DIFFERENTIAL GEOMETRY AND SYMMETRIC SPACES

Volume XII

Sigurdur Helgason

Differential
Geometry
and
Symmetric
Spaces

PURE AND APPLIED MATHEMATICS

A Series of Monographs and Textbooks

Edited by

PAUL A. SMITH and SAMUEL EILENBERG

Columbia University, New York

I: ARNOLD SOMMERFELD. Partial Differential Equations in Physics. 1949 (Lectures on Theoretical Physics, Volume VI)

II: REINHOLD BAER. Linear Algebra and Projective Geometry. 1952

III: HERBERT BUSEMANN AND PAUL J. KELLY. Projective Geometry and Projective Metrics. 1953

IV: STEFAN BERGMAN AND M. SCHIFFER. Kernel Functions and Elliptic Differential Equations in Mathematical Physics. 1953

V: RALPH PHILIP BOAS, JR. Entire Functions. 1954

VI: HERBERT BUSEMANN. The Geometry of Geodesics. 1955

VII: CLAUDE CHEVALLEY. Fundamental Concepts of Algebra. 1956

VIII: SZE-TSEN HU. Homotopy Theory. 1959

IX: A. OSTROWSKI. Solution of Equations and Systems of Equations. 1960

X: J. DIEUDONNÉ. Foundations of Modern Analysis. 1960

XI: S. I. GOLDBERG. Curvature and Homology. 1962

XII: SIGURÐUR HELGASON. Differential Geometry and Symmetric Spaces. 1962

In preparation

XIII: T. H. HILDEBRANDT. Introduction to the Theory of Integration.

Differential Geometry

and

Symmetric Spaces

Sigurður Helgason

Department of Mathematics, Massachusetts Institute of Technology, Cambridge, Massachusetts

1962

ACADEMIC PRESS
NEW YORK AND LONDON

ACADEMIC PRESS INC.
111 FIFTH AVENUE
NEW YORK 3, N. Y.

United Kingdom Edition
Published by
ACADEMIC PRESS INC. (LONDON) LTD.
Berkeley Square House, London, W. 1

Library of Congress Catalog Card Number 62-13107

PRINTED IN THE UNITED STATES OF AMERICA

To Artie

This Page Intentionally Left Blank

PREFACE

According to its original definition, a symmetric space is a Riemannian manifold whose curvature tensor is invariant under all parallel translations. The theory of symmetric spaces was initiated by É. Cartan in 1926 and was vigorously developed by him in the late 1920's. By their definition, symmetric spaces form a special topic in Riemannian geometry; their theory, however, has merged with the theory of semisimple Lie groups. This circumstance is the source of very detailed and extensive information about these spaces. They can therefore often serve as examples for the testing of general conjectures. On the other hand, symmetric spaces are numerous enough and their special nature among Riemannian manifolds so clear that a properly formulated extrapolation to general Riemannian manifolds often leads to good questions and conjectures.

The definition above does not immediately suggest the special nature of symmetric spaces (especially if one recalls that all Riemannian manifolds and all Kähler manifolds possess tensor fields invariant under the parallelism). However, the theory leads to the remarkable fact that symmetric spaces are locally just the Riemannian manifolds of the form $\boldsymbol{R}^n \times G/K$ where $\boldsymbol{R}^n$ is a Euclidean n-space, G is a semisimple Lie group which has an involutive automorphism whose fixed point set is the (essentially) compact group K, and G/K is provided with a G-invariant Riemannian structure. É. Cartan's classification of all real simple Lie algebras now led him quickly to an explicit classification of symmetric spaces in terms of the classical and exceptional simple Lie groups. On the other hand, the semisimple Lie group G (or rather the local isomorphism class of G) above is completely arbitrary; in this way valuable geometric tools become available to the theory of semisimple Lie groups. In addition, the theory of symmetric spaces helps to unify and explain in a general way various phenomena in classical geometries. Thus the isomorphisms which occur among the classical groups of low dimensions are geometrically interpreted by means of isometries; the analogy between the spherical geometries and the hyperbolic geometries is a special case of a general duality for symmetric spaces.

On a symmetric space with its well-developed geometry, global function theory becomes particularly interesting. Integration theory, Fourier analysis, and partial differential operators arise here in a canonical fashion by the requirement of geometric invariance. Although these subjects and their relationship are very well developed in Euclidean space (Lebesgue integral, Fourier integral, differential operators with constant coefficients) the extension to general symmetric spaces leads immediately to interesting unsolved problems. The two types of non-Euclidean symmetric spaces, the compact type and the .noncompact type, offer different sorts of function-theoretic problems. The symmetric spaces of the noncompact type present no topological difficulties (the spaces being homeomorphic to Euclidean spaces) and their function theory ties up with the theory of infinite-dimensional representations of arbitrary semisimple Lie groups, which has made great progress in recent years. For the symmetric spaces of the compact type, on the other hand, the classical theory of finite-dimensional representations of compact Lie groups provides a natural framework, but the geometry of the spaces enters now in a less trivial fashion into their function theory.

The objective of the present book is to provide a self-contained introduction to Cartan's theory, as well as to more recent developments in the theory of functions on symmetric spaces.

Chapter I deals with the differential-geometric prerequisites, and the basic geometric properties of symmetric spaces are developed in Chapter IV. From then on the subject is primarily Lie group theory, and in Chapter IX Cartan's classification of symmetric spaces is presented. Although this classification may be considered as the culmination of Cartan's theory, we have confined Chapter IX to proofs of general theorems involved in the classification and to a description of Cartan's list. The justification of this notable omission is first that the usefulness of the classification for experimentation is based on its existence rather than on the proof that it exhausts the class of symmetric spaces; secondly this omission enabled us to include Chapter X (on functions on symmetric spaces) where it is felt that more open questions present themselves. At some places we indicate connections with topics in classical analysis, such as Fourier analysis, theory of special functions (Bessel, Legendre), and integral theorems for invariant differential equations. However, no account is given of the role of symmetric spaces in the theory of automorphic functions and analytic number theory, nor have we found it possible to include more recent topological investigations of symmetric spaces.

Each chapter begins with a short summary and ends with an identification of sources as well as some comments on the historical development.

The purpose of these historical notes is primarily to orient the reader in the vast literature and secondly they are an attempt to give credit where it is due, but here we must apologize in advance for incompleteness as well as possible inaccuracies.†

This book grew out of lectures given at the University of Chicago 1958 and at Columbia University 1959–1960. At Columbia I had the privilege of many long and informative discussions with Professor Harish-Chandra; large parts of Chapters VIII and X are devoted to results of his. I am happy to express here my deep gratitude to him. I am also indebted to Professors A. Korányi, K. deLeeuw, E. Luft and H. Mirkil who read large portions of the manuscript and suggested many improvements. Finally I want to thank my wife who patiently helped with the preparation of the manuscript and did all the typing.

Suggestions to the Reader

Since this book is intended for readers with varied backgrounds we give here some suggestions for its use.

Chapter I, Chapter IV, § 1, and Chapter VIII, § 1 – § 3 can be read independently of the rest of the book. These parts would give the reader an incomplete but short and elementary introduction to modern differenctial geometry, with only advanced calculus and some point-set topology as prerequisites.

Chapter I, § 1 – § 6, Chapter II, and Chapter III can be read independently of the rest of the book as an introduction to semisimple Lie groups. However, Chapters II and III assume some familiarity with the elements of the theory of topological groups.

Chapters I – IX require no further prerequisites. Chapter X, however, makes use of a few facts from Hilbert space theory and assumes some knowledge of measure theory.

Each chapter ends with a few exercises. With a few possible exceptions (indicated with a star) the exercises can be worked out with methods developed in the text. The starred exercises are theorems which might have been included in the text, but were not found necessary for the subsequent chapters.

S. Helgason

† "En hvatki es missagt es í fræðum þessum, þá es skylt at hafa þat heldr, es sannara reynisk"; [Ari Fróði: Íslendingabók (1124)].

This Page Intentionally Left Blank

CONTENTS

CHAPTER III

Structure of Semisimple Lie Algebras

CHAPTER IV

Symmetric Spaces

CHAPTER V

Decomposition of Symmetric Spaces

CHAPTER VI

Symmetric Spaces of the Noncompact Type

CHAPTER VII

Symmetric Spaces of the Compact Type

CHAPTER VIII

Hermitian Symmetric Spaces

CHAPTER IX

On the Classification of Symmetric Spaces

CHAPTER X

Functions on Symmetric Spaces

CHAPTER I

ELEMENTARY DIFFERENTIAL GEOMETRY

This introductory chapter divides in a natural way into three parts: §1-§3 which deal with tensor fields on manifolds, §4-§8 which treat general properties of affine connections, and §9-§14 which give an introduction to Riemannian geometry with some emphasis on topics needed for the later treatment of symmetric spaces.

§1-§3. When a Euclidean space is stripped of its vector space structure and only its differentiable structure retained, there are many ways of piecing together domains of it in a smooth manner, thereby obtaining a so-called differentiable manifold. Local concepts like a differentiable function and a tangent vector can still be given a meaning whereby the manifold can be viewed "tangentially," that is, through its family of tangent spaces as a curve in the plane is, roughly speaking, determined by its family of tangents. This viewpoint leads to the study of tensor fields, which are important tools in local and global differential geometry. They form an algebra $\mathfrak{D}(M)$, the mixed tensor algebra over the manifold M. The alternate covariant tensor fields (the differential forms) form a submodule $\mathfrak{A}(M)$ of $\mathfrak{D}(M)$ which inherits a multiplication from $\mathfrak{D}(M)$, the exterior multiplication. The resulting algebra is called the Grassmann algebra of M. Through the work of É. Cartan the Grassmann algebra with the exterior differentiation d has become an indispensable tool for dealing with submanifolds, these being analytically described by the zeros of differential forms. Moreover, the pair $(\mathfrak{A}(M), d)$ determines the cohomology of M via de Rham's theorem, which however will not be dealt with here.

§4-§8. The concept of an affine connection was first defined by Levi-Civita for Riemannian manifolds, generalizing significantly the notion of parallelism for Euclidean spaces. On a manifold with a countable basis an affine connection always exists (see the exercises following this chapter). Given an affine connection on a manifold M there is to each curve $\gamma(t)$ in M associated an isomorphism between any two tangent spaces $M_{\gamma(t_1)}$ and $M_{\gamma(t_2)}$. Thus, an affine connection makes it possible to relate tangent spaces at distant points of the manifold. If the tangent vectors of the curve $\gamma(t)$ all correspond under these isomorphisms we have the analog of a straight line, the so-called geodesic. The theory of affine connections mainly amounts to a study of the mappings $\mathrm{Exp}_p : M_p \to M$ under which straight lines (or segments of them) through the origin in the tangent space M_p correspond to geodesics through p in M. Each mapping Exp_p is a diffeomorphism of a neighborhood of 0 in M_p into M, giving the so-called normal coordinates at p. Some other local properties of Exp_p are given in §6, the existence of convex neighborhoods and a formula for the differential of Exp_p.

An affine connection gives rise to two important tensor fields, the curvature tensor field and the torsion tensor field which in turn describe the affine connection through É. Cartan's structural equations [(6) and (7), §8)].

§9-§14. A particularly interesting tensor field on a manifold is the so-called Riemannian structure. This gives rise to a metric on the manifold in a canonical fashion. It also determines an affine connection on the manifold, the Riemannian connection; this affine connection has the property that the geodesic forms the shortest curve between any two (not too distant) points. The relation between the metric and geodesics is further developed in §9-§10. The treatment is mainly based on the structural equations of É. Cartan and is independent of the Calculus of Variations.

The higher-dimensional analog of the Gaussian curvature of a surface was discovered by Riemann. Riemann introduced a tensor field which for any pair of tangent vectors at a point measures the corresponding sectional curvature, that is, the Gaussian curvature of the surface generated by the geodesics tangent to the plane spanned by the two vectors. Of particular interest are Riemannian manifolds for which the sectional curvature always has the same sign. The irreducible symmetric spaces are of this type. Riemannian manifolds of negative curvature are considered in §13 owing to their importance in the theory of symmetric spaces. Much progress has been made recently in the study of Riemannian manifolds whose sectional curvature is bounded from below by a constant > 0. However, no discussion of these is given since it is not needed in later chapters. The last section deals with totally geodesic submanifolds which are characterized by the condition that a geodesic tangent to the submanifold at a point lies entirely in it. In contrast to the situation for general Riemannian manifolds, totally geodesic submanifolds are a common occurrence for symmetric spaces.

§1. Manifolds

Let $\boldsymbol{R}^m$ and $\boldsymbol{R}^n$ denote two Euclidean spaces of m and n dimensions, respectively. Let O and O' be open subsets, $O \subset \boldsymbol{R}^m$, $O' \subset \boldsymbol{R}^n$ and suppose φ is a mapping of O into O'. The mapping φ is called *differentiable* if the coordinates $y_j(\varphi(p))$ of $\varphi(p)$ are differentiable (that is, indefinitely differentiable) functions of the coordinates $x_i(p)$, $p \in O$. The mapping φ is called *analytic* if for each point $p \in O$ there exists a neighborhood U of p and n power series P_j $(1 \leqslant j \leqslant n)$ in m variables such that $y_j(\varphi(q)) = P_j(x_1(q) - x_1(p), \ldots, x_m(q) - x_m(p))$ $(1 \leqslant j \leqslant n)$ for $q \in U$. A differentiable mapping $\varphi : O \to O'$ is called *a diffeomorphism* of O onto O' if $\varphi(O) = O'$, φ is one-to-one, and the inverse mapping φ^{-1} is differentiable. In the case when $n = 1$ it is customary to replace the term "mapping" by the term "function."

An analytic function on $\boldsymbol{R}^m$ which vanishes on an open set is identically 0. For differentiable functions the situation is completely different. In fact, if A and B are disjoint subsets of $\boldsymbol{R}^m$, A compact and B closed, then there exists a differentiable function φ which is identically 1 on A and identically 0 on B. The standard procedure for constructing such a function φ is as follows:

Let $0 < a < b$ and consider the function f on $\boldsymbol{R}$ defined by

$$f(x) = \begin{cases} \exp\left(\dfrac{1}{x-b} - \dfrac{1}{x-a}\right) & \text{if } a < x < b, \\ 0 & \text{otherwise.} \end{cases}$$

Then f is differentiable and the same holds for the function

$$F(x) = \int_x^b f(t)\,dt \Big/ \int_a^b f(t)\,dt,$$

which has value 1 for $x \leqslant a$ and 0 for $x \geqslant b$. The function ψ on $\boldsymbol{R}^m$ given by

$$\psi(x_1, \ldots, x_m) = F(x_1^2 + \ldots + x_m^2)$$

is differentiable and has values 1 for $x_1^2 + \ldots + x_m^2 \leqslant a$ and 0 for $x_1^2 + \ldots + x_m^2 \geqslant b$. Let S and S' be two concentric spheres in $\boldsymbol{R}^m$, S' lying inside S. Starting from ψ we can by means of a linear transformation of $\boldsymbol{R}^m$ construct a differentiable function on $\boldsymbol{R}^m$ with value 1 in the interior of S' and value 0 outside S. Turning now to the sets A and B we can, owing to the compactness of A, find finitely many spheres S_i $(1 \leqslant i \leqslant n)$, such that the corresponding open balls B_i $(1 \leqslant i \leqslant n)$, form a covering of A (that is, $A \subset \bigcup_{i=1}^n B_i$) and such that the closed balls $\bar{B}_i$ $(1 \leqslant i \leqslant n)$ do not intersect B. Each sphere S_i can be shrunk to a concentric sphere S_i' such that the corresponding open balls B_i' still form a covering of A. Let ψ_i be a differentiable function on $\boldsymbol{R}^m$ which is identically 1 on B_i' and identically 0 in the complement of B_i. Then the function

$$\varphi = 1 - (1 - \psi_1)(1 - \psi_2) \ldots (1 - \psi_n)$$

is a differentiable function on $\boldsymbol{R}^m$ which is identically 1 on A and identically 0 on B.

Let M be a topological space. We assume that M satisfies the Hausdorf separation axiom which states that any two different points in M can be separated by disjoint open sets. An *open chart* on M is a pair (U, φ) where U is an open subset of M and φ is a homeomorphism of U onto an open subset of $\boldsymbol{R}^m$.

Definition. Let M be a Hausdorf space. A *differentiable structure* on M of dimension m is a collection of open charts $(U_\alpha, \varphi_\alpha)_{\alpha \in A}$ on M where $\varphi_\alpha(U_\alpha)$ is an open subset of $\boldsymbol{R}^m$ such that the following conditions are satisfied:

(M_1) $M = \bigcup_{\alpha \in A} U_\alpha$.

(M_2) For each pair $\alpha, \beta \in A$ the mapping $\varphi_\beta \circ \varphi_\alpha^{-1}$ is a differentiable mapping of $\varphi_\alpha(U_\alpha \cap U_\beta)$ onto $\varphi_\beta(U_\alpha \cap U_\beta)$.

(M_3) The collection $(U_\alpha, \varphi_\alpha)_{\alpha \in A}$ is a maximal family of open charts for which (M_1) and (M_2) hold.

A *differentiable manifold* (or C^∞ *manifold* or simply *manifold*) of dimension m is a Hausdorf space with a differentiable structure of dimension m. If M is a manifold, a *local chart* on M (or a *local coordinate system* on M) is by definition a pair $(U_\alpha, \varphi_\alpha)$ where $\alpha \in A$. If $p \in U_\alpha$ and $\varphi_\alpha(p) = (x_1(p), \ldots, x_m(p))$, the set U_α is called a *coordinate neighborhood* of p and the numbers $x_i(p)$ are called *local coordinates* of p. The mapping $\varphi_\alpha : q \to (x_1(q), \ldots, x_m(q))$, $q \in U_\alpha$, is often denoted $\{x_1, \ldots, x_m\}$.

Remark 1. Condition (M_3) will often be cumbersome to check in specific instances. It is therefore important to note that the condition (M_3) is not essential in the definition of a manifold. In fact, if only (M_1) and (M_2) are satisfied, the family $(U_\alpha, \varphi_\alpha)_{\alpha \in A}$ can be extended in a unique way to a larger family $\mathfrak{M}$ of open charts such that (M_1), (M_2), and (M_3) are all fulfilled. This is easily seen by defining $\mathfrak{M}$ as the set of all open charts (V, φ) on M satisfying: (1) $\varphi(V)$ is an open set in $\boldsymbol{R}^m$; (2) for each $\alpha \in A$, $\varphi_\alpha \circ \varphi^{-1}$ is a diffeomorphism of $\varphi(V \cap U_\alpha)$ onto $\varphi_\alpha(V \cap U_\alpha)$.

Remark 2. If we let $\boldsymbol{R}^m$ mean a single point for $m = 0$, the preceding definition applies. The manifolds of dimension 0 are then the discrete topological spaces.

Remark 3. A manifold is connected if and only if it is pathwise connected. The proof is left to the reader.

An *analytic structure* of dimension m is defined in a similar fashion. In (M_2) we just replace "differentiable" by "analytic." In this case M is called an *analytic manifold.*

In order to define a *complex manifold* of dimension m we replace $\boldsymbol{R}^m$ in the definition of differentiable manifold by the m-dimensional complex space $\boldsymbol{C}^m$. The condition (M_2) is replaced by the condition that the m coordinates of $\varphi_\beta \circ \varphi_\alpha^{-1}(p)$ should be holomorphic functions of the coordinates of p. Here a function $f(z_1, \ldots, z_m)$ of m complex variables is called *holomorphic* if at each point $(z_1^0, \ldots, z_m^0)$ there exists a power series

$$\sum a_{n_1 \ldots n_m} (z_1 - z_1^0)^{n_1} \ldots (z_m - z_m^0)^{n_m},$$

which converges absolutely to $f(z_1, \ldots, z_m)$ in a neighborhood of the point.

The manifolds dealt with in the later chapters of this book (mostly

Lie groups and their coset spaces) are analytic manifolds. From Remark 1 it is clear that we can always regard an analytic manifold as a differentiable manifold. It is often convenient to do so because, as pointed out before for $\boldsymbol{R}^m$, the class of differentiable functions is much richer than the class of analytic functions.

Let f be a real-valued function on a C^∞ manifold M. The function f is called *differentiable* at a point $p \in M$ if there exists a local chart $(U_\alpha, \varphi_\alpha)$ with $p \in U_\alpha$ such that the composite function $f \circ \varphi_\alpha^{-1}$ is a differentiable function on $\varphi_\alpha(U_\alpha)$. The function f is called *differentiable* if it is differentiable at each point $p \in M$. If M is analytic, the function f is said to be *analytic* at $p \in M$ if there exists a local chart $(U_\alpha, \varphi_\alpha)$ with $p \in U_\alpha$ such that $f \circ \varphi_\alpha^{-1}$ is an analytic function on the set $\varphi_\alpha(U_\alpha)$.

Let M be a differentiable manifold of dimension m and let $\mathfrak{F}$ denote the set of all differentiable functions on M. The set $\mathfrak{F}$ has the following properties:

($\mathfrak{F}_1$) Let $\varphi_1, \ldots, \varphi_r \in \mathfrak{F}$ and let u be a differentiable function on $\boldsymbol{R}^r$. Then $u(\varphi_1, \ldots, \varphi_r) \in \mathfrak{F}$.

($\mathfrak{F}_2$) Let f be a real function on M such that for each $p \in M$ there exists a function $g \in \mathfrak{F}$ which coincides with f in some neighborhood of p. Then $f \in \mathfrak{F}$.

($\mathfrak{F}_3$) For each $p \in M$ there exist m functions $\varphi_1, \ldots, \varphi_m \in \mathfrak{F}$ and an open neighborhood U of p such that the mapping $q \rightarrow (\varphi_1(q), \ldots, \varphi_m(q))$ $(q \in U)$ is a homeomorphism of U onto an open subset of $\boldsymbol{R}^m$. The set U and the functions $\varphi_1, \ldots, \varphi_m$ can be chosen in such a way that each $f \in \mathfrak{F}$ coincides on U with a function of the form $u(\varphi_1, \ldots, \varphi_m)$ where u is a differentiable function on $\boldsymbol{R}^m$.

The properties ($\mathfrak{F}_1$) and ($\mathfrak{F}_2$) are obvious. To establish ($\mathfrak{F}_3$) we pick a local chart $(U_\alpha, \varphi_\alpha)$ such that $p \in U_\alpha$ and write $\varphi_\alpha(q) = (x_1(q), \ldots, x_m(q)) \in \boldsymbol{R}^m$ for $q \in U_\alpha$. Let S be a compact neighborhood of $\varphi_\alpha(p)$ in $\boldsymbol{R}^m$ such that S is contained in the open set $\varphi_\alpha(U_\alpha)$. Then as shown earlier, there exists a differentiable function ψ on $\boldsymbol{R}^m$ such that ψ has compact support† contained in $\varphi_\alpha(U_\alpha)$ and such that $\psi(s) = 1$ for all $s \in S$. Let $U = \varphi_\alpha^{-1}(\mathring{S})$ where $\mathring{S}$ is the interior of S and define the function φ_i $(1 \leqslant i \leqslant m)$ on M by

$$\varphi_i(q) = \begin{cases} 0 & \text{if } q \notin U_\alpha, \\ x_i(q)\,\psi(\varphi_\alpha(q)) & \text{if } q \in U_\alpha. \end{cases}$$

Then the set U and the functions $\varphi_1, \ldots, \varphi_m$ have the property stated in ($\mathfrak{F}_3$). In fact, if $f \in \mathfrak{F}$, then the function $f \circ \varphi_\alpha^{-1}$ is differentiable on the set $\varphi_\alpha(U_\alpha)$.

† The *support* of a function is the closure of the set where the function is different from 0.

