Chapter 5

Lie Groups and Lie Algebras

5.1 The Exponential of a Matrix

Let $\mathbf{A}$ be a linear operator on the n-dimensional inner product space V. The **norm** $\|\mathbf{A}\|$ of $\mathbf{A}$ is defined by the expression

$$\|\mathbf{A}\| = \max_{\mathbf{v}\in V, \|\mathbf{v}\|=1} \{\|\mathbf{A}\mathbf{v}\|\}. \tag{1.1}$$

[Recall that the norm of a vector is $\|\mathbf{v}\| = (\mathbf{v}, \mathbf{v})^{1/2}$.] To see that this definition makes sense choose an ON basis $\{\mathbf{v}_j\}$ for V with respect to which $\mathbf{A}$ has matrix elements A_{ij} and $\mathbf{v}$ has vector components a_j. Then

$$\|\mathbf{A}\| = \max_{|a_1|^2+\cdots+|a_n|^2=1} \left[\sum_{i=1}^{n} \left|\sum_{j=1}^{n} A_{ij}a_j\right|^2\right]^{1/2} \tag{1.2}$$

This maximum clearly exists, so $\|\mathbf{A}\|$ is well-defined. (We have assumed that V is complex in this computation. There is an analogous result for V real.)

If $\mathbf{v}$ is a nonzero element of V then $\mathbf{v}/\|\mathbf{v}\| = \mathbf{w}$ has norm one. Thus $\|\mathbf{A}\mathbf{w}\| \leq \|\mathbf{A}\|$ from (1.1), so

$$\|\mathbf{A}\mathbf{v}\| \leq \|\mathbf{A}\| \cdot \|\mathbf{v}\|, \tag{1.3}$$

where we have used the fact $\|\alpha\mathbf{v}\| = \alpha\|\mathbf{v}\|$ for $\alpha \geq 0$. Expression (1.3) is also valid for $\mathbf{v} = \boldsymbol{\theta}$.

Lemma 5.1. Let $\mathbf{A}$ and $\mathbf{B}$ be linear operators on V. Then $\|\mathbf{AB}\| \leq \|\mathbf{A}\| \cdot \|\mathbf{B}\|$.

Proof. Let $\mathbf{v} \in V$ with $\|\mathbf{v}\| = 1$. By (1.3), $\|(\mathbf{AB})\mathbf{v}\| = \|\mathbf{A}(\mathbf{Bv})\| \leq \|\mathbf{A}\| \cdot \|\mathbf{Bv}\| \leq \|\mathbf{A}\| \cdot \|\mathbf{B}\|$. Therefore $\|\mathbf{AB}\| \leq \|\mathbf{A}\| \cdot \|\mathbf{B}\|$. Q.E.D.

Corollary 5.1. $\|\mathbf{A}^m\| \leq \|\mathbf{A}\|^m$, $m = 1, 2, \ldots$.

Proof. Lemma 5.1 with $\mathbf{A} = \mathbf{B}$, and induction on m.

Lemma 5.2. Let $\mathbf{A}$ and $\mathbf{B}$ be linear operators on V. Then $\|\alpha\mathbf{A}\| = \alpha\|\mathbf{A}\|$ for all $\alpha \geq 0$ and $\|\mathbf{A} + \mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|$.

Proof. These results follow directly from the definition (1.1) and the relations $\|\alpha\mathbf{Av}\| = \alpha\|\mathbf{Av}\|$ and $\|\mathbf{Av} + \mathbf{Bv}\| \leq \|\mathbf{Av}\| + \|\mathbf{Bv}\|$.

We can define the matrix norm $\|A\|$ of an $n \times n$ matrix $A = (A_{ij})$ by $\|A\| = \|\mathbf{A}\|$, where $\mathbf{A}$ is the linear operator on V determined by A and the ON basis $\{\mathbf{v}_j\}$. In particular, $\|A\|$ is given explicitly by (1.2). Thus, the results proved above for operator norms hold as well for matrix norms.

Recall that a sequence of complex numbers $\{\alpha_j\}$ is a **Cauchy sequence** if for any $\varepsilon > 0$ there is an integer $N_\varepsilon > 0$ such that $|\alpha_i - \alpha_j| < \varepsilon$ for all $i, j > N_\varepsilon$. Every Cauchy sequence converges, i.e., if $\{\alpha_j\}$ is Cauchy then there is an $\alpha \in \mathfrak{C}$ such that $\alpha_j \to \alpha$. A sequence $\{A_k\}$ of $n \times n$ matrices is said to be a **Cauchy sequence in the norm** $\|\cdot\|$ if for every $\varepsilon > 0$ there is an integer $N_\varepsilon > 0$ such that $\|A_i - A_j\| < \varepsilon$ for all $i, j > N_\varepsilon$. The sequence $\{A_k\}$ is Cauchy in the norm if and only if for each $i, j = 1, \ldots, n$ the matrix elements $\{A_{k,ij}, k = 1, 2, \ldots\}$ in the ith row and jth column form a Cauchy sequence of numbers. (Prove it.) Thus, if $\{A_k\}$ is a Cauchy sequence of matrices, the limits $A_{ij} = \lim_{k\to\infty} A_{k,ij}$ exist and $A_k \to A$ with $A = (A_{ij})$. Then for any $\varepsilon > 0$ there is an integer $M_\varepsilon > 0$ such that $\|A - A_k\| < \varepsilon$ whenever $k > M_\varepsilon$. We have sketched a proof of the fact that every Cauchy sequence of matrices converges. Passing from matrices to linear operators, we immediately obtain the result that every Cauchy sequence of operators converges. Most of the results obtained in this section are valid for both operators and matrices, but the formulation of the corresponding operator results is left to the reader.

The exponential $\exp A$ of an $n \times n$ matrix A is defined by the sum

$$\exp A = \sum_{j=0}^{\infty} \frac{A^j}{j!}, \tag{1.4}$$

where $A^0 = E_n$ (the $n \times n$ identity matrix). To justify this definition consider the finite sums

$$S_m = \sum_{j=0}^{m} \frac{A^j}{j!}, \qquad m = 0, 1, \ldots,$$

which are well-defined $n \times n$ matrices. For $k > m$ we have

$$\|S_k - S_m\| = \left\| \sum_{j=m+1}^{k} \frac{A^j}{j!} \right\| \leq \sum_{j=m+1}^{k} \frac{\|A\|^j}{j!}.$$

Since

$$e^{\|A\|} = \sum_{j=0}^{\infty} \frac{\|A\|^j}{j!} < \infty,$$

then for any $\varepsilon > 0$ there is an integer $N_\varepsilon > 0$ with the property $\sum_{j=q}^{\infty}(\|A\|^j/j!) < \varepsilon$ for every $q > N_\varepsilon$. This proves that the sequence of matrices $\{S_m\}$ is Cauchy and $\exp A = \lim_{m\to\infty} S_m$ exists.

Theorem 5.1. If A and B are $n \times n$ matrices and B is nonsingular then $\exp(BAB^{-1}) = B(\exp A)B^{-1}$.

Proof.

$$\exp(BAB^{-1}) = \sum_{j=0}^{\infty} \frac{(BAB^{-1})^j}{j!} = \sum_{j=0}^{\infty} \frac{BA^jB^{-1}}{j!} = B(\exp A)B^{-1},$$

since

$$\|BS_mB^{-1} - B(\exp A)B^{-1}\| \leq \|B\|\cdot\|S_m - \exp A\|\cdot\|B^{-1}\| \longrightarrow 0$$

as $m \longrightarrow \infty$. Q.E.D.

We review a few basic facts about the eigenvalues of matrices. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of the $n \times n$ matrix A, each eigenvalue repeated a number of times equal to its multiplicity. The λ_j are the solutions of the characteristic equation

$$P_A(\lambda) = \det(A - \lambda E_n) = 0.$$

If the matrix C is similar to A, i.e., if there exists a nonsingular $n \times n$ matrix B such that $C = BAB^{-1}$, then

$$\begin{aligned} P_C(\lambda) &= \det(BAB^{-1} - \lambda BE_nB^{-1}) = (\det B)\det(A - \lambda E_n)\det B^{-1} \\ &= \det(A - \lambda E_n) = P_A(\lambda). \end{aligned}$$

Thus, similar matrices have the same eigenvalues. The eigenvalues of a diagonal matrix are just the diagonal elements. We say that a matrix A can be diagonalized by similarity transformations if A is similar to a diagonal matrix D. The diagonal elements of D are just the eigenvalues of A.

Not all $n \times n$ matrices can be diagonalized. For example the matrix

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

has eigenvalues $\lambda_1 = \lambda_2 = 0$. Since A is obviously not similar to the zero matrix, it cannot be diagonalized. It *is* true that every complex matrix A is similar to a matrix C in upper triangular form. An $n \times n$ matrix C is **upper**

triangular if $C_{ij} = 0$ for all $1 \leq j < i \leq n$, i.e., if all matrix elements below the main diagonal are zero. The eigenvalues of a upper triangular matrix are just the diagonal elements:

$$C = \begin{pmatrix} \lambda_1 & & * & * & * \\ & \lambda_2 & & * & * \\ & & \cdot & & \\ & & & \cdot & * \\ & Z & & & \lambda_n \end{pmatrix}. \tag{1.5}$$

Theorem 5.2. Let A be a complex $n \times n$ matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$. Then A is similar to an upper triangular matrix C with diagonal elements $\lambda_1, \ldots, \lambda_n$.

Proof. Induction on n. The theorem is obvious for $n = 1$. Assume it is true for all $(n-1) \times (n-1)$ matrices and let A be a matrix of degree n. Since λ_1 is an eigenvalue of A there is a nonzero column vector $v = (v_1, \ldots, v_n)$ such that $Av = \lambda_1 v$, i.e., $\sum_{j=1}^{n} A_{ij} v_j = \lambda_1 v_i$. Let e_1 be the column vector defined by the n-tuple $(1, 0, \ldots, 0)$. Clearly, there exists an $n \times n$ nonsingular matrix B_1 such that $B_1 v = e_1$. Then the matrix $B_1 A B_1^{-1}$ is similar to A and takes the form

$$B_1 A B_1^{-1} = \begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & & & \\ \cdot & & A' & \\ \cdot & & & \\ \cdot & & & \\ 0 & & & \end{pmatrix},$$

where A' is an $(n-1) \times (n-1)$ matrix. Indeed, $B_1 A B_1^{-1} e_1 = B_1 A v = \lambda_1 B_1 v = \lambda_1 e_1$. Since $\det(A - \lambda E_n) = \det(B_1 A B_1^{-1} - \lambda E_n) = (\lambda_1 - \lambda) \times \det(A' - \lambda E_{n-1})$, the numbers $\lambda_2, \ldots, \lambda_n$ are the eigenvalues of A'. Furthermore, by the induction hypothesis there exists an $(n-1) \times (n-1)$ nonsingular matrix B_2' such that $B_2' A' (B_2')^{-1}$ is upper triangular. Defining the $n \times n$ matrix

$$B_2 = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \cdot & & B_2' & \\ \cdot & & & \\ \cdot & & & \\ 0 & & & \end{pmatrix},$$

we find $B_2(B_1 A B_1^{-1}) B_2^{-1} = C$, where C is an $n \times n$ upper triangular matrix, Eq. (1.5). Thus $BAB^{-1} = C$, where $B = B_2 B_1$. Q.E.D.

Corollary 5.2. Let A be an $n \times n$ matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$. Then $\exp A$ has eigenvalues $e^{\lambda_1}, \ldots, e^{\lambda_n}$.

Proof. The matrix A is similar to an upper triangular matrix C [Eq. (1.5)]. An elementary computation yields

$$C^m = \begin{pmatrix} \lambda_1{}^m & & * & * & * \\ & \lambda_2{}^m & & * & * \\ & & \ddots & & \\ & Z & & \ddots & \\ & & & & \lambda_n{}^m \end{pmatrix}, \qquad m = 1, 2, \ldots,$$

$$\exp C = \sum_{m=0}^{\infty} \frac{C^m}{m!} = \begin{pmatrix} e^{\lambda_1} & & * & * & * \\ & e^{\lambda_2} & & * & * \\ & & \ddots & & \\ & Z & & \ddots & \\ & & & & e^{\lambda_n} \end{pmatrix}.$$

Thus the eigenvalues of $\exp C$ are $e^{\lambda_1}, \ldots, e^{\lambda_n}$. Since A and C are similar there exists a nonsingular matrix B such that $BAB^{-1} = C$. By Theorem 5.1, $B(\exp A)B^{-1} = \exp(BAB^{-1}) = \exp C$, so $\exp A$ and $\exp C$ are similar. Similar matrices have the same eigenvalues. Q.E.D.

Corollary 5.3. $\det(\exp A) = \exp(\operatorname{tr} A) = e^{\operatorname{tr} A}$. In particular, $\exp A$ is a nonsingular matrix.

Proof. The determinant and trace of a matrix are the product and sum of its eigenvalues. By Theorem 5.2, $\det(\exp A) = \exp(\lambda_1 + \cdots + \lambda_n) = \exp(\operatorname{tr} A)$. Q.E.D.

We know from complex variable theory that corresponding to any power series

$$f(z) = \sum_{j=0}^{\infty} c_j z^j, \qquad c_j \in \mathfrak{C}, \tag{1.6}$$

there is associated a number $r \geq 0$, the **radius of convergence**, such that the power series converges absolutely if $|z| < r$ and diverges if $|z| > r$. In case the power series converges for all z, we set $r = +\infty$ (see Ahlfors [1].) By convergence of the power series for fixed z we mean that the partial sums

$$S_m(z) = \sum_{j=0}^{m} c_j z^j, \qquad m = 0, 1, \ldots, \tag{1.7}$$

form a Cauchy sequence of complex numbers. The limit of this Cauchy sequence is $f(z)$. The only properties of the absolute value needed to prove

these results are

$$(1.8) \qquad |a+b| \leq |a| + |b|, \qquad |\alpha a| = \alpha |a|, \qquad |ab| \leq |a| \cdot |b|,$$

for $a, b \in \mathfrak{C}$, $\alpha \geq 0$. Since the matrix norm $\|A\|$ shares these properties we can immediately carry over the fundamental results on power series of complex numbers to power series of matrices. Thus, if $f(z)$ is the power series (1.6) and A is an $n \times n$ matrix we define the matrix power series

$$(1.9a) \qquad f(A) = \sum_{j=0}^{\infty} c_j A^j.$$

This series converges to an $n \times n$ matrix provided the partial sums

$$(1.9b) \qquad S_m(A) = \sum_{j=0}^{m} c_j A^j, \qquad m = 0, 1, 2, \ldots,$$

form a Cauchy sequence with respect to the matrix norm $\|\cdot\|$. Then

$$f(A) = \lim_{m \to \infty} S_m(A).$$

It follows from the above remarks that there exists a nonegative number r such that (1.8) converges if $\|A\| < r$ and diverges if $\|A\| > r$. For example. the power series defining $\exp A$ corresponds to $r = +\infty$.

If $f_1(z) = \sum a_j z^j, f_2(z) = \sum b_k z^k$ have radii of convergence r_1, r_2, respectively, then

$$f_1(z) f_2(z) = (\sum a_j z^j)(\sum b_k z^k) = \sum c_l z^l, \qquad |z| < r,$$

where

$$c_l = \sum_{j=0}^{l} a_j b_{l-j}, \qquad r = \min\{r_1, r_2\}.$$

The proof of this result uses only the properties (1.8) (Ahlfors [1]), hence it carries over immediately to matrix functions $f_1(A) f_2(A)$. For example, the identity $e^{-z} e^z = 1$ for all z leads to the identity

$$(1.10) \qquad \exp(-A) \exp(A) = E_n$$

for all $n \times n$ matrices A. Similarly, the identity $(e^z)^m = e^{mz}$, $m = 0, \pm 1, \ldots$, leads to $[\exp A]^m = \exp(mA)$.

Since $\ln(1+z) = \sum_{j=1}^{\infty} (-1)^{j+1} z^j / j$ converges for $|z| < 1$ we can define the $n \times n$ matrix

$$(1.11) \qquad \ln(E_n + A) = \sum_{j=1}^{\infty} (-1)^{j+1} \frac{A^j}{j}, \qquad \|A\| < 1.$$

Furthermore, the formulas

$$\exp[\ln(1+z)] = 1 + z, \quad |z| < 1, \qquad \ln(\exp z) = z, \quad |z| < \ln 2,$$

which are proved by power series expansion, immediately lead to

(1.12)
$$\exp[\ln(E_n + A)] = E_n + A, \quad \|A\| < 1; \qquad \ln(\exp A) = A, \quad \|A\| < \ln 2.$$

Let L_n be the space of all $n \times n$ matrices, either real or complex, and let G_n be the group of all nonsingular matrices in L_n.

Theorem 5.3. The exponential mapping $A \longrightarrow \exp A$ transforms L_n into G_n. There exist positive constants ε, $\delta < 1$ such that the neighborhood $\mathfrak{A}$ of the zero matrix in L_n is mapped 1–1 onto the neighborhood $\mathfrak{B}$ of the identity matrix E_n in G_n. Here

$$\mathfrak{A} = \{A \in L_n : \|A\| < \varepsilon\}, \qquad \mathfrak{B} = \{B \in L_n : \|B - E_n\| < \delta\}.$$

Furthermore the mapping $A \longrightarrow \exp A$, $A \in \mathfrak{A}$, and its inverse $B \longrightarrow \ln B$, $B \in \mathfrak{B}$, are analytic, i.e,. the matrix elements of $\exp A$ and $\ln B$ are analytic functions of the matrix elements of A and B, respectively.

The proof is immediate from identities (1.12). The theorem shows that any matrix $B \in G_n$ in a sufficiently small neighborhood of E_n can be written **uniquely** in the form $B = \exp A$, with A in a sufficiently small neighborhood of Z.

Let t be a parameter, real or complex. We write $A(t)$ to denote an $n \times n$ matrix whose components $A_{ij}(t)$ are functions of t. If each of the matrix components is differentiable we say that the matrix $A(t)$ is differentiable with derivative $\dot{A}(t) = dA(t)/dt$, the matrix with components $\dot{A}_{ij}(t)$. If $A(t)$ and $B(t)$ are differentiable $n \times n$ matrices, the identities

(1.13)
$$\begin{aligned}(d/dt)[\alpha A(t) + \beta B(t)] &= \alpha \dot{A}(t) + \beta \dot{B}(t)\\ (d/dt)[A(t)B(t)] &= \dot{A}(t)B(t) + A(t)\dot{B}(t)\end{aligned}$$

hold, as the reader can easily prove.

If the power series $f(z) = \sum c_j z^j$ has radius of convergence $r > 0$ then $f(z)$ is differentiable and $f'(z)$ is obtained by differentiating the power series for $f(z)$ term by term:

$$f'(z) = df(z)/dz = \sum_{j=1}^{\infty} jc_j z^{j-1}, \qquad |z| < r.$$

Since the proof of this result employs only the properties (1.8) it can easily be extended to compute the derivative of the matrix function $\exp(tA)$, A a constant matrix. Thus,

(1.14)
$$\frac{d}{dt}\exp(tA) = \sum_{j=1}^{\infty} \frac{t^{j-1}}{(j-1)!} A^j = A\exp(tA) = [\exp(tA)]A.$$

We can use this differential formula to derive other interesting properties of the exponential function. Note that the general solution of the matrix differ-

ential equation $\dot{A}(t) = Z$ (the zero matrix) is $A(t) = A_0$, where A_0 is any constant matrix. To show the utility of this remark we will use it to find the general solution of the equation $\dot{A}(t) = AA(t)$, where A is a constant matrix. Suppose $A(t)$ is a solution and consider the matrix $B(t) = [\exp(-tA)]A(t)$. From (1.13) and (1.14), we have

$$\dot{B}(t) = -[\exp(-tA)]AA(t) + [\exp(-tA)]AA(t) = Z$$

so $B(t)$ is a constant matrix $A_0 = B(0)$. Thus $A(t) = [\exp(tA)]A_0$. Conversely, it is easy to show that $A(t) = [\exp(tA)]A_0$ satisfies the equation for all matrices A_0.

It is an elementary exercise to show that $B(\exp A) = (\exp A)B$ for any $n \times n$ matrices such that $AB = BA$.

Theorem 5.4. If $AB = BA$ then $\exp(A + B) = (\exp A)(\exp B)$.

Proof. We could obtain this result by explicit power series expansion. A more elegant proof is obtained from a consideration of the matrix function

$$C(t) = \exp[-t(A + B)](\exp tA)(\exp tB).$$

From (1.14) we have

$$\dot{C}(t) = \{\exp[-t(A + B)]\}(-A - B + A + B)(\exp tA) \exp tB = Z.$$

Therefore, $C(t)$ is a constant matrix. Since $C(0) = E_n$ we find $\{\exp[-t(A + B)]\}$ $(\exp tA)$ $\exp tB = E_n$ or $(\exp tA)$ $\exp$ $tB = \exp [t(A + B)]$ for all t. Q.E.D.

If the matrices A and B do not commute then Theorem 5.4 is no longer valid and the simple analogy between the matrix exponential and the usual exponential breaks down. However, from Theorem 5.3, if A and B are sufficiently close to Z then there must exist a unique matrix C sufficiently close to Z such that $(\exp A) \exp B = \exp C$. If A and B commute then $C = A + B$, but in the general case C is a complicated function $C(A, B)$ of A and B. There is an explicit formula for $C(A, B)$ called the Campbell–Baker–Hausdorff formula, which we will derive by means of several lemmas, some of which are useful in their own right.

Let L_n be the n^2-dimensional space of all complex $n \times n$ matrices. For $A \in L_n$ we define the linear transformation Ad A on L_n by

$$\text{Ad } A(B) = [A, B], \tag{1.15}$$

where

$$[A, B] = AB - BA \tag{1.16}$$

is the **commutator bracket.** The matrix of Ad A with respect to a basis for L_n is $n^2 \times n^2$. Clearly Ad A is the zero operator if and only if A commutes with

all $B \in L_n$. By $(\mathrm{Ad}\ A)^m$ we mean the operator

$$(\mathrm{Ad}\ A)^m(B) = [A, [A, \ldots, [A, B]], \ldots] \quad (m \quad \text{times}).$$

We will frequently write e^A for $\exp A$.

Lemma 5.3. Let $A, B \in L_n$. Then

$$e^A B e^{-A} = (\exp(\mathrm{Ad}\ A))B = \sum_{j=0}^{\infty} (j!)^{-1}(\mathrm{Ad}\ A)^j(B)$$
$$= E_n + [A, B] + \tfrac{1}{2}[A, [A, B]] + \cdots.$$

Proof. Set $B(t) = e^{tA}Be^{-tA}$. Clearly $B(0) = B$ and the matrix elements of $B(t)$ are entire functions of t. Thus, there exist matrices $C_j \in L_n$ such that $B(t) = \sum C_j t^j$, where

$$C_0 = B, \qquad C_j = \frac{1}{j!}\frac{d^j}{dt^j}B(t)\bigg|_{t=0}, \qquad j = 1, 2, \ldots.$$

Now

$$\dot{B}(t) = Ae^{tA}Be^{-tA} - e^{tA}Be^{-tA}A = [A, B(t)] = \mathrm{Ad}\ A(B(t))$$

and by induction on j,

$$d^jB(t)/dt^j = (\mathrm{Ad}\ A)^j(B(t)).$$

We conclude that $C_j = (j!)^{-1}(\mathrm{Ad}\ A)^jB$. Q.E.D.

