ with $\Gamma_{22} \in O(s)$, where $r, s > 0$ and $r+s = n$. We may identify H with $O(s)$. In this example $d = \frac{1}{2}n(n-1)$ and $d - m = \frac{1}{2}s(s-1)$. Let γ_{ij} be the (i, j) element of $\Gamma \in O(n)$.

Denote by Δ the set of (i,j) with $1 \le j < i \le n$ so that the γ_{ij} with$(i,j) \in \Delta$ are the elements of Γ below the diagonal. Likewise denote by Δ_1 and Δ_2 the (i,j) in Δ with $j \le r$ and $j > r$, respectively. Thus, the matrices in H are labeled by the γ_{ij} with $(i,j) \in \Delta_2$. A basis of $\mathfrak{g}$ is formed by the skew symmetric matrices $X_{ij} = E_{ij} - E_{ji}$ (see (5.2.3)) with $(i,j) \in \Delta$. The X_{ij} with $(i,j) \in \Delta_2$ form a basis of $\mathfrak{h}$. These X_{ij} are to be substituted for $X_1, \ldots, X_d$ of Theorem 5.8.1. Write the canonical coordinates $u_1, \ldots, u_d$ of Theorem 5.8.1 as u_{ij}, $(i,j) \in \Delta$, and denote the vector formed from the u_{ij} by u. Then (5.8.1) reads

$$(5.8.3)\cdot \qquad u \to \exp\left(\sum_{(i,j)\in\Delta_1} u_{ij}X_{ij}\right)\exp\left(\sum_{(i,j)\in\Delta_2} u_{ij}X_{ij}\right)$$

in which $X_{ij} = E_{ij} - E_{ji}$. The right hand side of (5.8.3) is the matrix $\Gamma = ((\gamma_{ij}))$ and (5.8.3) defines a bi-analytic relation between the u_{ij} and the γ_{ij}, $(i,j) \in \Delta$. Note that $u = 0$ corresponds to $\Gamma = I_n$, i.e., $g = e$. By differentiation of the function (5.8.3) it is easily found that $du_{ij} = d\gamma_{ij}$, $(i,j) \in \Delta$, at $g = e$. This will be used in Example 7.7.7. □

5.9. Coset space of a group as a product of coset spaces of subgroups.

5.9.1. ASSUMPTION. *Let G and H be closed Lie subgroups of a Lie group K such that*

$$(5.9.1) \qquad K = GH.$$

This assumption means that every $k \in K$ can be written (not necessarily uniquely) in the form $k = gh$, $g \in G$, $h \in H$. By taking inverses it follows immediately that (5.9.1) is equivalent to $K = HG$. More about this structure of K will be said in Section 7.6 where the interest lies in obtaining Haar measure on K from Haar measures on G and H. This section, on the other hand, is concerned with another question as a preparation for Theorem 8.12. There K acts transitively on a space and G_0, H_0, K_0 are the isotropy subgroups of G, H, K, respectively, at an arbitrarily chosen point of the space. The question is whether K/K_0 can be brought into 1-1 bi-analytic correspondence with $G/G_0 \times H/H_0$. Under certain assumptions this will indeed be possible. The method of proof makes use again of a canonical chart.

Of fundamental importance here, and in Section 7.6, is an ingenious device of Bourbaki (1963, VII §2.9) in which K appears as a homogeneous space under the action of $G \times H$. It has not been assumed that G and H have only e in common (although in most applications this will be the case). Let

$$F = G \cap H \tag{5.9.2}$$

as a subset of K, and

$$F^* = \{(g, g) : g \in F\}, \tag{5.9.3}$$

considered as a subset of $G \times H$. Now define the left action of $G \times H$ on K by

$$(g, h)k = gkh^{-1}, \quad g \in G, \ h \in H, \ k \in K. \tag{5.9.4}$$

This action is transitive, and the isotropy subgroup of $G \times H$ at $e \in K$ is F^*. Therefore, there is a 1-1 correspondence

$$(G \times H)/F^* \leftrightarrow K. \tag{5.9.5}$$

5.9.2. ASSUMPTION. *The 1-1 correspondence (5.9.5) is a homeomorphism.*

Assumption 5.9.2 will be satisfied if the action (5.9.4) of $G \times H$ on K is proper (Corollary 2.3.15) or if G and H are second countable (Lemma 2.3.17). The latter will be true if K is second countable, since G and H are subspaces of K.