Proposition 1.1. *Suppose M is a Hausdorf space and m an integer > 0. Assume $\mathfrak{F}$ is a collection of real-valued functions on M with the properties $\mathfrak{F}_1$, $\mathfrak{F}_2$, and $\mathfrak{F}_3$. Then there exists a unique collection $(U_\alpha, \varphi_\alpha)_{\alpha \in A}$ of open charts on M such that (M_1), (M_2), and (M_3) are satisfied and such that the differentiable functions on the resulting manifold are precisely the members of $\mathfrak{F}$.*

For the proof we select for each $p \in M$ the functions $\varphi_1, \ldots, \varphi_m$ and the neighborhood U of p given by $\mathfrak{F}_3$. Putting $U_\alpha = U$ and $\varphi_\alpha(q) = (\varphi_1(q), \ldots, \varphi_m(q))$ $(q \in U)$ we obtain a collection $(U_\alpha, \varphi_\alpha)_{\alpha \in A}$ of open charts on M satisfying (M_1). The condition (M_2) is also satisfied in virtue of $\mathfrak{F}_3$. As remarked earlier, the collection $(U_\alpha, \varphi_\alpha)_{\alpha \in A}$ can then be extended to a collection $(U_\alpha, \varphi_\alpha)_{\alpha \in A^*}$ which satisfies (M_1), (M_2), and (M_3). This induces a differentiable structure on M and each $g \in \mathfrak{F}$ is obviously a differentiable function. On the other hand, suppose that f is a differentiable function on the constructed manifold. If $p \in M$, there exists a local chart $(U_\alpha, \varphi_\alpha)$ where $\alpha \in A^*$ such that $p \in U_\alpha$ and such that $f \circ \varphi_\alpha^{-1}$ is a differentiable function on an open neighborhood of $\varphi_\alpha(p)$. Owing to (M_2) we may assume that $\alpha \in A$. There exists a differentiable function u on $\boldsymbol{R}^m$ such that $f \circ \varphi_\alpha^{-1}(x) = u(x_1, \ldots, x_m)$ for all points $x = (x_1, \ldots, x_m)$ in some open neighborhood of $\varphi_\alpha(p)$. This means (in terms of the φ_i above) that

$$f = u(\varphi_1, \ldots, \varphi_m)$$

in some neighborhood of p. Since $p \in M$ is arbitrary we conclude from $\mathfrak{F}_2$ and $\mathfrak{F}_1$ that $f \in \mathfrak{F}$. Finally, let $(V_\beta, \psi_\beta)_{\beta \in B}$ be another collection of open charts satisfying (M_1), (M_2), and (M_3) and giving rise to the same $\mathfrak{F}$. Writing $f \circ \varphi_\alpha^{-1} = f \circ \psi_\beta^{-1} \circ (\psi_\beta \circ \varphi_\alpha^{-1})$ for $f \in \mathfrak{F}$ we see that $\psi_\beta \circ \varphi_\alpha^{-1}$ is differentiable on $\varphi_\alpha(U_\alpha \cap V_\beta)$, so by the maximality (M_3), $(U_\alpha, \varphi_\alpha) \in (V_\beta, \psi_\beta)_{\beta \in B}$ and the uniqueness follows.

We shall often write $C^\infty(M)$ instead of $\mathfrak{F}$ and will sometimes denote by $C^\infty(p)$ the set of functions on M which are differentiable at p. The set $C^\infty(M)$ is an algebra over $\boldsymbol{R}$, the operations being

$$(\lambda f)(p) = \lambda f(p),$$
$$(f + g)(p) = f(p) + g(p),$$
$$(fg)(p) = f(p)\, g(p)$$

for $\lambda \in \boldsymbol{R}$, $p \in M$, f, $g \in C^\infty(M)$.

Lemma 1.2. *Let C be a compact subset of a manifold M and let V be an open subset of M containing C. Then there exists a function $\psi \in C^\infty(M)$ which is identically 1 on C, identically 0 outside V.*

This lemma has already been established in the case $M = \mathbf{R}^m$. We shall now show that the general case presents no additional difficulties.

Let $(U_\alpha, \varphi_\alpha)$ be a local chart on M and S a compact subset of U_α. There exists a differentiable function f on $\varphi_\alpha(U_\alpha)$ such that f is identically 1 on $\varphi_\alpha(S)$ and has compact support contained in $\varphi_\alpha(U_\alpha)$. The function F on M given by

$$F(q) = \begin{cases} f(\varphi_\alpha(q)) & \text{if } q \in U_\alpha, \\ 0 & \text{otherwise} \end{cases}$$

is a differentiable function on M which is identically 1 on S and identically 0 outside U_α. Since C is compact and V open, there exist finitely many coordinate neighborhoods $U_1, \ldots, U_n$ and compact sets $S_1, \ldots, S_n$ such that

$$C \subset \bigcup_1^n S_i, \qquad S_i \subset U_i$$

$$(\bigcup_1^n U_i) \subset V.$$

As shown previously, there exists a function $F_i \in C^\infty(M)$ which is identically 1 on S_i and identically 0 outside U_i. The function

$$\psi = 1 - (1 - F_1)(1 - F_2) \ldots (1 - F_n)$$

belongs to $C^\infty(M)$, is identically 1 on C and identically 0 outside V.

Let M be a C^∞ manifold and $(U_\alpha, \varphi_\alpha)_{\alpha \in A}$ a collection satisfying (M_1), (M_2), and (M_3). If U is an open subset of M, U can be given a differentiable structure by means of the open charts $(V_\alpha, \psi_\alpha)_{\alpha \in A}$ where $V_\alpha = U \cap U_\alpha$ and ψ_α is the restriction of φ_α to V_α. With this structure, U is called an *open submanifold* of M. In particular, since M is locally connected, each connected component of M is an open submanifold of M.

Let M and N be two manifolds of dimension m and n, respectively. Let $(U_\alpha, \varphi_\alpha)_{\alpha \in A}$ and $(V_\beta, \psi_\beta)_{\beta \in B}$ be collections of open charts on M and N, respectively, such that the conditions (M_1), (M_2), and (M_3) are satisfied. For $\alpha \in A$, $\beta \in B$, let $\varphi_\alpha \times \psi_\beta$ denote the mapping $(p, q) \to (\varphi_\alpha(p), \psi_\beta(q))$ of the product set $U_\alpha \times V_\beta$ into $\mathbf{R}^{m+n}$. Then the collection $(U_\alpha \times V_\beta, \varphi_\alpha \times \psi_\beta)_{\alpha \in A, \beta \in B}$ of open charts on the product space $M \times N$ satisfies (M_1) and (M_2) so by Remark 1, $M \times N$ can be turned into a manifold the *product* of M and N.

An immediate consequence of Lemma 1.2 is the following fact which will often be used: Let V be an open submanifold of M, f a function in $C^\infty(V)$, and p a point in V. Then there exists a function $\tilde{f} \in C^\infty(M)$ and an open neighborhood N, $p \in N \subset V$ such that f and $\tilde{f}$ agree on N.

Definition. Let M be a topological space. A *covering* of M is a collection of open subsets of M whose union if M. A covering $\{U_\alpha\}_{\alpha \in A}$ of M is said to be *locally finite* if each $p \in M$ has a neighborhood which intersects only finitely many of the sets U_α.

Definition. A Hausdorf space M is called *paracompact* if for each covering $\{U_\alpha\}_{\alpha \in A}$ of M there exists a locally finite covering $\{V_\beta\}_{\beta \in B}$ which is a refinement of $\{U_\alpha\}_{\alpha \in A}$ (that is, each V_β is contained in some U_α).

Definition. A topological space is called *normal* if for any two disjoint closed subsets A and B there exist disjoint open subsets U and V such that $A \subset U$, $B \subset V$.

It is known that a locally compact Hausdorf space which has a countable base is paracompact and every paracompact space is normal (see e.g., Kelley [1]; the notion of paracompactness is due to J. Dieudonné).

Theorem 1.3 (partition of unity). *Let M be a normal manifold and $\{U_\alpha\}_{\alpha \in A}$ a locally finite covering of M. Assume that each $\bar{U}_\alpha$ is compact. Then there exists a system $\{\varphi_\alpha\}_{\alpha \in A}$ of differentiable functions on M such that*

(i) *Each φ_α has compact support contained in U_α.*

(ii) $\varphi_\alpha \geqslant 0$, $\Sigma_{\alpha \in A}\, \varphi_\alpha = 1$.

We shall make use of the following fact (see, e.g., Kelley [1], p. 171): Let $\{U_\alpha\}_{\alpha \in A}$ be a locally finite covering of a normal space M. Then each set U_α can be shrunk to a set V_α, such that $\bar{V}_\alpha \subset U_\alpha$ and $\{V_\alpha\}_{\alpha \in A}$ is still a covering of M.

To prove Theorem 1.3 we first shrink the U_α as indicated and thus get a new covering $\{V_\alpha\}_{\alpha \in A}$. Owing to Lemma 1.2 there exists a function $\psi_\alpha \in C^\infty(M)$ of compact support contained in U_α such that ψ_α is identically 1 on V_α and $\psi_\alpha \geqslant 0$ on M. Owing to the local finiteness the sum $\Sigma_{\alpha \in A}\, \psi_\alpha = \psi$ exists. Moreover, $\psi \in C^\infty(M)$ and $\psi(p) > 0$ for each $p \in M$. The functions $\varphi_\alpha = \psi_\alpha/\psi$ have the desired properties (i) and (ii).

The system $\{\varphi_\alpha\}_{\alpha \in A}$ is called a partition of unity *subordinate to the covering* $\{U_\alpha\}_{\alpha \in A}$.

§ 2. Tensor Fields

1. Vector Fields and 1-Forms

Let A be an algebra over a field K. A *derivation* of A is a mapping $D: A \to A$ such that

(i) $D(\alpha f + \beta g) = \alpha Df + \beta Dg$ for $\alpha, \beta \in K$, $f, g \in A$;

(ii) $D(fg) = f(Dg) + (Df)g$ for $f, g \in A$.

Definition. A *vector field* X on a C^∞ manifold is a derivation of the algebra $C^\infty(M)$.

Let $\mathfrak{D}^1$ (or $\mathfrak{D}^1(M)$) denote the set of all vector fields on M. If $f \in C^\infty(M)$ and $X, Y \in \mathfrak{D}^1(M)$, then fX and $X + Y$ denote the vector fields

$$fX : g \to f(Xg), \qquad g \in C^\infty(M),$$
$$X + Y : g \to Xg + Yg, \qquad g \in C^\infty(M).$$

This turns $\mathfrak{D}^1(M)$ into a module over the ring $\mathfrak{F} = C^\infty(M)$. If X, $Y \in \mathfrak{D}^1(M)$, then $XY - YX$ is also a derivation of $C^\infty(M)$ and is denoted by the bracket $[X, Y]$. As is customary we shall often write $\theta(X)\,Y = [X, Y]$. The operator $\theta(X)$ is called the *Lie derivative* with respect to X. The bracket satisfies the *Jacobi identity* $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$ or, otherwise written $\theta(X)\,([Y, Z]) = [\theta(X)\,Y, Z] + [Y, \theta(X)\,Z]$.

It is immediate from (ii) that if f is constant and $X \in \mathfrak{D}^1$, then $Xf = 0$. Suppose now that a function $g \in C^\infty(M)$ vanishes on an open subset $V \subset M$. Let p be an arbitrary point in V. According to Lemma 1.2 there exists a function $f \in C^\infty(M)$ such that $f(p) = 0$, and $f = 1$ outside V. Then $g = fg$ so

$$Xg = f(Xg) + g(Xf),$$

which shows that Xg vanishes at p. Since p was arbitrary, $Xg = 0$ on V. We can now define Xf on V for every function $f \in C^\infty(V)$. If $p \in V$, select $\tilde{f} \in C^\infty(M)$ such that f and $\tilde{f}$ coincide in a neighborhood of p and put $(Xf)\,(p) = (X\tilde{f})\,(p)$. The consideration above shows that this is a valid definition, that is, independent of the choice of $\tilde{f}$. This shows that a vector field on a manifold induces a vector field on any open submanifold.

On the other hand, let Z be a vector field on an open submanifold $V \subset M$ and p a point in V. Then there exists a vector field $\tilde{Z}$ on M and an open neighborhood N, $p \in N \subset V$ such that $\tilde{Z}$ and Z induce the same vector field on N. In fact, let C be any compact neighborhood of p contained in V and let N be the interior of C. Choose $\psi \in C^\infty(M)$ of compact support contained in V such that $\psi = 1$ on C. For any $g \in C^\infty(M)$, let g_V denote its restriction to V and define $\tilde{Z}g$ by

$$\tilde{Z}(g)\,(q) = \begin{cases} \psi(q)\,(Zg_V)\,(q) & \text{for } q \in V, \\ 0 & \text{if } \ q \notin V. \end{cases}$$

Then $g \to \tilde{Z}g$ is the desired vector field on M.

Now, let (U, φ) be a local chart on M, X a vector field on U, and let p be an arbitrary point in U. We put $\varphi(q) = (x_1(q), \ldots, x_m(q))$ $(q \in U)$, and $f^* = f \circ \varphi^{-1}$ for $f \in C^\infty(M)$. Let V be an open subset of U such that $\varphi(V)$ is an open ball in $\boldsymbol{R}^m$ with center $\varphi(p) = (a_1, ..., a_m)$. If $(x_1, ..., x_m) \in \varphi(V)$, we have

$$\begin{aligned} f^*(x_1, ..., x_m) & \\ &= f^*(a_1, ..., a_m) + \int_0^1 \frac{\partial}{\partial t} f^*(a_1 + t(x_1 - a_1), ..., a_m + t(x_m - a_m))\, dt \\ &= f^*(a_1, ..., a_m) + \sum_{j=1}^m (x_j - a_j) \int_0^1 f_j^*(a_1 + t(x_1 - a_1), ..., a_m + t(x_m - a_m))dt. \end{aligned}$$

(Here f_j^* denotes the partial derivative of f^* with respect to the jth argument.) Transferring this relation back to M we obtain

$$f(q) = f(p) + \sum_{i=1}^m (x_i(q) - x_i(p))\, g_i(q) \qquad (q \in V), \tag{1}$$

where $g_i \in C^\infty(V)$, $(1 \leqslant i \leqslant m)$, and

$$g_i(p) = \left(\frac{\partial f^*}{\partial x_i}\right)_{\varphi(p)}.$$

It follows that

$$(Xf)(p) = \sum_{i=1}^m \left(\frac{\partial f^*}{\partial x_i}\right)_{\varphi(p)} (Xx_i)(p) \qquad \text{for } p \in U. \tag{2}$$

The mapping $f \to (\partial f^*/\partial x_i) \circ \varphi$ $(f \in C^\infty(U))$ is a vector field on U and is denoted $\partial/\partial x_i$. We write $\partial f/\partial x_i$ instead of $\partial/\partial x_i(f)$. Now, by (2)

$$X = \sum_{i=1}^m (Xx_i) \frac{\partial}{\partial x_i} \qquad \text{on } U. \tag{3}$$

Thus, $\partial/\partial x_i$ $(1 \leqslant i \leqslant m)$ is a basis of the module $\mathfrak{D}^1(U)$.

For $p \in M$ and $X \in \mathfrak{D}^1$, let X_p denote the linear mapping X_p: $f \to (Xf)(p)$ of $C^\infty(p)$ into $\boldsymbol{R}$. The set $\{X_p : X \in \mathfrak{D}^1(M)\}$ is called the *tangent space* to M at p; it will be denoted by $\mathfrak{D}^1(p)$ or M_p and its elements are called the *tangent vectors* to M at p. Relation (2) shows that M_p is a vector space over $\boldsymbol{R}$ spanned by the m linearly independent vectors

$$e_i : f \to \left(\frac{\partial f^*}{\partial x_i}\right)_{\varphi(p)}, \qquad f \in C^\infty(M).$$

This tangent vector e_i will often be denoted by $(\partial/\partial x_i)_p$. A linear mapping $L : C^\infty(p) \to \boldsymbol{R}$ is a tangent vector to M at p if and only if the condition

$L(fg) = f(p)\,L(g) + g(p)\,L(f)$ is satisfied for all $f, g \in C^\infty(p)$. In fact, the necessity of the condition is obvious and the sufficiency is a simple consequence of (1). Thus, a vector field X on M can be identified with a collection $X_p (p \in M)$ of tangent vectors to M with the property that for each $f \in C^\infty(M)$ the function $p \to X_p f$ is differentiable.

Suppose the manifold M is analytic. The vector field X on M is then called *analytic* at p if Xf is analytic at p whenever f is analytic at p.

Remark. Let V be a finite-dimensional vector space over $\boldsymbol{R}$. If $X_1, \ldots, X_n$ is any basis of V, the mapping $\sum_{i=1}^n x_i X_i \to (x_1, \ldots, x_n)$ is an open chart valid on the entire V. The resulting differentiable structure is independent of the choice of basis. If $X \in V$, the tangent space V_X is identified with V itself by the formula

$$(Yf)(X) = \left\{\frac{d}{dt} f(X + tY)\right\}_{t=0}, \qquad f \in C^\infty(V),$$

which to each $Y \in V$ assigns a tangent vector to V at X.

Let A be a commutative ring with identity element, E a module over A. Let E^* denote the set of all A-linear mappings of E into A. Then E^* is an A-module in an obvious fashion. It is called the *dual* of E.

Definition. Let M be a C^∞ manifold and put $\mathfrak{F} = C^\infty(M)$. Let $\mathfrak{D}_1(M)$ denote the dual of the $\mathfrak{F}$-module $\mathfrak{D}^1(M)$. The elements of $\mathfrak{D}_1(M)$ are called the *differential 1-forms* on M (or just 1-forms on M).

Let $X \in \mathfrak{D}^1(M)$, $\omega \in \mathfrak{D}_1(M)$. Suppose that X vanishes on an open set V. Then the function $\omega(X)$ vanishes on V. In fact, if $p \in V$, there exists a function $f \in C^\infty(M)$ such that $f = 0$ in a compact neighborhood of p and $f = 1$ outside V. Then $fX = X$ and since ω is $\mathfrak{F}$-linear, $\omega(X) = f\omega(X)$. Hence $(\omega(X))(p) = 0$. This shows also that a 1-form on M induces a 1-form on any open submanifold of M. Using (3) we obtain the following lemma.

Lemma 2.1. *Let $X \in \mathfrak{D}^1(M)$ and $\omega \in \mathfrak{D}_1(M)$. If $X_p = 0$ for some $p \in M$, then the function $\omega(X)$ vanishes at p.*

This lemma shows that given $\omega \in \mathfrak{D}_1(M)$, we can define the linear function ω_p on M_p by putting $\omega_p(X_p) = (\omega(X))(p)$ for $X \in \mathfrak{D}^1(M)$. The set $\mathfrak{D}_1(p) = \{\omega_p : \omega \in \mathfrak{D}_1(M)\}$ is a vector space over $\boldsymbol{R}$.

We have seen that a 1-form on M induces a 1-form on any open submanifold. On the other hand, suppose θ is a 1-form on an open submanifold V of M and p a point in V. Then there exists a 1-form $\tilde{\theta}$ on M, and an open neighborhood N of p, $p \in N \subset V$, such that θ and $\tilde{\theta}$ induce the same 1-form on N. In fact, let C be a compact neighborhood of p contained in V and let N be the interior of C. Select $\psi \in C^\infty(M)$ of compact support contained in V such that $\psi = 1$ on C. Then a 1-form $\tilde{\theta}$ with the desired property can be defined by

$$\tilde{\theta}(X) = \psi\theta(X_V) \text{ on } V, \qquad \tilde{\theta}(X) = 0 \text{ outside } V,$$

where $X \in \mathfrak{D}^1(M)$ and X_V denotes the vector field on V induced by X.

Lemma 2.2. *The space $\mathfrak{D}_1(p)$ coincides with M_p^*, the dual of M_p.*

We already know that $\mathfrak{D}_1(p) \subset M_p^*$. Now let $\{x_1, ..., x_m\}$ be a system of coordinates valid on an open neighborhood U of p. Owing to (3), there exist 1-forms ω^i on U such that† $\omega^i(\partial/\partial x_j) = \delta_{ij}$ $(1 \leqslant i, j \leqslant m)$. Let $L \in M_p^*$, $l_i = L((\partial/\partial x_i)_p)$ and $\theta = \sum_{i=1}^m l_i\omega_i$. Then there exists a 1-form $\tilde{\theta}$ on M and an open neighborhood N of p $(N \subset U)$ such that $\tilde{\theta}$ and θ induce the same form on N. Then $(\tilde{\theta})_p = L$ and the lemma is proved.

Each $X \in \mathfrak{D}^1(M)$ induces an $\mathfrak{F}$-linear mapping $\omega \rightarrow \omega(X)$ of $\mathfrak{D}_1(M)$ into $\mathfrak{F}$. If $X_1 \neq X_2$, the induced mappings are different (due to Lemma 2.2). Thus, $\mathfrak{D}^1(M)$ can be regarded as a subset of $(\mathfrak{D}_1(M))^*$.

Lemma 2.3. *The module $\mathfrak{D}^1(M)$ coincides with the dual of the module $\mathfrak{D}_1(M)$.*

Proof. Let $F \in \mathfrak{D}_1(M)^*$. Then $F(f\omega) = fF(\omega)$ for all $f \in C^\infty(M)$ and all $\omega \in \mathfrak{D}_1(M)$. This shows that if ω vanishes on an open set V, $F(\omega)$ also vanishes on V. Let $p \in M$ and $\{x_1, ..., x_m\}$ a system of local coordinates valid on an open neighborhood U of p. Each 1-form on U can be written $\sum_{i=1}^m f_i\omega^i$ where $f_i \in C^\infty(U)$ and ω^i has the same meaning as above. It follows easily that $F(\omega)$ vanishes at p whenever $\omega_p = 0$; consequently, the mapping $\omega_p \rightarrow (F(\omega))(p)$ is a well-defined linear function on $\mathfrak{D}_1(p)$. By Lemma 2.2 there exists a unique vector $X_p \in M_p$ such that $(F(\omega))(p) = \omega_p(X_p)$ for all $\omega \in \mathfrak{D}_1(M)$. Thus, F gives rise to a family X_p $(p \in M)$ of tangent vectors to M. For each $q \in U$ we can write

$$X_q = \sum_{i=1}^m a_i(q)\left(\frac{\partial}{\partial x_i}\right)_q,$$

where $a_i(q) \in \boldsymbol{R}$. For each i $(1 \leqslant i \leqslant m)$ there exists a 1-form $\tilde{\omega}^i$ on M which coincides with ω^i in an open neighborhood N_p of p, $(N_p \subset U)$. Then $(F(\tilde{\omega}^i))(q) = \tilde{\omega}_q^i(X_q) = a_i(q)$ for $q \in N_p$. This shows that the functions a_i are differentiable. If $f \in C^\infty(M)$ and we denote the function $q \rightarrow X_qf$ $(q \in M)$ by Xf, then the mapping $f \rightarrow Xf$ is a vector field on M which satisfies $\omega(X) = F(\omega)$ for all $\omega \in \mathfrak{D}_1(M)$. This proves the lemma.

† As usual, $\delta_{ij} = 0$ if $i \neq j$, $\delta_{ij} = 1$ if $i = j$.

2. The Tensor Algebra

Let A be a commutative ring with identity element. Let I be a set and suppose that for each $i \in I$ there is given an A-module E_i. The product set $\Pi_{i \in I} E_i$ can be turned into an A-module as follows: If $e = \{e_i\}$, $e' = \{e'_i\}$ are any two elements in ΠE_i (where $e_i, e'_i \in E_i$), and $a \in A$, then $e + e'$ and ae are given by

$$(e + e')_i = e_i + e'_i, \quad (ae)_i = ae_i \qquad \text{for } i \in I.$$

The module ΠE_i is called the *direct product* of the modules E_i. The *direct sum* $\Sigma_{i \in I} E_i$ is defined as the submodule of ΠE_i consisting of those elements $e = \{e_i\}$ for which all $e_i = 0$ except for finitely many i.

Suppose now the set I is finite, say $I = \{1, ..., s\}$. A mapping $f: E_1 \times ... \times E_s \to F$ where F is an A-module is said to be *A-multilinear* if it is A-linear in each argument. The set of all A-multilinear mappings of $E_1 \times ... \times E_s$ into F is again an A-module as follows:

$$(f + f')(e_1, ..., e_s) = f(e_1, ..., e_s) + f'(e_1, ..., e_s),$$
$$(af)(e_1, ..., e_s) = a(f(e_1, ..., e_s)).$$

Suppose that all the factors E_i coincide. The A-multilinear mapping f is called *alternate* if $f(X_1, ..., X_s) = 0$ whenever at least two X_i coincide.