Lemma 5.4. Let $A(t) \in L_n$ such that each of the matrix elements $A_{ij}(t)$ is an analytic function of the parameter t and let

$$f(z) = \frac{e^z - 1}{z} = 1 + \frac{z}{2!} + \frac{z^2}{3!} + \cdots,$$

(an entire function of z). Then

$$e^{A(t)}\, de^{-A(t)}/dt = -f(\mathrm{Ad}\ A(t))(\dot{A}(t)). \tag{1.17}$$

Recall that we can consider $\mathrm{Ad}\ A(t)$ as an $n^2 \times n^2$ matrix.

Proof. Set $B(s, t) = e^{sA(t)}\, d(e^{-sA(t)})/dt$. Since $B(s, t)$ is an entire function of s it can be written in the form

$$B(s, t) = \sum_{j=0}^{\infty} C_j(t)s^j, \qquad C_j(t) = \frac{1}{j!}\frac{\partial^j}{\partial s^j}B(s, t)\bigg|_{s=0}.$$

Now

(1.18)

$$\partial B/\partial s = A(t)B(s, t) - B(s, t)A(t) - \dot{A}(t) = (\mathrm{Ad}\ A(t))B(s, t) - \dot{A}(t).$$

Differentiating this expression successively with respect to s we find

$$\partial^j B(s, t)/\partial s^j = (\mathrm{Ad}\ A(t))^j B(s, t) - (\mathrm{Ad}\ A(t))^{j-1}\dot{A}(t), \qquad j = 1, 2, \ldots.$$

Since $B(0, t) = 0$ we have

$$C_0(t) = Z, \qquad C_j(t) = (-1/j!)(\mathrm{Ad}\, A(t))^{j-1}\dot{A}(t), \qquad j = 1, 2, \ldots.$$

Thus,

$$e^{sA(t)}\frac{de^{-sA(t)}}{dt} = -\sum_{j=1}^{\infty} s^j \frac{(\mathrm{Ad}\, A(t))^{j-1}}{j!}\dot{A}(t).$$

Setting $s = 1$ we obtain (1.17). Q.E.D.

If $A, B \in L_n$ have sufficiently small norms there exists a unique $C \in L_n$ sufficiently close to Z such that $e^A e^B = e^C$, or $C = \ln(e^A e^B)$, where the logarithm is defined by (1.11). That C must lie sufficiently close to Z for uniqueness follows from the identity $\exp(C + 2\pi i D) = \exp C$, where D is any diagonal matrix with integer coefficients. The logarithm exists provided $\| e^A e^B - E_n \| < 1$. The function

$$g(z) = \frac{\ln z}{z-1} = \sum_{j=0}^{\infty} \frac{(1-z)^j}{j+1} = 1 + \frac{1}{2}(1-z) + \frac{1}{3}(1-z)^2 + \cdots$$

is analytic for $|1 - z| < 1$. Thus the expression

$$g(F) = E_k + \tfrac{1}{2}(E_k - F) + \tfrac{1}{3}(E_k - F)^2 + \cdots$$

defines a matrix for any $k \times k$ matrix F with $\| E_k - F \| < 1$.

Theorem 5.5 (Campbell–Baker–Hausdorff). For A, B in a sufficiently small neighborhood of Z in L_n and $C = \ln(e^A e^B)$, then

$$C = B + \int_0^1 g[\exp(t\, \mathrm{Ad}\, A)\exp(\mathrm{Ad}\, B)](A)\, dt. \tag{1.19}$$

Indeed,

$$C = A + B + \tfrac{1}{2}[A, B] + \tfrac{1}{12}[A, [A, B]] - \tfrac{1}{12}[B, [B, A]] + \cdots \tag{1.20}$$

and the matrix elements of C are analytic functions of the matrix elements of A and B.

Proof. We will derive a differential equation for $C(t) = \ln(e^{tA}e^B)$, where $0 \le t \le 1$. Now, $e^{tA}e^B = e^{C(t)}$ and from Lemma 5.3,

$$\{\exp[\mathrm{Ad}\, C(t)]\}H = e^{tA}(e^B H e^{-B})e^{-tA} = [\exp(\mathrm{Ad}\, tA)](\exp \mathrm{Ad}\, B)H$$

for any $H \in L_n$. Thus

$$\exp[\mathrm{Ad}\, C(t)] = [\exp(\mathrm{Ad}\, tA)]\exp(\mathrm{Ad}\, B), \tag{1.21}$$

and for A, B such that $\| [\exp(\mathrm{Ad}\, tA)]\exp(\mathrm{Ad}\, B) - E_{n^2} \| < 1$ for all $0 \le t \le 1$, then

$$\mathrm{Ad}\, C(t) = \ln\{[\exp(\mathrm{Ad}\, tA)]\exp \mathrm{Ad}\, B\}. \tag{1.22}$$

By definition of $C(t)$,

$$e^{C(t)}\, de^{-C(t)}/dt = -e^{tA}e^{B}e^{-B}e^{-tA}A = -A.$$

Lemma 5.4 implies

$$f(\mathrm{Ad}\, C(t))\dot{C}(t) = A, \tag{1.23}$$

where $f(z) = (e^z - 1)/z$. We can solve this expression for $\dot{C}(t)$ by noting that $f(\ln z)g(z) = 1$ for $|1 - z| < 1$. Thus, $f(\ln F)g(F) = E_{n^2}$ for any $n^2 \times n^2$ matrix F with $\| E_{n^2} - F \| < 1$. In particular, $g(F) = f(\ln F)^{-1}$. Setting $F = [\exp(\mathrm{Ad}\, tA)] \exp(\mathrm{Ad}\, B)$ and using (1.22), we see that (1.23) becomes

$$\dot{C}(t) = g\{[\exp(\mathrm{Ad}\, tA)] \exp(\mathrm{Ad}\, B)\}(A) \tag{1.24}$$

or

$$C(t) = \int_0^t g\{[\exp(\mathrm{Ad}\, tA)] \exp(\mathrm{Ad}\, B)\}(A)\, dt + C_0, \tag{1.25}$$

where C_0 is a constant matrix. Since $C(0) = \ln e^B = B$, $C_0 = B$. Expression (1.19) follows from (1.25) by setting $t = 1$. Expression (1.20) is obtained by writing out the power series expansion of the integrand in (1.19) and integrating term by term. Q.E.D.

The preceding theorem will prove to be of great importance in the theory of Lie groups.

5.2 Local Lie Groups

Roughly speaking, a Lie group is an infinite group whose elements can be parametrized analytically. Thus, any group element g can be denoted $g(\lambda_1, \ldots, \lambda_n)$ in terms of parameters $\lambda_1, \ldots, \lambda_n$. The parameters of the product gg' are analytic functions of the parameters of g and g'. Using the notion of Lie group one can apply both calculus and algebra to group theory and obtain many results which neither discipline alone yields.

We denote by F the field of either the real numbers R or complex numbers $\mathbb{C}$. Let F_n be the vector space of n-tuples $\mathbf{g} = (g_1, \ldots, g_n)$, $g_i \in F$, and let $\mathbf{e} = (0, \ldots, 0)$ be the zero vector in F_n. Suppose V is an open set in F_n containing $\mathbf{e}$. (We adopt the usual topology for n-tuples. The reader can choose V as an open sphere in n-space.)

Definition. An n-dimensional **local Lie group** G in the neighborhood $V \subseteq F_n$ is determined by a function $\boldsymbol{\varphi}(\mathbf{g}, \mathbf{h})$ with the following properties:

(1) $\boldsymbol{\varphi}(\mathbf{g}, \mathbf{h}) \in F_n$ for all $\mathbf{g}, \mathbf{h} \in V$.
(2) $\boldsymbol{\varphi}(\mathbf{g}, \mathbf{h})$ is an analytic function of each of its $2n$ arguments.
(3) If $\boldsymbol{\varphi}(\mathbf{g}, \mathbf{h}) \in V$ and $\boldsymbol{\varphi}(\mathbf{h}, \mathbf{k}) \in V$ then $\boldsymbol{\varphi}(\boldsymbol{\varphi}(\mathbf{g}, \mathbf{h}), \mathbf{k}) = \boldsymbol{\varphi}(\mathbf{g}, \boldsymbol{\varphi}(\mathbf{h}, \mathbf{k}))$.
(4) $\boldsymbol{\varphi}(\mathbf{e}, \mathbf{g}) = \boldsymbol{\varphi}(\mathbf{g}, \mathbf{e}) = \mathbf{g}$ for all $\mathbf{g} \in V$.

To make the above definition more palatable we write $\boldsymbol{\varphi}(\mathbf{g}, \mathbf{h}) = \mathbf{gh}$. [We interpret $\boldsymbol{\varphi}(\mathbf{g}, \mathbf{h})$ as the product of the group elements $\mathbf{g}$ and $\mathbf{h}$.] Then property (3) is easily recognized as the associative law $(\mathbf{gh})\mathbf{k} = \mathbf{g}(\mathbf{hk})$. Property (4) becomes $\mathbf{eg} = \mathbf{ge} = \mathbf{g}$ and shows that $\mathbf{e}$ is the identity element for group multiplication.

We have not required the existence of an inverse for a group element. However, this fact follows from properties (2)–(4). An inverse for $\mathbf{g} \in V$ would be a group element $\mathbf{x}$ such that $\boldsymbol{\varphi}(\mathbf{g}, \mathbf{x}) = \mathbf{gx} = \mathbf{e}$. Let $\boldsymbol{\varphi}_j(\mathbf{g}, \mathbf{x})$ be the components of the n-tuple $\boldsymbol{\varphi}(\mathbf{g}, \mathbf{x})$. Then from property (4) we have $[\partial\varphi_j/\partial x_i(\mathbf{e}, \mathbf{x})]|_{\mathbf{x}=\mathbf{e}} = \delta_{ij}$. It follows that the Jacobian $\det[\partial\varphi_j/\partial x_i(\mathbf{e}, \mathbf{e})] = 1$. Therefore the inverse function theorem (Apostol [1]) guarantees the existence of a unique solution $\mathbf{x} \in V$ of the equation $\boldsymbol{\varphi}(\mathbf{g}, \mathbf{x}) = \mathbf{e}$ for all $\mathbf{g}$ in some open set V_r, $(\mathbf{e} \in V_r \subseteq V)$. Furthermore the components of $\mathbf{x}$ are analytic functions of the parameters of $\mathbf{g}$. This proves the existence of a right inverse $\mathbf{x}_r$ of $\mathbf{g}$ for $\mathbf{g} \in V_r$. Similarly, we could find a unique left inverse $\mathbf{x}_l$ for each $\mathbf{g} \in V_l$, an open set containing $\mathbf{e}$. Let $V^{-1} = V_r \cap V_l$. Then V^{-1} is an open set containing $\mathbf{e}$ and for each $\mathbf{g} \in V^{-1}$ we have

$$\mathbf{x}_l = \mathbf{x}_l\mathbf{e} = \mathbf{x}_l(\mathbf{gx}_r) = (\mathbf{x}_l\mathbf{g})\mathbf{x}_r = \mathbf{ex}_r = \mathbf{x}_r$$

so the left and right inverses coincide. We write $\mathbf{x}_l = \mathbf{x}_r = \mathbf{g}^{-1}$.

By the implicit function theorem for $\mathbf{h} \in V$ the mapping $\mathbf{x} \longrightarrow \mathbf{hx}$, $\mathbf{x} \in U$, is a homeomorphism of a suitably small neighborhood U of $\mathbf{e}$ onto a neighborhood U' of $\mathbf{h}$. Indeed $U' = \mathbf{h}U$.

A local Lie group G is not necessarily a group in the sense of Chapter 1. Indeed the group axioms are satisfied only for elements in a sufficiently small neighborhood of $\mathbf{e}$. It may not even make sense to write $\mathbf{gh}$ for n-tuples $\mathbf{g}$, $\mathbf{h}$ not both in V. Furthermore an n-tuple $\mathbf{g} \in V$ will not in general have a unique inverse unless $\mathbf{g} \in V^{-1} \subseteq V$. It is useful (and correct) to think of a local Lie group as a neighborhood of the identity element of a (global) group.

If $F = R$, then G is a **real** local Lie group; if $F = \mathfrak{C}$, G is a **complex** group. For the time being we shall develop the theory of real and complex Lie groups simultaneously.

As an example of a local Lie group consider the matrix group $GL(m, \mathfrak{C})$. We can uniquely write any $m \times m$ matrix $\mathbf{g}$ in the form

$$\mathbf{g} = (g_{ij}) = (\delta_{ij} + x_{ij}), \qquad 1 \leq i, j \leq m$$

where δ_{ij} is the Kronecker delta. For all (x_{ij}) in a suitably small neighborhood of the zero matrix, $\det(\delta_{ij} + x_{ij}) \neq 0$, so $\mathbf{g} \in GL(m, \mathfrak{C})$ and we can parametrize the group using the m^2 parameters x_{ij}. It is easy to find a smaller neighborhood V of the zero matrix such that for $(x_{ij}), (x'_{ij}) \in V$ the matrix product $\mathbf{g}(x_{ij})\mathbf{g}'(x'_{ij})$ can be expressed as $\mathbf{g}''(x''_{ij})$, where the parameters x''_{ij} are analytic functions of x_{ij} and x'_{ij}. Finally, the identity element $\mathbf{e}$ of $GL(m, \mathfrak{C})$ corresponds to the parameters $x_{ij} = 0$. Thus, $GL(m, \mathfrak{C})$ is an

m^2-dimensional complex local Lie group in the neighborhood V. Similarly, $GL(m, R)$ is a real local Lie group.

Our definition of $GL(m, \mathfrak{C})$ as a local Lie group depends on the selection of coordinates x_{ij} and neighborhood V. If we adopt different coordinates or choose a different neighborhood then we obtain a different local Lie group. However, we will show that all such local Lie groups are essentially the same, i.e., they are locally isomorphic. For this reason it makes sense to speak of *the* local Lie group $GL(m, \mathfrak{C})$ without explicitly mentioning the choice of local coordinates.

Most of the local Lie groups encountered in applications are groups of matrices. It is convenient, therefore, to introduce the concept of a local linear Lie group. Let W be an open, connected set containing $\mathbf{e}$ in the space F_n of all n-tuples $\mathbf{g} = (g_1, \ldots, g_n)$. (Without loss of generality the reader can assume W is an open sphere with center $\mathbf{e}$.)

Definition. An n-dimensional **local linear Lie group** G is a set of $m \times m$ nonsingular matrices $A(\mathbf{g}) = A(g_1, \ldots, g_n)$, defined for each $\mathbf{g} \in W$, such that

(1) $A(\mathbf{e}) = E_m$ (the identity matrix).
(2) The matrix elements of $A(\mathbf{g})$ are analytic functions of the parameters $g_1, \ldots, g_n$ and the map $\mathbf{g} \to A(\mathbf{g})$ is 1–1.
(3) The n matrices $\partial A(g)/\partial g_j$, $j = 1, \ldots, n$, are linearly independent for each $\mathbf{g} \in W$. That is, these matrices span an n-dimensional subspace of the m^2-dimensional space of all $m \times m$ matrices.
(4) There exists a neighborhood W' of $\mathbf{e}$ in F_n, $W' \subseteq W$, with the property that for every pair of n-tuples $\mathbf{g}, \mathbf{h}$ in W' there is an n-tuple $\mathbf{k}$ in W satisfying

$$A(\mathbf{g})A(\mathbf{h}) = A(\mathbf{k}) \tag{2.1}$$

where the operation on the left is matrix multiplication.

Every local linear Lie group G defines a local Lie group. Indeed we can identify the matrix $A(\mathbf{g})$ with the group element $\mathbf{g}$. It follows from (2), (3), and the implicit function theorem that the parameters g_i are analytic functions of the matrix elements of $A(\mathbf{g})$. Then (4) implies that there exists a nonzero neighborhood V of $\mathbf{e}$ such that $\mathbf{k} = \boldsymbol{\varphi}(\mathbf{g}, \mathbf{h})$ for all $\mathbf{g}, \mathbf{h} \in V$ where $\boldsymbol{\varphi}$ is an analytic vector-valued function of its $2n$ arguments and $\mathbf{g}, \mathbf{h}, \mathbf{k}$ are related by (2.1). (Clearly we can assume $V \subseteq W'$.) Thus,

$$A(\mathbf{g})A(\mathbf{h}) = A(\boldsymbol{\varphi}(\mathbf{g}, \mathbf{h})). \tag{2.2}$$

Since matrix multiplication is associative, $\boldsymbol{\varphi}(\mathbf{g}, \mathbf{h})$ satisfies the associative law. Finally, $A(\mathbf{e}) = E_m$, so $\boldsymbol{\varphi}(\mathbf{e}, \mathbf{g}) = \boldsymbol{\varphi}(\mathbf{g}, \mathbf{e}) = \mathbf{g}$ for all $\mathbf{g} \in V$. This proves that an n-dimensional local linear Lie group is an n-dimensional local Lie group.

It is now apparent that $GL(m, \mathfrak{C})$ is an m^2-dimensional local linear Lie group. However, $GL(m, \mathfrak{C})$ is not only a local Lie group, it is also a group in the abstract sense of Section 1.1, i.e., it is a global group. This remark leads us to the notion of a Lie group. A (global) **Lie group** G is (1) an abstract group and (2) an analytic manifold such that (3) group multiplication and group inversion are analytic with respect to the manifold structure. Unfortunately, it would take many pages to clarify this definition, particularly property (2). Rather than embark on such a topological digression, we refer the interested reader to Helgason [1] or Hausner and Schwartz [1], and merely show how one can construct a (global) **linear Lie group** from a local linear Lie group. Since the vast majority of Lie groups occurring in physics are linear Lie groups this simplification of the theory is worthwhile.

Let G be a local linear Lie group consisting of $m \times m$ matrices. We will define the (connected) global linear Lie group $\tilde{G}$ containing G. Algebraically, $\tilde{G}$ is the abstract subgroup of $GL(m, \mathfrak{C})$ generated by the matrices of G. That is, $\tilde{G}$ consists of all possible products of finite sequences of elements in G. In addition, the elements of $\tilde{G}$ can be parametrized analytically. If $B \in \tilde{G}$ we can introduce coordinates in a neighborhood of B by means of the map $\mathbf{g} \to BA(\mathbf{g})$ where $\mathbf{g}$ ranges over a suitably small neighborhood Z of $\mathbf{e}$ in F_n. In particular, the coordinates of B will be $\mathbf{e} = (0, \ldots, 0)$. Proceeding in this way for each $B \in \tilde{G}$ we can cover $\tilde{G}$ with local coordinate systems or "coordinate patches." The same group element C will have many different sets of coordinates depending on which coordinate patch containing C we happen to consider. Suppose C lies in the intersection of coordinate patches around B_1 and B_2, respectively. Then C will have coordinates $\mathbf{g}_1, \mathbf{g}_2$, respectively, where $C = B_1 A(\mathbf{g}_1) = B_2 A(\mathbf{g}_2)$. Since

$$A(\mathbf{g}_1) = B_1^{-1} B_2 A(\mathbf{g}_2), \qquad A(\mathbf{g}_2) = B_2^{-1} B_1 A(\mathbf{g}_1)$$

it follows that in a suitably small neighborhood of $\mathbf{e}$ the coordinates $\mathbf{g}_2$ are analytic functions $\mathbf{g}_2 = \boldsymbol{\rho}(\mathbf{g}_1)$ of the coordinates $\mathbf{g}_1$, $\boldsymbol{\rho}$ is 1–1, and the Jacobian of the coordinate transformation is nonzero (Freudenthal and De Vries [1], Cohn [1]). Thus, every element of $\tilde{G}$ is covered by coordinate neighborhoods and the transformation relating any two overlapping neighborhoods is (a) 1–1, (b) analytic, and (c) has a nonzero Jacobian. This makes $\tilde{G}$ into an analytic manifold. In addition to the coordinate neighborhoods described above we can always add more coordinate neighborhoods to $\tilde{G}$ provided they satisfy conditions (a)–(c) on the overlap with any of the original coordinate systems. We leave it to the reader to show that the Lie group $\tilde{G}$ is **connected.** That is, any two elements A, B in $\tilde{G}$ can be connected by an analytic curve $C(t)$ lying entirely in $\tilde{G}$. Here t is a real parameter, $0 \leq t \leq 1$, $C(0) = A$, $C(1) = B$, and whenever $C(t) \in \tilde{G}$ lies in a coordinate patch, its coordinates $\mathbf{g}(t)$ are analytic functions of t. (It is enough to show that any element A in $\tilde{G}$ can be connected to the identity by an analytic curve in $\tilde{G}$. For, if A is

connected to E_m and E_m is connected to B, then the composite curve connects A to B.)

In general an n-dimensional (global) **linear Lie group** K is an abstract matrix group which is also an n-dimensional local linear group. A linear Lie group need not be connected. Clearly $GL(m, \mathfrak{C})$ is an m^2-dimensional Lie group. A single coordinate patch suffices to cover $GL(m, \mathfrak{C})$, but this is not true for other Lie groups such as $O(m)$. Furthermore, we shall show later that $GL(m, \mathfrak{C})$ is connected, but $O(m)$ is not.

5.3 Lie Algebras

Let G be a local Lie group defined in the neighborhood $V \subseteq F_n$. An **analytic curve through the identity** on G is a mapping $t \longrightarrow \mathbf{g}(t) = (g_1(t), \ldots, g_n(t))$ of a neighborhood of $0 \in F$ into V such that $\mathbf{g}(0) = \mathbf{e}$ and the $g_j(t)$ are analytic in t. The **tangent vector** to $\mathbf{g}(t)$ at $\mathbf{e}$ is the vector

$$\alpha = [d\mathbf{g}(t)/dt]|_{t=0} = (\dot{g}_1(0), \ldots, \dot{g}_n(0)) \in F_n. \tag{3.1}$$

Every vector in F_n is the tangent vector at $\mathbf{e}$ for some analytic curve. Indeed, the curve

$$\alpha t = (\alpha_1 t, \alpha_2 t, \ldots, \alpha_n t) \tag{3.2}$$

clearly has the tangent vector $\alpha = (\alpha_1, \ldots, \alpha_n)$ at $\mathbf{e}$, so we can identify the set of tangent vectors with F_n. In particular the tangent vectors at $\mathbf{e}$ form an n-dimensional vector space.