5.9.3. ASSUMPTION. *G_0, H_0, K_0 are closed Lie subgroups of G, H, K, respectively, such that (i) $G \cap H = G_0 \cap H_0$, and (ii) $K_0 = G_0 H_0$.*

5.9.4. ASSUMPTION. *$hG_0h^{-1} = G_0$ for every $h \in H$.*

5.9.5. LEMMA. *Assume 5.9.1 and 5.9.3, and let $g_1, g_2 \in G$, $h_1, h_2 \in H$. Then $g_1h_1 = g_2h_2$ implies $g_2 \in g_1G_0$.*

PROOF. $g_1^{-1}g_2 = h_1h_2^{-1} \in G \cap H = G_0 \cap H_0$ (by 5.9.3(i)) $\subset G_0$. □

5.9.6. LEMMA. *Assume 5.9.1, 5.9.3, and 5.9.4, and let $g_1, g_2 \in G$, $h_1, h_2 \in H$. Then $g_2h_2 \in g_1h_1K_0$ implies $g_2 \in g_1G_0$, $h_2 \in h_1H_0$.*

PROOF. By 5.9.3(ii) and 5.9.4 we may write $g_2h_2 \in g_1h_1K_0$ in the form $g_2h_2 \in g_1G_0h_1H_0$. Thus, there exists $g_0 \in G_0$, $h_0 \in H_0$ such that $g_2h_2 = g_1g_0h_1h_0$. By Lemma 5.9.5, $g_2 \in g_1g_0G_0 = g_1G_0$. Therefore, $g_2h_2 \in g_1h_1K_0$ implies $h_2 \in (g_1G_0)^{-1}g_1h_1K_0 = G_0h_1K_0 =$ (by 5.9.4) $h_1G_0K_0 =$ (by 5.9.3(ii)) $h_1G_0H_0 = h_1H_0G_0$. Therefore, there exists $g_0' \in G_0$, $h_0' \in H_0$ such that $h_2 = h_1h_0'g_0'$. Then $(h_1h_0')^{-1}h_2 = g_0' \in G \cap H =$ (by 5.9.3(i)) $G_0 \cap H_0 \in H_0$ so that $h_2 \in h_1h_0'H_0 = h_1H_0$. □

Consider the coset spaces K/K_0, G/G_0, H/H_0. It will be convenient not to distinguish between the points of K/K_0 and the cosets kK_0 as subsets of K. Similarly, gG_0 and hH_0. Define the function $\phi : G/G_0 \times H/H_0 \to K/K_0$ by

$$\phi(gG_0, hH_0) = ghK_0. \tag{5.9.6}$$

5.9.7. LEMMA. *Assume 5.9.1, 5.9.3, and 5.9.4. Then ϕ of (5.9.6) is well-defined and bijective.*

PROOF. Since any $k \in K$ can be written in the form gh, by (5.9.1), any $kK_0 \in K/K_0$ is of the form ghK_0 and ϕ maps therefore onto. In order to show that ϕ is 1-1, suppose $\phi(g_1G_0, h_1H_0) = \phi(g_2G_0, h_2H_0)$, with $g_i \in G$, $h_i \in H$ for $i = 1, 2$. By (5.9.6) we have then $g_1h_1K_0 = g_2h_2K_0$ and it follows from Lemma 5.9.6 that $g_1G_0 = g_2G_0$, $h_1H_0 = h_2H_0$. □

The function ϕ is obviously analytic since group multiplication in K is analytic. It will be shown now that ϕ^{-1} is analytic, so that

ϕ establishes an analytic diffeomorphism between $G/G_0 \times H/H_0$ and K/K_0. Since G and H are submanifolds of K, they inherit their analytic structures from K. It was shown in Section 5.8 how unique analytic structures are placed on the coset spaces K/K_0, G/G_0, and H/H_0. Now K acts analytically and transitively on the left of K/K_0 so that a chart at one point of K/K_0 can be transferred to any other point. Similarly, $G \times H$ acts analytically and transitively on the left of $G/G_0 \times H/H_0$ so that a chart can be transferred from any point to any other point. Therefore, it suffices to prove the bi-analyticity of ϕ in a neighborhood of one point. In the proof we shall need the following lemma.