Now, let M be a C^∞ manifold and as usual we put $\mathfrak{F} = C^\infty(M)$. If s is an integer, $s \geqslant 1$, we consider the $\mathfrak{F}$-module

$$\mathfrak{D}^1 \times \mathfrak{D}^1 \times ... \times \mathfrak{D}^1 \qquad (s \text{ times})$$

and let $\mathfrak{D}_s$ denote the $\mathfrak{F}$-module of all $\mathfrak{F}$-multilinear mappings of $\mathfrak{D}^1 \times ... \times \mathfrak{D}^1$ into $\mathfrak{F}$. Similarly $\mathfrak{D}^r$ denotes the $\mathfrak{F}$-module of all $\mathfrak{F}$-multilinear mappings of

$$\mathfrak{D}_1 \times \mathfrak{D}_1 \times ... \times \mathfrak{D}_1 \qquad (r \text{ times})$$

into $\mathfrak{F}$. This notation is permissible since we have seen that the modules $\mathfrak{D}^1$ and $\mathfrak{D}_1$ are duals of each other. More generally, let $\mathfrak{D}^r_s$ denote the $\mathfrak{F}$-module of all $\mathfrak{F}$-multilinear mappings of

$$\mathfrak{D}_1 \times ... \times \mathfrak{D}_1 \times \mathfrak{D}^1 \times ... \times \mathfrak{D}^1 \qquad (\mathfrak{D}_1 \ r \text{ times}, \ \mathfrak{D}^1 \ s \text{ times})$$

into $\mathfrak{F}$. We often write $\mathfrak{D}^r_s(M)$ instead of $\mathfrak{D}^r_s$. We have $\mathfrak{D}^r_0 = \mathfrak{D}^r$, $\mathfrak{D}^0_s = \mathfrak{D}_s$ and we put $\mathfrak{D}^0_0 = \mathfrak{F}$.

A *tensor field* T on M of type (r, s) is by definition an element of $\mathfrak{D}^r_s(M)$. This tensor field T is said to be *contravariant* of degree r, *covariant* of degree s. In particular, the tensor fields of type $(0, 0)$, $(1, 0)$, and $(0, 1)$ on M are just the differentiable functions on M, the vector fields on M and the 1-forms on M, respectively.

If p is a point in M, we define $\mathfrak{D}^r_s(p)$ as the set of all $\boldsymbol{R}$-multilinear mappings of

$$M_p^* \times \ldots \times M_p^* \times M_p \times \ldots \times M_p \qquad (M_p^* \quad r \text{ times}, M_p \quad s \text{ times})$$

into $\boldsymbol{R}$. The set $\mathfrak{D}^r_s(p)$ is a vector space over $\boldsymbol{R}$ and is nothing but the tensor product

$$M_p \otimes \ldots \otimes M_p \otimes M_p^* \otimes \ldots \otimes M_p^* \qquad (M_p \quad r \text{ times}, M_p^* \quad s \text{ times})$$

or otherwise written

$$\mathfrak{D}^r_s(p) = \otimes^r M_p \otimes {}^s M_p^*.$$

We also put $\mathfrak{D}^0_0(p) = \boldsymbol{R}$. Consider now an element $T \in \mathfrak{D}^r_s(M)$. We have

$$T(g_1\theta_1, \ldots, g_r\theta_r, f_1Z_1, \ldots, f_sZ_s) = g_1 \ldots g_r f_1 \ldots f_s T(\theta_1, \ldots, \theta_r, Z_1, \ldots, Z_s)$$

for $f_i, g_j \in C^\infty(M)$, $Z_i \in \mathfrak{D}^1(M)$, $\theta_j \in \mathfrak{D}_1(M)$. It follows from Lemma 1.2 that if some θ_j or some Z_i vanishes on an open set V, then the function $T(\theta_1, \ldots, \theta_r, Z_1, \ldots, Z_s)$ vanishes on V. Let $\{x_1, \ldots, x_m\}$ be a system of coordinates valid on an open neighborhood U of p. Then there exist vector fields X_i $(1 \leqslant i \leqslant m)$ and 1-forms ω_j $(1 \leqslant j \leqslant m)$ on M and an open neighborhood N of p, $p \in N \subset U$ such that on N

$$X_i = \frac{\partial}{\partial x_i}, \qquad \omega_j(X_i) = \delta_{ij} \qquad (1 \leqslant i, j \leqslant m).$$

On N, Z_i and θ_j can be written

$$Z_i = \sum_{k=1}^m f_{ik} X_k, \qquad \theta_j = \sum_{l=1}^m g_{jl}\omega_l,$$

where $f_{ik}, g_{jl} \in C^\infty(N)$, and by the remark above we have for $q \in N$,

$$T(\theta_1, \ldots, \theta_r, Z_1, \ldots, Z_s)(q)$$
$$= \sum_{l_j=1, k_i=1}^m g_{1l_1} \ldots g_{rl_r} f_{1k_1} \ldots f_{sk_s} T(\omega_{l_1}, \ldots, \omega_{l_r}, X_{k_1}, \ldots, X_{k_s})(q).$$

This shows that $T(\theta_1, \ldots, \theta_r, Z_1, \ldots, Z_s)(p) = 0$ if some θ_j or some Z_i vanishes at p. We can therefore define an element $T_p \in \mathfrak{D}^r_s(p)$ by the condition

$$T_p((\theta_1)_p, \ldots, (\theta_r)_p, (Z_1)_p, \ldots, (Z_s)_p) = T(\theta_1, \ldots, \theta_r, Z_1, \ldots, Z_s)(p).$$

The tensor field T thus gives rise to a family $T_p, p \in M$, where $T_p \in \mathfrak{D}^r_s(p)$. It is clear that if $T_p = 0$ for all p, then $T = 0$. The element $T_p \in \mathfrak{D}^r_s(p)$ depends differentiably on p in the sense that if N is a coordinate neighborhood of p and T_q (for $q \in N$) is expressed as above in terms of bases for $\mathfrak{D}_1(N)$ and $\mathfrak{D}^1(N)$, then the coefficients are differentiable functions on N. On the other hand, if there is a rule $p \to T(p)$ which to each $p \in M$ assigns a member $T(p)$ of $\mathfrak{D}^r_s(p)$ in a differentiable manner (as described above), then there exists a tensor field T of type (r, s) such that $T_p = T(p)$ for all $p \in M$. In the case when M is analytic it is clear how to define analyticity of a tensor field T, generalizing the notion of an analytic vector field.

The vector spaces $\mathfrak{D}^r_s(p)$ and $\mathfrak{D}^s_r(p)$ are dual to each other under the nondegenerate bilinear form $\langle\ ,\ \rangle$ on $\mathfrak{D}^r_s(p) \times \mathfrak{D}^s_r(p)$ defined by the formula

$$\langle e_1 \otimes \dots \otimes e_r \otimes f_1 \otimes \dots \otimes f_s, e'_1 \otimes \dots \otimes e'_s \otimes f'_1 \otimes \dots \otimes f'_r \rangle = \prod_{i,j} f_j(e'_j) f'_i(e_i),$$

where e_i, e'_j are members of a basis of M_p, f_j, f'_i are members of a dual basis of M_p^*. It is then obvious that the formula holds if e_i, e'_j are arbitrary elements of M_p and f_j, f'_i are arbitrary elements of M_p^*. In particular, the form $\langle\ ,\ \rangle$ is independent of the choice of basis used in the definition.

Each $T \in \mathfrak{D}^r_s(M)$ induces an $\mathfrak{F}$-linear mapping of $\mathfrak{D}^s_r(M)$ into $\mathfrak{F}$ given by the formula

$$(T(S))(p) = \langle T_p, S_p \rangle \qquad \text{for } S \in \mathfrak{D}^s_r(M).$$

If $T(S) = 0$ for all $S \in \mathfrak{D}^s_r(M)$, then $T_p = 0$ for all $p \in M$, so $T = 0$. Consequently, $\mathfrak{D}^r_s(M)$ can be regarded as a subset of $(\mathfrak{D}^s_r(M))^*$. We have now the following generalization of Lemma 2.3.

Lemma 2.3′. *The module $\mathfrak{D}^r_s(M)$ is the dual of $\mathfrak{D}^s_r(M)$ $(r, s \geqslant 0)$.*

Except for a change in notation the proof is the same as that of Lemma 2.3. To emphasize the duality we sometimes write $\langle T, S \rangle$ instead of $T(S)$, $(T \in \mathfrak{D}^r_s,\ S \in \mathfrak{D}^s_r)$.

Let $\mathfrak{D}$ (or $\mathfrak{D}(M)$) denote the direct sum of the $\mathfrak{F}$-modules $\mathfrak{D}^r_s(M)$,

$$\mathfrak{D} = \sum_{r,s=0}^{\infty} \mathfrak{D}^r_s.$$

Similarly, if $p \in M$ we consider the direct sum

$$\mathfrak{D}(p) = \sum_{r,s=0}^{\infty} \mathfrak{D}^r_s(p).$$

The vector space $\mathfrak{D}(p)$ can be turned into an associative algebra over $\boldsymbol{R}$ as follows: Let $a = e_1 \otimes \dots \otimes e_r \otimes f_1 \otimes \dots \otimes f_s$, $b = e_1' \otimes \dots \otimes e_\rho' \otimes f_1' \dots \otimes f_\sigma'$, where e_i, e_i' are members of a basis for M_p, f_j, f_j' are members of a dual basis for M_p^*. Then $a \otimes b$ is defined by the formula

$$a \otimes b = e_1 \otimes \dots \otimes e_r \otimes e_1' \otimes \dots \otimes e_\rho' \otimes f_1 \otimes \dots \otimes f_s \otimes f_1' \otimes \dots \otimes f_\sigma'.$$

We put $a \otimes 1 = a$, $1 \otimes b = b$ and extend the operation $(a, b) \rightarrow a \otimes b$ to a bilinear mapping of $\mathfrak{D}(p) \times \mathfrak{D}(p)$ into $\mathfrak{D}(p)$. Then $\mathfrak{D}(p)$ is an associative algebra over $\boldsymbol{R}$. The formula for $a \otimes b$ now holds for arbitrary elements e_i, $e_i' \in M_p$ and f_j, $f_j' \in M_p^*$. Consequently, the multiplication in $\mathfrak{D}(p)$ is independent of the choice of basis.

The tensor product $\otimes$ in $\mathfrak{D}$ is now defined as the $\mathfrak{F}$-bilinear mapping $(S, T) \rightarrow S \otimes T$ of $\mathfrak{D} \times \mathfrak{D}$ into $\mathfrak{D}$ such that

$$(S \otimes T)_p = S_p \otimes T_p, \qquad S \in \mathfrak{D}_s^r,\ T \in \mathfrak{D}_\sigma^\rho,\ p \in M.$$

This turns the $\mathfrak{F}$-module $\mathfrak{D}$ into a ring satisfying

$$f(S \otimes T) = fS \otimes T = S \otimes fT$$

for $f \in \mathfrak{F}$, S, $T \in \mathfrak{D}$. In other words, $\mathfrak{D}$ is an associative algebra over the ring $\mathfrak{F}$. The algebras $\mathfrak{D}$ and $\mathfrak{D}(p)$ are called the *mixed tensor algebras* over M and M_p, respectively. The submodules

$$\mathfrak{D}^* = \sum_{r=0}^{\infty} \mathfrak{D}^r, \qquad \mathfrak{D}_* = \sum_{s=0}^{\infty} \mathfrak{D}_s$$

are subalgebras of $\mathfrak{D}$ (also denoted $\mathfrak{D}^*(M)$ and $\mathfrak{D}_*(M)$) and the subspaces

$$\mathfrak{D}^*(p) = \sum_{r=0}^{\infty} \mathfrak{D}^r(p), \qquad \mathfrak{D}_*(p) = \sum_{s=0}^{\infty} \mathfrak{D}_s(p)$$

are subalgebras of $\mathfrak{D}(p)$.

Now let r, s be two integers $\geqslant 1$, and let i, j be integers such that $1 \leqslant i \leqslant r$, $1 \leqslant j \leqslant s$. Consider the $\boldsymbol{R}$-linear mapping $C^i{}_j$: $\mathfrak{D}_s^r(p) \rightarrow \mathfrak{D}_{s-1}^{r-1}(p)$ defined by

$$C^i{}_j(e_1 \otimes \dots \otimes e_r \otimes f_1 \otimes \dots \otimes f_s) = \langle e_i, f_j \rangle (e_1 \otimes \dots \hat{e}_i \dots \otimes e_r \otimes f_1 \otimes \dots \hat{f}_j \dots \otimes f_s),$$

where $e_1, \dots, e_r$ are members of a basis of M_p, $f_1, \dots, f_s$ are members of the dual basis of M_p^*. (The symbol $\hat{}$ over a letter means that the letter is missing.) Now that the existence of $C^i{}_j$ is established, we note that the formula for C^i_j holds for arbitrary elements $e_1, \ldots, e_r \in M_p$, $f_1, \ldots, f_s \in M_p^*$. In particular, C^i_j is independent of the choice of basis.

There exists now a unique $\mathfrak{F}$-linear mapping C^i_j: $\mathfrak{D}^r_s(M) \to \mathfrak{D}^{r-1}_{s-1}(M)$ such that

$$(C^i_j(T))_p = C^i_j(T_p)$$

for all $T \in \mathfrak{D}^r_s(M)$ and all $p \in M$. This mapping satisfies the relation

$$\begin{aligned} &C^i_j(X_1 \otimes \ldots \otimes X_r \otimes \omega_1 \otimes \ldots \otimes \omega_s) \\ &\qquad = \langle X_i, \omega_j \rangle \, (X_1 \otimes \ldots \hat{X}_i \ldots \otimes X_r \otimes \omega_1 \otimes \ldots \hat{\omega}_j \ldots \otimes \omega_s) \end{aligned}$$

for all $X_1, \ldots, X_r \in \mathfrak{D}^1$, $\omega_1, \ldots, \omega_s \in \mathfrak{D}_1$. The mapping C^i_j is called the *contraction* of the ith contravariant index and the jth covariant index.

3. The Grassmann Algebra

As before, M denotes a C^∞ manifold and $\mathfrak{F} = C^\infty(M)$. If s is an integer $\geqslant 1$, let $\mathfrak{A}_s$ (or $\mathfrak{A}_s(M)$) denote the set of alternate $\mathfrak{F}$-multilinear mappings of $\mathfrak{D}^1 \times \ldots \times \mathfrak{D}^1$ (s times) into $\mathfrak{F}$. Then $\mathfrak{A}_s$ is a submodule of $\mathfrak{D}_s$. We put $\mathfrak{A}_0 = \mathfrak{F}$ and let $\mathfrak{A}$ (or $\mathfrak{A}(M)$) denote the direct sum $\mathfrak{A} = \sum_{s=0}^{\infty} \mathfrak{A}_s$ of the $\mathfrak{F}$-modules $\mathfrak{A}_s$. The elements of $\mathfrak{A}(M)$ are called *exterior differential forms* on M. The elements of $\mathfrak{A}_s$ are called differential s-forms (or just s-forms).

Let $\mathfrak{S}_s$ denote the group of permutations of the set $\{1, 2, \ldots, s\}$. Each $\sigma \in \mathfrak{S}_s$ induces an $\mathfrak{F}$-linear mapping of $\mathfrak{D}^1 \times \ldots \times \mathfrak{D}^1$ onto itself given by

$$(X_1, \ldots, X_s) \to (X_{\sigma^{-1}(1)}, \ldots, X_{\sigma^{-1}(s)}) \qquad (X_i \in \mathfrak{D}^1).$$

This mapping will also be denoted by σ. Since each $d \in \mathfrak{D}_s$ is a multilinear map of $\mathfrak{D}^1 \times \ldots \times \mathfrak{D}^1$ into $\mathfrak{F}$, the mapping $d \circ \sigma^{-1}$ is well defined. Moreover, the mapping $d \to d \circ \sigma^{-1}$ is a one-to-one $\mathfrak{F}$-linear mapping of $\mathfrak{D}_s$ onto itself. If we write $\sigma \cdot d = d \circ \sigma^{-1}$ we have $\sigma\tau \cdot d = \sigma \cdot (\tau \cdot d)$. Let $\epsilon(\sigma) = 1$ or -1 according to whether σ is an even or an odd permutation. Consider the linear transformation A_s: $\mathfrak{D}_s \to \mathfrak{D}_s$ given by

$$A_s(d_s) = \frac{1}{s!} \sum_{\sigma \in \mathfrak{S}_s} \epsilon(\sigma) \, \sigma \cdot d_s, \qquad d_s \in \mathfrak{D}_s.$$

If $s = 0$, we put $A_s(d_s) = d_s$. We extend A_s to an $\mathfrak{F}$-linear mapping A: $\mathfrak{D}_* \to \mathfrak{D}_*$ by putting $A(d) = \sum_{s=0}^{\infty} A_s(d_s)$ if $d = \sum_{s=0}^{\infty} d_s$, $d_s \in \mathfrak{D}_s$.

If $\tau \in \mathfrak{S}_s$, we have

$$\tau A_s(d_s) = \frac{1}{s!}\sum_{\sigma \in \mathfrak{S}_s} \epsilon(\sigma)\ \tau \cdot (\sigma \cdot d_s) = \frac{1}{s!}\sum_{\sigma \in \mathfrak{S}_s} \epsilon(\sigma)\,(\tau\sigma) \cdot d_s$$

$$= \epsilon(\tau)\,\frac{1}{s!}\sum_{\sigma \in \mathfrak{S}_s} \epsilon(\sigma)\,\sigma \cdot d_s.$$

Hence, $\tau \cdot (A_s(d_s)) = \epsilon(\tau)\,A_s(d_s)$. This shows that $A_s(\mathfrak{D}_s) \subset \mathfrak{A}_s$ and $A(\mathfrak{D}_*) \subset \mathfrak{A}$. On the other hand, if $d_s \in \mathfrak{A}_s$, then $\sigma \cdot d_s = \epsilon(\sigma)\,d_s$ for each $\sigma \in \mathfrak{S}_s$. Since $\epsilon(\sigma)^2 = 1$, we find that

$$A_s(d_s) = d_s \qquad \text{if } d_s \in \mathfrak{A}_s.$$

It follows that $A^2 = A$ and $A(\mathfrak{D}_*) = \mathfrak{A}_*$; in other words, A is a projection of $\mathfrak{D}_*$ onto $\mathfrak{A}_*$. The mapping A is called *alternation.*

Let N denote the kernel of A. Obviously N is a submodule of $\mathfrak{D}_*$.

Lemma 2.4.† *The module N is a two-sided ideal in $\mathfrak{D}_*$.*

It suffices to show that if $n_r \in N \cap \mathfrak{D}_r$, $d_s \in \mathfrak{D}_s$, then $A_{r+s}(n_r \otimes d_s) = A_{s+r}(d_s \otimes n_r) = 0$. Let $b_{r+s} = A_{r+s}(n_r \otimes d_s)$; then

$$(r+s)!\,b_{r+s} = \sum_{\sigma \in \mathfrak{S}_{r+s}} \epsilon(\sigma)\,\sigma \cdot (n_r \otimes d_s),$$

where

$$\sigma \cdot (n_r \otimes d_s)(X_1, \ldots, X_{r+s}) = n_r(X_{\sigma(1)}, \ldots, X_{\sigma(r)})\, d_s(X_{\sigma(r+1)}, \ldots, X_{\sigma(r+s)}).$$

The elements in $\mathfrak{S}_{r+s}$ which leave each number $r+1, \ldots, r+s$ fixed constitute a subgroup G of $\mathfrak{S}_{r+s}$, isomorphic to $\mathfrak{S}_r$. Let S be a subset of $\mathfrak{S}_{r+s}$ containing exactly one element from each left coset $\sigma_0 G$ of $\mathfrak{S}_{r+s}$. Then, since $\epsilon(\sigma_1\sigma_2) = \epsilon(\sigma_1)\,\epsilon(\sigma_2)$,

$$\sum_{\sigma \in \mathfrak{S}_{r+s}} \epsilon(\sigma)\,\sigma \cdot (n_r \otimes d_s) = \sum_{\sigma_0 \in S} \epsilon(\sigma_0) \sum_{\tau \in G} \epsilon(\tau)\,(\sigma_0\tau) \cdot (n_r \otimes d_s).$$

Let $X_i \in \mathfrak{D}^1$ $(1 \leqslant i \leqslant r+s)$, $(Y_1, \ldots, Y_{r+s}) = \sigma_0^{-1}(X_1, \ldots, X_{r+s})$. Then

$$\sum_{\tau \in G} \epsilon(\tau)\,((\sigma_0\tau) \cdot (n_r \otimes d_s))\,(X_1, \ldots, X_{r+s})$$
$$= d_s(Y_{r+1}, \ldots, Y_{r+s}) \sum_{\tau \in \mathfrak{S}_r} \epsilon(\tau)\,(\tau \cdot n_r)\,(Y_1, \ldots, Y_r) = 0.$$

This shows that $b_{r+s} = 0$. Similarly one proves $A_{s+r}(d_s \otimes n_r) = 0$.

† Chevalley [2], p. 142.

For any two $\theta, \omega \in \mathfrak{A}$ we can now define the *exterior product*

$$\theta \wedge \omega = A(\theta \otimes \omega).$$

This turns the $\mathfrak{F}$-module $\mathfrak{A}$ into an associative algebra, isomorphic to $\mathfrak{D}_*/N$. The module $\mathfrak{A}(M)$ of alternate $\mathfrak{F}$-multilinear functions with the exterior multiplication is called the *Grassmann algebra* of the manifold M.

We can also for each $p \in M$ define the Grassmann algebra $\mathfrak{A}(p)$ of the tangent space M_p. The elements of $\mathfrak{A}(p)$ are the alternate, $\boldsymbol{R}$-multilinear, real-valued functions on M_p and the product (also denoted $\wedge$) satisfies

$$\theta_p \wedge \omega_p = (\theta \wedge \omega)_p, \qquad \theta, \omega \in \mathfrak{A}.$$

This turns $\mathfrak{A}(p)$ into an associative algebra containing the dual space M_p^*. If $\theta, \omega \in M_p^*$, we have $\theta \wedge \omega = -\omega \wedge \theta$; as a consequence one derives easily the following rule:

Let $\theta^1, ..., \theta^l \in M_p^*$ and let $\omega^i = \sum_{j=1}^l a_{ij}\theta^j$, $1 \leqslant i, j \leqslant l$, $(a_{ij} \in \boldsymbol{R})$. Then

$$\omega^1 \wedge ... \wedge \omega^l = \det(a_{ij})\, \theta^1 \wedge ... \wedge \theta^l.$$

For convenience we write down the exterior multiplication explicitly. Let $f, g \in C^\infty(M)$, $\theta \in \mathfrak{A}_r$, $\omega \in \mathfrak{A}_s$, $X_i \in \mathfrak{D}^1$. Then

$$f \wedge g = fg,$$

$$(f \wedge \theta)(X_1, ..., X_r) = f\,\theta(X_1, ..., X_r),$$

$$(\omega \wedge g)(X_1, ..., X_s) = g\,\omega(X_1, ..., X_s),$$

$$(\theta \wedge \omega)(X_1, ..., X_{r+s}) \tag{4}$$

$$= \frac{1}{(r+s)!} \sum_{\sigma \in \mathfrak{S}_{r+s}} \epsilon(\sigma)\, \theta(X_{\sigma(1)}, ..., X_{\sigma(r)})\, \omega(X_{\sigma(r+1)}, ..., X_{\sigma(r+s)}).$$

We also have the relation

$$\theta \wedge \omega = (-1)^{rs} \omega \wedge \theta. \tag{5}$$

4. Exterior Differentiation

Let M be a C^∞ manifold, $\mathfrak{A}(M)$ the Grassmann algebra over M. The operator d, the *exterior differentiation*, is described in the following theorem.

Theorem 2.5. *There exists a unique* $\boldsymbol{R}$*-linear mapping* $d: \mathfrak{A}(M) \to \mathfrak{A}(M)$ *with the following properties*:

(i) $d\mathfrak{A}_s \subset \mathfrak{A}_{s+1}$ *for each* $s \geqslant 0$.

(ii) *If* $f \in \mathfrak{A}_0$ $(= C^\infty(M))$, *then* df *is the* 1*-form given by* $df(X) = Xf$, $X \in \mathfrak{D}^1(M)$.

(iii) $d \circ d = 0$.

(iv) $d(\omega_1 \wedge \omega_2) = d\omega_1 \wedge \omega_2 + (-1)^r \omega_1 \wedge d\omega_2$ *if* $\omega_1 \in \mathfrak{A}_r$, $\omega_2 \in \mathfrak{A}(M)$.