For $\mathbf{g}, \mathbf{h} \in V$ we have $\mathbf{gh} = \boldsymbol{\varphi}(\mathbf{g}, \mathbf{h})$, where $\boldsymbol{\varphi}$ is an analytic vector-valued function of its $2n$ arguments. Thus, the components $(\mathbf{gh})_j = \varphi_j(\mathbf{g}, \mathbf{h})$ can be expressed as Taylor series in g_l, h_l about $\mathbf{g} = \mathbf{h} = \mathbf{e}$. Since $\boldsymbol{\varphi}(\mathbf{e}, \mathbf{g}) = \boldsymbol{\varphi}(\mathbf{g}, \mathbf{e}) = \mathbf{g}$ there follows

$$\varphi_j(\mathbf{g}, \mathbf{h}) = g_j + h_j + \sum_{l,s=1}^{n} c_{j,ls} g_l h_s + r_j(\mathbf{g}, \mathbf{h}), \tag{3.3}$$

where r_j consists of terms of order greater than two in g_l, h_l and

$$c_{j,ls} = (\partial^2/\partial g_l \, \partial h_s)\varphi_j(\mathbf{g}, \mathbf{h})|_{\mathbf{g}=\mathbf{h}=\mathbf{e}}. \tag{3.4}$$

We now examine the relationship between analytic curves through $\mathbf{e}$ and vector space operations on tangent vectors. Let $\mathbf{g}(t)$, $\mathbf{h}(t)$ be analytic curves through $\mathbf{e}$ with tangent vectors α, β, respectively. Then for any constants $a, b \in F$, $\mathbf{k}(t) = \mathbf{g}(at)\mathbf{h}(bt) = \boldsymbol{\varphi}(\mathbf{g}(at), \mathbf{h}(bt))$ is also an analytic curve through $\mathbf{e}$. The tangent vector $\gamma = \dot{\mathbf{k}}(0)$ can be obtained by differentiating both sides of (3.3): $\gamma = a\alpha + b\beta$. [Note that only the first two terms on the right-hand side of (3.3) contribute to the result.] Here the plus sign (+) refers to vector addition in F_n.

The curve $\mathbf{e}(t) = \mathbf{e}$ obviously has tangent vector $\theta = (0, \ldots, 0)$ at $\mathbf{e}$.

Let α' be the tangent vector at $\mathbf{e}$ of the analytic curve $\mathbf{g}^{-1}(t)$. Since $\mathbf{e}(t) = \mathbf{g}(t)\mathbf{g}^{-1}(t)$ it follows that $\theta = \alpha + \alpha'$, or $\alpha' = -\alpha$.

The preceding results use only the first-order terms in the expansion for $\boldsymbol{\varphi}(\mathbf{g}, \mathbf{h})$. We now introduce an operation on tangent vectors at $\mathbf{e}$ which depends on the second-order terms. With $\mathbf{g}(t)$, $\mathbf{h}(t)$ as given above, we define the **commutator** $[\alpha, \beta]$ of α and β as the tangent vector at $\mathbf{e}$ of the curve

$$\mathbf{k}(t) = \mathbf{g}(\tau)\mathbf{h}(\tau)\mathbf{g}^{-1}(\tau)\mathbf{h}^{-1}(\tau), \qquad t = \tau^2. \tag{3.5}$$

Thus,

$$[\alpha, \beta] = [d/d(\tau^2)]\mathbf{g}(\tau)\mathbf{h}(\tau)\mathbf{g}^{-1}(\tau)\mathbf{h}^{-1}(\tau)|_{\tau=0}, \tag{3.6}$$

or more precisely, $[\alpha, \beta]$ is the coefficient of τ^2 in the Taylor series expansion for $\mathbf{k}$. (This definition makes sense even if $\mathbf{k}$ is not an analytic function of τ^2, because the coefficient of τ is zero.)

Theorem 5.6. $[\alpha, \beta]_j = \sum_{l,s=1}^{n} c_j^{ls}\alpha_l\beta_s$, $1 \le j \le n$, where $c_j^{ls} = c_{j,ls} - c_{j,sl}$.

Proof. We write $g_j(\tau)$, $h_j(\tau)$, $k_j(t)$ as Taylor series,

$$g_j(\tau) = \alpha_j\tau + b_j\tau^2 + \cdots, \qquad h_j(\tau) = \beta_j\tau + c_j\tau^2 + \cdots,$$
$$k_j(t) = \rho_j\tau + a_j\tau^2 + \cdots,$$

where the omitted terms are of order greater than two. Clearly $[\alpha, \beta]_j = a_j$ if $\rho_j = 0$. We will verify that $\rho_j = 0$. Since $\mathbf{k}(t)\mathbf{h}(\tau)\mathbf{g}(\tau) = \mathbf{g}(\tau)\mathbf{h}(\tau)$ we find by (3.3)

$$(\rho_j + \beta_j + \alpha_j)\tau + [a_j + c_j + b_j + \sum c_{j,ls}(\beta_l\alpha_s + \rho_l\alpha_s + \rho_l\beta_s)]\tau^2$$
$$= (\beta_j + \alpha_j)\tau + (c_j + b_j + \sum c_{j,ls}\alpha_l\beta_s)\tau^2,$$

where only the terms of order less than three are written explicitly. Comparing first-order terms in τ on both sides of the identity, we obtain $\rho_j = 0$. A comparison of second-order terms yields $a_j = \sum (c_{j,ls} - c_{j,sl})\alpha_l\beta_s$. Q.E.D.

The c_j^{ls} are called the **structure constants** of G. A knowledge of the structure constants yields information about the second-order terms in the expansion of $\boldsymbol{\varphi}(\mathbf{g}, \mathbf{h})$. It is, however, a remarkable fact that one can essentially determine the function $\boldsymbol{\varphi}(\mathbf{g}, \mathbf{h})$ from the structure constants, i.e., all information conveyed by the local Lie group G is already contained in the c_j^{ls}. For example, if G is abelian ($\mathbf{gh} = \mathbf{hg}$ for all $\mathbf{g}, \mathbf{h} \in V$) then it follows from (3.6) that $[\alpha, \beta] = \theta$ for all tangent vectors α, β at $\mathbf{e}$, and the structure constants are all zero. Conversely, we shall show that if the structure constants are zero then G is necessarily abelian.

The basic properties of the commutator are given by the following theorem.

Theorem 5.7. Let $\alpha, \beta, \gamma \in F_n$ and $a, b \in F$.

(1) $[\alpha, \beta] = -[\beta, \alpha]$.
(2) $[a\alpha + b\beta, \gamma] = a[\alpha, \gamma] + b[\beta, \gamma]$,
(3) $[[\alpha, \beta], \gamma] + [[\gamma, \alpha], \beta] + [[\beta, \gamma], \alpha] = \theta$ (Jacobi equality).

Proof. Properties (1) and (2) follow directly from the expression for the commutator in Theorem 5.6. The Jacobi equality is an expression of the associative law $(\mathbf{gh})\mathbf{k} = \mathbf{g}(\mathbf{hk})$ for local Lie groups. It is an identity for the structure constants which is obtained by substituting (3.3) into the associative law and equating terms of the same order. Q.E.D.

Definition. The **Lie algebra** $L(G)$ of a local Lie group G is the set of all tangent vectors at $\mathbf{e}$ equipped with the operations of scalar multiplication, vector addition, and commutator product.

Definition. An **abstract Lie algebra** $\mathcal{G}$ over F is a vector space over F together with a product $[\alpha, \beta] \in \mathcal{G}$ defined for all $\alpha, \beta \in \mathcal{G}$ such that for all $\alpha, \beta, \gamma \in \mathcal{G}$ and $a, b \in F$ (θ is the zero vector):

(1) $[\alpha, \beta] = -[\beta, \alpha]$.
(2) $[a\alpha + b\beta, \gamma] = a[\alpha, \gamma] + b[\beta, \gamma]$.
(3) $[[\alpha, \beta], \gamma] + [[\gamma, \alpha], \beta] + [[\beta, \gamma], \alpha] = \theta$.

Property (1) expresses the skew-symmetry of the commutator, while (1) and (2) imply that $[\alpha, \beta]$ is bilinear. The identity (3) is again called the Jacobi equality. We will always assume that $\mathcal{G}$ is a finite-dimensional vector space.

Clearly, $L(G)$ is an abstract Lie algebra. It is not clear (but true) that any abstract Lie algebra is in fact the Lie algebra of some local Lie group. The exact meaning of this statement will be clarified later.

The significance of our results concerning Lie algebras and commutators becomes much clearer when we consider linear Lie groups. Let G be an n-dimensional local linear Lie group of $m \times m$ matrices and let $A(t) = A(\mathbf{g}(t))$, $A(0) = E_m$, be an analytic curve through the identity. We can identify the tangent vector α at $\mathbf{e}$ with the matrix

$$\mathfrak{A} = \frac{d}{dt} A(\mathbf{g}(t))\Big|_{t=0} = \sum_{j=1}^{n} \frac{\partial}{\partial g_j} A(\mathbf{g})\Big|_{\mathbf{g}=\mathbf{e}} \dot{g}_j(0) = \sum_{j=1}^{n} \mathfrak{E}_j \alpha_j, \qquad \mathfrak{E}_j = \frac{\partial}{\partial g_j} A(\mathbf{g})\Big|_{\mathbf{g}=\mathbf{e}}. \tag{3.7}$$

Thus, we identify the tangent space at $\mathbf{e}$ with the space of all $m \times m$ matrices of the form (3.7). From the definition of linear Lie groups, the set $\{\mathfrak{E}_j\}$ is linearly independent and spans an n-dimensional subspace of the space of all $m \times m$ matrices. Since the vectors α are arbitrary we see again that the tangent space is n-dimensional.

Many of the properties which were obtained for general local Lie groups are much simpler to verify for local linear Lie groups. For example, let $A(t)$, $B(t)$ be analytic curves in G with tangent matrices $\mathcal{A}$, $\mathcal{B}$ at the identity. Since

$$(d/dt)(A(t)B(t)) = \dot{A}(t)B(t) + A(t)\dot{B}(t) \tag{3.8}$$

and $A(0) = B(0) = E_m$, it follows that $A(t)B(t)$ is an analytic curve with tangent matrix $\mathcal{A} + \mathcal{B}$ at E_m.

With $A(t)$, $B(t)$ as above we define the **commutator** of the tangent matrices $\mathcal{A}$, $\mathcal{B}$ by

$$[\mathcal{A}, \mathcal{B}] = (d/dt)[A(\tau)B(\tau)A^{-1}(\tau)B^{-1}(\tau)]_{t=0}, \tag{3.9}$$

where $t = \tau^2$, i.e,. $[\mathcal{A}, \mathcal{B}]$ is the tangent matrix at the identity of the curve $C(t) = A(\tau)B(\tau)A^{-1}(\tau)B^{-1}(\tau)$.

Theorem 5.8. $[\mathcal{A}, \mathcal{B}] = \mathcal{A}\mathcal{B} - \mathcal{B}\mathcal{A} \in L(G)$.

Proof. We copy the proof of Theorem 5.6 taking advantage of the simplifications afforded by matrix groups. Expanding $A(\tau)$, $B(\tau)$, and $C(t)$ as Taylor series in τ we find

$$A(\tau) = E_m + \mathcal{A}\tau + \mathcal{A}'\tau^2 + \cdots, \qquad B(\tau) = E_m + \mathcal{B}\tau + \mathcal{B}'\tau^2 + \cdots,$$
$$C(t) = E_m + \mathcal{C}\tau + \mathcal{C}'\tau^2 + \cdots.$$

We will show that $\mathcal{C} = Z$, so (3.9) is well-defined and $\mathcal{C}' = [\mathcal{A}, \mathcal{B}]$. Since $C(t)B(\tau)A(\tau) = A(\tau)B(\tau)$ there follows the identity

$$\begin{aligned} E_m &+ (\mathcal{A} + \mathcal{B} + \mathcal{C})\tau + (\mathcal{A}' + \mathcal{B}' + \mathcal{C}' + \mathcal{B}\mathcal{A} + \mathcal{C}\mathcal{B} + \mathcal{C}\mathcal{A})\tau^2 + \cdots \\ &= E_m + (\mathcal{A} + \mathcal{B})\tau + (\mathcal{A}' + \mathcal{B}' + \mathcal{A}\mathcal{B})\tau^2 + \cdots. \end{aligned}$$

Equating terms in τ and τ^2 we find $\mathcal{C} = Z$ and $\mathcal{C}' = \mathcal{A}\mathcal{B} - \mathcal{B}\mathcal{A}$. Q.E.D.

Clearly, $L(G)$ cannot be an arbitrary subspace of matrices, for we must have $[\mathcal{A}, \mathcal{B}] \in L(G)$ for all $\mathcal{A}, \mathcal{B} \in L(G)$. This requirement greatly restricts the possibilities for matrix Lie algebras.

Using Theorem 5.6 and (3.7), we can relate the structure constants c_j^{ls} of G to the matrix commutator:

$$[\mathcal{C}_l, \mathcal{C}_s] = \sum_{j=1}^{n} c_j^{ls}\mathcal{C}_j, \qquad 1 \le l, s \le n. \tag{3.10}$$

It follows directly from Theorem 5.7 that the matrix commutator is skew-symmetric, bilinear, and satisfies the Jacobi equality. However, it is worthwhile to verify these properties directly from the definition $[\mathcal{A}, \mathcal{B}] = \mathcal{A}\mathcal{B} - \mathcal{B}\mathcal{A}$. Skew-symmetry and bilinearity are trivial to check. Moreover, as the reader can easily show, the matrix commutator automatically satisfies the Jacobi equality. Thus, a set $\mathcal{G}$ of $m \times m$ matrices which is closed under addi-

tion of matrices, scalar multiplication, and matrix commutation is necessarily a Lie algebra. It is not necessary to check the Jacobi equality. (Note: We do not yet know that $\mathcal{G}$ is actually the Lie algebra of some local Lie group.) This fact makes matrix algebras much easier to study than general abstract Lie algebras.

Fortunately, we can restrict ourselves to the study of matrix Lie algebras with no loss of generality. To show this we need the concepts of homomorphism and isomorphism of Lie algebras. Let $\mathcal{G}, \mathcal{G}'$ be abstract Lie algebras over F with operations $+$, $[-,-]$ and $+'$, $[-,-]'$, respectively.

Definition. A **homomorphism** from $\mathcal{G}$ to $\mathcal{G}'$ is a map $\tau: \mathcal{G} \longrightarrow \mathcal{G}'$ such that

(1) $\tau(a\alpha + b\beta) = a\tau(\alpha) + b\tau(\beta)$.

(2) $\tau([\alpha, \beta]) = [\tau(\alpha), \tau(\beta)]'$, $\quad a, b \in F, \quad \alpha, \beta \in \mathcal{G}$.

A Lie algebra homomorphism which is a 1–1 map of $\mathcal{G}$ onto $\mathcal{G}'$ is an **isomorphism.** An isomorphism of $\mathcal{G}$ onto $\mathcal{G}$ is an **automorphism.**

Note that a homomorphism τ is a linear mapping of $\mathcal{G}$ into $\mathcal{G}'$ which preserves commutators. We can identify isomorphic Lie algebras $\mathcal{G}, \mathcal{G}'$ since they have the same structure as abstract Lie algebras.

Theorem 5.9 (Ado). Every abstract Lie algebra is isomorphic to a matrix Lie algebra.

The proof of Ado's Theorem is difficult and will not be given here (see Jacobson [1]). Moreover, there is no known general method to construct a matrix Lie algebra isomorphic to any given abstract Lie algebra, even though such a matrix algebra must exist. We quote Ado's theorem as moral support. Even though we limit ourselves to matrix Lie algebras in the following sections, we will actually be studying all Lie algebras (up to isomorphism).

Definition. A (local) **analytic homomorphism** of a local Lie group G into a local Lie group G' is a map $\mu: G \longrightarrow G'$, where $\mu(\mathbf{g})$ is defined for $\mathbf{g}$ in a suitably small neighborhood W of $\mathbf{e}$ such that

$$\mu(\mathbf{gh}) = \mu(\mathbf{g})\mu(\mathbf{h}), \qquad \mathbf{g}, \mathbf{h}, \mathbf{gh} \in G, \tag{3.11}$$

and μ is an analytic function of the coordinates of G. The group multiplication on the right-hand side of (3.11) takes place in G'. If μ is a 1–1 homomorphism of G onto a neighborhood of $\mathbf{e}'$ in G' such that the mapping $\mu^{-1}: G' \longrightarrow G$ is analytic, then μ is called an **isomorphism** and G is (locally) isomorphic to G'. An isomorphism of G onto G is an **automorphism.**

It follows from (3.11) that $\mu(\mathbf{e}) = \mathbf{e}'$, where $\mathbf{e}'$ is the identity element of G', and $\mu(\mathbf{g}^{-1}) = \mu(\mathbf{g})^{-1}$ for $\mathbf{g}$ in a suitably small neighborhood of $\mathbf{e}$. Note that μ is purely local. Even if G and G' are (global) Lie groups, it is not

necessary that μ be defined on all of G. Furthermore, μ need not be an abstract group homomorphism of G in the sense of Chapter 1, since (3.11) is required to hold only in a neighborhood of **e**. As usual, we identify isomorphic local Lie groups.

In Section 5.5 it will be shown that two local Lie groups are isomorphic if and only if they have isomorphic Lie algebras.

5.4 The Classical Groups

Here we introduce the **classical groups,** a family of linear Lie groups which is of fundamental importance in physics and geometry. The reader should note the significant role of Theorem 5.3 in the following analysis. Among the classical groups are the following.

$\boldsymbol{G_m = GL(m, \mathfrak{C})}$. We have already seen that G_m is an m^2-dimensional complex Lie group. (If we write the complex coordinates as sums of real and imaginary parts we can also consider G_m as a $2m^2$-dimensional real Lie group.) The Lie algebra of this group consists of all matrices of the form

$$\mathfrak{A} = (d/dt)A(t)|_{t=0}, \qquad A(0) = E_m,$$

where $A(t)$ is an analytic curve in G_m. Clearly $\mathfrak{A}$ is an $m \times m$ matrix. Conversely, if $\mathfrak{A}$ is any $m \times m$ matrix then $A(t) = \exp t\mathfrak{A}$ is an analytic curve in G_m with tangent matrix $\mathfrak{A}$. [See expression (1.14).] Thus, $L(G_m)$ is the m^2-dimensional space of all $m \times m$ matrices. There is an analogous result for the real Lie group $GL(m, R)$.

$\boldsymbol{SL(m) = SL(m, \mathfrak{C})}$. The special linear group

$$SL(m) = \{A \in GL(m, \mathfrak{C}): \det A = 1\}$$

is an $(m^2 - 1)$-dimensional complex Lie group. Indeed, by Theorem 5.3 and Corollary 5.3, every $A \in SL(m)$ sufficiently close to the identity can be written uniquely in the form $A = \exp \mathfrak{A}$, where $\mathfrak{A}$ is an $m \times m$ matrix with $\operatorname{tr} \mathfrak{A} = 0$. It is easy to see that $\mathfrak{A}$ has $m^2 - 1$ independent parameters which can be used as local coordinates for $SL(m)$. With these coordinates $SL(m)$ is a complex linear Lie group. [If we write the matrix elements of $\mathfrak{A}$ in terms of real and imaginary parts we can also consider $SL(m)$ as a $2(m^2 - 1)$-dimensional real Lie group.]

Theorem 5.10. The Lie algebra of $SL(m)$ is the space of all $m \times m$ matrices of trace zero.

Proof. Let $A(t)$, $A(0) = E_m$, be an analytic curve in $SL(m)$. For t sufficiently close to zero we can write $A(t) = \exp \mathfrak{A}(t)$, $\mathfrak{A}(0) = Z$, where $\mathfrak{A}(t)$ is an analytic

curve in the space of traceless matrices. Let $\mathcal{B} = \dot{\mathcal{A}}(0)$. Since $d/dt(\operatorname{tr} \mathcal{A}(t)) = \operatorname{tr} \dot{\mathcal{A}}(t)$ we have $\operatorname{tr} \mathcal{B} = 0$. A simple computation shows that $\mathcal{B}$ is the tangent matrix to $A(t)$ at E_m.

Conversely, let $\mathcal{B}$ be an $m \times m$ matrix with $\operatorname{tr} \mathcal{B} = 0$ and consider the analytic curve $A(t) = \exp t\mathcal{B}$. Since $\det A(t) = \exp(\operatorname{tr} t\mathcal{B}) = 1$, the curve $A(t)$ lies in $SL(m)$. Furthermore $A(0) = E_m$, and the tangent matrix to $A(t)$ at E_m is $\dot{A}(0) = \mathcal{B}$, so $\mathcal{B}$ is an element of the Lie algebra. Q.E.D.

There is a similar result for the real Lie groups $SL(m, R)$.

$\boldsymbol{O(m) = O(m, \mathfrak{C})}$. Recall that $O(m)$ is the group of all complex matrices A such that $A^t A = E_m$. For each $A \in O(m)$ sufficiently close to E_m we can find a unique matrix $\mathcal{A}$ sufficiently close to Z such that $A = \exp \mathcal{A}$. Since $A \in O(m)$, we have $(\exp \mathcal{A}^t)(\exp \mathcal{A}) = E_m = \exp Z$. For $\mathcal{A}$ sufficiently close to Z this is possible only if $\mathcal{A}^t + \mathcal{A} = Z$, i.e., only if $\mathcal{A}$ is skew-symmetric. Conversely, if $\mathcal{A}$ is skew-symmetric and $A = \exp \mathcal{A}$ then $A^t A = \exp \mathcal{A}^t \cdot \exp \mathcal{A} = E_m$, so $A \in O(m)$. We can use the independent matrix elements of $\mathcal{A}$ as local coordinates for $O(m)$. As the reader can check, there are $m(m-1)/2$ independent matrix elements in an $m \times m$ skew-symmetric matrix, so $O(m)$ is an $m(m-1)/2$-dimensional complex Lie group, [or an $(m^2 - m)$-dimensional real Lie group].

Theorem 5.11. The Lie algebra of $O(m)$ is the space of all $m \times m$ skew-symmetric matrices.

Proof. Let $A(t)$, $A(0) = E_m$, be an analytic curve in $O(m)$ with tangent matrix $\mathcal{A}$ at the identity. Differentiating the expression $A^t(t)A(t) = E_m$ and setting $t = 0$ we find $\mathcal{A}^t + \mathcal{A} = Z$. Thus, the elements of the Lie algebra are skew-symmetric matrices.

Conversely, let $\mathcal{A}$ be a skew-symmetric matrix. Then $A(t) = \exp t\mathcal{A}$ is an analytic curve in $O(m)$ with tangent matrix $\dot{A}(0) = \mathcal{A}$, so $\mathcal{A} \in L(O(m))$. Q.E.D.

There is a similar result for $O(m, R)$.

Note that the elements of $O(m)$ which lie in a small neighborhood of the identity all have determinant $+1$. Indeed if $A = \exp \mathcal{A}$ with $\mathcal{A}^t = -\mathcal{A}$ then $\det A = \exp(\operatorname{tr} \mathcal{A}) = +1$ since the trace of a skew-symmetric matrix is zero. Thus, the Lie algebra of $O(m)$ is exactly the same as the Lie algebra of $SO(m) = SO(m, \mathfrak{C})$:

$$SO(m) = \{A \in O(m) \colon \det A = +1\}.$$

We shall see later that $SO(m)$ is a connected Lie group. In particular, every $A \in SO(m)$ can be connected to the identity by an analytic curve $A(t)$,

$A(0) = E_m$, $A(1) = A$, $0 \leq t \leq 1$, which lies entirely in $SO(m)$. However, $O(m)$ is not connected. If it were, we could find an analytic curve $A(t)$ in $O(m)$ such that $A(0) = E_m$ and $A(1) = -E_m$. Then for m odd, $\det A(t)$ would be an analytic function of t taking only the values ± 1 with $\det A(0) = -\det A(1) = 1$. No such function exists. Thus, $O(m)$ consists of two connected components, the cosets $SO(m)$ and $-E_m \cdot SO(m)$. We leave to the reader a verification that $O(m)$ also has two connected components for m even.