5.9.8. LEMMA. *Assume 5.9.1, 5.9.2, and 5.9.3. Let* $\mathfrak{g}$, $\mathfrak{h}$, $\mathfrak{k}$, $\mathfrak{g}_0$, $\mathfrak{h}_0$, $\mathfrak{k}_0$ *be the Lie algebras of* G, H, K, G_0, H_0, K_0, *respectively. Then*

$$\mathfrak{g} + \mathfrak{h} = \mathfrak{k}, \tag{5.9.7}$$

$$\mathfrak{g}_0 + \mathfrak{h}_0 = \mathfrak{k}_0. \tag{5.9.8}$$

PROOF. Let $\dim K = d$, $\dim G = d_1$, $\dim H = d_2$, and $\dim F = d_3$ (F defined in (5.9.2)). Then also $\dim F^* = d_3$. Furthermore, $\dim(G \times H) = d_1 + d_2$. Then according to Section 5.8 $\dim((G \times H)/F^*) = d_1 + d_2 - d_3$. This must equal $\dim K$ by the homeomorphism assumption 5.9.2 (manifolds are locally Euclidean, and open subsets of Euclidean spaces of different dimensions cannot be homeomorphic; see, e.g., Dugundji, 1966, Chap. XVI, Theorem 6.3). This furnishes the equation

$$d_1 + d_2 - d_3 = d. \tag{5.9.9}$$

Observe that the Lie algebra of $F = G \cap H$ is $\mathfrak{g} \cap \mathfrak{h}$, that any Lie group and its Lie algebra have the same dimension, and that a Lie algebra is a finite dimensional vector space. Therefore, the left-hand side of (5.9.9) equals $\dim \mathfrak{g} + \dim \mathfrak{h} - \dim(\mathfrak{g} \cap \mathfrak{h}) = \dim(\mathfrak{g} + \mathfrak{h})$, and the right-hand side equals $\dim K$ so that (5.9.9) can be written as

$$\dim(\mathfrak{g} + \mathfrak{h}) = \dim \mathfrak{k}. \tag{5.9.10}$$

Furthermore, since both $\mathfrak{g}$ and $\mathfrak{h}$ are contained in $\mathfrak{k}$, the same is true of $\mathfrak{g}+\mathfrak{h}$. Then it follows from (5.9.10) that the linear spaces $\mathfrak{g}+\mathfrak{h}$ and $\mathfrak{k}$ must be equal, proving (5.9.7).

Equation (5.9.8) follows from (5.9.7) by applying the latter result to G_0, H_0, K_0 instead of G, H, K, taking into account Assumption 5.9.3(ii). Also needed is a 1-1 correspondence that results from (5.9.5) by replacing G, H, K by G_0, H_0, K_0, respectively,

$$(G_0 \times H_0)/F^* \leftrightarrow K_0. \tag{5.9.11}$$

Note that F^* in (5.9.11) is the same as in (5.9.5) by virtue of Assumption 5.9.3(i). Consequently, (5.9.11) is the restriction of (5.9.5) to G_0, H_0, K_0 and therefore the homeomorphism of (5.9.11) follows from that of (5.9.5). □