Proof. Assuming the existence of d for M as well as for open submanifolds of M, we first prove a formula for d ((9) below) which then has the uniqueness as a corollary. Let $p \in M$ and $\{x_1, ..., x_m\}$ a coordinate system valid on an open neighborhood U of p. Let V be an open subset of U such that $\bar{V}$ is compact and $p \in V$, $\bar{V} \subset U$. From (ii) we see that the forms dx_i $(1 \leqslant i \leqslant m)$ on U satisfy $dx_i(\partial/\partial x_j) = \delta_{ij}$ on U. Hence dx_i $(1 \leqslant i \leqslant m)$ is a basis of the $C^\infty(U)$-module $\mathfrak{D}_1(U)$; thus each element in $\mathfrak{D}_*(U)$ can be expressed in the form

$$\sum F_{i_1 \ldots i_r} dx_{i_1} \otimes \ldots \otimes dx_{i_r}, \qquad F_{i_1 \ldots i_r} \in C^\infty(U).$$

It follows that if $\theta \in \mathfrak{A}(M)$ and if θ_U denotes the form induced by θ on U, then θ_U can be written

$$\theta_U = \sum f_{i_1 \ldots i_r} dx_{i_1} \wedge \ldots \wedge dx_{i_r}, \qquad f_{i_1 \ldots i_r} \in C^\infty(U). \tag{6}$$

This is called an *expression* of θ_U on U. We shall prove the formula

$$d(\theta_V) = (d\theta)_V.$$

Owing to Lemma 1.2 there exist functions $\psi_{i_1 \ldots i_r} \in C^\infty(M)$, $\varphi_i \in C^\infty(M)$ $(1 \leqslant i \leqslant m)$ such that

$$\psi_{i_1 \ldots i_r} = f_{i_1 \ldots i_r}, \qquad \varphi_1 = x_1, ..., \varphi_m = x_m \text{ on } V.$$

We consider the form

$$\omega = \sum \psi_{i_1 \ldots i_r} d\varphi_{i_1} \wedge \ldots \wedge d\varphi_{i_r}$$

on M. We have obviously $\omega_V = \theta_V$. Moreover, since $d(f(\theta - \omega)) = df \wedge (\theta - \omega) + fd(\theta - \omega)$ for each $f \in C^\infty(M)$, we can, choosing f identically 0 outside V, identically 1 on an open subset of V, deduce that $(d\theta)_V = (d\omega)_V$.

Since

$$d\omega = \sum d\psi_{i_1 \ldots i_r} \wedge d\varphi_{i_1} \wedge \ldots \wedge d\varphi_{i_r}$$

owing to (iii) and (iv), and since $d(f_V) = (df)_V$ for each $f \in C^\infty(M)$, we conclude that

$$(d\omega)_V = \sum df_{i_1 \ldots i_r} \wedge dx_{i_1} \wedge \ldots \wedge dx_{i_r}. \tag{7}$$

This proves the relation

$$(d\theta)_V = d(\theta_V) = \sum df_{i_1 \ldots i_r} \wedge dx_{i_1} \wedge \ldots \wedge dx_{i_r}. \tag{8}$$

On M itself we have the formula

$$(p+1)\, d\omega(X_1, \ldots, X_{p+1}) = \sum_{i=1}^{p+1} (-1)^{i+1} X_i \cdot \omega(X_1, \ldots, \hat{X}_i, \ldots, X_{p+1})$$
$$+ \sum_{i<j} (-1)^{i+j} \omega([X_i, X_j], X_1, \ldots, \hat{X}_i, \ldots, \hat{X}_j, \ldots, X_{p+1}) \tag{9}$$

for $\omega \in \mathfrak{A}_p(M)$ $(p \geqslant 1)$, $X_i \in \mathfrak{D}^1(M)$. In fact, it suffices to prove it in a coordinate neighborhood of each point; in that case it is a simple consequence of (8). The uniqueness of d is now obvious.

On the other hand, to prove the existence of d, we *define* d by (9) and (ii). Using the relation $[X, fY] = f[X, Y] + (Xf)\, Y$, $(f \in \mathfrak{F};\ X, Y \in \mathfrak{D}^1)$, it follows quickly that the right-hand side of (9) is $\mathfrak{F}$-linear in each variable X_i and vanishes whenever two variables coincide. Hence $d\omega \in \mathfrak{A}_{p+1}$ if $\omega \in \mathfrak{A}_p$. If $X \in \mathfrak{D}^1$, let X_V denote the vector field induced on V. Then $[X, Y]_V = [X_V, Y_V]$ and therefore the relation $(d\theta)_V = d(\theta_V)$ follows from (9). Next we observe that (8) follows from (9) and (ii). Also

$$d(fg) = f dg + g df \tag{10}$$

as a consequence of (ii). To show that (iii) and (iv) hold, it suffices to show that they hold in a coordinate neighborhood of each point of M. But on V, (iv) is a simple consequence of (10) and (8). Moreover, (8) and (ii) imply $d(dx_i) = 0$; consequently (using (iv)),

$$d(df) = d\Big(\sum_j \frac{\partial f}{\partial x_j}\, dx_j\Big) = \sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}\, dx_i \wedge dx_j = 0$$

for each $f \in C^\infty(U)$. The relation (iii) now follows from (8) and (iv).

§ 3. Mappings

1. The Interpretation of the Jacobian

Let M and N be C^∞ manifolds and Φ a mapping of M into N. Let $p \in M$. The mapping Φ is called *differentiable at* p if $g \circ \Phi \in C^\infty(p)$ for each $g \in C^\infty(\Phi(p))$. The mapping Φ is called *differentiable* if it is differentiable at each $p \in M$. Similarly *analytic* mappings are defined. Let $\psi : q \to (x_1(q), ..., x_m(q))$ be a system of coordinates on a neighborhood U of $p \in M$ and $\psi' : r \to (y_1(r), ..., y_n(r))$ a system of coordinates on a neighborhood U' of $\Phi(p)$ in N. Assume $\Phi(U) \subset U'$. The mapping $\psi' \circ \Phi \circ \psi^{-1}$ of $\psi(U)$ into $\psi'(U')$ is given by a system of n functions

$$y_j = \varphi_j(x_1, ..., x_m) \qquad (1 \leqslant j \leqslant n), \tag{1}$$

which we call the *expression* of Φ in coordinates. The mapping Φ is differentiable at p if and only if the functions φ_i have partial derivatives of all orders in some fixed neighborhood of $(x_1(p), ..., x_m(p))$.

The mapping Φ is called a *diffeomorphism* of M onto N if Φ is a one-to-one differentiable mapping of M onto N and Φ^{-1} is differentiable. If in addition M, N, Φ, and Φ^{-1} are analytic, Φ is called an *analytic diffeomorphism.*

If Φ is differentiable at $p \in M$ and $A \in M_p$, then the linear mapping $B : C^\infty(\Phi(p)) \to \boldsymbol{R}$ given by $B(g) = A(g \circ \Phi)$ for $g \in C^\infty(\Phi(p))$ is a tangent vector to N at $\Phi(p)$. The mapping $A \to B$ of M_p into $N_{\Phi(p)}$ is denoted $d\Phi_p$ (or just Φ_p) and is called the *differential* of Φ at p. We have seen that the vectors

$$e_i : f \to \left(\frac{\partial f^*}{\partial x_i}\right)_{\psi(p)} \qquad (1 \leqslant i \leqslant m), \qquad f^* = f \circ \psi^{-1},$$

$$\bar{e}_j : g \to \left(\frac{\partial g^*}{\partial y_j}\right)_{\psi'(\Phi(p))} \qquad (1 \leqslant j \leqslant n), \qquad g^* = g \circ (\psi')^{-1}$$

form a basis of M_p and $N_{\Phi(p)}$, respectively. Then

$$d\Phi_p(e_i)\, g = e_i(g \circ \Phi) = \left(\frac{\partial (g \circ \Phi)^*}{\partial x_i}\right)_{\psi(p)}.$$

But

$$(g \circ \Phi)^*(x_1, ..., x_m) = g^*(y_1, ..., y_n),$$

where $y_j = \varphi_j(x_1, ..., x_m)$ $(1 \leqslant j \leqslant n)$. Hence

$$d\Phi_p(e_i) = \sum_{j=1}^{n} \left(\frac{\partial \varphi_j}{\partial x_i}\right)_{\psi(p)} \bar{e}_j. \tag{2}$$

This shows that if we use the bases e_i $(1 \leqslant i \leqslant m)$, $\bar{e}_j$ $(1 \leqslant j \leqslant n)$ to express the linear transformation $d\Phi_p$ in matrix form, then the matrix we obtain is just the Jacobian of the system (1). From a standard theorem on the Jacobian (the inverse function theorem), we can conclude:

Proposition 3.1. *If $d\Phi_p$ is an isomorphism of M_p onto $N_{\Phi(p)}$, then there exist open submanifolds $U \subset M$ and $V \subset N$ such that $p \in U$ and Φ is a diffeomorphism of U onto V.*

Remark. If $N = \boldsymbol{R}$, $N_{\Phi(p)}$ is identified with $\boldsymbol{R}$ (Remark, §2) and thus $d\Phi_p$ becomes a linear function on M_p. This is the same linear function as we obtain by considering $d\Phi$ as a differential form on M. In fact, if $X \in M_p$, the tangent vector $d\Phi_p(X)$ and the tangent vector

$$f \to \left\{\frac{d}{dt} f(\Phi(p) + tX\Phi)\right\}_{t=0}, \qquad f \in C^\infty(\boldsymbol{R}),$$

both assign to f the number $f'(\Phi(p))\,(X\Phi)$.

Definition.

Let M and N be differentiable (or analytic) manifolds.

(a) A mapping $\Phi : M \to N$ is called *regular* at $p \in M$ if Φ is differentiable (analytic) at $p \in M$ and $d\Phi_p$ is a one-to-one mapping of M_p into $N_{\Phi(p)}$.

(b) M is called a *submanifold* of N if (1) $M \subset N$ (set theoretically); (2) the identity mapping I of M into N is regular at each point of M.

For example, the sphere $x_1^2 + x_2^2 + x_3^2 = 1$ is a submanifold of $\boldsymbol{R}^3$ and a topological subspace as well. However, a submanifold M of a manifold N is not necessarily a topological subspace of N. For example, let N be a torus and let M be a curve on N without double points, dense in N (Chapter II, §2). Proposition 3.1 shows that a submanifold M of a manifold N is an open submanifold of N if and only if $\dim M = \dim N$.

Proposition 3.2. *Let M be a submanifold of a manifold N and let $p \in M$. Then there exists a coordinate system $\{x_1, ..., x_n\}$ valid on an open neighborhood V of p in N such that $x_1(p) = ... = x_n(p) = 0$ and such that the set*

$$U = \{q \in V : x_j(q) = 0 \text{ for } m + 1 \leqslant j \leqslant n\}$$

together with the restrictions of $(x_1, ..., x_m)$ to U form a local chart on M containing p.

Proof. Let $\{y_1, ..., y_m\}$ and $\{z_1, ..., z_n\}$ be coordinate systems valid on open neighborhoods of p in M and N, respectively, such that $y_i(p) = z_j(p) = 0$, $(1 \leqslant i \leqslant m,\ 1 \leqslant j \leqslant n)$. The expression of the identity mapping $I\colon M \to N$ is (near p) given by a system of functions $z_j = \varphi_j(y_1, \ldots, y_m)$, $1 \leqslant j \leqslant n$. The Jacobian matrix $(\partial\varphi_j/\partial y_i)$ of this system has rank m at p since I is regular at p. Without loss of generality we may assume that the square matrix $(\partial\varphi_j/\partial y_i)_{1\leqslant i,j\leqslant m}$ has determinant $\neq 0$ at p. In a neighborhood of $(0, \ldots, 0)$ we have therefore $y_i = \psi_i(z_1, \ldots, z_m)$, $1 \leqslant i \leqslant m$, where each ψ_i is a differentiable function. If we now put

$$x_i = z_i, \qquad 1 \leqslant i \leqslant m,$$

$$x_j = z_j - \varphi_j(\psi_1(z_1, \ldots, z_m), \ldots, \psi_m(z_1, \ldots, z_m)), \qquad m+1 \leqslant j \leqslant n,$$

it is clear that

$$\det\left(\frac{\partial x_i}{\partial y_l}\right)_{1\leqslant i,\ l\leqslant m} \neq 0, \qquad \det\left(\frac{\partial x_j}{\partial z_k}\right)_{1\leqslant j,\ k\leqslant n} \neq 0.$$

Therefore $\{x_1, \ldots, x_n\}$ gives the desired coordinate system.

2. Transformation of Vector Fields

Let M and N be C^∞ manifolds and Φ a differentiable mapping of M into N. Let X and Y be vector fields on M and N, respectively; X and Y are called *Φ-related* if

$$d\Phi_p(X_p) = Y_{\Phi(p)} \qquad \text{for all } p \in M. \tag{3}$$

It is easy to see that (3) is equivalent to

$$(Yf)\circ\Phi = X(f\circ\Phi) \qquad \text{for all } f \in C^\infty(N). \tag{4}$$

It is convenient to write $d\Phi \cdot X = Y$ or $X^\Phi = Y$ instead of (3).

Proposition 3.3.

(i) *Suppose* $d\Phi \cdot X_i = Y_i$ $(i = 1, 2)$. *Then*

$$d\Phi \cdot [X_1, X_2] = [Y_1, Y_2].$$

(ii) *Suppose Φ is a diffeomorphism of M onto itself and put $f^\Phi = f\circ\Phi^{-1}$ for $f \in C^\infty(M)$. Then if $X \in \mathfrak{D}^1(M)$,*

$$(fX)^\Phi = f^\Phi X^\Phi, \qquad (Xf)^\Phi = X^\Phi f^\Phi.$$

Proof. From (4) we have $(Y_1(Y_2f))\circ\Phi = X_1(Y_2f\circ\Phi) = X_1(X_2(f\circ\Phi))$, so (i) follows. The last relation in (ii) is also an immediate consequence of (4). As to the first one, we have for $g \in C^\infty(M)$

$$((fX)^\Phi g)\circ\Phi = (fX)(g\circ\Phi) = f((X^\Phi g)\circ\Phi),$$

so

$$(fX)^\Phi g = f^\Phi(X^\Phi g).$$

Remark. Since $X^\Phi f = (Xf^{\Phi^{-1}})^\Phi$ it is natural to make the following definition. Let Φ be a diffeomorphism of M onto M and A a mapping of $C^\infty(M)$ into itself. The mapping A^Φ is defined by $A^\Phi f = (Af^{\Phi^{-1}})^\Phi$ for $f \in C^\infty(M)$. We also write $[Af](p)$ for the value of the function Af at $p \in M$. If Φ and Ψ are two diffeomorphisms of M, then $f^{\Phi\Psi} = (f^\Psi)^\Phi$ and $A^{\Phi\Psi} = (A^\Psi)^\Phi$.

3. Effect on Differential Forms

Let M and N be C^∞ manifolds and $\Phi: M \to N$ a differentiable mapping. Let ω be an r-form on N. Then there is a unique r-form $\Phi^*\omega$ on M which satisfies

$$\Phi^*\omega(X_1, \ldots, X_r) = \omega(Y_1, \ldots, Y_r) \circ \Phi$$

whenever the vector fields X_i and Y_i $(1 \leqslant i \leqslant r)$ are Φ-related. It suffices to put

$$(\Phi^*\omega)_p(A_1, \ldots, A_r) = \omega_{\Phi(p)}(d\Phi_p(A_1), \ldots, d\Phi_p(A_r))$$

for each $p \in M$, and $A_i \in M_p$. If $f \in C^\infty(N)$, we put $\Phi^*f = f \circ \Phi$ and by linearity $\Phi^*\theta$ is defined for each $\theta \in \mathfrak{A}(M)$. Then the following formulas hold:

$$\Phi^*(\omega_1 \wedge \omega_2) = \Phi^*(\omega_1) \wedge \Phi^*(\omega_2), \qquad \omega_1, \omega_2 \in \mathfrak{A}(M); \tag{5}$$

$$d(\Phi^*\omega) = \Phi^*(d\omega). \tag{6}$$

In fact, (5) follows from (4), § 2, and (6) follows from (9), § 2. In the same way we can define Φ^*T for an arbitrary covariant tensor field $T \in \mathfrak{D}_*(M)$. If $M = N$ and Φ is a diffeomorphism of M onto itself such that $\Phi^*T = T$, we say that T is *invariant* under Φ.

The computation of $\Phi^*\omega$ in coordinates is very simple. Suppose U and V are open sets in M and N, respectively, where the coordinate systems

$$\xi: q \to (x_1(q), \ldots, x_m(q)), \qquad \eta: r \to (y_1(r), \ldots, y_n(r))$$

are valid. Assume $\Phi(U) \subset V$. On U, Φ has a coordinate expression

$$y_j = \varphi_j(x_1, \ldots, x_m) \qquad (1 \leqslant j \leqslant n).$$

If $\omega \in \mathfrak{A}(N)$, the form ω_V has an expression

$$\omega_V = \sum g_{j_1 \ldots j_s}\, dy_{j_1} \wedge \ldots \wedge dy_{j_s} \tag{7}$$

where $g_{j_1 \ldots j_s} \in C^\infty(V)$. The form $\Phi^*\omega$ induces the form $(\Phi^*\omega)_U$ on U, which has an expression

$$(\Phi^*\omega)_U = \sum f_{i_1 \ldots i_r}\, dx_{i_1} \wedge \ldots \wedge dx_{i_r}.$$

This expression is obtained just by substituting

$$y_j = \varphi_j(x_1, \ldots, x_m), \qquad dy_j = \sum_{i=1}^{m} \frac{\partial \varphi_j}{\partial x_i} dx_i \qquad (1 \leqslant j \leqslant n)$$

into (7). This follows from (5) if we observe that

$$\Phi^*(dy_j) = \sum_{i=1}^{m} \left(\frac{\partial \varphi_j}{\partial x_i} \circ \xi \right) dx_i.$$

§ 4. Affine Connections

Definition. An *affine connection* on a manifold M is a rule ∇ which assigns to each $X \in \mathfrak{D}^1(M)$ a linear mapping ∇_X of the vector space $\mathfrak{D}^1(M)$ into itself satisfying the following two conditions:

$$(\nabla_1) \qquad \nabla_{fX+gY} = f\nabla_X + g\nabla_Y;$$

$$(\nabla_2) \qquad \nabla_X(fY) = f\nabla_X(Y) + (Xf)\, Y$$

for $f, g \in C^\infty(M)$, $X, Y \in \mathfrak{D}^1(M)$. The operator ∇_X is called *covariant differentiation* with respect to X.

Lemma 4.1. *Suppose M has the affine connection $X \to \nabla_X$ and let U be an open submanifold of M. Let $X, Y \in \mathfrak{D}^1(M)$. If X or Y vanishes identically on U, then so does $\nabla_X(Y)$.*

Proof. Suppose Y vanishes on U. Let $p \in U$ and $g \in C^\infty(M)$. To prove that $(\nabla_X(Y)g)(p) = 0$, we select $f \in C^\infty(M)$ such that $f(p) = 0$ and $f = 1$ outside U (Lemma 1.2). Then $fY = Y$ and

$$\nabla_X(Y)g = \nabla_X(fY)g = (Xf)(Yg) + f(\nabla_X(Y)g)$$

which vanishes at p. The statement about X follows similarly.

An affine connection ∇ on M induces an affine connection ∇_U on an arbitrary open submanifold U of M. In fact, let X, Y be two vector fields on U. For each $p \in U$ there exist vector fields X', Y' on M which agree with X and Y in an open neighborhood V of p. We then put $((\nabla_U)_X(Y))_q = (\nabla_{X'}(Y'))_q$ for $q \in V$. By Lemma 4.1, the right-hand side of this equation is independent of the choice of X', Y'. It follows immediately that the rule ∇_U: $X \to (\nabla_U)_X$, $(X \in \mathfrak{D}^1(U))$ is an affine connection on U.

In particular, suppose U is a coordinate neighborhood where a coordinate system $\varphi : q \to (x_1(q), ..., x_m(q))$ is valid. For simplicity, we write ∇_i instead of $(\nabla_U)_{\partial/\partial x_i}$. We define the functions $\Gamma_{ij}{}^k$ on U by

$$\nabla_i \left(\frac{\partial}{\partial x_j}\right) = \sum_k \Gamma_{ij}{}^k \frac{\partial}{\partial x_k}. \tag{1}$$

For simplicity of notation we write also $\Gamma_{ij}{}^k$ for the function $\Gamma_{ij}{}^k \circ \varphi^{-1}$. If $\{y_1, ..., y_m\}$ is another coordinate system valid on U, we get another set of functions $\Gamma'_{\alpha\beta}{}^\gamma$ by

$$\nabla_\alpha \left(\frac{\partial}{\partial y_\beta}\right) = \sum_\gamma \Gamma'_{\alpha\beta}{}^\gamma \frac{\partial}{\partial y_\gamma}.$$

Using the axioms ∇_1 and ∇_2 we find easily

$$\Gamma'_{\alpha\beta}{}^\gamma = \sum_{i,j,k} \frac{\partial x_i}{\partial y_\alpha} \frac{\partial x_j}{\partial y_\beta} \frac{\partial y_\gamma}{\partial x_k} \Gamma_{ij}{}^k + \sum_j \frac{\partial^2 x_j}{\partial y_\alpha \partial y_\beta} \frac{\partial y_\gamma}{\partial x_j}. \tag{2}$$

On the other hand, suppose there is given a covering of a manifold M by open coordinate neighborhoods U and in each neighborhood a system of functions $\Gamma_{ij}{}^k$ such that (2) holds whenever two of these neighborhoods overlap. Then we can define ∇_i by (1) and thus we get an affine connection ∇_U in each coordinate neighborhood U. We finally define an affine connection $\tilde{\nabla}$ on M as follows: Let $X, Y \in \mathfrak{D}^1(M)$ and $p \in M$. If U is a coordinate neighborhood containing p, let

$$(\tilde{\nabla}_X(Y))_p = ((\nabla_U)_{X_1}(Y_1))_p$$

if X_1 and Y_1 are the vector fields on U induced by X and Y, respectively. Then $\tilde{\nabla}$ is an affine connection on M which on each coordinate neighborhood U induces the connection ∇_U.

Lemma 4.2. *Let $X, Y \in \mathfrak{D}^1(M)$. If X vanishes at a point p in M, then so does $\nabla_X(Y)$.*

Let $\{x_1, ..., x_m\}$ be a coordinate system valid on an open neighborhood U of p. On the set U we have $X = \Sigma_i f_i(\partial/\partial x_i)$ where $f_i \in C^\infty(U)$ and $f_i(p) = 0$, $(1 \leqslant i \leqslant m)$. Using Lemma 4.1 we find $(\nabla_X(Y))_p = \Sigma_i f_i(p)(\nabla_i(Y))_p = 0$.

Definition. Suppose ∇ is an affine connection on M and that Φ is a diffeomorphism of M. A new affine connection ∇' can be defined on M by

$$\nabla'_X(Y) = (\nabla_{X^\Phi}(Y^\Phi))^{\Phi^{-1}}, \qquad X, Y \in \mathfrak{D}^1(M).$$

That ∇' is indeed an affine connection on M is best seen from Prop. 3.3.

The affine connection ∇ is called *invariant* under Φ if $\nabla' = \nabla$. In this case Φ is called an *affine transformation* of M. Similarly one can define an affine transformation of one manifold onto another.

§ 5. Parallelism

Let M be a C^∞ manifold. A *curve* in M is a regular mapping of an open interval $I \subset \boldsymbol{R}$ into M. The restriction of a curve to a closed subinterval is called a *curve segment*. The curve segment is called finite if the interval is finite.

Let $\gamma : t \rightarrow \gamma(t)$ $(t \in I)$ be a curve in M. Differentiation with respect to the parameter will often be denoted by a dot $(\cdot)$. In particular, $\dot{\gamma}(t)$ stands for the tangent vector $d\gamma(d/dt)_t$. Suppose now that to each $t \in I$ is associated a vector $Y(t) \in M_{\gamma(t)}$. Assuming $Y(t)$ to vary differentiably with t, we shall now define what it means for the family $Y(t)$ to be parallel with respect to γ. Let J be a compact subinterval of I such that the finite curve segment $\gamma_J : t \rightarrow \gamma(t)$ $(t \in J)$ has no double points and such that $\gamma(J)$ is contained in a coordinate neighborhood U. Owing to the regularity of γ each point of I is contained in such an interval J with nonempty interior. Let $\{x_1, ..., x_m\}$ be a coordinate system on U.