U(m). The **unitary group** is the set

$$U(m) = \{A \in GL(m, \mathfrak{C}) \colon \bar{A}^t A = E_m\}$$

of all $m \times m$ unitary matrices. Given $A \in U(m)$ in a sufficiently small neighborhood of E_m we can find a unique matrix $\mathfrak{A}$ in a sufficiently small neighborhood of Z such that $A = \exp \mathfrak{A}$. The requirement $\bar{A}^t A = E_m$ implies $\exp \bar{\mathfrak{A}}^t \exp \mathfrak{A} = E_m$ or $\exp \bar{\mathfrak{A}}^t = \exp(-\mathfrak{A})$. For $\mathfrak{A}$ sufficiently close to Z this implies $\bar{\mathfrak{A}}^t = -\mathfrak{A}$, i.e., $\mathfrak{A}$ is skew-Hermitian. On the other hand if $\mathfrak{A}$ is a skew-Hermitian matrix and $A = \exp \mathfrak{A}$ then $\bar{A}^t A = \exp \bar{\mathfrak{A}} \exp \mathfrak{A} = E_m$, so $A \in U(m)$. Thus we can use the matrix elements of $\mathfrak{A}$ to provide local coordinates in $U(m)$. As the reader can check, an $m \times m$ skew-Hermitian matrix has m^2 *real* independent parameters. (Write the matrix elements of $\mathfrak{A}$ in terms of real and imaginary parts.) Using these parameters as local coordinates we see that $U(m)$ is an m^2-dimensional *real* Lie group. Here $U(m)$ is a real Lie group even though it consists of complex matrices.

Theorem 5.12. The Lie algebra of $U(m)$ is the space of all $m \times m$ skew-Hermitian matrices.

Proof. Analogous to Theorem 5.11.

If $\mathfrak{A}$ is skew-Hermitian then the matrix $\mathfrak{H} = -i\mathfrak{A}$ is Hermitian, i.e., $\bar{\mathfrak{H}}^t = \mathfrak{H}$. Thus, any unitary matrix A sufficiently close to the identity can be expressed $A = \exp(i\mathfrak{H})$, where $\mathfrak{H}$ is Hermitian. It is of considerable interest that such an expression is valid for all unitary matrices A, not just those sufficiently close to the identity.

Theorem 5.13. If $A \in U(m)$ there exists a Hermitian matrix $\mathfrak{H}$ (not unique) such that $A = \exp(i\mathfrak{H})$.

Proof. We know from linear algebra that the eigenvalues ε_j of A take the form $\varepsilon_j = e^{i\theta_j}$, θ_j real, $j = 1, \ldots, m$, and that A can be diagonalized by a unitary similarity transformation. Thus there exists $B \in U(m)$ and a diagonal matrix D with diagonal entries, $e^{i\theta_1}, \ldots, e^{i\theta_m}$, such that $BAB^{-1} = D$ (Fink-

beiner [1]). Clearly, $D = \exp(i\mathfrak{D})$, where $\mathfrak{D}$ is the diagonal matrix with matrix elements $\theta_1, \ldots, \theta_m$ down the diagonal. Thus $A = B^{-1}DB = \exp(i\mathcal{H})$, where $\mathcal{H} = B^{-1}\mathfrak{D}B$. The matrix $\mathcal{H}$ is Hermitian because $\bar{\mathcal{H}}^t = \bar{B}^t\bar{\mathfrak{D}}^t(\bar{B}^t)^{-1} = B^{-1}\mathfrak{D}B = \mathcal{H}$. Q.E.D.

We can use the above result to show that $U(m)$ is connected. Indeed, for any $A \in U(m)$ the matrix function $A(t) = \exp(it\mathcal{H})$, $A(1) = A$, $0 \leq t \leq 1$, defines an analytic curve in $U(m)$ connecting A to E_m.

SU(m). The **special unitary group** is the set

$$SU(m) = \{A \in U(m) \colon \det A = +1\},$$

a subgroup of $U(m)$. Recall that $|\det A| = 1$ for $A \in U(m)$. It is left to the reader to show that $SU(m)$ is an $(m^2 - 1)$-dimensional real Lie group whose Lie algebra consists of all skew-Hermitian matrices with trace zero.

Sp(m). Let J be the $2m \times 2m$ skew-symmetric matrix

$$J = \begin{pmatrix} Z & E_m \\ -E_m & Z \end{pmatrix}. \tag{4.1}$$

Definition. A **symplectic** matrix is a $2m \times 2m$ complex matrix A such that $A^tJA = J$. The symplectic group $Sp(m) = Sp(m, \mathfrak{C})$ is the set of all $2m \times 2m$ symplectic matrices.

Theorem 5.14. $Sp(m)$ is a matrix group.

Proof. Clearly the identity matrix belongs to $Sp(m)$. Suppose $A, B \in Sp(m)$. Then $(AB)^tJ(AB) = B^t(A^tJA)B = B^tJB = J$, so $AB \in Sp(m)$. Thus, we need only show that A^{-1} belongs to $Sp(m)$ for each $A \in Sp(m)$. Taking the determinant of both sides of the identity $A^tJA = J$ and using the fact $\det J = 1$, we obtain $(\det A)^2 = 1$, so A and A^t are invertible. Then $(A^{-1})^t JA^{-1} = (A^t)^{-1}(A^tJA)A^{-1} = J$ and $A^{-1} \in Sp(m)$. Q.E.D.

Since $J^{-1} = -J$ the relation $A^tJA = J$ implies $A^{-1} = -JA^tJ$.

For $A \in Sp(m)$ in a sufficiently small neighborhood of E_{2m} we can find a unique matrix $\mathcal{A}$ in a sufficiently small neighborhood of Z such that $A = \exp \mathcal{A}$. Since $A^{-1} = J^{-1}A^tJ$ it follows that $\exp(-\mathcal{A}) = J^{-1}(\exp \mathcal{A}^t)J = \exp(J^{-1}\mathcal{A}^tJ)$, or $-\mathcal{A} = J^{-1}\mathcal{A}^tJ$ for $\mathcal{A}$ in a sufficiently small neighborhood of Z. This result is usually expressed in the form $J\mathcal{A} + \mathcal{A}^tJ = Z$. Conversely, if $A = \exp \mathcal{A}$, where $\mathcal{A}$ satisfies the preceding equality, then it is easy to show that $A \in Sp(m)$. Thus, $Sp(m)$ is a complex Lie group with the independent matrix elements of $\mathcal{A}$ as parameters. As the reader can check, there are

$2m^2 + m$ independent parameters. Note that we can also consider $Sp(m)$ as a $(4m^2 + 2m)$-dimensional real Lie group.

Theorem 5.15. The Lie algebra of $Sp(m)$ is the space of all $2m \times 2m$ matrices $\mathfrak{A}$ such that $J\mathfrak{A} + \mathfrak{A}^t J = Z$.

Proof. Analogous to Theorem 5.11.

In each of the above examples, every element A of the Lie group G sufficiently close to the identity can be expressed uniquely in the form $A = \exp \mathfrak{A}$, where $\mathfrak{A}$ is an element of the corresponding Lie algebra $L(G)$ sufficiently close to Z. Thus the structure of G as a local Lie group is uniquely determined by $L(G)$. In the next section we shall show that this is a general property of local Lie groups: G is uniquely determined by $L(G)$. Thus, all questions concerning the structure of G can be reduced to purely algebraic questions about the structure of the Lie algebra $L(G)$.

5.5 The Exponential Map of a Lie Algebra

Let G be a local Lie group, not necessarily a matrix group. If $\mathbf{g}, \mathbf{h} \in V$ then

$$\mathbf{gh} = \boldsymbol{\varphi}(\mathbf{g}, \mathbf{h}) = (\varphi_1(\mathbf{g}, \mathbf{h}), \ldots, \varphi_n(\mathbf{g}, \mathbf{h}))$$

is an analytic function of $\mathbf{g}$ and $\mathbf{h}$. For fixed $\mathbf{h}$ we expand this function in a Taylor series about $\mathbf{g} = \mathbf{e}$:

$$\varphi_i(\mathbf{g}, \mathbf{h}) = h_i + \sum_{j=1}^{n} g_j F_{ij}(\mathbf{h}) + \cdots, \tag{5.1}$$

$$F_{ij}(\mathbf{h}) = [\partial\varphi_i(\mathbf{g}, \mathbf{h})/\partial g_j]|_{\mathbf{g}=\mathbf{e}},$$

where the omitted terms are of order two or more in the g_j. By the associative law, $\mathbf{g}(\mathbf{hk}) = (\mathbf{gh})\mathbf{k}$ for $\mathbf{g}, \mathbf{h}, \mathbf{k} \in V$ sufficiently close to $\mathbf{e}$. Expanding each side of this equality about $\mathbf{g} = \mathbf{e}$, we find

$$\varphi_i(\mathbf{g}, \mathbf{hk}) = \varphi_i(\mathbf{h}, \mathbf{k}) + \sum_{j=1}^{n} g_j F_{ij}(\mathbf{hk}) + \cdots, \tag{5.2}$$

$$\begin{aligned} \varphi_i(\mathbf{gh}, \mathbf{k}) &= \varphi_i(\varphi_l(\mathbf{g}, \mathbf{h}), k_l) = \varphi_i(h_l + \sum_{j=1}^{n} g_j F_{lj}(\mathbf{h}) + \cdots, k_l) \\ &= \varphi_i(\mathbf{h}, \mathbf{k}) + \sum_{j,l=1}^{n} [\partial\varphi_i(\mathbf{h}, \mathbf{k})/\partial h_l] F_{lj}(\mathbf{h}) g_j + \cdots. \end{aligned} \tag{5.3}$$

A comparison of first order terms in g_j leads to the identity

$$F_{ij}(\mathbf{hk}) = \sum_{l=1}^{n} [\partial\varphi_i(\mathbf{h}, \mathbf{k})/\partial h_l] F_{lj}(\mathbf{h}) \tag{5.4}$$

for the functions $F_{ij}(\mathbf{h})$.

Let $\mathbf{g}(t)$ be the analytic curve in G defined in some neighborhood W of $0 \in F$ such that $\mathbf{g}(0) = \mathbf{e}$ and

$$\mathbf{g}(s)\mathbf{g}(t) = \boldsymbol{\varphi}(\mathbf{g}(s), \mathbf{g}(t)) = \mathbf{g}(s+t), \qquad s, t, s+t \in W. \tag{5.5}$$

Such a curve is called a **one-parameter subgroup** of G. Suppose $\alpha \in L(G)$ is the tangent vector to $\mathbf{g}(t)$ at $\mathbf{e}$. Differentiating (5.5) with respect to s, setting $s = 0$ and making use of (5.1), we obtain the system of differential equations

$$dg_i(t)/dt = \sum_{j=1}^{n} \alpha_j F_{ij}(\mathbf{g}), \qquad i = 1, \cdots, n. \tag{5.6}$$

Thus, $\mathbf{g}(t)$ is a solution of the differential system (5.6) such that $\mathbf{g}(0) = \mathbf{e}$, i.e., $g_i(0) = 0$.

Now (5.6) is a system of first-order analytic ordinary differential equations. There is a standard existence and uniqueness theorem for such systems which states that for given constants $\mathbf{a} = (a_1, \ldots, a_n)$ in V, Eq. (5.6) have one and only one solution $\mathbf{g}(t)$ with $\mathbf{g}(0) = \mathbf{a}$. This solution is defined and analytic for all $|t| < \varepsilon$, where $\varepsilon > 0$ depends on the F_{ij} but not on α (see Ince [1]).

We derived (5.6) by assuming $\mathbf{g}(t)$ was a one-parameter subgroup with tangent vector α. Now we choose an arbitrary $\alpha \in L(G)$, and look for a solution $\mathbf{g}(t)$ of (5.6) such that $\mathbf{g}(0) = \mathbf{e}$. [Recall that the functions $F_{ij}(\mathbf{g})$ are known.] By the theorem quoted above, there exists a unique solution, defined and analytic for $|t| < \varepsilon$. We write $\mathbf{g}(t) = \exp \alpha t$.

Theorem 5.16. The integral curve $\mathbf{g}(t) = \exp \alpha t$ is a one-parameter subgroup of G with tangent vector α at $\mathbf{e}$. In particular, $\exp \alpha(s+t) = (\exp \alpha s)(\exp \alpha t)$ for suitably small values of $|s|, |t|$.

Proof. We have shown that $\exp \alpha t$ is an analytic curve in G such that $\exp \alpha 0 = \mathbf{e}$. By (5.1), $F_{ij}(\mathbf{e}) = \delta_{ij}$. Setting $t = 0$ in (5.6) we find $\dot{g}_i(0) = \alpha_i$, $1 \le i \le n$, so α is the tangent vector to $\exp \alpha t$ at $\mathbf{e}$. Let $\mathbf{h}(t) = \mathbf{g}(s+t) = \exp \alpha(s+t)$ for fixed s sufficiently close to zero. Then

$$\dot{h}_i(t) = \frac{dg_i(s+t)}{dt} = \sum_{j=1}^{n} \alpha_j F_{ij}(\mathbf{h}(t)), \qquad \mathbf{h}(0) = \mathbf{g}(s), \tag{5.7}$$

as follows easily from (5.6). On the other hand the function $\mathbf{k}(t) = \mathbf{g}(t)\mathbf{g}(s)$ satisfies

$$\begin{aligned} \dot{k}_i(t) &= \frac{d\varphi_i(\mathbf{g}(t), \mathbf{g}(s))}{dt} = \sum_{j,l=1}^{n} \frac{\partial \varphi_i(\mathbf{g}(t), \mathbf{g}(s))}{\partial h_l} \alpha_j F_{lj}(\mathbf{g}(t)) \\ &= \sum_{j=1}^{n} \alpha_j F_{ij}(\mathbf{k}(t)), \qquad \mathbf{k}(0) = \mathbf{g}(s). \end{aligned} \tag{5.8}$$

Here we have made use of (5.4) and (5.6). The expressions $\partial\varphi_i/\partial h_l$ refer to the derivatives of φ_i with respect to its first n arguments. Thus $\mathbf{h}(t)$ and $\mathbf{k}(t)$

satisfy the same differential system and have the same value at $t = 0$. According to the uniqueness theorem these solutions must be equal: $\mathbf{h}(t) = \mathbf{k}(t)$. Therefore, $(\exp \alpha s)(\exp \alpha t) = \exp \alpha(s + t)$. Q.E.D.

This proof shows that for each $\alpha \in L(G)$ there is a unique one-parameter subgroup with tangent vector α at the identity. Furthermore, every one-parameter subgroup (5.5) is of the form $\mathbf{g}(t) = \exp \alpha t$. Our notation suggests the relation

$$\exp (a\alpha)t = \exp \alpha(at) \tag{5.9}$$

for $a \in F$, $\alpha \in L(G)$, and sufficiently small values of $|t|$. That is, if $\mathbf{g}_\alpha(t) = \exp \alpha t$ then (5.9) asserts $\mathbf{g}_{a\alpha}(t) = \mathbf{g}_\alpha(at)$. The reader can check this equality by verifying that both sides of (5.9) are one-parameter subgroups of G with tangent vector $a\alpha$ at $\mathbf{e}$.

It follows that we can extend the definition of $\exp \alpha t$ to *all* $\alpha \in L(G)$ and $t \in F$ in such a way that

$$(\exp \alpha s)(\exp \alpha t) = \exp \alpha(s + t). \tag{5.10}$$

Indeed $\exp \alpha t$ is originally defined for $|t| < \varepsilon$ and all α. If $|t| \geq \varepsilon$ we define $\exp \alpha t = \exp a\alpha(t/a)$, where a is any constant such that $|t/a| < \varepsilon$. The reader can verify that this definition is independent of a and leads to an entire function of t satisfying the group property (5.10).

For fixed t, $\mathbf{g}(t) = \exp \alpha t$ is an entire function of $\alpha_1, \ldots, \alpha_n$. To see this we expand $\mathbf{g}(t)$ in a Taylor series about $t = 0$:

$$g_k(t) = \sum_{m=1}^{\infty} \frac{d^m g_k(s)}{ds^m}\bigg|_{s=0} \frac{t^m}{m!}, \qquad 1 \leq k \leq n.$$

Now $\dot{g}_k(0) = \alpha_k$, and in general $g_k^{(m)}(0)$ is a homogeneous polynomial of order m in $\alpha_1, \ldots, \alpha_n$. [This follows from (5.6), the chain rule, and a simple induction argument.] Since for suitably small fixed $t \neq 0$ the Taylor series converges for all $\alpha_1, \ldots, \alpha_n$ it follows that $\mathbf{g}(t)$ is an entire function of these variables.

Now set $t = 1$ and consider the **exponential map** $\alpha \longrightarrow \exp \alpha(1) = \exp \alpha$. If G is a global Lie group then $\exp \alpha$ maps all of $L(G)$ into G. If G is only local then $\exp \alpha$ maps a neighborhood of θ in $L(G)$ into G. The n coordinates α_j of α can be used to parametrize a neighborhood of $\mathbf{e}$ in G. Indeed, by (5.1), $F_{ij}(\mathbf{e}) = \delta_{ij}$. It follows from (5.6) that

$$(\partial/\partial\alpha_j)(\exp \alpha)_i|_{\alpha=\theta} = \delta_{ij},$$

so the Jacobian of the transformation $\alpha \longrightarrow \exp \alpha = \mathbf{g}$ is nonzero in a neighborhood of $\theta \in L(G)$. Therefore, by the inverse function theorem the exponential map defines an analytic coordinate transformation on some neighborhood of $\mathbf{e}$. We can use a neighborhood of the zero vector in $L(G)$

to parametrize the local Lie group G. The coordinates α are called **canonical coordinates of the first kind** or just **canonical** coordinates (Pontrjagin [1]). In canonical coordinates the one-parameter subgroups are the straight lines αt through θ. Group multiplication in a one-parameter subgroup is given by $\alpha t + \alpha s = \alpha(t + s)$, i.e., by vector addition. Note, however, that it is in general *not* true that $(\exp \alpha)(\exp \beta) = \exp(\alpha + \beta)$ if α and β are linearly independent vectors.

In the special case where G is a local linear Lie group of $m \times m$ matrices the above analysis becomes much more transparent, From (5.5) a one-parameter subgroup of G is an analytic curve $A(t)$ such that

$$A(s)A(t) = A(s + t) \tag{5.11}$$

for sufficiently small values of $|s|$, $|t|$. Differentiating this expression with respect to s and setting $s = 0$, we obtain the matrix equation

$$dA(t)/dt = \mathfrak{A}A(t), \qquad A(0) = E_m, \tag{5.12}$$

where $\mathfrak{A} \in L(G)$ is the tangent matrix to $A(t)$ at the identity. Clearly, (5.12) is the analogy of (5.6) for matrix groups. It is a consequence of our general theory that the solutions of (5.12) define the exponential mapping of $L(G)$ into G, as $\mathfrak{A}$ ranges over $L(G)$. We have already seen that (5.6) always has a unique solution $\exp \alpha t$ for each $\alpha \in L(G)$. However, since we did not know the functions $F_{ij}(\mathbf{g})$ explicitly we were not able to write down the general form of the solution. In the special case where G is a linear group, however, this difficulty vanishes. Indeed from the remarks following (1.14), (5.12) has the unique solution

$$A(t) = \exp \mathfrak{A}t = \sum_{j=0}^{\infty} \mathfrak{A}^j t^j / j!, \tag{5.13}$$

so the exponential map for linear Lie groups is the ordinary matrix exponential. Thus the transformation

$$\mathfrak{A} \longrightarrow \exp \mathfrak{A}, \qquad \mathfrak{A} \in L(G),$$

is a 1–1 analytic map of a neighborhood of Z in $L(G)$ onto a neighborhood of E_m in G. We can introduce canonical coordinates in G by choosing a basis $\{\mathfrak{B}_j\}$ for $L(G)$. If $\mathfrak{A} \in L(G)$, $\mathfrak{A} = \sum \alpha_j \mathfrak{B}_j$, then the transformation $\alpha \longrightarrow \exp(\sum \alpha_j \mathfrak{B}_j)$ defines canonical coordinates in G. Every one-parameter subgroup in G is of the form $\exp \mathfrak{A}t$, $\mathfrak{A} \in L(G)$.

5.6 Local Homomorphisms and Isomorphisms

In Section 5.3 we defined local analytic homomorphisms of Lie groups and homomorphisms of Lie algebras. Here we study the relationship between these two concepts.

Let G, G' be local Lie groups of dimensions n and n', respectively, with corresponding Lie algebras $L(G)$, $L(G')$. Suppose μ is an analytic homomorphism of G into G'. We will show that μ induces a homomorphism μ^* of $L(G)$ into $L(G')$.

Theorem 5.17. An analytic homomorphism $\mu: G \longrightarrow G'$ induces a Lie algebra homomorphism $\mu^*: L(G) \longrightarrow L(G')$ defined by

$$\mu^*(\alpha) = (d/dt)\mu(\mathbf{g}(t))|_{t=0} \tag{6.1}$$

where $\mathbf{g}(t)$ is an analytic curve in G with tangent vector α at $\mathbf{e}$ The μ^* is an isomorphism (automorphism) if μ is an isomorphism (automorphism).

Proof. First, we show that the right-hand side of (6.1) is well-defined, i.e., that it depends only on $\alpha \in L(G)$, not on the particular curve $\mathbf{g}(t)$. Since $\mu(\mathbf{g}) = (\mu_1(\mathbf{g}), \ldots, \mu_{n'}(\mathbf{g}))$ is analytic with $\mu(\mathbf{e}) = \mathbf{e}'$, we have the Taylor series expansion

$$\mu_i(\mathbf{g}) = \sum_{j=1}^{n} \mu_i{}^j g_j + \cdots, \qquad 1 \leq i \leq n',$$

where only the first-order term in g_j has been given explicitly. Now $\mathbf{g}(0) = \mathbf{e}$ and $\dot{\mathbf{g}}(0) = \alpha$, so (6.1) yields

$$\mu_i{}^*(\alpha) = \sum_{j=1}^{n} \mu_i{}^j \alpha_j, \qquad 1 \leq i \leq n' \tag{6.2}$$

where $\mu^*(\alpha) = (\mu_1{}^*(\alpha), \ldots, \mu_n^*(\alpha))$ and the constants $\mu_i{}^j$ are determined completely by μ. Thus (6.1) is well-defined. It is clear from (6.2) that μ^* is linear, i.e.,

$$\mu^*(a\alpha + b\beta) = a\mu^*(\alpha) + b\mu^*(\beta), \qquad \alpha, \beta \in L(G), \quad a, b \in F.$$

If $\alpha \in L(G)$ then by (6.1) and the definition of μ, $\mu(\exp \alpha t)$ is the one-parameter subgroup of G' with tangent vector $\mu^*(\alpha)$. Thus,

$$\mu(\exp \alpha t) = \exp(\mu^*(\alpha)t). \tag{6.3}$$

Suppose μ is an isomorphism of G onto G'. If $\alpha \neq \theta$ then $\mu(\exp \alpha t)$ will be a nontrivial one-parameter subgroup of G', i.e., $\mu(\exp \alpha t) \not\equiv \mathbf{e}'$. If $\mu^*(\alpha) = \theta'$ then by (6.3) we would have $\mu(\exp \alpha t) \equiv \mathbf{e}'$, a contradiction. Therefore μ^* is 1–1. Furthermore, μ^* is onto because for each $\alpha' \in L(G')$ there is a unique one-parameter subgroup $\exp \alpha t$ in G such that $\mu(\exp \alpha t) = \exp \alpha' t$. According to (6.3), $\alpha' = \mu^*(\alpha)$.