5.9.9. THEOREM. *Under Assumptions 5.9.1 through 5.9.4 the function ϕ of (5.9.6) is an analytic diffeomorphism.*

PROOF. As a consequence of Lemma 5.9.7 and the discussion following its proof the only thing left to prove is that the correspondence furnished by ϕ between some neighborhood of $([e],[e])$ in $G/G_0 \times H/H_0$ and some neighborhood of $[e]$ in K/K_0 is bi-analytic. These neighborhoods will be defined by an appropriately chosen canonical chart at $e \in K$, whose construction will now be given. Denote by $\mathfrak{f}$ the Lie algebra of $F = G \cap H = G_0 \cap H_0$ (Assumption 5.9.3(i)). The Lie algebras $\mathfrak{g}$, $\mathfrak{h}$, etc. were defined in the hypothesis of Lemma 5.9.8. Then $F \subset G_0 \subset G$ implies $\mathfrak{f} \subset \mathfrak{g}_0 \subset \mathfrak{g}$. Choose a basis of $\mathfrak{f}$ consisting of vectors (actually, vector fields) $W_1, \dots , W_{d_3}$, where $d_3 = \dim F$. Denote these vectors generically by W_m $(m = 1, \dots , d_3)$, and, for short, refer to this basis as W. Extend W to a basis (U, W) of $\mathfrak{g}_0$ by choosing additional vectors U_k, and then extend (U, W) to a basis (X, U, W) of G by choosing additional vectors X_i. Similarly, we have $F \subset H_0 \subset H$ and by choosing vectors V_ℓ in $\mathfrak{h}_0$ and Y_j in $\mathfrak{h}$ appropriately we have a basis (V, W) of $\mathfrak{h}_0$ and (Y, V, W) of $\mathfrak{h}$. Then (X, Y, U, V, W) is a basis of $\mathfrak{g}+\mathfrak{h}$ and therefore of $\mathfrak{k}$, by (5.9.7). Furthermore, (U, V, W) is a basis of $\mathfrak{g}_0 + \mathfrak{h}_0$ and therefore of $\mathfrak{k}_0$, by (5.9.8). The various charts that we need will be defined in terms of the above vectors via the exponential map. This involves linear combinations of the vectors, but in order to avoid bulky notation with sums and subscripts we shall make the convention that an expression such as, for instance, $\sum w_m W_m$ (sum over m from 1 to d_3) will be abbreviated wW. Similarly, xX, etc. With this convention there is on G at e a canonical chart

$$(x, u, w) \to \exp(xX)\exp(uU + wW). \tag{5.9.12}$$

By the proof of Theorem 5.8.1, x forms a coordinate system of a chart on a neighborhood, say A, of G/G_0 at $[e]$. Similarly,

$$(y, v, w) \to \exp(yY)\exp(vV + wW) \tag{5.9.13}$$

is a chart on H at e, so y is a coordinate system of a chart on a neighborhood, say B, of H/H_0 at $[e]$. Finally, as a chart on K at e choose

$$(x, y, u, v, w) \to \exp(xX)\exp(yY)\exp(uU + vV + wW), \tag{5.9.14}$$

and use the fact that (U, V, W) is a basis of $\mathfrak{k}_0$ to conclude that (x, y) is a coordinate system of a chart on a neighborhood, say C, of K/K_0 at $[e]$. Thus, there is the same coordinate system (x, y) on $A \times B$ as on C. Restrict ϕ of (5.9.6) to $A \times B$, which is possible by taking x, etc. in (5.9.12)–(5.9.14) small enough. In (5.9.6) let g and h be given by the right-hand sides of (5.9.12) and (5.9.13), respectively (but not necessarily with the same value of w). Write this as

$$g = \exp(xX)g_0, \quad h = \exp(yY)h_0, \tag{5.9.15}$$

for some $g_0 \in G_0$, $h_0 \in H_0$. Then on the right-hand side of (5.9.6) we have $gh = \exp(xX)g_0\exp(yY)h_0 = \exp(xX)\exp(yY)g_0'h_0$ for some $g_0' \in G_0$ by Assumption 5.9.4. Furthermore, $g_0'h_0 = k_0 \in K_0$ by Assumption 5.9.3(ii). Therefore,

$$gh = \exp(xX)\exp(yY)k_0 \tag{5.9.16}$$

for some $k_0 = K_0$. We have now that gG_0 and hH_0 in (5.9.6) have coordinates x, y, respectively, and ghK_0 has coordinates (x, y) by comparing (5.9.16) to the chart (5.9.14). Hence, if ϕ on $A \times B$ is written as a function, say ϕ^*, of (x, y), then

$$\phi^*(x, y) = (x, y). \tag{5.9.17}$$

This is the identity function, therefore trivially an analytic diffeomorphism. □

5.9.10. REMARK. In Theorem 5.9.9 the groups G and H are treated asymmetrically. It is true that Assumptions 5.9.1, 5.9.2, and 5.9.3 are symmetric in G and H, but Assumption 5.9.4 is not, and the order of g and h on the right-hand side of (5.9.6) also matters. The theorem remains of course true if in Assumption 5.9.4 G and H are interchanged and on the right-hand side of (5.9.6) gh is replaced by hg. This will be used in Chapter 11. □