Lemma 5.1. *Let $g(t)$ be a differentiable function on an open interval containing J. Then there exists a function $G \in C^\infty(M)$ such that*

$$G(\gamma(t)) = g(t) \qquad (t \in J).$$

Proof. Fix $t_0 \in J$. There exists an index i such that the mapping $t \rightarrow x_i(\gamma(t))$ has nonzero derivative when $t = t_0$. Thus there exists a function η_i of one variable, differentiable in a neighborhood of $x_i(\gamma(t_0))$, such that $t = \eta_i(x_i(\gamma(t)))$ for all t in an interval around t_0. The function $q \rightarrow g(\eta_i(x_i(q)))$ is defined and differentiable for all $q \in U$ sufficiently near $\gamma(t_0)$. Select $G^* \in C^\infty(U)$ such that $G^*(q) = g(\eta_i(x_i(q)))$ for all q in some neighborhood of $\gamma(t_0)$. Then

$$G^*(\gamma(t)) = g(t)$$

for all t in some interval around t_0. Owing to the compactness of J there exist finitely many relatively compact open subsets $U_1, ..., U_n$ of U covering $\gamma(J)$ and functions $G_i \in C^\infty(U)$ such that $G_i(\gamma(t)) = g(t)$ if $\gamma(t) \in U_i$ $(1 \leqslant i \leqslant n)$. Since U has a countable base it is paracompact and the sequence $U_1, ..., U_n$ can be completed to a locally finite covering $\{U_\alpha\}_{\alpha \in A}$ of U. We may assume that each U_α is relatively compact and that $U_\alpha \cap \gamma(J) = \emptyset$ if α is none of the numbers $1, ..., n$. Let $\{\varphi_\alpha\}_{\alpha \in A}$

be a partition of unity subordinate to this covering. Then the function $G' = \sum_{i=1}^n G_i\varphi_i$ belongs to $C^\infty(U)$ and $G'(\gamma(t)) = g(t)$ for each $t \in J$. Finally, let ψ be a function in $C^\infty(M)$ of compact support contained in U such that $\psi = 1$ on $\gamma(J)$. The function G given by $G(q) = \psi(q)\, G'(q)$ if $q \in U$ and $G(q) = 0$ if $q \notin U$ then has the required properties.

We put $X(t) = \dot{\gamma}(t)$ $(t \in I)$. Using Lemma 5.1 it is easy to see that there exist vector fields $X, Y \in \mathfrak{D}^1(M)$ such that ($Y(t)$ being as before)

$$X_{\gamma(t)} = X(t), \qquad Y_{\gamma(t)} = Y(t) \qquad (t \in J).$$

The family $Y(t)$ $(t \in J)$ is said to be *parallel* with respect to γ_J(or parallel along γ_J) if

$$\nabla_X(Y)_{\gamma(t)} = 0 \qquad \text{for all } t \in J. \tag{1}$$

To show that this definition is independent of the choice of X and Y, we express (1) in the coordinates $\{x_1, ..., x_m\}$. There exist functions X^i, Y^j $(1 \leqslant i, j \leqslant m)$ on U such that

$$X = \sum_i X^i \frac{\partial}{\partial x_i}, \qquad Y = \sum_j Y^j \frac{\partial}{\partial x_j} \qquad \text{on } U.$$

For simplicity we put $x_i(t) = x_i(\gamma(t))$, $X^i(t) = X^i(\gamma(t))$, and $Y^i(t) = Y^i(\gamma(t))$, $(t \in J)$, $(1 \leqslant i \leqslant m)$. Then $X^i(t) = \dot{x}_i(t)$ and since

$$\nabla_X(Y) = \sum_k \left(\sum_i X^i \frac{\partial Y^k}{\partial x_i} + \sum_{i,j} X^i Y^j \Gamma_{ij}{}^k\right) \frac{\partial}{\partial x_k} \qquad \text{on } U$$

we obtain

$$\frac{dY^k}{dt} + \sum_{i,j} \Gamma_{ij}{}^k \frac{dx_i}{dt} Y^j = 0 \qquad (t \in J). \tag{2}$$

This equation involves X and Y only through their values *on* the curve. Consequently, condition (1) for parallelism is independent of the choice of X and Y. It is now obvious how to define parallelism with respect to any finite curve segment γ_J and finally with respect to the entire curve γ.

Definition. Let $\gamma : t \to \gamma(t)$ $(t \in I)$ be a curve in M. The curve γ is called a *geodesic* if the family of tangent vectors $\dot{\gamma}(t)$ is parallel with respect to γ. A geodesic γ is called maximal if it is not a proper restriction of any geodesic.

Suppose γ_J is a finite geodesic segment without double points contained in a coordinate neighborhood U where the coordinates $\{x_1, ..., x_m\}$ are valid. Then (2) implies

$$\frac{d^2x_k}{dt^2} + \sum_{i,j} \Gamma_{ij}{}^k \frac{dx_i}{dt} \frac{dx_j}{dt} = 0 \qquad (t \in J). \tag{3}$$

If we change the parameter on the geodesic and put $t = f(s)$, ($f'(s) \neq 0$), then we get a new curve $s \to \gamma_J(f(s))$. This curve is a geodesic if and only if f is a linear function, as (3) shows.

Proposition 5.2. *Let p, q be two points in M and γ a curve segment from p to q. The parallelism τ with respect to γ induces an isomorphism of M_p onto M_q.*

Proof. Without loss of generality we may assume that γ has no double points and lies in a coordinate neighborhood U. Let $\{x_1, ..., x_m\}$ be a system of coordinates on U. Suppose the curve segment γ is given by the mapping $t \to \gamma(t)$ $(a \leqslant t \leqslant b)$ such that $\gamma(a) = p$, $\gamma(b) = q$. As before we put $x_i(t) = x_i(\gamma(t))$, $(a \leqslant t \leqslant b)$, $(1 \leqslant i \leqslant m)$.

Consider the system (2). From the theory of systems of ordinary, linear differential equations of first order we can conclude:

There exist m functions $\varphi_i(t, y_1, ..., y_m)$, $(1 \leqslant i \leqslant m)$ defined and differentiable† for $a \leqslant t \leqslant b$, $-\infty < y_i < \infty$ such that

(i) For each m-tuple $(y_1, ..., y_m)$, the functions $Y^i(t) = \varphi_i(t, y_1, ..., y_m)$ satisfy the system (2).

(ii) $\varphi_i(a, y_1, ..., y_m) = y_i \qquad (1 \leqslant i \leqslant m).$

The functions φ_i are uniquely determined by these properties.

The properties (i) and (ii) show that the family of vectors $Y(t) = \sum_i Y^i(t) (\partial/\partial x_i)$ $(a \leqslant t \leqslant b)$ is parallel with respect to γ and that $Y(a) = \sum_i y_i(\partial/\partial x_i)_p$. The mapping $Y(a) \to Y(b)$ is a linear mapping of M_p into M_q since the functions φ_i are linear in the variables $y_1, ..., y_m$. This mapping is one-to-one owing to the uniqueness of the functions φ_i. Consequently, it is an isomorphism.

Proposition 5.3. *Let M be a differentiable manifold with an affine connection. Let p be any point in M and let $X \neq 0$ in M_p. Then there exists a unique maximal geodesic $t \to \gamma(t)$ in M such that*

$$\gamma(0) = p, \qquad \dot{\gamma}(0) = X. \tag{4}$$

† A function on a closed interval I is called differentiable on I if it is extendable to a differentiable function on some open interval containing I.

Proof. Let $\varphi : q \to (x_1(q), \ldots, x_m(q))$ be a system of coordinates on a neighborhood U of p such that $\varphi(U)$ is a cube $\{(x_1, \ldots, x_m) : |x_i| < c\}$ and $\varphi(p) = 0$. Then X can be written $X = \sum_i \alpha_i(\partial/\partial x_i)_p$ where $\alpha_i \in \boldsymbol{R}$. We consider the system of differential equations

$$\frac{dx_i}{dt} = z_i \qquad (1 \leqslant i \leqslant m), \tag{5}$$

$$\frac{dz_k}{dt} = -\sum_{i,j=1}^{m} \Gamma_{ij}{}^{k}(x_1, \ldots, x_m)\, z_i z_j \qquad (1 \leqslant k \leqslant m), \tag{5'}$$

with the initial conditions

$$(x_1, \ldots, x_m, z_1, \ldots, z_m)_{t=0} = (0, \ldots, 0, \alpha_1, \ldots, \alpha_m).$$

Let c_1, K satisfy $0 < c_1 < c$, $0 < K < \infty$. In the interval $|x_i| < c_1$, $|z_i| < K$ $(1 \leqslant i \leqslant m)$, the right-hand sides of the foregoing equations satisfy a Lipschitz condition.

From the existence and uniqueness theorem (see, e.g., Miller and Murray [1], p. 42) for a system of ordinary differential equations we conclude:

There exists a constant $b_1 > 0$ and differentiable functions $x_i(t)$, $z_i(t)$ $(1 \leqslant i \leqslant m)$ in the interval $|t| \leqslant b_1$ such that

(i) $$\frac{dx_i(t)}{dt} = z_i(t) \qquad (1 \leqslant i \leqslant m), \qquad |t| < b_1,$$

$$\frac{dz_k(t)}{dt} = -\sum_{i,j=1}^{m} \Gamma_{ij}{}^{k}(x_1(t), \ldots, x_m(t))\, z_i(t)\, z_j(t) \qquad (1 \leqslant k \leqslant m), \qquad |t| < b_1;$$

(ii) $(x_1(t), \ldots, x_m(t), z_1(t), \ldots, z_m(t))_{t=0} = (0, \ldots, 0, \alpha_1, \ldots, \alpha_m)$;

(iii) $|x_i(t)| < c_1$, $|z_i(t)| < K \qquad (1 \leqslant i \leqslant m), \qquad |t| < b_1$;

(iv) $x_i(t)$, $z_i(t) \qquad (1 \leqslant i \leqslant m)$ is the only set of functions satisfying the conditions (i), (ii), and (iii).

This shows that there exists a geodesic $t \to \gamma(t)$ in M satisfying (4) and that two such geodesics coincide in some interval around $t = 0$. Moreover, we can conclude from (iv) that if two geodesics $t \to \gamma_1(t)$ $(t \in I_1)$, $t \to \gamma_2(t)$ $(t \in I_2)$ coincide in some open interval, then they coincide for all $t \in I_1 \cap I_2$. Proposition 5.3 now follows immediately.

Definition. The geodesic with the properties in Prop. 5.3 will be denoted γ_X. If $X = 0$, we put $\gamma_X(t) = p$ for all $t \in \boldsymbol{R}$.

§ 6. The Exponential Mapping

Suppose again M is a C^∞ manifold with an affine connection. Let $p \in M$. We use the notation from the proof of Prop. 5.3. We shall now study the solutions of (5) and (5′) and their dependence on the initial values. From the existence and uniqueness theorem (see, e.g., Miller and Murray [1], p. 64) for the system (5), (5′), we can conclude:

There exists a constant b $(0 < b < c)$ and differentiable functions $\varphi_i(t, \xi_1, ..., \xi_m, \zeta_1, ..., \zeta_m)$ for $|t| \leqslant 2b$, $|\xi_i| \leqslant b$, $|\zeta_j| \leqslant b\,(1 \leqslant i, j \leqslant m)$ such that:

(i) For each fixed set $(\xi_1, ..., \xi_m, \zeta_1, ..., \zeta_m)$ the functions

$$x_i(t) = \varphi_i(t, \xi_1, ..., \xi_m, \zeta_1, ..., \zeta_m)$$

$$z_i(t) = \left[\frac{\partial \varphi_i}{\partial t}\right](t, \xi_1, ..., \xi_m, \zeta_1, ..., \zeta_m), \qquad 1 \leqslant i \leqslant m, \qquad |t| \leqslant 2b,$$

satisfy (5) and (5′) and $|x_i(t)| < c_1$, $|z_i(t)| < K$.

(ii) $(x_1(t), ..., x_m(t), z_1(t), ..., z_m(t))_{t=0} = (\xi_1, ..., \xi_m, \zeta_1, ..., \zeta_m)$.

(iii) The functions φ_i are uniquely determined by the above properties.

Theorem 6.1. *Let M be a manifold with an affine connection. Let p be any point in M. Then there exists an open neighborhood N_0 of 0 in M_p and an open neighborhood N_p of p in M such that the mapping $X \to \gamma_X(1)$ is a diffeomorphism of N_0 onto N_p.*

Proof. Using the notation above, we put

$$\psi_i(t, \zeta_1, ..., \zeta_m) = \varphi_i(t, 0, ..., 0, \zeta_1, ..., \zeta_m)$$

for $1 \leqslant i \leqslant m$, $|t| \leqslant 2b$, $|\zeta_i| \leqslant b$. Then

$$\psi_i(0, \zeta_1, ..., \zeta_m) = 0,$$

$$\left[\frac{\partial \psi_i}{\partial t}\right](0, \zeta_1, ..., \zeta_m) = \zeta_i.$$

Since $\gamma_X(st) = \gamma_{sX}(t)$, the uniqueness (iii) implies

$$\psi_i(st, \zeta_1, ..., \zeta_m) = \psi_i(t, s\zeta_1, ..., s\zeta_m) \tag{1}$$

for $|s| \leqslant 1$, $|t| \leqslant 2b$, $|\zeta_i| \leqslant b$. Now let D_i denote partial derivative with respect to the ith argument; from (1) we get by differentiating with respect to s,

$$t[D_1\psi_i](st, \zeta_1, ..., \zeta_m) = \sum_{k=1}^{m} \zeta_k[D_{k+1}\psi_i](t, s\zeta_1, ..., s\zeta_m),$$

$$t^2[D_1^2\psi_i](st, \zeta_1, ..., \zeta_m) = \sum_{j,k=1}^{m} \zeta_j\zeta_k[D_{j+1}D_{k+1}\psi_i](t, s\zeta_1, ..., s\zeta_m).$$

From Taylor's formula we find

$$\psi_i(b, \zeta_1, ..., \zeta_m) = \zeta_i b + \tfrac{1}{2}[D_1^2\psi_i](b^*, \zeta_1, ..., \zeta_m)\, b^2 \qquad (0 \leqslant b^* \leqslant b),$$

so

$$\psi_i(b, \zeta_1, ..., \zeta_m) = \zeta_i b + \tfrac{1}{2}\sum_{j,k=1}^{m} \zeta_j\zeta_k[D_{j+1}D_{k+1}\psi_i]\left(b, \frac{b^*}{b}\zeta_1, ..., \frac{b^*}{b}\zeta_m\right).$$

This shows that the mapping

$$\Psi : (\zeta_1, ..., \zeta_m) \to (\psi_1(b, \zeta_1, ..., \zeta_m), ..., \psi_m(b, \zeta_1, ..., \zeta_m))$$

has Jacobian at the origin equal to b^m. The mapping Ψ is just the mapping $X \to \gamma_X(b)$ expressed in coordinates (§ 3, No. 1). Since $\gamma_{bX}(1) = \gamma_X(b)$, the theorem follows.

Definition. The mapping $X \to \gamma_X(1)$ described in Theorem 6.1 is called the *Exponential mapping* at p and will be denoted by Exp (or Exp_p).

Definition. Let M be a manifold with an affine connection and p a point in M. An open neighborhood N_0 of the origin in M_p is said to be *normal* if: (1) the mapping Exp is a diffeomorphism of N_0 onto an open neighborhood N_p of p in M; (2) if $X \in N_0$, and $0 \leqslant t \leqslant 1$, then $tX \in N_0$.

The last condition means that N_0 is *"star-shaped."* A neighborhood N_p of p in M is called a *normal neighborhood* of p if $N_p = \mathrm{Exp}\, N_0$ where N_0 is a normal neighborhood of 0 in M_p. Assuming this to be the case, and letting $X_1, ..., X_m$ denote some basis of M_p, the inverse mapping†

$$\mathrm{Exp}_p\,(a_1X_1 + ... + a_mX_m) \to (a_1, ..., a_m)$$

of N_p into $\boldsymbol{R}^m$ is called a system of *normal coordinates* at p.

We shall now prove a useful refinement of Theorem 6.1.

† Here and sometimes in the sequel we allow ourselves to denote the inverse of a one-to-one mapping $X \to \phi(X)$ by $\phi(X) \to X$.

Theorem 6.2. *Let M be a C^∞ manifold with an affine connection. Then each point $p \in M$ has a normal neighborhood N_p which is a normal neighborhood of each of its points. (In particular, two arbitrary points in N_p can be joined by exactly one*† *geodesic segment contained in N_p.)*

We shall use the notation from the proof of Prop. 5.3 and consider again the functions $\varphi_i(t, \xi_1, ..., \xi_m, \zeta_1, ..., \zeta_m)$ above. If $q \in U$ and $0 < \delta \leqslant c - \max_i | x_i(q) |$, then the subset of U given by

$$V_\delta(q) = \left\{ r \in U : \sum_{i=1}^{m} (x_i(r) - x_i(q))^2 < \delta^2 \right\}$$

will be called a *spherical neighborhood* of q with radius δ.

Now consider an m-tuple $(\xi_1, ..., \xi_m)$ where $| \xi_i | < b$, $1 \leqslant i \leqslant m$. Let $q \in U$ be determined by $x_i(q) = \xi_i$ $(1 \leqslant i \leqslant m)$. By the proof above, the mapping

$$\Phi : (\zeta_1, ..., \zeta_m) \to (\varphi_1(b, \xi_1, ..., \xi_m, \zeta_1, ..., \zeta_m), ..., \varphi_m(b, \xi_1, ..., \xi_m, \zeta_1, ..., \zeta_m))$$

has Jacobian b^m at the origin $(\zeta_1, ..., \zeta_m) = (0, ..., 0)$. Hence $\varphi^{-1} \circ \Phi$ is a diffeomorphism of a neighborhood $\zeta_1^2 + ... + \zeta_m^2 < r^2$ $(r \leqslant b)$ of the origin in $\boldsymbol{R}^m$ onto an open neighborhood N_q of q in M. We can suppose r taken as large as possible with this property. For reasons of continuity there exists a $\delta_0 > 0$ such that, if $\xi_1^2 + ... + \xi_m^2 < \delta_0^2$, then the corresponding N_q all have a spherical neighborhood of p in common. By taking δ_0 small enough we may assume that this spherical neighborhood is $V_{4\delta_0}(p)$. Since N_q is normal, this proves:

Lemma 6.3. *There exists a number $\delta_0 > 0$ such that for each $q \in V_{\delta_0}(p)$, the spherical neighborhood $V_{2\delta_0}(q)$ is contained in a normal neighborhood of q.*

The neighborhood $V_{\delta_0}(p)$ therefore has the following property: For each pair of points $q_1, q_2 \in V_{\delta_0}(p)$ there exists *at most* one geodesic segment contained in $V_{\delta_0}(p)$, joining q_1 and q_2. A neighborhood with this property will be called *simple.* It is obvious that if $0 < \delta \leqslant \delta_0$, then $V_\delta(p)$ is also simple. We shall now show that for all sufficiently small δ, $V_\delta(p)$ is also *convex*, that is, two arbitrary points in $V_\delta(p)$ can be joined by a geodesic segment contained in $V_\delta(p)$.

Let δ^* be a number satisfying the following two conditions:

(i) $0 < \delta^* < \delta_0$;

(ii) The matrix $(\delta_{ij} - \sum_k x_k \Gamma_{ij}{}^k)$ is strictly positive definite for $\sum_k x_k^2 \leqslant (\delta^*)^2$. (Here $\delta_{ij} = 1$ if $i = j$, 0 otherwise.)

† Except for a linear change of parameter on the geodesic.

It is obvious that such a number δ^* does exist. Theorem 6.2 is contained in the following lemma.

Lemma 6.4. *If* $0 < \delta \leqslant \delta^*$, *the neighborhood* $V_\delta(p)$ *is a normal neighborhood of each of its points. In particular,* $V_\delta(p)$ *is simple and convex.*

Proof. The boundary D of $V_\delta(p)$ is a submanifold of $V_{\delta_0}(p)$. We first prove that if a geodesic $\gamma : t \to \gamma(t)$ is tangent to D at a point $q_0 = \gamma(t_0)$, then for all $t \neq t_0$ sufficiently close to t_0, the point $\gamma(t)$ lies outside D. As before, we put $x_k(t) = x_k(\gamma(t))$ $(1 \leqslant k \leqslant m)$. Then the functions $x_k(t)$ satisfy (3), § 5, in a neighborhood of t_0. In Taylor's formula

$$F(t_0 + \Delta t) = F(t_0) + \Delta t\dot{F}(t_0) + \tfrac{1}{2}(\Delta t)^2 \ddot{F}(t_0) + 0(\Delta t)^3$$

for the function $F(t) = \Sigma_{k=1}^m (x_k(t))^2 - \delta^2$, we have

$$\dot{F}(t_0) = 2\sum_{k=1}^m x_k(t_0)\,\dot{x}_k(t_0) = 0, \qquad F(t_0) = 0,$$

$$\ddot{F}(t) = 2\sum_{k=1}^m (\dot{x}_k(t)\,\dot{x}_k(t) + x_k(t)\,\ddot{x}_k(t))$$

$$= 2\sum_{i,j=1}^m \Big(\delta_{ij} - \sum_k x_k(t)\,\Gamma_{ij}{}^k\Big)\,\dot{x}_i(t)\,\dot{x}_j(t).$$

Using (ii), it follows that $F(t_0 + \Delta t) > 0$ provided Δt is sufficiently small and $\neq 0$. This proves the statement concerning γ.

For a pair $P, Q \in V_\delta(p)$ we have therefore only two possibilities:

1) There is no geodesic segment inside $V_\delta(p)$ which joins P and Q. In this case, the unique geodesic segment inside $V_{2\delta_0}(P)$ which joins P and Q will contain points outside the boundary D.

2) There exists a geodesic segment inside $V_\delta(p)$ which joins P and Q. In this case P and Q are said to be *mutually visible* inside $V_\delta(p)$.

Let S denote the subset of $V_\delta(p) \times V_\delta(p)$ consisting of all point-pairs which are mutually visible inside $V_\delta(p)$. The set S is nonempty and we shall now show that S is open and closed in the relative topology of $V_\delta(p) \times V_\delta(p)$. In view of the connectedness of $V_\delta(p)$, this will prove Lemma 6.4.

I. *S is closed.* Let (p_n, q_n) be a sequence in S which converges to $(p^*, q^*) \in V_\delta(p) \times V_\delta(p)$. We join p_n and q_n by a geodesic segment $t \to \gamma_n(t)$ in $V_\delta(p)$ such that $\gamma_n(0) = p_n$, $\gamma_n(b) = q_n$. Similarly, p^* and q^* are joined by a geodesic segment $\gamma^*(t)$ $(0 \leqslant t \leqslant b)$ inside $V_{2\delta_0}(p^*)$. Consider the mapping Φ above for $(\xi_1^{(n)}, \ldots, \xi_m^{(n)}) = (x_1(p_n), \ldots, x_m(p_n))$.

Under this mapping the point q_n corresponds to a certain m-tuple $(\zeta_1^{(n)}, ..., \zeta_m^{(n)})$. Since these m-tuples are bounded ($|\zeta_i| \leqslant b$), we can, passing to a subsequence if necessary, assume that $(\zeta_1^{(n)}, ..., \zeta_m^{(n)})$ converges to a limit $(\zeta_1^*, ..., \zeta_m^*)$ as $n \to \infty$. Then for $1 \leqslant i \leqslant m$ and $0 \leqslant t \leqslant b$ the sequence

$$\varphi_i(t, \xi_1^{(n)}, ..., \xi_m^{(n)}, \zeta_1^{(n)}, ..., \zeta_m^{(n)})$$

converges to

$$\varphi_i(t, x_1(p^*), ..., x_m(p^*), \zeta_1^*, ..., \zeta_m^*)$$

which represents a geodesic inside $V_{2\delta_0}(p^*)$ joining p^* to q^*. Owing to the uniqueness, it follows that

$$x_i(\gamma^*(t)) = \varphi_i(t, x_1(p^*), ..., x_m(p^*), \zeta_1^*, ..., \zeta_m^*)$$

for $0 \leqslant t \leqslant b$, $1 \leqslant i \leqslant m$; in other words $\gamma_n(t) \to \gamma^*(t)$ for $0 \leqslant t \leqslant b$. Since $\gamma_n(t) \in V_\delta(p)$ $(0 \leqslant t \leqslant b)$ it follows that γ^* contains no points outside the boundary D. Owing to 1) above, we have $(p^*, q^*) \in S$; hence S is closed.