Now assume only that μ is a homomorphism. Let $g(t) = \exp(\alpha\tau)\exp(\beta\tau)\exp(-\alpha\tau)\exp(-\beta\tau)$, $t = \tau^2$, with $\alpha, \beta \in L(G)$. Then

$$\mu(\mathbf{g}(t)) = \exp(\mu^*(\alpha)\tau)\exp(\mu^*(\beta)\tau)\exp(-\mu^*(\alpha)\tau)\exp(-\mu^*(\beta)\tau), \tag{6.4}$$

where we have used (6.3) and the homomorphism property of μ. According

to (3.6) the tangent vector to $\mathbf{g}(t)$ at $\mathbf{e}$ is $[\alpha, \beta]$. Thus, differentiating both sides of (6.4) with respect to t and setting $t = 0$, we obtain

$$\mu^*([\alpha, \beta]) = [\mu^*(\alpha), \mu^*(\beta)]'$$

where the primed commutator refers to $L(G')$. Q.E.D.

We now prove the converse of this theorem for local **linear** Lie groups.

Theorem 5.18. Let G, G' be local linear Lie groups and $\rho: L(G) \to L(G')$ a Lie algebra homomorphism. There exists a unique local analytic homomorphism μ of G into G' such that $\mu^* = \rho$. If ρ is an isomorphism (automorphism) then μ is an isomorphism (automorphism).

Proof. By the preceding theorem, if μ exists then

$$\mu(\exp \mathfrak{A}) = \exp \rho(\mathfrak{A}), \qquad \mathfrak{A} \in L(G), \tag{6.5}$$

where exp is the matrix exponential. Since $\mathfrak{A} \to \exp \mathfrak{A}$ is a 1–1 analytic mapping of a neighborhood of Z in $L(G)$ onto a neighborhood of E_m in G, expression (6.5) uniquely determines μ. Just as in the computation of (6.2), we can show that μ^* as defined by

$$\mu^*(\mathfrak{A}) = (d/dt)\mu(A(t))|_{t=0}, \qquad \dot{A}(0) = \mathfrak{A},$$

depends only on $\mathfrak{A}$, not on the particular analytic curve $A(t)$. Then, a direct computation with $A(t) = \exp t\mathfrak{A}$ yields

$$\mu^*(\mathfrak{A}) = \frac{d}{dt}\,\mu(\exp t\mathfrak{A})\bigg|_{t=0} = \frac{d}{dt}\exp t\rho(\mathfrak{A})\bigg|_{t=0} = \rho(\mathfrak{A}),$$

so $\mu^* = \rho$.

We now show that μ is a local homomorphism. From Theorem 5.5, for $\mathfrak{A}, \mathfrak{B} \in L(G)$ sufficiently close to Z we have $(\exp \mathfrak{A})(\exp \mathfrak{B}) = \exp \mathfrak{C}$, where

$$\mathfrak{C}(\mathfrak{A}, \mathfrak{B}) = \mathfrak{B} + \int_0^1 g(\exp(t \operatorname{Ad} \mathfrak{A}) \exp(\operatorname{Ad} \mathfrak{B}))\mathfrak{A}\, dt. \tag{6.6}$$

Since the integral on the right-hand side of (6.6) is a convergent sum of commutators of elements in $L(G)$ and $L(G)$ is finite-dimensional, it follows that $\mathfrak{C} \in L(G)$. (Prove it!) Furthermore $\mathfrak{C}$ is an analytic function of the coordinates of $\mathfrak{A}$ and $\mathfrak{B}$. By definition, $(\operatorname{Ad} \mathfrak{A})\mathfrak{B} = [\mathfrak{A}, \mathfrak{B}] = \mathfrak{A}\mathfrak{B} - \mathfrak{B}\mathfrak{A}$. Now ρ is a homomorphism, so

$$\rho((\operatorname{Ad} \mathfrak{A})\mathfrak{B}) = \rho([\mathfrak{A}, \mathfrak{B}]) = [\rho(\mathfrak{A}), \rho(\mathfrak{B})] = (\operatorname{Ad} \rho(\mathfrak{A}))\rho(\mathfrak{B}).$$

Applying ρ to both sides of (6.6) and expanding the integral in a power series, we find $\rho(\mathfrak{C}(\mathfrak{A}, \mathfrak{B})) = \mathfrak{C}(\rho(\mathfrak{A}), \rho(\mathfrak{B})) \in L(G')$. (Note that ρ can be

applied term-by-term in the infinite series.) Thus,

$$\mu(\exp \mathcal{A})\mu(\exp \mathcal{B}) = \exp \rho(\mathcal{A}) \exp \rho(\mathcal{B}) = \exp \mathcal{C}(\rho(\mathcal{A}), \rho(\mathcal{B}))$$
$$= \exp \rho(\mathcal{C}(\mathcal{A}, \mathcal{B})) = \mu(\exp \mathcal{C}) = \mu(\exp \mathcal{A} \exp \mathcal{B}),$$

so μ is a local homomorphism. Q.E.D.

Corollary 5.4. Let G and G' be local Linear Lie groups. Then $L(G)$ is isomorphic to $L(G')$ if and only if G is locally isomorphic to G'.

Even though the above results were proved only for **linear** groups they are actually true for any local Lie groups G and G'. One can show that the Campbell–Baker–Hausdorff formula (1.19) is valid for any local Lie group, not just for matrix groups (see Hausner and Schwartz [1]). Once this formula is established, the proof of Theorem 5.18 for general Lie groups follows almost exactly as we have given it.

We know that every local Lie group uniquely defines a Lie algebra. Now suppose $\mathcal{G}$ is an abstract Lie algebra. Does there exist a local Lie group G such that $L(G) = \mathcal{G}$ [or such that $L(G)$ is isomorphic to $\mathcal{G}$]? We establish the affirmative answer to this question in two steps. First we show rigorously that every matrix Lie algebra $\tilde{\mathcal{G}}$ is the Lie algebra $L(G)$ of a local linear Lie group. Then we note by Ado's theorem that any abstract Lie algebra $\mathcal{G}$ is isomorphic to a matrix Lie algebra $\tilde{\mathcal{G}}$. Thus, $\mathcal{G}$ is isomorphic to the Lie algebra $\tilde{\mathcal{G}} = L(G)$ of the local Lie group G.

Theorem 5.19. Let $\tilde{\mathcal{G}}$ be a matrix Lie algebra. There exists a local linear Lie group G such that $L(G) = \tilde{\mathcal{G}}$.

Proof. The obvious candidate for G is the set of matrices $\exp \mathcal{A}$, where $\mathcal{A}$ runs over $\tilde{\mathcal{G}}$. Let the matrices $\{\mathcal{B}_j : j = 1, \ldots, n\}$ form a basis for $\tilde{\mathcal{G}}$. We introduce coordinates $\alpha_1, \ldots, \alpha_n$ in G by writing $A(\alpha) = \exp(\sum \alpha_j \mathcal{B}_j)$ and letting α run over a sufficiently small neighborhood of θ such that the mapping $\alpha \longrightarrow A(\alpha)$ is 1–1. From the Campbell–Baker–Hausdorff formula (1.19),

$$\exp(\sum \alpha_j \mathcal{B}_j) \exp(\sum \beta_k \mathcal{B}_k) = \exp(\sum \gamma_l \mathcal{B}_l),$$

where the γ_l are analytic functions of α_j, β_k for α, β sufficiently close to θ. Thus G is a local linear Lie group. It is easy to verify that $L(G) = \tilde{\mathcal{G}}$. Q.E.D.

Let H be a local Lie group with Lie algebra $L(H)$. By Ado's theorem, $L(H)$ is isomorphic to a matrix Lie algebra $\mathcal{G}$. By Theorem 5.18, $\mathcal{G} = L(G)$, where G is a local linear Lie group. Finally, by the extension of Theorem 5.18 to all local Lie groups, we see that H is isomorphic to G. Thus, every

local Lie group is isomorphic to a local linear Lie group! (There do exist global Lie groups which are not globally isomorphic to linear groups.) In the remainder of this book we study only local linear Lie groups, but this restriction leads to no loss of generality.

Theorem 5.18 is valid only for local groups. For example, the one-dimensional real Lie groups

$$R = \{x: -\infty < x < +\infty\}, \qquad U(1) = \{e^{ix}: x \in R\}$$

have isomorphic Lie algebras, so they are locally isomorphic. Indeed the isomorphism is given by $x \to e^{ix}$. However, this isomorphism cannot be extended to a global isomorphism since the distinct elements x and $x + 2\pi$ of R are mapped into the same element of $U(1)$. Note that R is isomorphic to the linear Lie group of matrices

$$\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}, \qquad x \in R.$$

Since every local Lie group is uniquely determined by its corresponding Lie algebra, to compute all local Lie groups up to isomorphism one need only compute all Lie algebras up to isomorphism. This latter problem is purely algebraic and for low-dimensional Lie algebras at least, it can be solved explicitly (see Jacobson [1]). In general, any problem concerning the structure of a local group G can be reduced to a purely algebraic problem concerning the structure of $L(G)$.

As an example of this relationship recall that if G is commutative then $[\mathcal{A}, \mathcal{B}] = Z$ for all $\mathcal{A}, \mathcal{B} \in L(G)$, i.e., the Lie algebra $L(G)$ is **commutative.** Conversely, if $L(G)$ is commutative it follows immediately from the Campbell–Baker–Hausdorff formula that

$$\exp(\mathcal{A} + \mathcal{B}) = \exp \mathcal{A} \exp \mathcal{B} = \exp \mathcal{B} \exp \mathcal{A}$$

for all $\mathcal{A}, \mathcal{B} \in L(G)$, i.e., G is commutative. Thus G is commutative if and only if $L(G)$ is commutative. An n-dimensional commutative local Lie group is locally isomorphic to F_n.

We now examine some special automorphisms of a local linear Lie group G of $m \times m$ matrices. [Since any local linear Lie group can be uniquely extended to a global group, one can assume without loss of generality that G is a (global) Lie group.] For fixed $B \in G$, the map μ_B,

$$\mu_B(A) = BAB^{-1}, \qquad A \in G,$$

is clearly an automorphism of G, an **inner automorphism.** The set G of all inner automorphisms of G is itself a group since

$$\mu_{B_1B_2}(A) = B_1(B_2AB_2^{-1})B_1^{-1} = \mu_{B_1}\mu_{B_2}(A), \qquad \mu_{E_m}(A) = A. \tag{6.7}$$

Moreover, $\tilde{G}$ is a linear group, the **adjoint group** of G. Indeed we can think of the elements of $\tilde{G}$ as $m^2 \times m^2$ matrices acting on an m^2-dimensional vector

space. The map $\mu: G \longrightarrow \tilde{G}$ given by $A \longrightarrow \mu_A$ defines an analytic homomorphism from G onto $\tilde{G}$.

By Theorem 5.17, for fixed $B \in G$ the inner automorphism μ_B of G induces an automorphism $\mu_B{}^*$ of $L(G)$:

$$\mu_B{}^*(\mathfrak{A}) = (d/dt)\mu_B(\exp t\mathfrak{A})|_{t=0}.$$

Thus,

$$\mu_B{}^*(\mathfrak{A}) = (d/dt)[B(\exp t\mathfrak{A})B^{-1}]|_{t=0} = B\mathfrak{A}B^{-1}. \tag{6.8}$$

It is easy to check that $\mu^*_{B_1B_2} = \mu^*_{B_1}\mu^*_{B_2}$ and $\mu^*_{B^{-1}} = (\mu_B{}^*)^{-1}$, so the operators $\mu_B{}^*$ form a linear group, also called the **adjoint group** $\tilde{G}$. (This makes sense because the two adjoint groups introduced above are locally isomorphic.) We will show that $\tilde{G}$ is a local linear Lie group by explicitly introducing canonical coordinates. Think of $L(G)$ as an n-dimensional vector space on which the $n^2 \times n^2$ matrices $\mu_A{}^*$ are acting. Let $B(t)$ be an analytic curve in G with tangent matrix $\mathfrak{B}$ at the identity. We define the linear operator $\operatorname{Ad}\mathfrak{B}: L(G) \longrightarrow L(G)$ as follows:

$$\operatorname{Ad}\mathfrak{B}(\mathfrak{A}) = \frac{d}{dt}\,\mu^*_{B(t)}(\mathfrak{A})\bigg|_{t=0} = \frac{d}{dt}\,B(t)\mathfrak{A}B^{-1}(t)\bigg|_{t=0} = [\mathfrak{B}, \mathfrak{A}]. \tag{6.9}$$

The operators $\{\operatorname{Ad}\mathfrak{B}: \mathfrak{B} \in L(G)\}$ form a Lie algebra $\widetilde{L(G)}$, the **adjoint Lie algebra,** and the map $\operatorname{Ad}: \mathfrak{B} \longrightarrow \operatorname{Ad}\mathfrak{B}$ is a homomorphism of $L(G)$ onto $\widetilde{L(G)}$. Indeed

(6.10)

$$\operatorname{Ad}(a\mathfrak{A} + b\mathfrak{B}) = a\operatorname{Ad}\mathfrak{A} + b\operatorname{Ad}\mathfrak{B}, \qquad \operatorname{Ad}([\mathfrak{A}, \mathfrak{B}]) = [\operatorname{Ad}\mathfrak{A}, \operatorname{Ad}\mathfrak{B}].$$

The kernel of the map Ad, i.e., the set of $\mathfrak{B} \in L(G)$ such that $[\mathfrak{B}, \mathfrak{A}] = Z$ for all $\mathfrak{A} \in L(G)$ is called the **center** of $L(G)$. Clearly, the center is a commutative Lie algebra.

For any $\mathfrak{B} \in L(G)$ we can consider $\operatorname{Ad}\mathfrak{B}$ as an $n \times n$ matrix acting on the n-dimensional space $L(G)$. Then by Lemma 5.3 we have

$$(\exp(\operatorname{Ad}\mathfrak{B}))\mathfrak{A} = (\exp\mathfrak{B})\mathfrak{A}(\exp\mathfrak{B})^{-1} = B\mathfrak{A}B^{-1}, \tag{6.11}$$

where $B = \exp\mathfrak{B} \in G$, so we can identify $\tilde{G}$ with $\exp\widetilde{L(G)}$ in a neighborhood of the identity matrix E_n. This shows that $\tilde{G}$ is a local linear Lie group with Lie algebra $\widetilde{L(G)}$. If the center of $L(G)$ has dimension n' then $\dim\widetilde{L(G)} = n - n'$ and $\tilde{G}$ is an $(n - n')$-dimensional Lie group.

5.7 Subgroups and Subalgebras

Let G be an n-dimensional local Lie group with elements $\mathbf{g} = (g_1, \ldots, g_n)$ defined in some connected neighborhood $V \in F_n$ of $\mathbf{e} = (0, \ldots, 0)$. Suppose H is an m-dimensional local Lie group over F with elements $\mathbf{h} = (h_1, \ldots, h_m)$ and identity $\mathbf{e}' = (0, \ldots, 0)$.

Definition. H is a **local Lie subgroup** of G if there is an analytic homomorphism $\mu: H \longrightarrow G$ such that (1) $\mu(\mathbf{h})$ is locally 1–1; (2) the vectors $\{\partial\mu(\mathbf{h})/\partial h_j|_{\mathbf{h}=\mathbf{e}'}, 1 \leq j \leq m\}$ are linearly independent.

We can consider the subgroup H as embedded in G. Indeed the image $\mu(H) \subseteq G$ is itself a local Lie group which is isomorphic to H. It is clear that the tangent vectors at the identity of analytic curves in $\mu(H)$ are contained in $L(G)$. Every such analytic curve with tangent vector β is of the form $\mu(\mathbf{h}(t))$, where $\mathbf{h}(t)$ is an analytic curve through $\mathbf{e}'$ in H. Furthermore, if $\mathbf{h}(t)$ has tangent vector α at $\mathbf{e}'$ then $\beta = \mu^*(\alpha)$ by (6.1). Since μ^* is a homomorphism, for every $\beta = \mu^*(\alpha)$, $\beta' = \mu^*(\alpha')$, we have

$$[\beta, \beta'] = [\mu^*(\alpha), \mu^*(\alpha')] = \mu^*([\alpha, \alpha']).$$

Thus the tangent vectors at $\mathbf{e}$ to analytic curves in $\mu(H)$ form a Lie algebra $L(\mu(H))$ contained in $L(G)$. Furthermore, $L(\mu(H))$ is isomorphic to $L(H)$ since μ^* is 1–1.

Definition. Let $\mathcal{G}$ be a Lie algebra and $\mathcal{K}$ a subset of $\mathcal{G}$. We say $\mathcal{K}$ is a **subalgebra** of $\mathcal{G}$ if $\mathcal{K}$ is itself a Lie algebra under the operations in $\mathcal{G}$, i.e., if $\mathcal{K}$ is a vector subspace which is closed under the commutator operation.

Since $L(H)$ is isomorphic to the subalgebra $L(\mu(H))$, we can consider $L(H)$ as a subalgebra of $L(G)$.

Theorem 5.20. If H is a Lie subgroup of the local Lie group G then $L(H)$ is a Lie subalgebra of $L(G)$.

Theorem 5.21. Let G be a local linear Lie group and let $\mathcal{H}$ be a subalgebra of $L(G)$. There exists a unique local Lie subgroup H of G such that $L(H) = \mathcal{H}$.

Proof. Let $\{\mathcal{B}_j\}$, $1 \leq j \leq n$, be a basis for $L(G)$ such that $\{\mathcal{B}_j\}$, $1 \leq j \leq m$, is a basis for $\mathcal{H}$. (Note that $m \leq n$.) By Theorem 5.18 the map

$$(\alpha_1, \ldots, \alpha_n) \longrightarrow \exp(\sum_{j=1}^{n} \alpha_j \mathcal{B}_j) \in G \tag{7.1}$$

defines canonical coordinates on G. Furthermore, by Theorem 5.19 the map

$$(\alpha_1, \ldots, \alpha_m) \longrightarrow \exp(\sum_{j=1}^{m} \alpha_j \mathcal{B}_j) \tag{7.2}$$

defines coordinates in a local linear Lie group (which we call H) as $(\alpha_1, \ldots, \alpha_m)$ ranges over a suitably small neighborhood of θ in F_m. Clearly, $L(H) = \mathcal{H}$ and H is a local Lie subgroup of G. (Here we are considering H as embedded in G). Indeed, in the canonical coordinate system we have chosen, the elements of G that lie in H are just those with coordinates $(\alpha_1, \ldots, \alpha_m, 0, \ldots, 0)$.

If H' is a Lie subgroup of G with $L(H') = \mathcal{H}$, then H' can be described locally by the exponential mapping (7.2). Thus $H' = H$. Q.E.D.

Even though we have proved the above result only for local linear Lie groups it is valid for all local Lie groups. The preceding theorems establish a 1–1 correspondence between Lie subgroups of G and Lie subalgebras of $L(G)$. Any question involving the structure of local Lie subgroups can be reduced to a purely algebraic problem concerning Lie subalgebras.

As an example we work out the relationship between normal subgroups of G and ideals in $L(G)$.

Definition. A local Lie subgroup H of G is a **normal** subgroup if there exists a neighborhood V of the identity in G such that $\mathbf{ghg}^{-1} \in H$ for all $\mathbf{h} \in H \cap V$ and $\mathbf{g} \in G \cap V$.

This definition of normal subgroup agrees with that given in Chapter 1, except that here the normality property is only local.

Definition. A subalgebra $\mathcal{H}$ of the Lie algebra $\mathcal{G}$ is an **ideal** if $[\alpha, \beta] \in \mathcal{H}$ for all $\alpha \in \mathcal{G}$, $\beta \in \mathcal{H}$. We write $[\mathcal{G}, \mathcal{H}] \subseteq \mathcal{H}$.

Theorem 5.22. Let H be a local Lie subgroup of the local linear Lie group G. Then H is normal if and only if $L(H)$ is an ideal.

Proof. Suppose H is a normal subgroup. Given $\mathcal{B} \in L(H)$, we have $A(\exp t\mathcal{B})A^{-1} \in H$ for all $A \in G$ sufficiently close to the identity and all $t \in F$ sufficiently close to zero. Since

$$(d/dt)A(\exp \mathcal{B})A^{-1}|_{t=0} = A\mathcal{B}A^{-1}$$

it follows that $A\mathcal{B}A^{-1} \in L(H)$. Given $\mathcal{A} \in L(G)$, let $A(s) = \exp s\mathcal{A}$. Since $A(s)\mathcal{B}A^{-1}(s)$ is an analytic curve in $L(H)$ and $L(H)$ is finite-dimensional we have

$$[\mathcal{A}, \mathcal{B}] = (d/ds)(\exp(s\mathcal{A})\cdot\mathcal{B}\cdot\exp(-s\mathcal{A}))|_{s=0} \in L(H),$$

so $L(H)$ is an ideal.

Conversely, if $L(H)$ is an ideal and $\mathcal{B} \in L(H)$, $\mathcal{A} \in L(G)$ are chosen in suitably small neighborhoods of Z then $B(t) = \exp \mathcal{A} \exp t\mathcal{B} \exp(-\mathcal{A})$ is a one-parameter subgroup of G with tangent matrix $(\exp \mathcal{A})\mathcal{B}(\exp -\mathcal{A})$ at the identity. However, from Lemma 5.3,

$$(\exp \mathcal{A})\mathcal{B}(\exp -\mathcal{A}) = \sum_{j=0}^{\infty} (\operatorname{Ad} \mathcal{A})^j\mathcal{B}/j! \in L(H)$$

since $L(H)$ is an ideal. Thus $B(t) \in H$ for sufficiently small t and H is normal. Q.E.D.

As usual, the above result holds for all local Lie groups, not just linear groups.

5.8 Representations of Lie Groups

We have seen that any local linear Lie group can be uniquely extended to a connected (global) linear Lie group. Here we assume this has been done, so G is not only a local linear Lie group, but also a matrix group in the abstract sense.

Let V be a finite-dimensional vector space over F and let $GL(V)$ be the group of all nonsingular linear transformations of V onto V.

Definition. A **representation** of a linear Lie group G with **representation space** V is an analytic homomorphism $\mathbf{T}: A \longrightarrow \mathbf{T}(A)$ of G into $GL(V)$.

By an **analytic** homomorphism we mean a homomorphism such that the matrix elements $T_{ij}(A)$ with respect to any basis in V are analytic functions of the local coordinates of A in G. (If the matrix elements are analytic with respect to one basis then they will be analytic with respect to every basis.) This definition agrees with that of Section 3.1 except for the analyticity requirement, which is added so that we can take advantage of the Lie structure of G. Note that $GL(V)$ is isomorphic to $GL(m, F)$ where $m = \dim V$ and an isomorphism is defined by choosing a basis for V.