II. *S is open.* In fact, the same argument as in I shows that the complement $(V_\delta(p) \times V_\delta(p)) - S$ is closed.

Definition. Let M be a manifold with an affine connection. Let p be a point in M and N_p a normal neighborhood of p. Let $X \in M_p$ and for each $q \in N_p$ put $(X^*)_q = \tau_{pq}X$ where τ_{pq} is the parallel translation along the unique geodesic segment in N_p which joins p and q. It is clear from (2), § 5, that $(X^*)_q$ depends differentiably on q. The vector field X^* on N_p, thus defined, is said to be *adapted* to the tangent vector X. As before, let $\theta(X^*)$ denote the Lie derivative with respect to X^*.

Definition. An affine connection ∇ on an analytic manifold M is called *analytic* if for each $p \in M$, $\nabla_X(Y)$ is analytic at p whenever the vector fields X and Y are analytic at p.

Theorem 6.5. *Let M be an analytic manifold with an analytic affine connection ∇. Let $p \in M$ and $X \neq 0$ in M_p. Then there exists an $\epsilon > 0$ such that the differential of Exp ($=$ Exp$_p$) is given by*

$$(d\,\mathrm{Exp})_{tX}(Y) = \left\{\frac{1 - e^{\theta(-tX^*)}}{\theta(tX^*)}(Y^*)\right\}_{\mathrm{Exp}\,tX}, \qquad Y \in M_p,$$

for $|t| < \epsilon$.

Here $(1 - e^{-A})/A$ stands for $\sum_0^\infty (-A)^m/(m+1)!$ and as usual (Remark, § 2, No. 1) M_p is identified with its tangent space at each point.

Proof. The mapping Exp is analytic at the origin in M_p. Let f be an anlytic function at p. Then there exists a star-shaped neighborhood U_0 of 0 in M_p such that

$$f(\operatorname{Exp} Z) = P(z_1, ..., z_m), \qquad Z \in U_0,$$

where P is an absolutely convergent power series and $z_1, ..., z_m$ are the coordinates of Z with respect to some basis of M_p. It follows that for fixed $Z \in U_0$

$$f(\operatorname{Exp} tZ) = P(tz_1, ..., tz_m) = \sum_0^\infty \frac{1}{n!} a_n t^n \qquad (a_n \in \boldsymbol{R})$$

for $0 \leqslant t \leqslant 1$. If t is sufficiently small

$$[Z^*f](\operatorname{Exp} tZ) = \left\{\frac{d}{du} f(\operatorname{Exp}(t+u)Z)\right\}_{u=0} = \frac{d}{dt} f(\operatorname{Exp} tZ)$$

and by induction

$$[(Z^*)^n f](\operatorname{Exp} tZ) = \frac{d^n}{dt^n} f(\operatorname{Exp} tZ).$$

On putting $t = 0$ we find that $[(Z^*)^n f](p) = a_n$; hence

$$f(\operatorname{Exp} Z) = \sum_0^\infty \frac{1}{n!} [(Z^*)^n f](p) \qquad (Z \in U_0). \tag{2}$$

Now suppose $Y \in M_p$. Then

$$d \operatorname{Exp}_{tX}(Y) f = Y_{tX}(f \circ \operatorname{Exp}) = \left\{\frac{d}{du} f(\operatorname{Exp}(tX + uY))\right\}_{u=0}.$$

If t and u are sufficiently small we get from (2)

$$f(\operatorname{Exp}(tX+uY)) = \sum_0^\infty \frac{1}{r!} [(tX^*+uY^*)^r f](p) = \sum_{m,n \geq 0} \frac{t^n u^m}{(n+m)!} [S_{n,m} f](p) \tag{3}$$

where $S_{n,m}$ is the coefficient to $t^n u^m$ in $(tX^* + uY^*)^{n+m}$. In particular,

$$S_{n,1} = (X^*)^n Y^* + (X^*)^{n-1} Y^* X^* + ... + Y^*(X^*)^n.$$

We differentiate the expansion (3) with respect to u and put $u = 0$. We obtain

$$d \operatorname{Exp}_{tX}(Y) f = \sum_0^\infty \frac{t^n}{(n+1)!} [((X^*)^n Y^* + ... + Y^*(X^*)^n) f](p).$$

Let N_p be a normal neighborhood of p. Let $\mathbf{D}(N_p)$ denote the algebra of operators on $C^\infty(N_p)$ generated by the vector fields Z^*, as Z varies through M_p. Let L_{X^*} and R_{X^*} denote the linear transformations of $\mathbf{D}(N_p)$ given by $L_{X^*} : A \to X^*A$ and $R_{X^*} : A \to AX^*$. Since $\theta(X^*)\, Y^* = X^*Y^* - Y^*X^*$, we put $\theta(X^*)\, A = X^*A - AX^*$ for each $A \in \mathbf{D}(N_p)$. Then $\theta(X^*) = L_{X^*} - R_{X^*}$ so $\theta(X^*)$ and L_{X^*} commute; hence we have

$$(R_{X^*})^m = (L_{X^*} - \theta(X^*))^m = \sum_{p=0}^{m} (-1)^p \binom{m}{p} (L_{X^*})^{m-p}(\theta(X^*))^p.$$

On using the relation

$$\sum_{p=0}^{n-k} \binom{n-p}{k} = \binom{n+1}{k+1}$$

we find

$$S_{n,1} = \sum_{p=0}^{n} (X^*)^p \sum_{k=0}^{n-p} (-1)^k \binom{n-p}{k} (X^*)^{n-p-k}\, \theta(X^*)^k\, (Y^*)$$

$$= \sum_{k=0}^{n} \binom{n+1}{k+1} (X^*)^{n-k}\, \theta(-X^*)^k\, (Y^*)$$

so

$$d\,\mathrm{Exp}_{tX}(Y) f = \sum_{n=0}^{\infty} \Big[\sum_{k=0}^{n} \Big\{\frac{(tX^*)^{n-k}}{(n-k)!} \frac{\theta(-tX^*)^k}{(k+1)!} (Y^*)\Big\} f\Big] (p).$$

For sufficiently small t, the right-hand side can be rewritten by the formula

$$\sum_{n=0}^{\infty} \Big[\sum_{k=0}^{n} \Big\{\frac{(tX^*)^{n-k}}{(n-k)!} \frac{(\theta(-tX^*))^k}{(k+1)!} (Y^*)\Big\} f\Big] (p) \tag{4}$$

$$= \sum_{r=0}^{\infty} \Big[\frac{(tX^*)^r}{r!} \Big(\sum_{m=0}^{\infty} \Big\{\frac{(\theta(-tX^*))^m}{(m+1)!} (Y^*)\Big\} f\Big)\Big] (p).$$

In order to justify (4) we first prove the statements (i) and (ii) below.

(i) There exists an interval $I_\delta : -\delta < t < \delta$ and an open neighborhood U of p such that the series

$$G(t, q) = \sum_{m=0}^{\infty} \Big\{\frac{\theta(-tX)^m}{(m+1)!} (Y^*)\Big\}_q f \tag{5}$$

converges absolutely and represents an analytic function for $(t, q) \in I_\delta \times U$, and such that the operator X^* can be applied to the series (5) term by term.

(ii) There exists a subinterval $I_{\delta_1}: -\delta_1 < t < \delta_1$ of I_δ such that the series

$$\sum_{r=0}^{\infty} \frac{t^r}{r!} [(X^*)^r G] (t, p)$$

converges uniformly for $t \in I_{\delta_1}$.

In order to prove (i) and (ii) we can assume that in a suitable coordinate system $\{x_1, ..., x_m\}$ valid near p we have $x_1(p) = ... = x_m(p) = 0$ and $X^* = \partial/\partial x_1$ (see Chevalley, [2], p. 89 for a simple proof). We may also assume that $Y^* = g\,\partial/\partial x_2$ where g is analytic. Then

$$\theta(X^*)^m(Y^*) = \frac{\partial^m g}{\partial x_1^m} \frac{\partial}{\partial x_2}$$

and the series (5) reduces to

$$\sum_{s=0}^{\infty} \frac{(-t)^s}{(s+1)!} \frac{\partial^s g}{\partial x_1^s} \frac{\partial f}{\partial x_2}.$$

Since g is analytic its derivatives of sth order are bounded by $Cs!$ (C = constant) uniformly in a neighborhood of the origin. Hence (i) follows. For (ii) we expand $G(t, q)$ in a power series

$$G(t, q) = \sum a_{i_0 i_1 \ldots i_m} t^{i_0} x_1^{i_1} \ldots x_m^{i_m}.$$

Then

$$[(X^*)^r G] (t, p) = \sum_{i=0}^{\infty} (r!\, a_{i,r,0\ldots0})\, t^i.$$

Let $\rho > 0$ be a number such that the series for $G(t, q)$ converges for $(t, x_1, ..., x_m) = (\rho, \rho, ..., \rho)$. Then there exists a constant K such that

$$| a_{i_0 i_1 \ldots i_m} | \leqslant K\rho^{-(i_0 + \ldots + i_m)}.$$

It follows that

$$| [(X^*)^r G] (t, p) | \leqslant K\rho^{-r} r! \sum_{i=0}^{\infty} (t/\rho)^i \leqslant 2K\rho^{-r} r!$$

provided $| t | \leqslant \frac{1}{2}\rho$. Statement (ii) is now obvious.

We put now

$$a_{m+r,r} = \left[\left\{\frac{(X^*)^r}{r!}\frac{\theta(-X^*)^m}{(m+1)!}(Y^*)\right\}f\right](p).$$

Then from (i) and (ii) we know that there exists a disk D around 0 in the complex plane such that the series

$$\sum_{n=r}^{\infty} a_{n,r}t^n \quad \text{and} \quad \sum_{r=0}^{\infty}\left(\sum_{n=r}^{\infty} a_{n,r}t^n\right)$$

converge absolutely and uniformly for $t \in D$. By Weierstrass' theorem on double series we can interchange the summation so that we have

$$\sum_{r=0}^{\infty}\sum_{n=r}^{\infty} a_{n,r}t^n = \sum_{n=0}^{\infty}\sum_{r=0}^{n} a_{n,r}t^n.$$

This, however, is precisely the relation (4). In view of (2), we have proved: There exists a number $\epsilon_f > 0$ such that

$$d\,\mathrm{Exp}_{tX}(Y)f = \left\{\frac{1 - e^{-\theta(tX^*)}}{\theta(tX^*)}(Y^*)\right\}_{\mathrm{Exp}\,tX} f$$

for $|t| < \epsilon_f$. Using this relation on the coordinate functions $f = x_1, \ldots, f = x_m$, the theorem follows with $\epsilon = \min(\epsilon_{x_1}, \ldots, \epsilon_{x_m})$.

§ 7. Covariant Differentiation

In § 5, parallelism was defined by means of the covariant differentiation ∇_X. Theorem 7.1 below shows that it is also possible to go the other way and describe the covariant derivative by means of parallel translation. This makes it possible to define the covariant derivative of other objects.

Definition. Let X be a vector field on a manifold M. A curve $s \to \varphi(s)$ $(s \in I)$ is called an *integral curve* of X if

$$\dot{\varphi}(s) = X_{\varphi(s)}, \qquad s \in I. \tag{1}$$

Assuming $0 \in I$, let $p = \varphi(0)$ and let $\{x_1, \ldots, x_m\}$ be a system of coordinates valid in a neighborhood U of p. There exist functions

$X^i \in C^\infty(U)$ such that $X = \sum_i X^i \, \partial/\partial x_i$ on U. For simplicity let $x_i(s) = x_i(\varphi(s))$ and write X^i instead of $(X^i)^*$ (§ 2, No. 1). Then (1) is equivalent to

$$\frac{dx_i(s)}{ds} = X^i(x_1(s), \ldots, x_m(s)) \qquad (1 \leqslant i \leqslant m). \tag{2}$$

Therefore if $X_p \neq 0$ there exists an integral curve of X through p.

Theorem 7.1. *Let M be a manifold with an affine connection. Let $p \in M$ and let X, Y be two vector fields on M. Assume $X_p \neq 0$. Let $s \to \varphi(s)$ be an integral curve of X through $p = \varphi(0)$ and τ_t the parallel translation from p to $\varphi(t)$ with respect to the curve φ. Then*

$$(\nabla_X(Y))_p = \lim_{s \to 0} \frac{1}{s} (\tau_s^{-1} Y_{\varphi(s)} - Y_p).$$

Proof. We shall use the notation introduced above. Consider a fixed $s > 0$ and the family $Z_{\varphi(t)}$ $(0 \leqslant t \leqslant s)$ which is parallel with respect to the curve φ such that $Z_{\varphi(0)} = \tau_s^{-1} Y_{\varphi(s)}$. We can write

$$Z_{\varphi(t)} = \sum_i Z^i(t) \left(\frac{\partial}{\partial x_i}\right)_{\varphi(t)}, \quad Y_{\varphi(t)} = \sum_i Y^i(t) \left(\frac{\partial}{\partial x_i}\right)_{\varphi(t)},$$

and have the relations

$$\dot{Z}^k(t) + \sum_{i,j} \Gamma_{ij}{}^k \, \dot{x}_i(t) \, Z^j(t) = 0 \qquad (0 \leqslant t \leqslant s)$$

$$Z^k(s) = Y^k(s) \qquad (1 \leqslant k \leqslant m).$$

By the mean value theorem

$$Z^k(s) = Z^k(0) + s\dot{Z}^k(t^*)$$

for a suitable number t^* between 0 and s. Hence the kth component of $(1/s)(\tau_s^{-1} Y_{\varphi(s)} - Y_p)$ is

$$\begin{aligned} \frac{1}{s}(Z^k(0) - Y^k(0)) &= \frac{1}{s}\{Z^k(s) - s\dot{Z}^k(t^*) - Y^k(0)\} \\ &= \sum_{i,j} \Gamma_{ij}{}^k(\varphi(t^*)) \, \dot{x}_i(t^*) \, Z^j(t^*) + \frac{1}{s}(Y^k(s) - Y^k(0)). \end{aligned}$$

As $s \to 0$ this expression has the limit

$$\frac{dY^k}{ds} + \sum_{i,j} \Gamma_{ij}{}^k \frac{dx_i}{ds} Y^j.$$

Let this last expression be denoted by A_k. It was shown earlier that

$$\nabla_X(Y)_p = \sum_k A_k \left(\frac{\partial}{\partial x_k}\right)_p.$$

This proves the theorem.

By using Theorem 7.1 it is now possible to define covariant derivatives of arbitrary tensor fields. Let p and q be two points in M and γ a curve segment in M from p to q. Let τ denote the parallel translation along γ. If $F \in M_p^*$ we define $\tau \cdot F \in M_q^*$ by the formula $(\tau \cdot F)(A) = F(\tau^{-1} \cdot A)$ for each $A \in M_q$. If T is a tensor field on M of type (r, s) where $r + s > 0$, we define $\tau \cdot T_p \in \mathfrak{D}_s^r(q)$ by

$$(\tau \cdot T_p)(F_1, \ldots, F_r, A_1, \ldots, A_s) = T_p(\tau^{-1}F_1, \ldots, \tau^{-1}F_r, \tau^{-1}A_1, \ldots, \tau^{-1}A_s)$$

for $A_i \in M_q$, $F_j \in M_q^*$. Now, let $X \in \mathfrak{D}^1(M)$ and let p be any point in M where $X_p \neq 0$. With the notation of Theorem 7.1 we put

$$(\nabla_X T)_p = \lim_{s \to 0} \frac{1}{s} (\tau_s^{-1} T_{\varphi(s)} - T_p). \tag{3}$$

For each point $q \in M$ where $X_q = 0$ we put $(\nabla_X T)_q = 0$ in accordance with Lemma 4.2. For a function $f \in C^\infty(M)$ we put

$$(\nabla_X f)_p = \lim_{s \to 0} \frac{1}{s} (f(\varphi(s)) - f(p)),$$

if $X_p \neq 0$, otherwise we put $(\nabla_X f)_p = 0$. Then we have $\nabla_X f = Xf$. Finally ∇_X is extended to a linear mapping of $\mathfrak{D}$ into $\mathfrak{D}$.

Proposition 7.2. *The operator ∇_X has the following properties:*

(i) *∇_X is a derivation of the mixed tensor algebra $\mathfrak{D}(M)$ (considered as an algebra over $\boldsymbol{R}$).*

(ii) *∇_X preserves type of tensors.*

(iii) *∇_X commutes with all contractions C^i_j.*

The verification of these properties is quite straightforward. For a simple application, let $X, Y \in \mathfrak{D}^1(M)$, $\omega \in \mathfrak{D}_1(M)$. Then by (i)

$$\nabla_X(Y \otimes \omega) = \nabla_X(Y) \otimes \omega + Y \otimes \nabla_X \omega, \tag{4}$$

so (iii) implies

$$(\nabla_X \omega)(Y) = X \cdot \omega(Y) - \omega(\nabla_X(Y)). \tag{5}$$

The tensor field $Y \otimes \omega$ can be regarded as an $\mathfrak{F}$-linear mapping of $\mathfrak{D}^1$ into itself given by

$$Y \otimes \omega : Z \to \omega(Z)\, Y \qquad (Z \in \mathfrak{D}^1).$$

If A, B are two mappings of $\mathfrak{D}^1$ into itself we put $[A, B] = AB - BA$. Then

$$[\nabla_X, Y \otimes \omega] = \nabla_X(Y \otimes \omega). \tag{6}$$

In fact,

$$\begin{aligned}[\nabla_X, Y \otimes \omega]\,(Z) &= \nabla_X(\omega(Z)Y) - (Y \otimes \omega)\,(\nabla_X Z)\\ &= \omega(Z)\,\nabla_X(Y) + (X \cdot \omega(Z))\,Y - \omega(\nabla_X Z)\,Y.\end{aligned}$$

On the other hand, (4) and (5) imply

$$\begin{aligned}\nabla_X(Y \otimes \omega)\,(Z) &= \omega(Z)\,\nabla_X(Y) + (\nabla_X\omega)\,(Z)\,Y\\ &= \omega(Z)\,\nabla_X(Y) + (X \cdot \omega(Z))\,Y - \omega(\nabla_X(Z))\,Y,\end{aligned}$$

proving (6). We have therefore

$$[\nabla_X, B] = \nabla_X \cdot B \tag{7}$$

if B is a tensor field of type (1,1). On the left-hand side, B is to be considered as the linear mapping of $\mathfrak{D}^1$ into $\mathfrak{D}^1$ given by

$$Z \to C^1{}_1(Z \otimes B),$$

where $C^1{}_1$ is the contraction of the first contravariant and first covariant index.

§ 8. The Structural Equations

Let M be a manifold with an affine connection ∇. We put

$$T(X, Y) = \nabla_X(Y) - \nabla_Y(X) - [X, Y],$$

$$R(X, Y) = \nabla_X\nabla_Y - \nabla_Y\nabla_X - \nabla_{[X,Y]}$$

for all $X, Y \in \mathfrak{D}^1$. Note that $T(X, Y) = -T(Y, X)$ and $R(X, Y) = -R(Y, X)$. It is easy to verify that $T(fX, gY) = fgT(X, Y)$ and $R(fX, gY) \cdot hZ = fghR(X, Y) \cdot Z$ for all $f, g, h \in C^\infty(M)$, $X, Y, Z \in \mathfrak{D}^1$. The mapping

$$(\omega, X, Y) \to \omega(T(X, Y))$$

is an $\mathfrak{F}$-multilinear mapping of $\mathfrak{D}_1 \times \mathfrak{D}^1 \times \mathfrak{D}^1$ into $\mathfrak{F}$ and therefore is an element of $\mathfrak{D}^1_2(M)$. This element is called the *torsion tensor field* and is also denoted by T. Similarly, the mapping

$$(\omega, Z, X, Y) \to \omega(R(X, Y) \cdot Z)$$

is an $\mathfrak{F}$-multilinear mapping of $\mathfrak{D}_1 \times \mathfrak{D}^1 \times \mathfrak{D}^1 \times \mathfrak{D}^1$ into $\mathfrak{F}$ and therefore is an element of $\mathfrak{D}^1_3(M)$. This element is called the *curvature tensor field* and is also denoted by R. The tensor fields T and R have type $(1, 2)$ and $(1, 3)$, respectively.

Let $p \in M$ and suppose $X_1, \ldots, X_m$ is a basis for the vector fields in some open neighborhood N_p of p, that is, each vector field X on N_p can be written $X = \sum_i f_i X_i$ where $f_i \in C^\infty(N_p)$. We define the functions $\Gamma_{ij}{}^k$, $T^k{}_{ij}$, $R^k{}_{lij}$ on N_p by the formulas

$$\nabla_{X_i}(X_j) = \sum_k \Gamma_{ij}{}^k X_k,$$

$$T(X_i, X_j) = \sum_k T^k{}_{ij} X_k,$$

$$R(X_i, X_j) \cdot X_l = \sum_k R^k{}_{lij} X_k.$$

Let ω^i, $\omega^i{}_j$ $(1 \leqslant i, j \leqslant m)$ be the 1-forms on N_p determined by

$$\omega^i(X_j) = \delta^i{}_j, \qquad \omega^i{}_j = \sum_k \Gamma_{kj}{}^i \omega^k.$$

It is clear that the forms $\omega^i{}_j$ determine the functions $\Gamma_{kj}{}^i$ on N_p and thereby the connection ∇. On the other hand, as the next theorem shows, the forms $\omega^i{}_j$ are described by the torsion and curvature tensor fields.

Theorem 8.1 (*the structural equations of Cartan*).

$$d\omega^i = -\sum_p \omega^i{}_p \wedge \omega^p + \tfrac{1}{2}\sum_{j,k} T^i{}_{jk}\omega^j \wedge \omega^k, \tag{1}$$

$$d\omega^i{}_l = -\sum_p \omega^i{}_p \wedge \omega^p{}_l + \tfrac{1}{2}\sum_{j,k} R^i{}_{ljk}\omega^j \wedge \omega^k. \tag{2}$$

Both sides of (1) represent a 2-form on N_p. We apply both sides of that equation to (X_j, X_k) and evaluate by means of the rules (4) and (9) in § 2. If we define the functions $c^i{}_{jk}$ by $[X_j, X_k] = \Sigma_i\, c^i{}_{jk}\, X_i$, the left-hand side of (1) is

$$d\omega^i(X_j, X_k) = \tfrac{1}{2}\{X_j \cdot \omega^i(X_k) - X_k \cdot \omega^i(X_j) - \omega^i([X_j, X_k])\}$$

$$= \tfrac{1}{2}\{0 - 0 - c^i{}_{jk}\}.$$

As for the right-hand side we have first

$$T(X_j, X_k) = \nabla_{X_j}(X_k) - \nabla_{X_k}(X_j) - [X_j, X_k]$$

$$= \sum_i (\Gamma_{jk}{}^i - \Gamma_{kj}{}^i - c^i{}_{jk})\, X_i$$

so

$$T^i{}_{jk} = \Gamma_{jk}{}^i - \Gamma_{kj}{}^i - c^i{}_{jk}. \tag{3}$$

Similarly, we find

$$R^k{}_{lij} = \sum_p (\Gamma_{jl}{}^p \Gamma_{ip}{}^k - \Gamma_{il}{}^p \Gamma_{jp}{}^k) + X_i \cdot \Gamma_{jl}{}^k - X_j \cdot \Gamma_{il}{}^k - \sum_p c^p{}_{ij} \Gamma_{pl}{}^k; \tag{4}$$

furthermore,

$$(-\sum_p \omega^i{}_p \wedge \omega^p)\,(X_j, X_k) = -\sum_p \tfrac{1}{2}\{\omega^i{}_p(X_j)\, \omega^p(X_k) - \omega^p(X_j)\, \omega^i{}_p(X_k)\}$$

$$= \tfrac{1}{2}(\Gamma_{kj}{}^i - \Gamma_{jk}{}^i),$$

$$\Big(\tfrac{1}{2}\sum_{r,s} T^i{}_{rs}\omega^r \wedge \omega^s\Big)\,(X_j, X_k) = \tfrac{1}{4}\sum_{r,s} (\Gamma_{rs}{}^i - \Gamma_{sr}{}^i - c^i{}_{rs})\,(\delta_{jr}\delta_{ks} - \delta_{js}\delta_{kr})$$

$$= \tfrac{1}{2}(\Gamma_{jk}{}^i - \Gamma_{kj}{}^i - c^i{}_{jk}),$$

so (1) follows immediately; in the same way (2) follows if we use formula (4).