For every $\mathfrak{A} \in L(G)$ we define the **infinitesimal operator** $\boldsymbol{\mathfrak{A}}$ on V by

$$\boldsymbol{\mathfrak{A}} = (d/dt)\mathbf{T}(A(t))|_{t=0}, \tag{8.1}$$

where $A(t)$ is an analytic curve in G with tangent matrix $\mathfrak{A}$ at the identity. Since $\mathbf{T}$ is an analytic homomorphism, $\boldsymbol{\mathfrak{A}}$ depends on $\mathfrak{A}$ alone, not on the particular curve $A(t)$. Furthermore, the operators $\boldsymbol{\mathfrak{A}}$ form a Lie algebra which is a homomorphic image of $L(G)$.

Definition. A **representation** of a Lie algebra $\mathcal{G}$ with **representation space** V is a map $\boldsymbol{\rho}$ from $\mathcal{G}$ to the space of all linear operators on V such that ($a, b \in F$ and $\alpha, \beta \in \mathcal{G}$):

(1) $\boldsymbol{\rho}(a\alpha + b\beta) = a\boldsymbol{\rho}(\alpha) + b\boldsymbol{\rho}(\beta)$.

(2) $\boldsymbol{\rho}([\alpha, \beta]) = \boldsymbol{\rho}(\alpha)\boldsymbol{\rho}(\beta) - \boldsymbol{\rho}(\beta)\boldsymbol{\rho}(\alpha) = [\boldsymbol{\rho}(\alpha), \boldsymbol{\rho}(\beta)]$.

The operators $\boldsymbol{\mathfrak{A}} = \boldsymbol{\rho}(\mathfrak{A})$ define a rep of $L(G)$ on V, so every rep of G induces a rep of $L(G)$. Conversely, suppose $\boldsymbol{\rho}$ is a rep of $L(G)$ on V. Set $\boldsymbol{\rho}(\mathfrak{A}) = \boldsymbol{\mathfrak{A}}$ and define the map $\mathbf{T}$ by

$$\mathbf{T}(\exp \mathfrak{A}) = \exp \boldsymbol{\mathfrak{A}}, \mathfrak{A} \in L(G), \tag{8.2}$$

where $\mathfrak{A}$ ranges over a suitably small neighborhood $\mathcal{U}$ of Z. By Theorem

5.18, $\mathbf{T}$ is a local analytic homomorphism of G into $GL(V)$. We refer to $\mathbf{T}$ as a **local rep** of G since it is defined only in some neighborhood of the identity element. If $\boldsymbol{\rho}$ is the Lie algebra rep induced from a (global) rep $\mathbf{T}'$ of G it follows from Theorem 5.18 and (8.2) that $\mathbf{T}(A) = \mathbf{T}'(A)$ for all A in a suitably small neighborhood of the identity. Furthermore, if G is connected then any $A \in G$ can be written as a finite chain

$$A = \exp \mathfrak{A}_1 \exp \mathfrak{A}_2 \cdots \exp \mathfrak{A}_k, \qquad \mathfrak{A}_i \in \mathcal{U},$$

so

$$\mathbf{T}'(A) = \mathbf{T}'(\exp \mathfrak{A}_1) \cdots \mathbf{T}'(\exp \mathfrak{A}_k) = \exp \boldsymbol{\mathfrak{A}}_1 \cdots \exp \boldsymbol{\mathfrak{A}}_k$$

and $\mathbf{T}'$ is **uniquely** determined by $\boldsymbol{\rho}$. Thus, any rep of a connected linear Lie group is uniquely determined by its associated Lie algebra rep. If G is not connected, e.g., $G = O(m, R)$, then only the action of the rep in the connected component containing the identity is determined by the Lie algebra rep.

Warning. If $\boldsymbol{\rho}$ is an arbitrary Lie algebra rep it may not be possible to extend the local rep (8.2) of G to a global rep. For example, let $\{\mathfrak{A}_1\}$ be a basis for the one-dimensional Lie algebra of $U(1) = \{e^{ix}\}$. The choice $\boldsymbol{\rho}(\mathfrak{A}_1) = 1$ defines a one-dimensional rep of $L(U(1))$. From (8.2) we get $\mathbf{T}(e^{ix}) = e^{x}$. Since $\exp i(x + 2\pi) = \exp ix$ this local rep does not extend globally.

Thus the correspondence between (global) group reps and Lie algebra reps is *not* 1–1. It is possible to make the correspondence 1–1 by considering only **simply connected** Lie groups. However, the study of such groups leads to topological complications beyond the scope of this book. The interested reader can find relevant material in the work of Pontrjagin [1] or Hausner and Schwartz [1].

The definitions of invariant and irred subspaces as well as the Schur lemmas given in Sections 3.1–3.3 carry over immediately to reps of Lie groups and algebras.

Theorem 5.23. Let $\mathbf{T}$ be a rep of the connected linear Lie group G on the vector space V and let $\boldsymbol{\rho}$ be the associated rep of $L(G)$. Suppose W is a subspace of V. Then, (1) W is $\mathbf{T}$-invariant if and only if it is $\boldsymbol{\rho}$-invariant; (2) W is $\mathbf{T}$-irred if and only if it is $\boldsymbol{\rho}$-irred.

Proof. A straightforward exercise for the reader.

Thus, the problem of computing all reps (irred reps) $\mathbf{T}$ of G is equivalent to the purely algebraic problem of computing all reps (irred reps) $\boldsymbol{\rho}$ of $L(G)$ except that we must check every algebra rep $\boldsymbol{\rho}$ to see that it actually determines a global group rep. In the following chapters we shall usually determine the reps $\boldsymbol{\rho}$ first and then obtain the group reps $\mathbf{T}$ by exponentiation.

5.9 Local Transformation Groups

A knowledge of abstract group theory is not sufficient for the application of group-theoretic methods to physics. In physical problems the groups usually occur as transformation groups acting on some mathematical structure. Here we study the important case of a local Lie group acting on a coordinate manifold. The following definitions and theorems will be stated for complex groups acting on complex manifolds, but the results for real groups acting on real or complex manifolds are completely analogous.

Let U be an open connected set in $\mathfrak{C}_m$. Any $\mathbf{x} \in U$ can be designated by its coordinates $\mathbf{x} = (x_1, \ldots, x_m)$, $x_i \in \mathfrak{C}$. Let G be an n-dimensional local Lie group defined in some connected neighborhood of $\mathbf{e} = (0, \ldots, 0)$ in $\mathfrak{C}_n$. Finally let $\mathbf{Q}$ be a mapping which associates to each pair $(\mathbf{x}, g)$, $\mathbf{x} \in U$, $g \in G$, an element $\mathbf{Q}(\mathbf{x}, g)$ in $\mathfrak{C}_m$. We write $\mathbf{Q}(\mathbf{x}, g) = \mathbf{x}g \in \mathfrak{C}_m$.

Definition. G acts on the manifold U as a **local Lie transformation group** if $\mathbf{Q}$ satisfies the following properties:

(1) $\mathbf{x}g$ is analytic in the $n + m$ coordinates of $\mathbf{x}$ and g;
(2) $\mathbf{x}\mathbf{e} = \mathbf{x}$, all $\mathbf{x} \in U$;
(3) if $\mathbf{x}g \in U$ then $(\mathbf{x}g)h = \mathbf{x}(gh)$, $g, h, gh \in G$.

With the exception of the identity $\mathbf{e}$ the elements of G are printed in lightface type to distinguish them from elements of U. Conditions (2) and (3) express the transformation group property of G. Condition (1) is necessary if we are to make use of the analyticity of G. Note that the group operates on the right $(\mathbf{x}g)$ rather than the left $(g\mathbf{x})$ as in Chapter 1. The reason for this notational change will become apparent later.

If $\mathbf{x} \in U$ and g is in a sufficiently small neighborhood of $\mathbf{e}$, properties (2) and (3) imply $(\mathbf{x}g)g^{-1} = \mathbf{x}\mathbf{e} = \mathbf{x}$. Thus the map $\mathbf{x} \to \mathbf{x}g$ is locally analytic and 1–1 for fixed g, and the inverse mapping is also analytic.

Let $\exp \alpha t$, $\alpha \in L(G)$, be a one-parameter subgroup of G. For fixed $\mathbf{x}^0 \in U$ we call the curve $\mathbf{x}(t) = \mathbf{x}^0 \exp \alpha t \in U$ the **trajectory** of $\mathbf{x}^0$ under $\exp \alpha t$. [Clearly, $\mathbf{x}(t)$ is defined for sufficiently small values of $|t|$.] The vector-valued function $\mathbf{x}(t)$ can be expanded in a Taylor series in t about $t = 0$:

$$x_i(t) = x_i^0 + t(\partial Q_i/\partial t)(\mathbf{x}^0, \exp \alpha t)|_{t=0} + \cdots$$

or

$$x_i(t) = x_i^0 + t \sum_{j=1}^{n} P_{ij}(\mathbf{x}^0)\alpha_j + \cdots, \qquad 1 \le i \le m, \tag{9.1}$$

where

$$P_{ij}(\mathbf{x}^0) = (\partial Q_i/\partial g_j)(\mathbf{x}^0, g)|_{g=\mathbf{e}} \tag{9.2}$$

and $\mathbf{Q} = (Q_1, \ldots, Q_m)$ Thus, $\dot{\mathbf{x}}(0)$ is tangent to the trajectory of $\exp \alpha t$ through $\mathbf{x}^0$.

If $f(\mathbf{x})$ is analytic in a neighborhood of $\mathbf{x}^0$ we can define analytic functions $[(\exp \alpha t)f](\mathbf{x}) = f(\mathbf{x} \exp \alpha t)$ for $\alpha \in L(G)$ and suitably small values of $|t|$. More generally, let $\mathcal{A}_{\mathbf{x}^0}$ be the space of all functions f analytic in a neighborhood of $\mathbf{x}^0$, where the neighborhood is allowed to vary with the function. We define operators $\mathbf{T}(g)\colon \mathcal{A}_{\mathbf{x}^0} \longrightarrow \mathcal{A}_{\mathbf{x}^0}$ by

$$[\mathbf{T}(g)f](\mathbf{x}) = f(\mathbf{x}g), \qquad \mathbf{x} \in U, \quad g \in G. \tag{9.3}$$

For a given $f \in \mathcal{A}_{\mathbf{x}^0}$ the right-hand side will be well-defined for $\mathbf{x}$ suitably close to $\mathbf{x}^0$ and g suitably close to $\mathbf{e}$. We say that (9.3) holds **locally.** Since G is a local transformation group, we have

$$[\mathbf{T}(g_1g_2)f](\mathbf{x}) = f(\mathbf{x}(g_1g_2)) = \{\mathbf{T}(g_1)[\mathbf{T}(g_2)f]\}(\mathbf{x}), \tag{9.4}$$

so $\mathbf{T}(g_1g_2) = \mathbf{T}(g_1)\mathbf{T}(g_2)$ for $g_1, g_2 \in G$. Again the homomorphism property (9.4) holds locally. (If we let G act on U to the left, $\mathbf{x} \longrightarrow g\mathbf{x}$, then to obtain the homomorphism property for $\mathbf{T}$-operators we have to set $[\mathbf{T}(g)f](\mathbf{x}) = f(g^{-1}\mathbf{x})$. To avoid the inverse we have written $\mathbf{x} \longrightarrow \mathbf{x}g$.)

The operators $\mathbf{T}(g)$ define a local rep of G on the infinite-dimensional vector space $\mathcal{A}_{\mathbf{x}^0}$. In analogy with (8.1) we can define infinitesimal operators corresponding to this rep.

Definition. The **Lie derivative** $L_\alpha f$ of an analytic function $f \in \mathcal{A}_{\mathbf{x}^0}$ is

$$L_\alpha f(\mathbf{x}) = (d/dt)f(\mathbf{x}g(t))|_{t=0},$$

where $g(t)$ is an analytic curve in G with tangent vector α at $\mathbf{e}$.

A direct computation yields

$$L_\alpha f(\mathbf{x}) = \sum_{i=1}^{m} \sum_{j=1}^{n} (\partial f(\mathbf{x})/\partial x_i) P_{ij}(\mathbf{x}) \alpha_j,$$

where $P_{ij}(\mathbf{x})$ is given by (9.2). In other words, we have

$$L_\alpha = \sum_{i=1}^{m} \sum_{j=1}^{n} P_{ij}(\mathbf{x}) \alpha_j (\partial/\partial x_i). \tag{9.5}$$

It is clear from this expression that L_α depends only on α, not on the particular analytic curve $g(t)$.

Example 1. Let G be the real line R, a one-dimensional real Lie group. For $a \in R$ the mapping $x \longrightarrow x + a$ for all $x \in R$ defines an action of R as a transformation group on itself. Then $[\mathbf{T}(a)f](x) = f(x + a)$. The constant 1 forms a basis for the one-dimensional Lie algebra $L(R)$. Since $g(t) = t$

is an analytic curve in R with tangent vector 1 we have

$$L_1 f(x) = (d/dt) f(x+t)|_{t=0} = df(x)/dx,$$

or $L_1 = d/dx$. If $\alpha = \alpha \cdot 1 \in L(R)$ for $\alpha \in R$ then $L_\alpha = \alpha\, d/dx$.

Example 2. The proper Euclidean group in the plane $E^+(2)$ has the matrix realization

$$(9.6) \qquad g(x_1, x_2, \varphi) = \begin{pmatrix} \cos\varphi & -\sin\varphi & x_1 \\ \sin\varphi & \cos\varphi & x_2 \\ 0 & 0 & 1 \end{pmatrix}, \qquad x_1, x_2, \varphi \in R$$

where φ is determined up to a multiple of 2π. The group multiplication rule is

$$(9.7) \qquad g(x_1, x_2, \varphi) g(x_1', x_2', \varphi') = g(x_1' \cos\varphi - x_2' \sin\varphi + x_1, \\ x_1' \sin\varphi + x_2' \cos\varphi + x_2, \varphi + \varphi').$$

Thus $\mathbf{a} = (x_1, x_2)$ and

$$\mathbf{O} = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$$

where $\mathbf{a}$ and $\mathbf{O}$ are analogous to (2.4), Section 2.2. We can consider $E^+(2)$ as a transformation group in the plane R_2, where the group action is

$$(9.8) \qquad (y_1, y_2) \longrightarrow ((y_1 - x_1)\cos\varphi + (y_2 - x_2)\sin\varphi, \\ -(y_1 - x_1)\sin\varphi + (y_2 - x_2)\cos\varphi), \\ \mathbf{y} = (y_1, y_2) \in R_2.$$

In terms of the action of $E^+(2)$ as a transformation group defined as in Section 2.2, this reads

$$\mathbf{y} \longrightarrow \{\mathbf{a}, \mathbf{O}\}^{-1}\mathbf{y} = \{-\mathbf{O}^{-1}\mathbf{a}, \mathbf{O}^{-1}\}\mathbf{y}.$$

It follows that $E^+(2)$ is a two-dimensional real Lie group which acts as a Lie transformation group on R_2. As the reader can check, $L(E^+(2))$ is the Lie algebra of all matrices

$$(9.9) \qquad \mathcal{J}(a, b, c) = \begin{pmatrix} 0 & -c & a \\ c & 0 & b \\ 0 & 0 & 0 \end{pmatrix}, \qquad a, b, c \in R.$$

A basis is provided by the matrices $\mathcal{J}(1, 0, 0) = \mathcal{J}_1$, $\mathcal{J}(0, 1, 0) = \mathcal{J}_2$, $\mathcal{J}(0, 0, 1) = \mathcal{J}_3$, with commutation relations

$$[\mathcal{J}_1, \mathcal{J}_2] = Z, \qquad [\mathcal{J}_3, \mathcal{J}_1] = \mathcal{J}_2, \qquad [\mathcal{J}_3, \mathcal{J}_2] = -\mathcal{J}_1.$$

Note that $\exp \mathcal{J}_1 t = g(t, 0, 0)$, $\exp \mathcal{J}_2 t = g(0, t, 0)$, $\exp \mathcal{J}_3 t = g(0, 0, t)$.

The Lie derivatives are

$$L_k f(\mathbf{y}) = (d/dt) f(\mathbf{y} \exp \mathcal{J}_k t)|_{t=0}, \qquad k = 1, 2, 3,$$

or

$$(9.10) \qquad L_1 = -\partial/\partial y_1, \quad L_2 = -\partial/\partial y_2, \quad L_3 = y_2(\partial/\partial y_1) - y_1(\partial/\partial y_2).$$

The Lie derivative corresponding to the Lie algebra element $\mathcal{J}(a, b, c)$ is just $aL_1 + bL_2 + cL_3$. Note that

$$[L_1, L_2]f(\mathbf{y}) = 0, \quad [L_3, L_1]f(\mathbf{y}) = L_2 f(\mathbf{y}), \quad [L_3, L_2]f(\mathbf{y}) = -L_1 f(\mathbf{y}),$$

where $[L_\alpha, L_\beta] = L_\alpha L_\beta - L_\beta L_\alpha$ is the **commutator** of the Lie derivatives L_α, L_β. Thus the L_k satisfy the same commutations relations as the generators of $L(E^+(2))$. In particular the Lie derivatives themselves form a Lie algebra.

Returning to the study of a general transformation group G, we let $\mathbf{x}^0 \in U$. Then for $\alpha \in L(G)$ and sufficiently small values of $|t|, |s|$, $\mathbf{x}(t+s) = \mathbf{x}^0 \cdot \exp \alpha(t+s) = (\mathbf{x}^0 \exp \alpha t) \exp \alpha s = \mathbf{Q}(\mathbf{x}(t), \exp \alpha s)$. Differentiating this expression with respect to s and setting $s = 0$, we find

$$(9.11) \qquad (d/dt)x_i(t) = \sum_{j=1}^{n} P_{ij}(\mathbf{x}(t))\alpha_j, \qquad 1 \le i \le n, \qquad \mathbf{x}(0) = \mathbf{x}^0,$$

where we have used (9.1) and (9.2). From (9.5) this system can be rewritten as $\dot{x}_i(t) = L_\alpha x_i$, or

$$(9.12) \qquad d\mathbf{x}/dt = L_\alpha \mathbf{x}, \qquad \mathbf{x}(0) = \mathbf{x}^0.$$

According to the fundamental existence and uniqueness theorem for differential equations, system (9.12) has a unique solution $\mathbf{x}(t)$. Clearly, this solution is $\mathbf{x}(t) = \mathbf{x}^0 \exp \alpha t$. Since every $g \in G$ sufficiently close to $\mathbf{e}$ lies on a one-parameter subgroup $\exp \alpha t$, the action of G on U is uniquely determined by the solutions of (9.12). Thus a knowledge of the Lie derivatives L_α determines the action of G.

Theorem 5.24. The unique solution of (9.12) is the trajectory $\mathbf{x}(t) = \mathbf{x}^0 \cdot \exp \alpha t$. The Lie derivatives L_α uniquely determine the local Lie transformation group G.

If $f \in \mathcal{A}_{\mathbf{x}^0}$ and $\mathbf{x}(t) = \mathbf{x}^0 \exp \alpha t$ then $f(\mathbf{x}(t))$ satisfies the differential equation

$$\frac{d}{dt} f(\mathbf{x}(t)) = \sum_{i=1}^{m} \frac{\partial f}{\partial x_i} \frac{\partial x_i}{\partial t} = \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{\partial f}{\partial x_i} P_{ij}(\mathbf{x}(t))\alpha_j,$$

or

$$(9.13) \qquad \frac{d}{dt} f(\mathbf{x}(t)) = L_\alpha f(\mathbf{x}(t), \qquad f(\mathbf{x}(0)) = f(\mathbf{x}^0).$$

Again we conclude from the existence and uniqueness theorem that (9.13) has a unique solution $f(\mathbf{x}(t)) \equiv f(\mathbf{x}^0 \exp \alpha t)$ analytic in t. Since $L_\alpha f(\mathbf{x}(t))$ is itself an analytic function of $\mathbf{x}(t)$ it follows that $(d/dt)(L_\alpha f(\mathbf{x}(t))) = L_\alpha(L_\alpha f(\mathbf{x}(t))) = L_\alpha{}^2 f(\mathbf{x}(t))$. Similarly

$$(d/dt)^j f(\mathbf{x}(t)) = L_\alpha{}^j f(\mathbf{x}(t)), \qquad j = 1, 2, \ldots,.$$

Now $f(\mathbf{x}(t))$ has a Taylor series expansion in t about $t = 0$:

$$f(\mathbf{x}(t)) = \sum_{j=0}^{\infty} \left(\frac{d}{ds}\right)^j f(\mathbf{x}(s))\bigg|_{s=0} \frac{t^j}{j!} = \sum_{j=0}^{\infty} (L_\alpha{}^j f)(\mathbf{x}^0)\frac{t^j}{j!}$$
$$= (\exp tL_\alpha) f(\mathbf{x}^0),$$

where $\exp tL_\alpha$ is defined by its formal power series expansion.

Theorem 5.25. $f(\mathbf{x} \exp \alpha t) = (\exp tL_\alpha) f(\mathbf{x}) = \sum_{j=0}^{\infty} (t^j/j!) L_\alpha{}^j f(\mathbf{x})$.

Corollary 5.5. $\mathbf{x} \exp \alpha t = (\exp tL_\alpha)\mathbf{x}$.

We apply these results to the Lie derivative $L_1 = d/dx$ of Example 1. Equation (9.12) becomes

$$dx/dt = 1, \qquad x(0) = x^0,$$

or $x(t) = x^0 + t$, so the action of R as a transformation group on itself is just $x \to x + t$. Furthermore, by Theorem 5.25, if f is analytic near x then

$$f(x \exp 1t) = f(x + t) = \left(\exp t\frac{d}{dx}\right) f(x) = \sum_{j=0}^{\infty} \frac{t^j}{j!} f^{(j)}(x),$$

which is the usual Taylor series expansion.

In Examples 1 and 2 the Lie derivatives themselves form a Lie algebra. In particular the commutator $[L_\alpha, L_\beta]$ is again a Lie derivative. We show that this is true in general.

Theorem 5.26. The set of all Lie derivatives of a local Lie transformation group G forms a Lie algebra which is a homomorphic image of $L(G)$. In fact, (1) $L_{(a\alpha+b\beta)} = aL_\alpha + bL_\beta$; (2) $L_{[\alpha,\beta]} = L_\alpha L_\beta - L_\beta L_\alpha = [L_\alpha, L_\beta]$ for all $a, b \in \mathfrak{C}$ and $\alpha, \beta \in L(G)$.