Suppose $Y_1, \ldots, Y_m$ is a basis for the tangent space M_p. Let N_0 be a normal neighborhood of the origin in M_p and let N_p denote the normal neighborhood $\operatorname{Exp} N_0$ of p in M. Let $Y_1^*, \ldots, Y_m^*$ be the vector fields on N_p that are adapted to the tangent vectors $Y_1, \ldots, Y_m$. Then $Y_1^*, \ldots, Y_m^*$ is a basis for the vector fields on N_p, due to Prop. 5.2. Suppose ω^i, $\omega^i{}_l$, $\Gamma_{ij}{}^k$, $R^k{}_{lij}$, and $T^k{}_{ij}$ are defined by means of this basis. Let V be the set of points $(t, a_1, \ldots, a_m) \in \boldsymbol{R} \times \boldsymbol{R}^m$ for which $ta_1 Y_1 + \ldots + ta_m Y_m \in N_0$. We consider now the mapping $\Phi : V \to N_p$ given by

$$\Phi : (t, a_1, \ldots, a_m) \to \operatorname{Exp}(ta_1 Y_1 + \ldots + ta_m Y_m). \tag{5}$$

We shall then prove that the dual forms $\Phi^*\omega^i$ and $\Phi^*\omega^i{}_l$ are given by the formulas

$$\Phi^*\omega^i = a_i dt + \bar{\omega}^i, \qquad \Phi^*\omega^i{}_l = \bar{\omega}^i{}_l, \tag{5'}$$

where $\bar{\omega}^i$ and $\bar{\omega}^i{}_l$ are 1-forms in $da_1, ..., da_m$ (and do not contain dt). In fact, we can write

$$\Phi^*\omega^i = f_i(t, a_1, ..., a_m)\, dt + \bar{\omega}^i,$$

$$\Phi^*\omega^i{}_l = g_{il}(t, a_1, ..., a_m)\, dt + \bar{\omega}^i{}_l.$$

For a fixed point $(a_1, ..., a_m) \in \boldsymbol{R}^m$ we consider the mapping

$$\tau : t \to \operatorname{Exp}(ta_1 Y_1 + ... + ta_m Y_m),$$

which maps an open subset of $\boldsymbol{R}$ into M. It is easy to see that

$$\tau^*\omega^i = f_i(t, a_1, ..., a_m)\, dt, \qquad \tau^*\omega^i{}_l = g_{il}(t, a_1, ..., a_m)\, dt,$$

$$\dot{\tau}(t) = \Big(\sum_j a_j Y_j^*\Big)_{\tau(t)}.$$

By using the duality of τ^* and $d\tau$ we obtain

$$f_i(t, a_1, ..., a_m) = a_i,$$

$$g_{il}(t, a_1, ..., a_m) = \sum_q \Gamma_{ql}{}^i\, a_q.$$

This last sum, however, vanishes, because $Y^* = \Sigma_j\, a_j Y_j^*$ is the tangent vector field to the geodesic $t \to \operatorname{Exp}(ta_1 Y_1 + ... + ta_m Y_m)$ and consequently the expression

$$\nabla_{Y^*}(Y_l^*) = \sum_k \Big(\sum_q a_q\, \Gamma_{ql}{}^k\Big)\, Y_k^*$$

vanishes along that geodesic. This proves (5').

The forms $\bar{\omega}^i$ vanish for $t = 0$. In fact, let A be the point $(0, a_1, ..., a_m)$ in V. Then

$$\bar{\omega}^i{}_A\Big(\frac{\partial}{\partial a_j}\Big) = \omega^i\Big(d\Phi_A\Big(\frac{\partial}{\partial a_j}\Big)\Big),$$

and if f is differentiable in a neighborhood of p

$$d\Phi_A\Big(\frac{\partial}{\partial a_j}\Big) f = \Big\{\frac{\partial}{\partial a_j}(f \circ \Phi)\Big\}_{t=0} = \frac{\partial}{\partial a_j}\Big((f \circ \Phi)_{t=0}\Big) = 0.$$

Similarly, the forms $\bar{\omega}^i{}_l$ vanish for $t = 0$.

For the exterior derivatives of the forms (5′) we have

$$\Phi^*(d\omega^i) = d(\Phi^*\omega^i) = da_i \wedge dt + dt \wedge \frac{\partial\bar{\omega}^i}{\partial t} + \dots$$

$$\Phi^*(d\omega^i{}_l) = d(\Phi^*\omega^i{}_l) = dt \wedge \frac{\partial\bar{\omega}^i{}_l}{\partial t} + \dots,$$

where the terms which are not written do not contain dt. On the other hand, since Φ^* is a homomorphism with respect to exterior products ((5), § 3), we can evaluate $\Phi^*(d\omega^i)$ and $\Phi^*(d\omega^i{}_l)$ by means of the structural equations. Equating the coefficients to dt, we obtain the system of differential equations on V:

$$\frac{\partial\bar{\omega}^i}{\partial t} = da_i + \sum_k a_k\bar{\omega}^i{}_k + \sum_{j,k} T^i{}_{jk}a_j\bar{\omega}^k, \qquad \bar{\omega}^i(t; a_j; da_k)_{t=0} = 0; \tag{6}$$

$$\frac{\partial\bar{\omega}^i{}_l}{\partial t} = \sum_{j,k} R^i{}_{ljk}a_j\bar{\omega}^k, \qquad \bar{\omega}^i{}_l(t; a_j; da_k)_{t=0} = 0, \tag{7}$$

which will be useful later. In the derivation of (6) and (7) the antisymmetry of R and T in the two last indices was used. Note that in (6) and (7) we have written for simplicity $T^i{}_{jk}$ and $R^i{}_{ljk}$ in place of $(T^i{}_{jk} \circ \Phi)$ and $(R^i{}_{ljk} \circ \Phi)$. These equations, which represent the structural equations in "polar coordinates," are particularly important in É. Cartan's treatment of Riemannian geometry (É. Cartan [22]).

§ 9. The Riemannian Connection

Definition. Let M be a C^∞-manifold. A *pseudo-Riemannian structure* on M is a tensor field g of type (0, 2) which satisfies

(a) $g(X, Y) = g(Y, X)$ for all $X, Y \in \mathfrak{D}^1(M)$.

(b) For each $p \in M$, g_p is a nondegenerate bilinear form on $M_p \times M_p$.

A *pseudo-Riemannian manifold* is a *connected* C^∞-manifold with a pseudo-Riemannian structure. If (and only if) g_p is positive definite for each $p \in M$, we drop the prefix "pseudo" and speak of a Riemannian structure and Riemannian manifold. A Riemannian structure on a manifold induces in an obvious manner a Riemannian structure on any submanifold. The analogous statement does not hold for a general pseudo-Riemannian structure.

Theorem 9.1. *On a pseudo-Riemannian manifold there exists one and only one affine connection satisfying the following two conditions*:

(i) *The torsion tensor T is* 0.

(ii) *The parallel displacement preserves the inner product on the tangent spaces.*

Proof. Conditions (i) and (ii) can be written:

(i′) $\nabla_X Y - \nabla_Y X = [X, Y], \qquad X, Y \in \mathfrak{D}^1;$
(ii′) $\nabla_Z g = 0, \qquad Z \in \mathfrak{D}^1.$

We apply the derivation ∇_Z to the tensor field $X \otimes Y \otimes g$ and use the fact that ∇_Z commutes with contractions. In view of (ii′) we obtain

$$Z\,g(X, Y) = g(\nabla_Z X, Y) + g(X, \nabla_Z Y),$$
$$Z\,g(X, Y) = g(\nabla_X Z, Y) + g(X, \nabla_Z Y) + g([Z, X], Y). \tag{1}$$

In (1) we permute the letters cyclically and eliminate ∇_X and ∇_Y from the obtained relations. This gives

$$2g(X, \nabla_Z Y) = Z\,g(X, Y) + g(Z, [X, Y]) + Y\,g(X, Z) + g(Y, [X, Z]) - Xg(Y, Z) - g(X, [Y, Z]), \tag{2}$$

and this relation shows (g being nondegenerate) that there can be at most one affine connection satisfying (i) and (ii). On the other hand, we can define $\nabla_Z Y$ by (2) and a routine computation shows that the axioms ∇_1 and ∇_2 for an affine connection are satisfied. Moreover, carrying out the computations above in reverse order, one verifies (i′) and (ii′) on the basis of (2).

The connection ∇ given by (2) is called the *pseudo-Riemannian* (*or Riemannian*) connection. If M is analytic and the tensor field g is analytic, M is called an *analytic pseudo-Riemannian manifold.* In this case, the pseudo-Riemannian connection is analytic.

Suppose now M is a Riemannian manifold. Then g_p is positive definite for each $p \in M$. If $X \in M_p$, we sometimes write $\|X\|$ instead of $g_p(X, X)^{1/2}$. There exists a basis $Y_1, \ldots, Y_m$ of M_p such that $g_p(Y_i, Y_j) = \delta_{ij}$ $(1 \leqslant i, j \leqslant m)$. Let N_0 denote a normal neighborhood of 0 in M_p and let $N_p = \operatorname{Exp} N_0$. We shall now apply (6) and (7) in § 8 to the adapted basis $Y_1^*, \ldots, Y_m^*$ of vector fields on the normal neighborhood N_p. Since g is invariant under parallelism, we have

$$g(Y_i^*, Y_j^*) = g(Y_i, Y_j) = \delta_{ij}$$

on N_p. Since $\omega^i(Y_j^*) = \delta^i{}_j$, we have

$$g = \sum_i (\omega^i)^2 \qquad \text{on } N_p,$$

where the symmetric product $\alpha\beta$ of two 1-forms is given by $\alpha\beta = \frac{1}{2}(\alpha \otimes \beta + \beta \otimes \alpha)$.

Let S denote the unit sphere $\| X \| = 1$ in M_p, and let U denote the set of all pairs $(t, X) \in \boldsymbol{R} \times S$ for which $tX \in N_0$. Then U is open in $\boldsymbol{R} \times S$. Let Ψ denote the mapping $\Psi = \Phi \circ I$ where Φ is the mapping (5), § 8, and I is the identity mapping of U into $\boldsymbol{R} \times M_p$, M_p being identified with $\boldsymbol{R}^m$ by means of the basis $Y_1, \ldots, Y_m$. Thus $\Psi(t, X) = \operatorname{Exp} tX$ if $(t, X) \in U$. Then Ψ^*g (§ 3, No. 3) is an element of $\mathfrak{D}_2^0(U)$ and we have

$$\Psi^*g = \sum_i (\Psi^*\omega^i)^2.$$

Lemma 9.2. *The tensor field Ψ^*g on U is given by*

$$\Psi^*g = (dt)^2 + \sum_{i=1}^m (\bar{\omega}^i)^2.$$

Proof. For the Riemannian connection we have $\omega^l{}_i = -\omega^i{}_l$ so (by (5′), § 8) $\bar{\omega}^i{}_l = -\bar{\omega}^l{}_i$. In fact, this is an immediate consequence of the relation

$$g(\nabla_{Y_i^*}(Y_j^*), Y_k^*) + g(Y_j^*, \nabla_{Y_i^*}(Y_k^*)) = Y_i^* g(Y_j^*, Y_k^*) = 0.$$

Using (5′), § 8, we have

$$\Phi^*g = \sum_i (\Phi^*\omega^i)^2 = \sum_i a_i^2(dt)^2 + \sum_i (\bar{\omega}^i)^2 + 2\sum_i a_i\bar{\omega}^i dt.$$

From (6), § 8, we obtain

$$\frac{\partial}{\partial t}\left(\sum_i a_i\bar{\omega}^i\right) = \sum_i a_i da_i + \sum_{i,k} a_i a_k \bar{\omega}^i{}_k = \sum_i a_i da_i$$

due to the skew symmetry of $\bar{\omega}^i{}_k$. Moreover, $I^*(\sum_i a_i da_i) = 0$. Now, N_0 is star-shaped. Therefore, if $X \in S$, the set of $t \in \boldsymbol{R}$ such that $(t, X) \in U$ is an interval containing $t = 0$. Inasmuch as $\bar{\omega}^i = 0$ for $t = 0$, we obtain $I^*(\sum_i a_i\bar{\omega}^i) = 0$. Moreover

$$I^*\left(\sum_i a_i^2\right) = 1,$$

so, using $\Psi^* = I^* \circ \Phi^*$ the lemma follows.

We shall now introduce the Riemannian metric on the Riemannian manifold M. Let $t \to \gamma(t)$ $(\alpha \leqslant t \leqslant \beta)$ be a curve segment in M. The arc length of γ is defined by

$$L(\gamma) = \int_\alpha^\beta \{g_{\gamma(t)}(\dot{\gamma}(t), \dot{\gamma}(t))\}^{1/2}\, dt. \tag{3}$$

It is clear from (3) that two curve segments which are the same except for a change of parameter have the same arc length.

For a Riemannian manifold, we write for simplicity "geodesic" instead of "geodesic segment." It will also be convenient not always to distinguish between two curves which coincide after a change of parameter.

Lemma 9.3. *Let M be a Riemannian manifold and p any point in M. Let N_0 be any normal neighborhood of 0 in M_p and put $N_p = Exp\ N_0$. For each $q \in N_p$, let γ_{pq} denote the unique geodesic in N_p joining p to q. Then*

$$L(\gamma_{pq}) < L(\gamma)$$

for each curve segment $\gamma \neq \gamma_{pq}$ in N_p which joins p to q. If, in particular, the normal neighborhood N_0 is an open ball $0 \leqslant \| X \| < \delta$ in M_p, the inequality $L(\gamma_{pq}) < L(\gamma)$ holds for each curve segment $\gamma \neq \gamma_{pq}$ in M which joins p to q.

Proof. Let $s \to \gamma(s)$ $(0 \leqslant s \leqslant 1)$ be any curve segment in N_p joining p to q. For the purpose of proving the inequality above, we can assume that $\gamma(s) \neq p$ for $s \neq 0$. We can then write $\gamma = \Psi \circ \gamma_0$ where γ_0 is a curve segment in $\boldsymbol{R} \times S$,

$$\gamma_0 : s \to (t(s), X(s)) \qquad (0 \leqslant s \leqslant 1),$$

such that $t(s) > 0$ for $0 < s \leqslant 1$. The curve segment γ_0 is contained in the set U of Lemma 9.2. We have

$$\dot{\gamma}_0(s) = d\gamma_0\left(\frac{d}{ds}\right) = \left(\dot{t}(s)\frac{d}{dt}, \dot{X}(s)\right),$$

t denoting the coordinate on $\boldsymbol{R}$. Consequently,

$$g(\dot{\gamma}(s), \dot{\gamma}(s)) = (\Psi^* g)(\dot{\gamma}_0(s), \dot{\gamma}_0(s)) = (\dot{t}(s))^2 + \sum_{i=1}^m (\bar{\omega}^i(\dot{X}(s)))^2. \tag{4}$$

Then

$$L(\gamma) = \int_0^1 \Big\{ \dot{t}(s))^2 + \sum_i (\bar{\omega}^i(\dot{X}(s)))^2 \Big\}^{1/2} ds \geqslant \int_0^1 | \dot{t}(s) |\, ds = L(\gamma_{pq}). \tag{5}$$

The equality sign holds if and only if $\bar{\omega}^i(\dot{X}(s)) = 0$ for all i and all s. In view of (5′), § 8, this is equivalent to $\dot{X}(s) = 0$ or $X(s) =$ constant, which means that $\gamma = \gamma_{pq}$ (up to change of parameter).

Finally, let us consider the case when N_0 is an open ball $0 \leqslant \| X \| < \delta$ in M_p. Let $s \to \gamma(s)$ be a curve segment in M joining p to q such that γ does not lie in N_p. Let X_1 be the element in N_0 such that $\operatorname{Exp} X_1 = q$ and suppose δ^* satisfies the inequalities $\| X_1 \| < \delta^* < \delta$. Put

$$N^* = \{\operatorname{Exp} X : X \in M_p, \| X \| < \delta^*\}.$$

Let s_0 be the infimum of the set of parameter values s for which $\gamma(s) \notin N^*$. Then the point $q_0 = \gamma(s_0)$ lies on the boundary of N^* and, by the first part of the proof, the length of γ from p to q_0 is $\geqslant \delta^*$. Since $L(\gamma_{pq}) = \| X_1 \|$, it follows that

$$L(\gamma) > L(\gamma_{pq}).$$

This proves the lemma.

The Riemannian manifold M can now be turned into a metric space. Since M is connected, each pair of points $p, q \in M$ can be joined by a curve segment. The *distance* of p and q is defined by

$$d(p, q) = \inf_{\gamma} L(\gamma), \tag{6}$$

where γ runs over all curve segments joining p and q. Then we have

(a) $d(p, q) = d(q, p)$,
(b) $d(p, q) \leqslant d(p, r) + d(r, q)$,
(c) $d(p, q) = 0$ if and only if $p = q$.

The two first are obvious and the last one is a direct consequence of Lemma 9.3. Thus d is a metric on the set M. For $p \in M$ we put

$$B_r(p) = \{q \in M : d(p, q) < r\}, \qquad 0 \leqslant r \leqslant \infty,$$

$$S_r(p) = \{q \in M : d(p, q) = r\}, \qquad 0 \leqslant r < \infty;$$

$B_r(p)$ is called the *open ball* around p with radius r and $S_r(p)$ is called the *sphere* around p with radius r.

Proposition 9.4. *Suppose that the open ball*

$$V_r(0) = \{X \in M_p : 0 \leqslant \| X \| < r\}$$

is a normal neighborhood of 0 *in* M_p. *Then*

$$B_r(p) = \operatorname{Exp} V_r(0).$$

Proof. It is obvious that Exp $V_r(0) \subset B_r(p)$. On the other hand, if q were a point in $B_r(p)$, not belonging to Exp $V_r(0)$, then each curve segment joining p to q must intersect the boundary of each set Exp $V_\rho(p)$ $(\rho < r)$. Lemma 9.3 then implies that $d(p, q) \geqslant \rho$ for each $\rho < r$. Hence $d(p, q) \geqslant r$. This contradiction shows that $B_r(p) = \text{Exp } V_r(0)$.

Corollary 9.5. *The topology of M given by the metric d coincides with the original topology of M.*

In fact, the sets $B_r(p)$ $(r > 0)$, form a fundamental system of neighborhoods of p in the metric topology of M. On the other hand, Theorem 6.1 shows that in the original topology of M, the sets Exp $V_r(0)$ $(r > 0)$ form a fundamental system of neighborhoods of the point p.

Proposition 9.6. *Every Riemannian manifold is separable.*

This proposition is just a special case of the following theorem of P. Alexandroff: *A connected, locally compact metric space is separable.* For a simple proof of this elementary result, see Pfluger [1], pp. 22-23.

Definition. Under the assumption of Prop. 9.4, the neighborhoods $B_r(p)$ and $V_r(0)$ are called *spherical normal neighborhoods* of p in M and of 0 in M_p, respectively.

Suppose V is a connected submanifold of a Riemannian manifold M. The Riemannian structure on M induces a Riemannian structure on V. Let d_M and d_V be the distances in M and V given by the Riemannian structures. Then $d_V(p, q) \geqslant d_M(p, q)$ for each pair $p, q \in V$. Examples show that in general we do not have equality sign; however, in the case when V is an open submanifold of M and $p \in V$, the equality sign holds, owing to Prop. 9.4, if q is sufficiently close to p.

Suppose $B_\delta(p)$ is a spherical normal neighborhood of p; if $r < \delta$, then $S_r(p)$ is the image of the sphere $\| X \| = r$ in M_p under the diffeomorphism Exp_p. Thus $S_r(p)$ is a submanifold of M.

Lemma 9.7. *Let $r < \delta$. Then each geodesic γ emanating from p is perpendicular to $S_r(p)$ at the first point of intersection.*

Proof. Assuming the geodesic γ parametrized by its arc length measured from p, let X denote its tangent vector at p. Then $\| X \| = 1$ and the segment of γ from p to the first point of intersection with $S_r(p)$ is $\gamma(s) = \text{Exp } sX$ $(0 \leqslant s \leqslant r)$. Let Y be any tangent vector to $S_r(p)$ at the point $\gamma(r)$. Then there exists a unique tangent vector Y_0 to S at X such that $(d\Psi)_{(r,X)}(0, Y_0) = Y$. Then

$$g\left(d\gamma_r\left(\frac{d}{ds}\right), Y\right) = (\Psi^*g)\left(\left(\frac{d}{ds}, 0\right), (0, Y_0)\right) = 0,$$

owing to Lemma 9.2.

Remark. In general the geodesic γ will intersect $S_r(p)$ $(r < \delta)$ more than once. The intersection does not always take place at a right angle. An example is provided by an everywhere dense geodesic on a flat two-dimensional torus. (For the definition of "flat" see Chapter V, § 6.)

Lemma 9.8. *Let M be a Riemannian manifold with metric d given by (6). Let p and q be two points in M and γ_{pq} a curve segment joining p and q. If $L(\gamma_{pq}) = d(p, q)$, then γ_{pq} is a geodesic.*

Proof. There exists a finite sequence of points $r_0, r_1, ..., r_n$ (where $r_0 = p$, $r_n = q$) on γ_{pq} such that each segment $\gamma_{r_i r_{i+1}}$ lies in a spherical normal neighborhood $B_{\delta_i}(r_i)$ (see Theorem 6.2). Then

$$\sum_{i=0}^{n-1} L(\gamma_{r_i r_{i+1}}) = L(\gamma_{pq}), \qquad d(p, q) \leqslant \sum_{i=0}^{n-1} d(r_i, r_{i+1})$$

and $d(r_i, r_{i+1}) \leqslant L(\gamma_{r_i r_{i+1}})$. Assuming now $L(\gamma_{pq}) = d(p, q)$ it follows that

$$d(r_i, r_{i+1}) = L(\gamma_{r_i r_{i+1}}).$$

By Lemma 9.3, $\gamma_{r_i r_{i+1}}$ is a geodesic and the lemma follows.

Remark. The conclusion of the lemma holds even if γ_{pq} is only assumed piecewise differentiable.

The following result which combines Lemmas 6.4 and 9.3 will often be useful.

Theorem 9.9. *Let M be a Riemannian manifold with metric d. To each $p \in M$ corresponds a number $r(p) > 0$ such that if $0 < \rho \leqslant r(p)$, then $B_\rho(p)$ has the properties:*

(A) *$B_\rho(p)$ is a normal neighborhood of each of its points.*

(B) *Let $a, b \in B_\rho(p)$ and let γ_{ab} be the unique geodesic in $B_\rho(p)$ joining a and b. Then γ_{ab} is the only curve segment in M of length $d(a, b)$ which joins a and b.*

Proof. Let $X_1, ..., X_m$ be an orthonormal basis of M_p and let $x_1, ..., x_m$ be normal coordinates at p with respect to this basis, valid on a neighborhood U of p. If $\delta > 0$ is sufficiently small we have

$$B_\delta(p) = \left\{ q \in U : \sum_{i=1}^{m} x_i(q)^2 < \delta^2 \right\}.$$

Using Lemma 6.4 we conclude that there exists a number $\delta^* > 0$ such that for $0 < \delta \leqslant \delta^*$, $B_\delta(p)$ is a normal neighborhood of each of its points. We put $r(p) = \frac{1}{4}\delta^*$. If $0 < \rho \leqslant r(p)$, then the neighborhood $B_\rho(p)$ clearly has property (A). It has also property (B). In fact, since $B_{\delta^*}(p)$ is a normal neighborhood of a, γ_{ab} is the only shortest curve segment in $B_{\delta^*}(p)$ which joins a and b. On the other hand, a curve segment which joins a and b but does not lie entirely in $B_{\delta^*}(p)$ has obviously length $> 3\rho$. Since $L(\gamma_{ab}) \leqslant d(a, p) + d(p, b) \leqslant 2\rho$, property (B) is also verified.