Proof. Property (1) follows from expression (9.5) for L_α. Property (2) is a little more complicated. Using Corollary 5.5 we can copy the proof of the analogous result for local linear Lie groups. Let $g(t) = (\exp \alpha\tau)(\exp \beta\tau)(\exp -\alpha\tau)(\exp -\beta\tau)$, where $t = \tau^2$. Then $g(t)$ is a curve in G with tangent vector $[\alpha, \beta]$ at $\mathbf{e}$. Thus, for any $f \in \mathfrak{A}_\mathbf{x}$ we have

$$L_{[\alpha,\beta]} f(\mathbf{x}) = (d/dt) f(\mathbf{x}g(t))|_{t=0}. \tag{9.14}$$

However,

$$f(\mathbf{x}g(t)) = \{(\exp \tau L_\alpha)(\exp \tau L_\beta)(\exp -\tau L_\alpha)(\exp -\tau L_\beta)\}f(\mathbf{x})$$
$$= (1 + \tau^2[L_\alpha, L_\beta] + \cdots)f(\mathbf{x}).$$

It follows from this result and (9.14) that $L_{[\alpha,\beta]} = [L_\alpha, L_\beta]$. Q.E.D.

The reader should verify explicitly from (9.5) that $[L_\alpha, L_\beta]$ is a first-order differential operator even though $L_\alpha L_\beta$ and $L_\beta L_\alpha$ are second-order operators.

Two differential operators L_1, L_2 are **equal** on U, $L_1 = L_2$, if $L_1 f \equiv L_2 f$ for all analytic functions f on U. A set $\{L_1, \ldots, L_k\}$ of differential operators is **linearly dependent** on U if there exist constants $a_j \in \mathfrak{C}$ not all zero such that $\sum a_j L_j f \equiv 0$ for all analytic functions f on U. If the set $\{L_j\}$ is not linearly dependent then it is **linearly independent.**

With these definitions, we can discuss the structure of the Lie algebra $\mathfrak{L}(G)$ formed by the Lie derivatives $\{L_\alpha\}$. The map $\alpha \longrightarrow L_\alpha$ is a homomorphism of $L(G)$ onto $\mathfrak{L}(G)$. If this mapping is an isomorphism, i.e., if dim $L(G) =$ dim $\mathfrak{L}(G)$, we say G acts **effectively** as a transformation group.

Every local Lie group G acts effectively as a transformation group on itself: $g \longrightarrow g_0 g, g_0, g \in G$. Indeed it follows from Section 5.2 that the Lie derivatives corresponding to this action are

$$(9.15) \qquad L_\alpha = \sum_{i=1}^{n} \sum_{j=1}^{n} R_{ij}(\mathbf{g})\alpha_j \frac{\partial}{\partial g_i}, \qquad R_{ij}(\mathbf{g}) = \left.\frac{\partial \varphi_i(\mathbf{g}, \mathbf{h})}{\partial h_j}\right|_{\mathbf{h}=\mathbf{e}}.$$

Since $R_{ij}(\mathbf{e}) = \delta_{ij}$ the map $\alpha \longrightarrow L_\alpha$ is an isomorphism. We can now use Theorem 5.25 to gain some insight into the exponential mapping:

$$(9.16) \qquad \mathbf{g} \exp \alpha t = (\exp tL_\alpha)\mathbf{g} = \sum_{j=0}^{\infty} (t^j/j!)L_\alpha{}^j\mathbf{g}, \qquad \mathbf{g} \in G.$$

Suppose G is a local Lie transformation group which does not act effectively, i.e., the map $\alpha \longrightarrow L_\alpha$ is not 1–1. Let

$$L' = \{\alpha \in L(G): L_\alpha \equiv 0\}.$$

Then L' is a nontrivial subspace of $L(G)$. Furthermore, if $\alpha \in L'$ and $\beta \in L(G)$, then

$$L_{[\alpha,\beta]} = [L_\alpha, L_\beta] = [0, L_\beta] = 0,$$

so $[\alpha, \beta] \in L'$. This shows that L' is an ideal in $L(G)$. Thus if G does not act effectively then $L(G)$ contains a proper ideal. If a local Lie group G has a Lie algebra $L(G)$ with no proper ideals then G must always act effectively.

The following result is of basic importance for applications of Lie theory. It states that *any* Lie algebra of differential operators is the algebra of Lie derivatives of a local Lie transformation group.

Theorem 5.27. Let

$$L_j = \sum_{i=1}^{m} P_{ij}(\mathbf{x})(\partial/\partial x_i), \qquad j = 1, \cdots, n, \tag{9.17}$$

be n linearly independent differential operators defined and analytic in a connected open set $U \subset \mathfrak{C}_m$. If there exist constants c^l_{jk} such that

$$[L_j, L_k] = L_j L_k - L_k L_j = \sum_{l=1}^{n} c^l_{jk} L_l, \qquad 1 \le j, k \le n, \tag{9.18}$$

then the n-dimensional Lie algebra $\mathcal{G}$ generated by the L_j is the algebra of Lie derivatives of a local Lie transformation group G acting effectively on U. The action of G on $\mathbf{x}^0 \in U$ is obtained by solving the equations

$$d\mathbf{x}/dt = (\sum_{j=1}^{n} \alpha_j L_j)\mathbf{x}, \qquad \mathbf{x}(0) = \mathbf{x}^0. \tag{9.19}$$

Indeed the unique solution of these equations is $\mathbf{x}^0 \exp \alpha t = \mathbf{Q}(\mathbf{x}^0, \exp \alpha t)$.

Proof. By (9.18), the L_j form a basis for an n-dimensional Lie algebra $\mathcal{G}$. Any element α in $\mathcal{G}$ can be written uniquely as $L_\alpha = \sum \alpha_j L_j$, where $\alpha = (\alpha_1, \ldots, \alpha_n)$. (Note: The Jacobi equality is automatically satisfied by any set of linear operators, so $\mathcal{G}$ is indeed a Lie algebra.) By Theorems 5.9 and 5.19, $\mathcal{G}$ is the Lie algebra of some local Lie group G, unique up to isomorphism. We label the elements of G by means of the exponential mapping: $g = \exp \alpha$, $\alpha \in \mathcal{G}$, i.e., we use canonical coordinates in G.

By Theorem 5.24, the action $\mathbf{x}(g) = \mathbf{x} \exp \alpha = \mathbf{Q}(\mathbf{x}, \exp \alpha)$ of G on U must be given by the solutions of (9.19). We need only verify that G is a transformation group on U, i.e., that $(\mathbf{x}g)h = \mathbf{x}(gh)$, where $g = \exp \alpha$ and $h = \exp \beta$.

Let

$$\begin{aligned} \mathbf{y}(t) &= (\mathbf{x}g)h(t) = \mathbf{Q}(\mathbf{x}g, h(t)) = \exp tL_\beta(\mathbf{x}g) \\ \mathbf{z}(t) &= \mathbf{x}(gh(t)) = \mathbf{Q}(\mathbf{x}, gh(t)), \end{aligned} \tag{9.20}$$

where $h(t) = \exp \beta t$ and $(\mathbf{x}k)_i = Q_i(x_1, \ldots, x_m; k_1, \ldots, k_n.)$ We will show that $\mathbf{y}(t) \equiv \mathbf{z}(t)$. The proof of this fact is not trivial and will be accomplished through a chain of lemmas.

Lemma 5.5. Let G be an n-dimensional local Lie transformation group acting on U. Denote the action of G on U by $\mathbf{x}g = \mathbf{Q}(\mathbf{x}, g)$. Then the function $\mathbf{Q}(\mathbf{x}, g)$ satisfies the identity

$$\sum_{j=1}^{n} \frac{\partial Q_i(\mathbf{x}, g)}{\partial g_j} R_{js}(g) = P_{is}(\mathbf{x}g) = P_{is}(\mathbf{Q}), \qquad 1 \le i \le m, \quad 1 \le s \le n, \tag{9.21}$$

where

$$P_{is}(\mathbf{x}) = \left.\frac{\partial Q_i(\mathbf{x}, h)}{\partial h_s}\right|_{h=\mathbf{e}}, \qquad R_{js}(g) = \left.\frac{\partial \varphi_j(g, k)}{\partial k_s}\right|_{k=\mathbf{e}}.$$

Proof. This result follows from the identity $Q_i(\mathbf{x}, g \exp \alpha t) = Q_i(\mathbf{x}g, \exp \alpha t)$ by differentiating with respect to t and setting $t = 0$ exactly as in the derivation of (5.4) and (5.6). Q.E.D.

Since $R_{js}(\mathbf{e}) = \delta_{is}$, the matrix $(R_{js}(g))$ is invertible for all g in a suitably small neighborhood of $\mathbf{e}$. Denoting the inverse matrix by $S(g)$, we can write (9.21) in the form

$$\partial Q_i(\mathbf{x}, g)/\partial g_j = \sum_{k=1}^{n} P_{ik}(Q) S_{kj}(g), \qquad 1 \le i \le m, \quad 1 \le j \le n. \tag{9.22}$$

Lemma 5.6. Let $P(\mathbf{x})$ be an analytic $m \times n$ matrix of rank n and $S(g) = R^{-1}(g)$ an analytic $n \times n$ matrix with the properties $S(\mathbf{e}) = E_n$ and

(9.23)

$$\sum_{i=1}^{m} \left(P_{lj} \frac{\partial P_{lk}}{\partial x_i} - P_{ik} \frac{\partial P_{lj}}{\partial x_i} \right) = \sum_{r=1}^{n} c^r_{jk} P_{lr}, \qquad 1 \le l \le m, \quad 1 \le j, k \le n,$$

$$\sum_{i=1}^{n} \left(R_{lj} \frac{\partial R_{lk}}{\partial g_i} - R_{ik} \frac{\partial R_{lj}}{\partial g_i} \right) = \sum_{r=1}^{n} c^r_{jk} R_{lr}, \qquad 1 \le j, k, l \le n, \tag{9.24}$$

where the c^r_{jk} are constants. Then (9.22), regarded as a system of equations for $Q_i(\mathbf{x}, g)$, has a unique solution satisfying the initial condition $\mathbf{Q}(\mathbf{x}, \mathbf{e}) = \mathbf{x}$.

Proof. According to the fundamental existence and uniqueness theorem for first-order systems of partial differential equations, the system (9.22) has a unique solution $\mathbf{Q}$ such that $\mathbf{Q}(\mathbf{x}, \mathbf{e}) = \mathbf{x}$ provided the **integrability conditions** $\partial_{g_j} \partial_{g_l} Q_i = \partial_{g_l} \partial_{g_j} Q_i$, or

$$\sum_{q=1}^{m} \sum_{r,s=1}^{n} \left(P_{qr} \frac{\partial P_{is}}{\partial Q_q} - P_{qs} \frac{\partial P_{ir}}{\partial Q_q} \right) S_{rl} S_{sj} = \sum_{s=1}^{n} P_{is} \left(\frac{\partial S_{sl}}{\partial g_j} - \frac{\partial S_{sj}}{\partial g_l} \right) \tag{9.25}$$

are satisfied (Cohen [1], Pontrjagin [1]). It follows from (9.23) and the assumption rank $P = n$ that the integrability conditions become

$$\sum_{r,s=1}^{n} c^k_{rs} S_{rl} S_{sj} = \frac{\partial S_{kl}}{\partial g_j} - \frac{\partial S_{kj}}{\partial g_l}. \tag{9.26}$$

Now $RS = E_n$ (matrix multiplication), so

$$\frac{\partial R}{\partial g_j} S + R \frac{\partial S}{\partial g_j} = Z \qquad \text{or} \qquad \frac{\partial R}{\partial g_j} = -R \frac{\partial S}{\partial g_j} R.$$

Substituting this identity into (9.24), we find (9.24) is equivalent to (9.26). Therefore, the integrability conditions (9.25) are satisfied. Q.E.D.

Now we return to the proof of the theorem. If the function $\mathbf{Q}(\mathbf{x}, g)$ defines the action of a local transformation group on U then it satisfies the system (9.22). Thus, we will construct $\mathbf{Q}(\mathbf{x}, g)$ by requiring that it be a solution of

(9.22) satisfying the initial condition $\mathbf{Q}(\mathbf{x}, \mathbf{e}) = \mathbf{x}$. The matrix functions P and $S = R^{-1}$ are already determined since P is given by (9.17) and R is obtained from G. Furthermore, it follows from (9.15) and (9.18) that P and R satisfy conditions (9.23) and (9.24), where the c^r_{jk} are the structure constants of G. By Lemma 5.6 there is a unique solution $\mathbf{Q}(\mathbf{x}, g)$ of (9.22). This is the same solution we would get by solving (9.19), Indeed $\mathbf{x}(t) = \mathbf{Q}(\mathbf{x}^0, \exp \alpha t)$ satisfies the conditions $\mathbf{x}(0) = \mathbf{x}^0$ and

$$\dot{x}_i(t) = \sum_{jl} \frac{\partial Q_i}{\partial g_j} R_{jl}(\exp \alpha t)\alpha_l = \sum_{sjl} P_{is}(\mathbf{x}(t))S_{sj}(\exp \alpha t)R_{jl}(\exp \alpha t)\alpha_l$$

$$= \sum_{l=1}^{n} P_{il}(\mathbf{x}(t))\alpha_l, \qquad 1 \leq i \leq m.$$

At this point we can verify the equality $\mathbf{y}(t) = \mathbf{z}(t)$ in (9.20). Indeed

$$\dot{y}_i(t) = L_\beta y_i = \sum_{j=1}^{n} P_{ij}(\mathbf{y})\beta_j$$

$$\dot{z}_i(t) = \sum_{qlj} \frac{\partial Q_i(\mathbf{x}, gh(t))}{\partial k_q} \frac{\partial \varphi_q(g, h(t))}{\partial h_l} R_{lj}(h(t))\beta_j$$

$$= \sum_{sqj} P_{is}(\mathbf{z})S_{sq}(gh(t))R_{qj}(gh(t))\beta_j = \sum_{j=1}^{n} P_{ij}(\mathbf{z})\beta_j$$

and $\mathbf{y}(0) = \mathbf{z}(0) = \mathbf{x}^0 \exp \alpha$. [We have used the identity

$$\sum_{q=1}^{n} \frac{\partial \varphi_q(g, h)}{\partial h_l} R_{lj}(h) = R_{qj}(gh),$$

whose derivation is analogous to (5.4).] Since $\mathbf{y}(t)$ and $\mathbf{z}(t)$ satisfy the same system of equations and the same initial conditions, we have $\mathbf{y}(t) \equiv \mathbf{z}(t)$. Setting $t = 1$ we conclude that $\mathbf{x}^0(\exp \alpha \exp \beta) = (\mathbf{x}^0 \exp \alpha) \exp \beta$. Finally the reader can verify from Theorem 5.25 that $\mathbf{Q}(\mathbf{x}, g)$ is analytic in its $m + n$ arguments. Q.E.D.

Thus, the study of local Lie transformation groups is completely equivalent to the study of Lie algebras of differential operators. We shall need a slight extension of the above result: the equivalence of local multiplier reps and Lie algebras of generalized Lie derivatives.

Let L_j, $1 \leq j \leq n$, be linearly independent differential operators defined on the neighborhood $U \subset \mathfrak{C}_{m+1}$ and such that $[L_j, L_k] = \sum c^l_{jk} L_l$ for $i, k = 1, \ldots, n$, where the c^l_{jk} are constants, Suppose also that

$$[\partial/\partial x_{m+1}, L_j] = 0, \qquad 1 \leq j \leq n,$$

i.e., $(\partial/\partial x_{m+1})(P_{ij}(\mathbf{x})) \equiv 0$ for $1 \leq j \leq n$, $1 \leq i \leq m + 1$, where

$$L_j = \sum_{i=1}^{m+1} P_{ij}(\mathbf{x})(\partial/\partial x_i).$$

We emphasize the distinction between x_{m+1} and the remaining coefficients of

$\mathbf{x}$ by writing $\mathbf{x} = (\mathbf{x}', w)$, where $\mathbf{x}' = (x_1, \ldots, x_m)$, $w = x_{m+1}$. Theorem 5.27 states that the L_j generate a Lie algebra which is the algebra of Lie derivatives of a local group G acting on U. The action of G is given by integration of the differential system

(9.27)
$$\dot{x}_i(t) = \sum_{j=1}^{n} P_{ij}(\mathbf{x}'(t))\alpha_j, \qquad \dot{w}(t) = \sum_{j=1}^{n} P_{m+1,j}(\mathbf{x}'(t))\alpha_j, \qquad \mathbf{x}(0) = \mathbf{x}.$$

We write $P_{ij}(\mathbf{x}) = P_{ij}(\mathbf{x}')$ since the P_{ij} are independent of $\mathbf{w}$. The right-hand sides of Eq. (9.27) are independent of w, so the solution takes the form $\mathbf{x}(t) = \mathbf{x} \exp \alpha t = (\mathbf{x}' \exp \alpha t, q(\mathbf{x}', \exp \alpha t) + w)$, where q is an analytic scalar-valued function of its $m + n$ arguments. The action of G on U is completely determined by the one-parameter subgroups

(9.28)
$$\mathbf{x}g = (\mathbf{x}'g, q(\mathbf{x}', g) + w), \qquad g \in G.$$

If $\mathbf{x}g_1 \in U$ and $g_2 \in G$, then $(\mathbf{x}g_1)g_2 = \mathbf{x}(g_1g_2)$, which leads to

$$(\mathbf{x}'g_1)g_2 = \mathbf{x}'(g_1g_2), \qquad q(\mathbf{x}', g_1g_2) = q(\mathbf{x}', g_1) + q(\mathbf{x}'g_1, g_2),$$

or

(9.29)
$$\nu(\mathbf{x}', g_1g_2) = \nu(\mathbf{x}', g_1)\nu(\mathbf{x}'g_1, g_2),$$

where $\nu(\mathbf{x}', g) = \exp q(\mathbf{x}, g)$. If $g = \mathbf{e}$ then $q(\mathbf{x}', \mathbf{e}) = 0$ and $\nu(\mathbf{x}', \mathbf{e}) = 1$ for all $x \in U$.

Let $\mathfrak{A}_w$ be the set of all functions $f(\mathbf{x})$ on $\mathfrak{C}_{m+1}$, analytic in a neighborhood of $\mathbf{x}^0 \in U$ and such that $f(\mathbf{x}) = \mathbf{e}^w h(\mathbf{x}')$, i.e., $e^{-w}f(\mathbf{x})$ is independent of w. Without loss of generality we can assume $\mathbf{x}^0 = \boldsymbol{\theta}$. Given $g \in G$, we define operators $\mathbf{T}_w{}^g$ on $\mathfrak{A}_w$ by

$$\mathbf{T}_w{}^g f(\mathbf{x}) = f(\mathbf{x}g), \qquad f \in \mathfrak{A}_w.$$

By (9.28), $\mathbf{T}_w{}^g f \in \mathfrak{A}_w$ and these operators define a local rep of G on $\mathfrak{A}_w$.

Let $\mathfrak{A}$ be the set of all functions $f(\mathbf{x}')$ on $\mathfrak{C}_m$ analytic in some neighborhood of $\boldsymbol{\theta}' = (0, \ldots, 0)$. There is a 1–1 mapping ρ of $\mathfrak{A}$ onto $\mathfrak{A}_w$ given by $\rho: f(\mathbf{x}') \to e^w f(\mathbf{x}')$. Since this mapping is invertible, the operators $\mathbf{T}_w{}^g$ on $\mathfrak{A}_w$ induce operators $\mathbf{T}^g = \rho^{-1}\mathbf{T}_w{}^g\rho$ on $\mathfrak{A}$ with the properties

(9.30)
$$\begin{aligned} &(1) \quad [\mathbf{T}^g f](\mathbf{x}') = \nu(\mathbf{x}', g) f(\mathbf{x}'g) \\ &(2) \quad [\mathbf{T}^e f](\mathbf{x}') = f(\mathbf{x}') \\ &(3) \quad [\mathbf{T}^{g_1g_2} f](\mathbf{x}') = [\mathbf{T}^{g_1}(\mathbf{T}^{g_2} f)](\mathbf{x}') \end{aligned}$$

valid for all $f \in \mathfrak{A}$, $\mathbf{x}' \in \mathfrak{C}_m$, and $g_1, g_2 \in G$ such that both sides of expressions (9.30) make sense. The $\mathbf{T}^g$ define a local rep of G on $\mathfrak{A}$.

Definition. Let G be a local Lie transformation group acting on a neighborhood $U \subset \mathfrak{C}_m$, $\mathbf{x}^0 \in U$, and let $\mathfrak{A}$ be the set of all functions analytic in a neighborhood of $\mathbf{x}^0$. A **(local) multiplier representation** $\mathbf{T}$ of G on $\mathfrak{A}$ with

multiplier ν consists of a mapping $\mathbf{T}(g)$ of $\mathfrak{A}$ onto $\mathfrak{A}$ defined for $g \in G$ and $f \in \mathfrak{A}$ by

$$[\mathbf{T}(g)f](\mathbf{x}) = \nu(\mathbf{x}, g)f(\mathbf{x}g).$$

Here $\nu(\mathbf{x}, g)$ is a scalar-valued function analytic in $\mathbf{x}$ and g such that (1) $\nu(\mathbf{x}, \mathbf{e}) = 1$ and (2) $\nu(\mathbf{x}, g_1g_2) = \nu(\mathbf{x}, g_1)\nu(\mathbf{x}g_1, g_2)$.

Property (2) is equivalent to (3) of expressions (9.30). The mappings (9.30) define a multiplier rep and every multiplier rep can be so obtained. Indeed, if $\mathbf{T}$ is a multiplier rep of G on $U \subset \mathfrak{C}_m$ we can define an action of G as a transformation group on a neighborhood of $\mathfrak{C}_{m+1}$ by $(\mathbf{x}, w) \longrightarrow (\mathbf{x}g, w + \ln[\nu(\mathbf{x}, g)])$, where $\mathbf{x} \in U$ and $w \in \mathfrak{C}$. The mapping $\mathbf{T}^g$, (9.30), induced by this action is just $\mathbf{T}(g)$.

Let $\mathbf{T}$ be a multiplier rep of $G, f \in \mathfrak{A}$, and $\alpha \in L(G)$.

Definition. The **generalized Lie derivative** $D_\alpha f$ of f under the one-parameter group $\exp \alpha t$ is the analytic function

$$D_\alpha f(\mathbf{x}) = (d/dt)[\mathbf{T}(\exp \alpha t)f](\mathbf{x})|_{t=0}. \tag{9.31}$$

For $\nu \equiv 1$ the generalized Lie derivative becomes the ordinary Lie derivative L_α.

Direct computation from (9.31) yields

$$D_\alpha f(\mathbf{x}) = \sum_{j=1}^{n} \sum_{i=1}^{m} P_{ij}(\mathbf{x})\alpha_j \frac{\partial f}{\partial x_i} + \sum_{j=1}^{n} \alpha_j P_j(\mathbf{x}) f(\mathbf{x}), \tag{9.32}$$

where the analytic functions $P_{ij}(\mathbf{x})$ are defined by (9.2) and

$$\sum_{j=1}^{n} \alpha_j P_j(\mathbf{x}) = (d/dt)\nu(\mathbf{x}, \exp \alpha t)|_{t=0}.$$

Multiplier reps are a particular type of ordinary local reps, so Theorems 5.24–5.27 have immediate analogies for multiplier reps and generalized Lie derivatives. The elementary proofs of the following results are left to the reader.

Theorem 5.28. The generalized Lie derivatives of a local multiplier rep form a Lie algebra under the operations of addition of derivatives and Lie bracket

$$[D_\alpha, D_\beta] = D_\alpha D_\beta - D_\beta D_\alpha.$$

This algebra is a homomorphic image of $L(G)$:

$$D_{a\alpha+b\beta} = aD_\alpha + bD_\beta, \qquad D_{[\alpha,\beta]} = [D_\alpha, D_\beta].$$

Definition. A Lie group G acts **effectively** in a local multiplier rep $\mathbf{T}$ if $L(G)$ is isomorphic to the algebra of generalized Lie derivatives.

Theorem 5.29. A local multiplier rep is completely determined by its generalized Lie derivatives.

Theorem 5.30. $[\mathbf{T}(\exp \alpha t)f](\mathbf{x}) = \sum_{j=0}^{\infty} (t^j/j!)D_\alpha{}^j f(\mathbf{x}) = \exp(tD_\alpha)f(\mathbf{x})$.

Theorem 5.31. Let

$$D_j = \sum_{i=1}^{m} P_{ij}(\mathbf{x})(\partial/\partial x_i) + P_j(\mathbf{x}), \qquad j = 1, \cdots, n,$$

be n linearly independent differential operators defined and analytic in an open set $U \subset \mathfrak{C}_m$. If there exist constants c^l_{jk} such that

$$[D_j, D_k] = \sum_{l=1}^{n} c^l_{jk} D_l, \qquad 1 \leq j, k \leq n,$$

then the D_j form a basis for a Lie algebra which is the algebra of generalized Lie derivatives of an effective local multiplier rep $\mathbf{T}$. The action of the group G is obtained by integration of the equations

$$(9.33) \qquad \dot{x}_i(t) = \sum_{j=1}^{n} P_{ij}(\mathbf{x}(t))\alpha_j, \qquad (d/dt) \ln v(\mathbf{x}^0, \exp \alpha t) = \sum_{j=1}^{n} \alpha_j P_j(\mathbf{x}(t)),$$

where $\mathbf{x}(0) = \mathbf{x}^0$, $v(\mathbf{x}^0, \mathbf{e}) = 1$, $\mathbf{x}(t) = \mathbf{x}^0 \exp \alpha t$, and $1 \leq i \leq m$.

5.10 Examples of Transformation Groups

Consider the group $SL(2) = SL(2, \mathfrak{C})$ of all 2×2 complex matrices

$$(10.1) \qquad g = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

with $\det g = 1$. As we showed in Section 5.4, the Lie algebra $sl(2)$ of $SL(2)$ is the three-dimensional space of all 2×2 complex matrices α with $\operatorname{tr} \alpha = 0$:

$$(10.2) \qquad \alpha = \begin{pmatrix} \alpha_1 & \alpha_2 \\ \alpha_3 & -\alpha_1 \end{pmatrix}.$$

The elements

$$(10.3) \qquad \mathcal{J}^+ = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix}, \quad \mathcal{J}^- = \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix}, \quad \mathcal{J}^3 = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix}$$

define a basis for $sl(2)$ with commutation relations

$$(10.4) \qquad [\mathcal{J}^3, \mathcal{J}^\pm] = \pm\mathcal{J}^\pm, \qquad [\mathcal{J}^+, \mathcal{J}^-] = 2\mathcal{J}^3.$$

An explicit computation shows that $\mathcal{J}^\pm$, $\mathcal{J}^3$ generate one-parameter subgroups

(10.5)

$$\exp a\mathcal{J}^3 = \begin{pmatrix} e^{a/2} & 0 \\ 0 & e^{-a/2} \end{pmatrix}, \quad \exp b\mathcal{J}^+ = \begin{pmatrix} 1 & -b \\ 0 & 1 \end{pmatrix}, \quad \exp c\mathcal{J}^- = \begin{pmatrix} 1 & 0 \\ -c & 1 \end{pmatrix}.$$

In terms of canonical coordinates we can write every $g \in SL(2)$ uniquely

in the form

$$g = \exp(a\mathcal{J}^3 + b\mathcal{J}^+ + c\mathcal{J}^-)$$

for g sufficiently close to the identity. However, we shall adopt another method of parametrizing $SL(2)$. Note that

(10.6)

$$(\exp b'\mathcal{J}^+)(\exp c'\mathcal{J}^-)\exp \tau'\mathcal{J}^3 = \begin{pmatrix} (\exp \frac{1}{2}\tau')(1 + b'c') & -b' \exp -\frac{1}{2}\tau' \\ -c' \exp \frac{1}{2}\tau' & \exp -\frac{1}{2}\tau' \end{pmatrix}.$$

Thus, if $g \in SL(2)$, (10.1), is sufficiently close to the identity we can write it uniquely in the form

(10.7) $$g = (\exp b'\mathcal{J}^+)(\exp c'\mathcal{J}^-)\exp \tau'\mathcal{J}^3,$$

where $\exp \frac{1}{2}\tau' = d^{-1}$, $b' = -b/d$, and $c' = -cd$.

Consider the differential operators

(10.8) $$J^+ = -2uz + z^2(d/dz), \qquad J^- = -d/dz, \qquad J^3 = -u + z(d/dz)$$

acting on a neighborhood of $0 \in \mathfrak{C}$. Here $2u$ is a complex number. These operators satisfy the commutation relations

(10.9) $$[J^3, J^\pm] = \pm J^\pm, \qquad [J^+, J^-] = 2J^3.$$

Comparing these relations with (10.4), we see that the J-operators generate a Lie algebra isomorphic to $sl(2)$. Hence the J-operators are generalized Lie derivatives corresponding to a local multiplier rep of $SL(2)$ on $\mathfrak{C}$. We will use Theorem 5.31 to compute this multiplier rep.

According to (9.33) the action of the one-parameter subgroup $\exp \tau\mathcal{J}^3$ on $\mathfrak{C}$ is obtained by integration of the equations

(10.10)

$$\frac{dz}{d\tau} = z, \qquad \frac{d}{d\tau}\ln \nu(z^0, \exp \tau\mathcal{J}^3) = -u, \qquad z(0) = z^0, \quad \nu(z^0, \mathbf{e}) = 1.$$

The solution is

$$z(\tau) = z^0 e^\tau, \qquad \nu(z^0, \exp \tau\mathcal{J}^3) = e^{-u\tau},$$

so if f is analytic in a neighborhood of $z^0 \in \mathfrak{C}$ then

(10.11) $$[\mathbf{T}(\exp \tau\mathcal{J})f](z^0) = e^{-u\tau} f(z^0 e^\tau).$$

To compute the action of $\exp b\mathcal{J}^+$ we must integrate the equations

(10.12)

$$\frac{dz}{db} = z^2, \qquad \frac{d}{db}\ln \nu(z^0, \exp b\mathcal{J}^+) = -2uz, \qquad z(0) = z^0, \quad \nu(z^0, \mathbf{e}) = 1.$$

The solution is

$$z(b) = z^0(1 - bz^0)^{-1}, \qquad \nu(z^0, \exp b\mathcal{J}^+) = (1 - bz^0)^{2u}, \qquad |bz^0| < 1.$$

Thus,

$$(10.13)\qquad [\mathbf{T}(\exp b\mathcal{J}^+)f](z^0) = (1 - bz^0)^{2u} f\left(\frac{z^0}{1 - bz^0}\right), \qquad |bz^0| < 1.$$

A similar computation yields

$$(10.14)\qquad [\mathbf{T}(\exp c\mathcal{J}^-)f](z^0) = f(z^0 - c).$$

Note. This action of $SL(2)$ is purely local. Unless $2u$ is a nonnegative integer, it is not possible to extend our local multiplier rep to a global rep of $SL(2)$. To give a precise meaning to Eqs. (10.11), (10.13), and (10.14) it would be necessary to state explicitly the values of τ, b, c, and z^0 and the functions f for which the equations are defined. However, these values are easily determined by inspection once the domain of f is given, so ordinarily we will not bother to list them. Similar remarks hold for all computations in this book involving local transformation groups.

We can now determine the local multiplier rep $\mathbf{T}(g)$ defined by the generalized Lie derivatives (10.8). Expressing g by (10.7) we have

$$\begin{aligned}[\mathbf{T}(g)f](z) &= [\mathbf{T}(\exp b'\mathcal{J}^+)\mathbf{T}(\exp c'\mathcal{J}^-)\mathbf{T}(\exp \tau'\mathcal{J}^3)f](z)\\ &= (\exp -u\tau')(1 - b'z)^{2u} f\left(\frac{z(\exp \tau')(1 + c'b') - c' \exp \tau'}{1 - b'z}\right)\\ &= (bz + d)^{2u} f\left(\frac{az + c}{bz + d}\right).\end{aligned}$$

Thus, if $g \in SL(2)$, (10.1), then $zg = (az + c)/(bz + d)$, $v(z, g) = (bz + d)^{2u}$, and

$$(10.15)\qquad [\mathbf{T}(g)f](z) = (bz + d)^{2u} f\left(\frac{az + c}{bz + d}\right)$$

for g in a sufficiently small neighborhood of $\mathbf{e}$. The operators (10.8) are the generalized Lie derivatives of (10.15). For example, by (10.5),

(10.16)

$$J^+ f(z) = \frac{d}{db}\mathbf{T}(\exp b\mathcal{J}^+)f(z)\bigg|_{b=0} = -2uzf(z) + z^2 \frac{d}{dz} f(z).$$

Our general theory shows that the operators (10.15) define a local multiplier rep of $SL(2)$: $\mathbf{T}(g_1)\mathbf{T}(g_2) = \mathbf{T}(g_1 g_2)$ for g_1, g_2 in a sufficiently small neighborhood of $\mathbf{e}$, a fact which the reader can also verify directly.

There is one case in which this local rep can actually be extended to a global group rep of $SL(2)$. Suppose $2u$ is a nonnegative integer and let $\mathcal{V}^{(u)}$ be the $(2u + 1)$-dimensional vector space consisting of all polynomials

$$f(z) = \sum_{j=0}^{2u} c_j z^j.$$

Now $\mathcal{V}^{(u)}$ is invariant under the operators $\mathbf{T}(g)$ for *all* $g \in SL(2)$ since by (10.15) each such operator maps a polynomial of order $2u$ into a polynomial of order $2u$. Furthermore, it is easy to check that on $\mathcal{V}^{(u)}$ the group property is valid for *all* $g \in SL(2)$. Indeed, if $g, g' \in SL(2)$ then

$$(10.17)\quad \mathbf{T}(g)[\mathbf{T}(g')f](z) = [\mathbf{T}(gg')f](z)$$
$$= [(b'a + bd')z + (b'c + dd')]^{2u} f\left(\frac{(a'a + c'b)z + (a'c + c'd)}{(b'a + d'b)z + (b'c + d'd)}\right),$$

where

$$gg' = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} = \begin{pmatrix} aa' + bc', & ab' + bd' \\ ca' + dc', & cb' + dd' \end{pmatrix}.$$

Thus, we have obtained a class of finite-dimensional reps $\mathbf{D}^{(u)}$, $2u = 0, 1, 2, \ldots$, of the global group $SL(2)$, i.e., reps in the sense of Section 5.8.

Let us show that the reps $\mathbf{D}^{(u)}$ are irred. By Theorem 5.23 it is enough to show that the induced Lie algebra reps (which we also denote $\mathbf{D}^{(u)}$) are irred. Choose the natural basis $h_j(z) = z^j, j = 0, 1, \ldots, 2u$, for $\mathcal{V}^{(u)}$. The action of the generalized Lie derivatives on this basis is

(10.18)

$$J^3 h_j = \left(-u + z\frac{d}{dz}\right)z^j = (j - u)h_j, \qquad J^- h_j = -\frac{dz^j}{dz} = -jh_{j-1},$$

$$J^+ h_j = \left(-2uz + z^2\frac{d}{dz}\right)z^j = (j - 2u)h_{j+1}.$$

This action defines the Lie algebra rep $\mathbf{D}^{(u)}$ of $sl(2)$. We will use one of the Schur lemmas to prove that $\mathbf{D}^{(u)}$ is irred. (Note the remarks at the end of Section 5.8). Let $\mathbf{A}$ be a linear operator on $\mathcal{V}^{(u)}$ which commutes with $J^{\pm}$, J^3. Then $\mathbf{A}$ must commute with all of the J-operators in the rep $\mathbf{D}^{(u)}$. Now $J^3\mathbf{A}h_j = \mathbf{A}J^3 h_j = (j - u)\mathbf{A}h_j$, so $\mathbf{A}h_j$ is a multiple of h_j: $\mathbf{A}h_j = \alpha_j h_j$, $\alpha_j \in \mathbb{C}$. (Each eigenvalue $u - j$ of J^3 has multiplicity one.) Furthermore,

$$J^+\mathbf{A}h_j = (j - 2u)\alpha_j h_{j+1}, \qquad \mathbf{A}J^+ h_j = (j - 2u)\alpha_{j+1} h_{j+1},$$

so $\alpha_j = \alpha_{j+1}$, $0 \le j \le 2u - 1$. Setting $\alpha_0 = \alpha$ we obtain $\mathbf{A}h_j = \alpha h_j$ or $\mathbf{A} = \alpha\mathbf{E}$, where $\mathbf{E}$ is the identity operator. Therefore, $\mathbf{D}^{(u)}$ is irred. We will see later that the $\{\mathbf{D}^{(u)}\}$ constitute a complete set of nonequivalent irred reps of the complex Lie group $SL(2)$.

Now that we have constructed the reps $\mathbf{D}^{(u)}$ let us examine their relationship to special function theory. The **matrix elements** $D_{lk}(g)$ of $\mathbf{T}(g)$ with respect to the basis $\{h_j\}$ are defined by

$$(10.19)\quad [\mathbf{T}(g)h_k](z) = \sum_{l=0}^{2u} D_{lk}(g)h_l(z), \qquad g \in SL(2), \quad 0 \le k \le 2u,$$

or explicitly,

$$(10.20)\qquad (az+c)^k(bz+d)^{2u-k} = \sum_{l=0}^{2u} D_{lk}(g)z^l.$$

Since $\mathbf{T}(g_1g_2) = \mathbf{T}(g_1)\mathbf{T}(g_2)$ the matrix elements satisfy the **addition theorems**

$$(10.21)\qquad D_{lk}(g_1g_2) = \sum_{j=0}^{2u} D_{lj}(g_1)D_{jk}(g_2), \qquad l, k = 0, 1, \ldots, 2u.$$

Formula (10.20) is a **generating function** for the matrix elements. Using the binomial theorem to expand the left-hand side of this expression we find that the matrix elements are given by

(10.22)

$$D_{lk}(g) = \begin{cases} \dfrac{a^l d^{2u-k} c^{k-l} k!}{l!(k-l)!}\, {}_2F_1\left(-l, -2u+k; k-l+1; \dfrac{bc}{ad}\right), \\ \qquad\qquad\qquad\qquad\qquad\qquad 2u \geq k \geq l \geq 0, \\ \dfrac{a^k d^{2u-l} b^{l-k}(2u-k)!}{(2u-l)!(l-k)!}\, {}_2F_1\left(-k, -2u+l; l-k+1; \dfrac{bc}{ad}\right), \\ \qquad\qquad\qquad\qquad\qquad\qquad 2u \geq l \geq k \geq 0, \end{cases}$$

where the ${}_2F_1$ are hypergeometric polynomials. (See the Symbol Index.) Relations (10.21) yield addition theorems for the ${}_2F_1$ which are not easy to prove directly. We shall show that this relation between group reps and special functions occurs frequently: Many special functions appear as matrix elements of Lie group reps and the group property leads to addition theorems obeyed by the functions.

We return to the case where $2u$ is an arbitrary complex number, not a nonnegative integer. Then (10.15) and (10.17) make sense only for group elements in a small neighborhood of the identity and functions f analytic in a neighborhood of $z = 0$. Since $2u$ is not a nonnegative integer, the multiplier $\nu(z, g) = (bz+d)^{2u}$ is defined by its power series expansion in z about $z = 0$. This series converges only if $|bz/d| < 1$. Furthermore if f is analytic near zero then $f(zg)$ is analytic near zero only if $(az+c)/(bz+d)$ is in the domain of f and $|az/c| < 1$ and $|bz/d| < 1$. Let $\mathfrak{A}$ be the space of all functions f analytic in a neighborhood of zero. There is now no finite-dimensional subspace of $\mathfrak{A}$ which is invariant under the operators $\mathbf{T}(g)$, (10.15). However, $\mathfrak{A}$ itself is invariant since if $f \in \mathfrak{A}$ we have $\mathbf{T}(g)f \in \mathfrak{A}$ for g is sufficiently close to the identity. Since the elements of $\mathfrak{A}$ are just those functions with convergent power series expansions about $z = 0$, the action of $\mathbf{T}(g)$ on $\mathfrak{A}$ can be determined from a knowledge of $\mathbf{T}(g)$ on the basis functions $h_l(z) = z^l$, $l = 0, 1, 2, \ldots$. We define the matrix elements $B_{lk}(g)$ by

$$(10.23)\qquad [\mathbf{T}(g)h_k](z) = \sum_{l=0}^{\infty} B_{lk}(g)h_l(z), \qquad k = 0, 1, 2, \ldots,$$

or

$$(10.24)\qquad (az+c)^k(bz+d)^{2u-k} = \sum_{l=0}^{\infty} B_{lk}(g)z^l.$$

Since **T** is a multiplier rep we have

$$(10.25)\qquad B_{lk}(g_1g_2) = \sum_{j=0}^{\infty} B_{lj}(g_1)B_{jk}(g_2), \qquad l, k = 0, 1, 2, \ldots,$$

for g_1, g_2 in a sufficiently small neighborhood of **e**. Thus, (10.24) is a generating function and (10.25) is an addition theorem for the special functions $B_{lk}(g)$. Computing the coefficient of z^l in (10.24) we find

(10.26)

$$B_{lk}(g) = \begin{cases} \dfrac{a^l d^{2u-k} c^{k-l} k!}{l!(k-l)!}\, {}_2F_1(-l, -2u+k; k-l+1; bc/ad), & k \geq l \geq 0, \\ \dfrac{a^k d^{2u-l} b^{l-k} \Gamma(2u-k+1)}{\Gamma(2u-l+1)(l-k)!}\, {}_2F_1(-k, -2u+l; l-k+1; bc/ad), & \\ & l \geq k \geq 0, \end{cases}$$

where $\Gamma(z)$ is the gamma function (see the Symbol Index). Such local group reps are of importance in special function theory, as we shall see.

There are exactly three local transformation groups on the line. The corresponding algebras of differential operators are as follows:

(10.27)

(1) $L_1 = d/dz$, one-dimensional,

(2) $L_1 = d/dz$, $L_2 = z\, d/dz$, $[L_1, L_2] = L_1$, two-dimensional

(3) $L_- = d/dz$, $L_3 = z\, d/dz$, $L_+ = z^2\, d/dz$, three-dimensional

$$[L_3, L_\pm] = \pm L_\pm, \quad [L_+, L_-] = 2L_3.$$

Any local group on the (real or complex) line can be transformed to one of these three by an analytic change of variable. Note that the operators (3) define the action of $SL(2)$ as a transformation group on the line. Any one-parameter group can always be expressed in the form (1) by a change of variable. For a proof of this classification see Lie [1] or Campbell [1].

The classification of local multiplier reps on the line is more complicated. Here it is necessary to classify the Lie algebras of generalized Lie derivatives in one complex variable, e.g., (10.8). There are an infinite number of such algebras. They are listed in the work of Miller [1].

Lie [1] has classified all local transformation groups in the plane. Again there are an infinite number of such groups. Some results on multiplier reps in the plane are given by Miller [1].

Problems

5.1 Let B be an $n \times n$ matrix with distinct eigenvalues $\lambda_1, \ldots, \lambda_n$. Show that Ad B acting on the space of all $n \times n$ complex matrices has the n^2 eigenvalues $\lambda_i - \lambda_j$, $1 \leq i, j \leq n$.

5.2 Verify directly that the matrix commutator satisfies the Jacobi identity.

5.3 Verify that the set of all real matrices

$$\begin{pmatrix} 1 & a & b \\ 0 & c & d \\ 0 & 0 & 1 \end{pmatrix}, \quad c > 0,$$

forms a Lie group, and compute the associated Lie algebra. In particular, determine the commutation relations.

5.4 Show that the following spaces of $m \times m$ matrices form Lie algebras: (a) all upper triangular matrices; (b) all upper triangular matrices with trace zero; (c) all upper triangular matrices with diagonal elements zero. Determine the corresponding Lie groups.

5.5 Compute all real Lie algebras of dimensions one and two (identifying isomorphic algebras). For each algebra compute the associated local linear Lie group.

5.6 Repeat Problem 5.5 for all complex Lie algebras of dimension three (see Jacobson [1, Chapter 1]).

5.7 Show that the element

$$\begin{pmatrix} a & 0 \\ 0 & -a^{-1} \end{pmatrix}, \quad a > 0, \quad a \neq 1,$$

of $GL(2, R)$ cannot be expressed as $\exp \mathfrak{A}$ for any $\mathfrak{A} \in gl(2, R)$. Thus the exponential mapping may not cover a Lie group.

5.8 Let $\mu: G \longrightarrow G'$ be an analytic homomorphism of local Lie groups. Show that, in terms of canonical coordinates in G and G', μ is a linear mapping.

5.9 Let G be a linear Lie group with Lie algebra $\mathfrak{G}$. Let $\mathfrak{G}'$ be the subspace of $\mathfrak{G}$ spanned by all elements of the form $[\mathfrak{A}, \mathfrak{B}]$, $\mathfrak{A}, \mathfrak{B} \in \mathfrak{G}$. Show that $\mathfrak{G}'$ is a Lie algebra, the **derived algebra** of $\mathfrak{G}$, and that $\mathfrak{G}' = L(G_c)$, where G_c is the commutator subgroup of G.

5.10 Verify that $L_1 = z\partial_y + y\partial_z$, $L_2 = x\partial_z + z\partial_x$, and $L_3 = y\partial_x - x\partial_y$, x, y, z real, generate the algebra of Lie derivatives of a local Lie transformation group G. Compute the action of G on R_3.

5.11 For $j = 1, 2$ let $\mathbf{T}_j$ be an analytic rep of the Lie group G on the vector space V_j with associated Lie algebra rep ρ_j. Show that the Lie algebra rep corresponding to $\mathbf{T}_1 \otimes \mathbf{T}_2$ takes the form $\boldsymbol{\rho}_1(\boldsymbol{\alpha}) \otimes \mathbf{E}_2 + \mathbf{E}_1 \otimes \boldsymbol{\rho}_2(\boldsymbol{\alpha})$ on $V_1 \otimes V_2$ where $\mathbf{E}_j$ is the identity operator on V_j and $\boldsymbol{\alpha} \in L(G)$.

5.12 Let G be a local transformation group acting on a neighborhood of $\mathbf{x}^0 \in F_m$. A function f analytic near $\mathbf{x}^0$ is an **invariant** of G if $f(\mathbf{x}g) = f(\mathbf{x})$ for each $\mathbf{x}$ sufficiently close to $\mathbf{x}^0$ and $g \in G$. Prove: A function f is an invariant of G if and only if $Lf(\mathbf{x}) = 0$ for all Lie derivatives L of G.

5.13 Compute the invariants of $E^+(2)$, (9.10), and of the group in Problem 5.10.