Definition. A ball $B_\rho(p)$ which is a normal neighborhood of each of its points will be called a *convex normal ball*. It will be called *minimizing* if it also has the property (B) above.

Remark. It is easy to show by examples that a convex normal ball is not necessarily minimizing.

Proposition 9.10. *In the notation of Theorem 9.9 let A and B be the unique points in M_p satisfying the relations*

$$\operatorname{Exp}_p A = a, \qquad \operatorname{Exp}_p B = b, \qquad \| A \| < r(p), \qquad \| B \| < r(p).$$

Then

$$\frac{\| A - B \|}{d(a, b)} \to 1$$

as $(a, b) \to (p, p)$.

Proof. We may assume that the straight line segment joining A and B does not pass through the origin. Consider now Eqs. (6) and (7) in § 8. The forms $\bar{\omega}^i - t da_i$ and $\bar{\omega}^i{}_l$ and their first derivatives with respect to t all vanish for $t = 0$. Using Taylor's formula with remainder we conclude that

$$\bar{\omega}^i = t da_i + t^2\theta^i, \qquad \bar{\omega}^i{}_l = t^2\theta^i{}_l,$$

where θ^i and $\theta^i{}_l$ are 1-forms. Now let $\Gamma : s \to \Gamma(s)$ $(0 \leqslant s \leqslant 1)$ be any curve segment in $B_{r(p)}(p)$ joining a and b and not passing through p. Let Γ_0 be the curve segment in the ball $\| X \| < r(p)$ in M_p joining A and B such that $\Gamma = \operatorname{Exp} \circ \Gamma_0$. Then $\Gamma_0(s)$ can be written

$$\Gamma_0(s) = t(s)\, X(s) \qquad (0 \leqslant s \leqslant 1),$$

where $t(s) > 0$ for all s and $s \to X(s)$ is a curve segment on S. Then we have as before

$$g(\dot{\Gamma}(s), \dot{\Gamma}(s)) = \dot{t}(s)^2 + \sum_{i=1}^{m} (\bar{\omega}^i(\dot{X}(s)))^2$$

and

$$L(\Gamma) = \int_0^1 \left\{ \dot{t}(s)^2 + t(s)^2 \sum_i \left[da_i(\dot{X}(s)) + t(s)\, \theta^i(\dot{X}(s)) \right]^2 \right\}^{1/2} ds. \tag{7}$$

For the Riemannian manifold M_p we have $R = 0$ and $\bar{\omega}^i = t da_i$. Hence

$$L(\Gamma_0) = \int_0^1 \left\{ \dot{t}(s)^2 + (t(s))^2 \sum_i (da_i(\dot{X}(s)))^2 \right\}^{1/2} ds. \tag{8}$$

If Γ_0 is the straight line joining A and B, then $t(s) \to 0$ uniformly in s as $(a, b) \to (p, p)$. It follows from (7) and (8) that

$$\lim_{(a,b)\to(p,p)} \frac{L(\Gamma_0)}{L(\Gamma)} = 1. \tag{9}$$

This relation holds for the same reason for γ_{ab}, the unique geodesic in $B_{r(p)}(p)$ joining a and b, and γ_{AB}, the corresponding curve segment in the ball $\| X \| < r(p)$. In other words,

$$\lim_{(a,b)\to(p,p)} \frac{L(\gamma_{AB})}{L(\gamma_{ab})} = 1. \tag{10}$$

Now $L(\gamma_{AB}) \geqslant L(\Gamma_0)$ and $L(\Gamma) \geqslant L(\gamma_{ab})$ so

$$\frac{L(\gamma_{AB})}{L(\gamma_{ab})} \geqslant \frac{L(\Gamma_0)}{L(\gamma_{ab})} \geqslant \frac{L(\Gamma_0)}{L(\Gamma)}.$$

Since $L(\Gamma_0) = \| A - B \|$ and $L(\gamma_{ab}) = d(a, b)$, the proposition follows from (9) and (10).

§ 10. Complete Riemannian Manifolds

Definition. A Riemannian manifold M is said to be *complete* if every Cauchy sequence† in M is convergent.

Lemma 10.1. *For each point p_0 in a Riemannian manifold M there exists a convex normal ball $B_\rho(p_0)$ around p_0 with the following property: Let p and q be two points in $B_\rho(p_0)$, γ the unique geodesic in $B_\rho(p_0)$ joining p and q. Let $L(\gamma)$ denote the length of γ. Let Y and Z denote the unit tangent vectors to γ at p and q, respectively.*

† A sequence (x_n) in a metric space with metric d is called a Cauchy sequence if to each $\epsilon > 0$ there exists an integer N such that $d(x_n, x_m) < \epsilon$ for $n, m > N$.

Then:

(i) *If* (p', Y') *is sufficiently close to* (p, Y) *and* Y' *is a unit vector in* $M_{p'}$, *there exists a geodesic in* $B_\rho(p_0)$ *of length* $L(\gamma)$, *starting at* p' *with tangent vector* Y'.

(ii) *The pair* (q, Z) *depends differentiably on* p, Y *and* $L(\gamma)$.

This lemma follows directly from the existence and uniqueness theorem stated in the beginning of § 6.

Lemma 10.2. *Let* p *be a point in a Riemannian manifold* M *and* $\gamma_n(t)$ $(t \in I_n)$ *a sequence of geodesics emanating from* p, t *being the arc length measured from* p. *Let* X_n *be a tangent vector to* γ_n *at* p *and suppose the sequence* (X_n) *converges to* $X \in M_p$. *Let* $\gamma_X(t)$ $(t \in I)$ *be the maximal geodesic tangent to* X *such that* $\gamma_X(0) = p$. *Assume* $t^* \in I$ *is a limit* $t^* = \lim t_n$ *where* $t_n \in I_n$. *Then* $\gamma_X(t^*) = \lim_n \gamma_n(t_n)$.

In fact, the segment $\gamma_X(t)$, $0 \leqslant t \leqslant t^*$, can be broken into finitely many segments each of which lies in a convex normal ball B_i with the property of Lemma 10.1. The first part of Lemma 10.1 implies that for sufficiently large n, all $\gamma_n(t)$, $0 \leqslant t \leqslant t_n$, lie in the union of the balls B_i. Now Lemma 10.2 follows by repeated application of the second part of Lemma 10.1.

The following two theorems show clearly the importance of the completeness condition.

Theorem 10.3. *Let* M *be a Riemannian manifold. The following conditions are equivalent.*

(i) M *is complete.*

(ii) *Each bounded closed subset of* M *is compact.*

(iii) *Each maximal geodesic in* M *has the form* $\gamma_X(t)$, $-\infty < t < \infty$ (*"has infinite length"*).

Theorem 10.4. *In a complete Riemannian manifold* M *with metric* d *each pair* $p, q \in M$ *can be joined by a geodesic of length* $d(p, q)$.

These two theorems will be proved simultaneously.

(i) $\Rightarrow$ (iii). Let $\gamma_X(t)$ $(t \in I)$ be a maximal geodesic in M, $|t|$ being the arc length measured from the point of origin of X. If t_0 were a boundary point of the (open) interval I, say on the right, select a sequence $(t_n) \subset I$ converging to t_0. Then $(\gamma_X(t_n))$ is a Cauchy sequence in M, hence converges to a limit $p \in M$, which is clearly independent of the choice of (t_n). Let $B_\rho(p)$ be a convex normal ball around p and let $\{x_1, \ldots, x_m\}$ be a system of normal coordinates at p valid on $B_\rho(p)$. Let $J = \{t \in I : \gamma_X(t) \in B_\rho(p)\}$, put $x_i(t) = x_i(\gamma_X(t))$ $(t \in J)$ and $x_i(t_0) = \lim_{t \to t_0} x_i(t)$, knowing that this limit exists.

Now we have

$$\ddot{x}_i(t) + \sum_{j,k} \Gamma_{jk}{}^i \, \dot{x}_j(t) \, \dot{x}_k(t) = 0 \tag{1}$$

for $t \in J$. The functions $\dot{x}_i(t)$ are bounded $(\sum_i \dot{x}_i(t)^2 = 1)$ and the differential equation shows that each $\ddot{x}_i(t)$ is bounded. In particular, $\dot{x}_i(t)$ is uniformly continuous near t_0 and thus has a limit as $t \to t_0$. From the mean value theorem we have

$$(x_i(t) - x_i(t_0))/(t - t_0) = \dot{x}_i(t_0 + \theta(t - t_0)) \qquad (0 \leqslant \theta \leqslant 1)$$

which implies for the left derivative $\dot{x}_i(t_0)$,

$$\dot{x}_i(t_0) = \lim_{t \to t_0} \dot{x}_i(t) \qquad (1 \leqslant i \leqslant m). \tag{2}$$

From the differential equation (1) follows the existence of the limit $\lim_{t \to t_0} \ddot{x}_i(t)$; the mean value theorem again implies for the left derivative

$$\ddot{x}_i(t_0) = \lim_{t \to t_0} \ddot{x}_i(t). \tag{3}$$

The vector $Z = (\dot{x}_1(t_0), \ldots, \dot{x}_m(t_0))$ in the tangent space M_p has length 1 and we can form the geodesic $\gamma_Z(t)$ for $t_0 \leqslant t < t_0 + \rho$. The mapping $t \to \Gamma(t)$ where

$$\Gamma(t) = \begin{cases} \gamma_X(t), & t \in I, \\ \gamma_Z(t), & t_0 \leqslant t < t_0 + \rho, \end{cases}$$

satisfies (1) for $t \in J$ and for $t_0 < t < t_0 + \rho$. Moreover, (1) is satisfied for the *right derivatives* at $t = t_0$. Equations (2) and (3) show that (1) is also satisfied for the left derivatives at $t = t_0$. Thus, $\Gamma(t)$ is a geodesic, contradicting the maximality of $\gamma_X(t)$, $t \in I$.

(ii) $\Rightarrow$ (i). Let (x_n) be a Cauchy sequence in M. The closure of the set (x_n) is bounded, hence compact by (ii). Thus a subsequence of (x_n) is convergent; being a Cauchy sequence, the sequence (x_n) itself is convergent.

Using a procedure of de Rham [1] we next prove that if the condition (iii) is satisfied, then each pair $p, q \in M$ can be joined by a geodesic of length $d(p, q)$. For each $r \geqslant 0$ let $\bar{B}_r$ denote the closure of the open ball $B_r(p)$ and let E_r denote the set of points x in $\bar{B}_r$ which can be joined to p by a geodesic of length $d(p, x)$. It suffices to prove

$$E_r = \bar{B}_r \tag{4}$$

for each $r \geqslant 0$. In view of Lemma 10.2, (iii) implies separability of M and compactness of E_r. The relation (4) is valid for $r = 0$, and if it is valid for $r = r_0 > 0$, it is obviously valid for $r < r_0$. On the other hand, if (4) holds for $r < r_0$, it is also valid for $r = r_0$; in fact, each $x \in \bar{B}_{r_0}$ is a limit of a sequence of points x_n each of which has distance from p less than r_0. Thus, $x_n \in E_{r_0}$ and $x \in E_{r_0}$, the set E_{r_0} being closed. Thus, it suffices to prove that if (4) holds for $r = R$, it holds also for some larger value $r = R + \rho$. For this we may assume $B_R \neq M$.

By compactness of $E_R = \bar{B}_R$ there exist finitely many points $x_1, \ldots, x_N \in E_R$ and positive numbers $\rho_1, \ldots, \rho_N$ such that the balls $B_{\rho_i}(x_i)$ $(1 \leqslant i \leqslant N)$ cover E_R and such that each $B_{2\rho_i}(x_i)$ is a relatively compact, minimizing convex normal ball. Since the set $\bigcup_{i=1}^N B_{\rho_i}(x_i)$ is relatively compact there exists a point in its complement at shortest distance from p. Since this distance is $> R$, there exists a number ρ such that $0 < \rho < \min(\rho_1, \ldots, \rho_N)$ and such that $\bar{B}_{R+\rho} \subset \bigcup_{i=1}^N B_{\rho_i}(x_i)$.

Suppose now y is a point in M such that $R < d(p, y) \leqslant R + \rho$. Let x be a point on the (compact) sphere $S_R(p)$ at smallest distance from y. Since every curve segment joining p and y must intersect $S_R(p)$, it follows that

$$d(p, x) + d(x, y) = d(p, y).$$

Consequently, $d(x, y) = d(p, y) - d(p, x) \leqslant R + \rho - R = \rho$ so y and x lie in the same ball $B_{2\rho_i}(x_i)$. Combining the shortest curve joining x and y with a curve of shortest length joining p and x, we obtain a curve of length $d(p, y)$ joining p and y. By Lemma 9.8 this curve is a geodesic so (4) is proved for all $r \geqslant 0$.

(iii) $\Rightarrow$ (ii). Let S be a bounded closed subset of M. Let (q_n) be a sequence of points in S, and p any fixed point in M. We know now that M is separable and that there exists a geodesic γ_n of length $d(p, q_n)$ joining p and q_n. Passing to a subsequence if necessary, we can assume that the unit tangent vectors to γ_n at p form a convergent sequence and that the sequence $(d(p, q_n))$ converges. Lemma 10.2 now shows that (iii) $\Rightarrow$ (ii). This concludes the proof of Theorems 10.3 and 10.4.

Remark. Call a Riemannian manifold M *complete at a point* $p \in M$ if Exp_p is defined on the entire M_p. Then the proof above shows that completeness at a single point $p \in M$ implies completeness of M.

Proposition 10.5. *For each point p in a complete Riemannian manifold M the Exponential mapping* Exp_p *is a differentiable mapping of* M_p *onto M. If M is analytic, then* Exp_p *is analytic.*

Proof. Lemma 10.2 expresses the continuity of Exp_p. The differentiability is proved in the same way (applying Lemma 10.1) since we now know that Exp_p is defined on the entire M_p. If M is analytic we first observe that the existence and uniqueness theorem in § 6 also holds for the analytic case giving analytic solutions. Thus we can replace "differentiably" in Lemma 10.1 by "analytically" and proceed as before.

Definition. Let M be a complete Riemannian manifold, p a point in M and Exp the Exponential mapping at p. Let $C(p)$ denote the set of vectors $X \in M_p$ for which the linear mapping $d\,\mathrm{Exp}_X$ is singular. A point in M (or in M_p) is said to be *conjugate* to p if it lies in Exp $C(p)$ (or $C(p)$).

In view of Prop. 10.5 the set $C(p)$ is a closed subset of M_p. It plays an important role in global differential geometry. In general, Exp $C(p)$ is not a submanifold of M. For the sphere S^2, the set Exp $C(p)$ consists of two antipodal points. In Chapter VII we shall prove the inequality

$$\dim (\mathrm{Exp}\ C(p)) \leqslant \dim M - 2$$

for a symmetric space of the compact type, dim denoting topological dimension. However, this inequality fails to hold for general Riemannian manifolds.

Let M and N be connected topological spaces and π a continuous mapping of M into N. The pair (M, π) is called a *covering space* of N if each point $n \in N$ has an open neighborhood U such that each component of $\pi^{-1}(U)$ is homeomorphic to U under π.

Suppose N is a differentiable manifold and that (M, π) is a covering space of N. Then there is a unique differentiable structure on M such that the mapping π is regular. If M is given this differentiable structure, we say that (M, π) is a *covering manifold* of N.

We shall require the following standard theorem from the theory of covering spaces. We state it only for manifolds although it holds under suitable local connectedness hypotheses.

Let (M, π) be a covering manifold of N and let $\Gamma : [a, b] \to N$ be a path in N. If m is any point in M such that $\pi(m) = \Gamma(a)$, there exists a unique path $\Gamma^* : [a, b] \to M$ such that $\Gamma^*(a) = m$ and $\pi \circ \Gamma^* = \Gamma$.

The path Γ^* is called the *lift* of Γ through m.

Proposition 10.6. *Let N be a Riemannian manifold with a Riemannian structure g. Let (M, π) be a covering manifold of N. Then π^*g is a Riemannian structure on M. Moreover, M is complete if and only if N is complete.*

Proof. The mapping π is regular, so obviously π^*g is a Riemannian structure on M. If γ is a curve segment in M, then $\pi \circ \gamma$ is a curve segment in N. Using the characterization of geodesics by means of differential equations (3), (§ 5), it is clear that γ is a geodesic if and only if $\pi \circ \gamma$ is a geodesic. But completeness is equivalent to the infiniteness of each maximal geodesic. The proposition follows immediately.

Remark. Let N be a Riemannian manifold, (M, π) a covering manifold of N, M taken with Riemannian structure induced by π. Given a point $a \in N$, there exists a convex, minimizing normal ball $B_r(a)$ such that each component B of $\pi^{-1}(B_r(a))$ is diffeomorphic to $B_r(a)$ under π. Since the geodesics correspond under π, it follows that π is a distance-preserving mapping of B onto $B_r(a)$.

Definition. A geodesic $\gamma(t)$, $0 \leqslant t < \infty$, in a Riemannian manifold is called a *ray* if it realizes the shortest distance between any two of its points. The point $\gamma(0)$ is called the *initial point* of the ray.

Proposition 10.7. *Let o be a point in a complete, noncompact Riemannian manifold M. Then M contains a ray with initial point o.*

It follows from Theorem 10.3 that M is not bounded. Let (p_n) be a sequence in M such that $d(o, p_n) \to \infty$. Let γ_n be a geodesic of length $d(o, p_n)$ joining o and p_n. We parametrize γ_n by arc length measured from o. Let X_n be the unit tangent vector to γ_n at o. We can assume that the sequence (X_n) converges to a limit $X \in M_o$. Then $t \to \gamma_X(t)$ $(0 \leqslant t < \infty)$ is a ray. In fact, let $t_0 \geqslant 0$. There exists an integer N such that $d(o, p_n) > t_0$ for $n \geqslant N$. We have from Lemma 10.2

$$\lim_{n > N, n \to \infty} \gamma_n(t_0) = \gamma_X(t_0),$$

and consequently

$$\lim_{n > N, n \to \infty} d(o, \gamma_n(t_0)) = d(o, \gamma_X(t_0)).$$

On the other hand, $t_0 = d(o, \gamma_n(t_0))$ for $n \geqslant N$. The last relation therefore shows that $t_0 = d(o, \gamma_X(t_0))$. This implies that $d(\gamma_X(t_1), \gamma_X(t_0)) = t_0 - t_1$ for $0 \leqslant t_1 \leqslant t_0$, proving the proposition.

§ 11. Isometries

Definition. Let M and N be two C^∞ manifolds with pseudo-Riemannian structures g and h, respectively. Let φ be a mapping of M into N.

(i) φ is called an *isometry* if φ is a diffeomorphism of M onto N and $\varphi^* h = g$.

(ii) φ is called a *local isometry* if for each $p \in M$ there exist open neighborhoods U of p and V of $\varphi(p)$ such that φ is an isometry of U onto V.

It is obvious that if φ is an isometry of a Riemannian manifold M onto itself, then φ preserves distances, i. e., $d(\varphi(p), \varphi(q)) = d(p, q)$ for $p, q \in M$. On the other hand, we have:

Theorem 11.1. *Let M be a Riemannian manifold and φ a distance-preserving mapping of M onto itself. Then φ is an isometry.*

Proof. Let p be an arbitrary point in M and put $q = \varphi(p)$. Let $B_r(p)$ and $B_r(q)$ be spherical normal neighborhoods of $p \in M$ and $q \in M$, respectively. Then φ gives a one-to-one mapping of $B_r(p)$ onto $B_r(q)$. For each $X \in M_p$ we consider the geodesic $\operatorname{Exp} tX$, $(-r/\|X\| < t < r/\|X\|)$. The image $\Gamma(t) = \varphi(\operatorname{Exp} tX)$ lies in $B_r(q)$ and has the property that $d(\Gamma(t), \Gamma(t')) = |t - t'| \, \|X\|$ for all t, t' in the interval of definition. To see that Γ is a geodesic we consider the point $q = \Gamma(0)$ and an arbitrary point Q on Γ. They can be joined by a unique geodesic γ of length $d(q, Q)$. Let $B_R(Q)$ be a spherical normal neighborhood of Q and let m be any point of Γ between q and Q such that $m \in B_R(Q)$. Then $d(q, m) + d(m, Q) = d(q, Q)$. If we join q and m by the shortest geodesic, and then join m and Q by the shortest geodesic, we get a broken curve of length $d(q, Q)$ joining q and Q. This curve must coincide with γ due to Lemma 9.8 (and the subsequent remark). Since Q was arbitrary on Γ, this proves that Γ is a geodesic; in particular, Γ is differentiable. Let X' denote the tangent vector to Γ at the point q. We have obtained a mapping $X \to X'$ of M_p into M_q. Denoting this mapping by φ' we have $\|X\| = \|\varphi'(X)\|$ and $\varphi'(\alpha X) = \alpha\varphi'(X)$ for $\alpha \in \boldsymbol{R}$, $X \in M_p$. Let $A, B \in M_p$ and select ρ such that $\|\rho A\|$ and $\|\rho B\|$ are both less than r. Let $a_t = \operatorname{Exp} tA$, $b_t = \operatorname{Exp} tB$ for $0 \leqslant t \leqslant \rho$. Then Prop. 9.10 shows that

$$\frac{2g_p(A, B)}{\|A\| \, \|B\|} = \frac{\|A\|^2 + \|B\|^2}{\|A\| \, \|B\|} - \frac{\|tA - tB\|^2}{\|tA\| \, \|tB\|}$$

$$= \frac{\|A\|^2 + \|B\|^2}{\|A\| \, \|B\|} - \lim_{t \to 0} \frac{d(a_t, b_t)^2}{\|tA\| \, \|tB\|}.$$

Since the right-hand side is preserved by the mapping φ, it follows that

$$g_p(A, B) = g_q(\varphi' A, \varphi' B).$$

But $A + B$ is determined by the quantities $\|A\|$, $\|B\|$, and $g_p(A, B)$, all of which are preserved by φ'. It follows that $\varphi'(A + B) = \varphi' A + \varphi' B$ which together with the previous properties of φ' shows that it is a diffeomorphism of M_p onto M_q. On $B_r(p)$ we have

$$\varphi = \operatorname{Exp}_q \circ \varphi' \circ \operatorname{Exp}_p^{-1} \tag{1}$$

and the theorem follows.

Lemma 11.2. *Let M be a Riemannian manifold, φ and ψ two isometries of M onto itself. Suppose there exists a point $p \in M$ for which $\varphi(p) = \psi(p)$ and $d\varphi_p = d\psi_p$. Then $\varphi = \psi$.*

Proof. Considering $\varphi \circ \psi^{-1}$ it is clear that we may assume that $\varphi(p) = p$ and that $d\varphi_p$ is the identity mapping. It is then obvious that all points in an arbitrary normal neighborhood of p are left fixed by φ. Since M is connected, each point $q \in M$ can be connected to p by a chain of overlapping normal neighborhoods. It follows that $\varphi(q) = q$.

Let M and N be Riemannian manifolds, U_i a domain in M $(i = 1, 2)$, and φ_i an isometry of U_i onto a domain in N $(i = 1, 2)$. It follows from Lemma 11.2 that if $(d\varphi_1)_p = (d\varphi_2)_p$ and $\varphi_1(p) = \varphi_2(p)$ for some point $p \in U_1 \cap U_2$, then φ_1 and φ_2 coincide on the component of p in $U_1 \cap U_2$. The isometries φ_1 and φ_2 are called *immediate continuations* if $U_1 \cap U_2 \neq \emptyset$ and $\varphi_1 = \varphi_2$ on $U_1 \cap U_2$.

Let φ be an isometry of a domain $U \subset M$ onto a domain in N. Let $\gamma(t)$, $0 \leqslant t \leqslant 1$, be a continuous curve in M such that $\gamma(0) \in U$. The isometry φ is said to be *extendable along* γ if for each t $(0 \leqslant t \leqslant 1)$ there exists an isometry φ_t of a domain U_t containing $\gamma(t)$ onto an open subset of N such that $\varphi_0 = \varphi$ and $U_0 = U$ and such that φ_t and $\varphi_{t'}$ are immediate continuations whenever $|t - t'|$ is sufficiently small. The family φ_t, $0 \leqslant t \leqslant 1$, is called a *continuation of* φ *along* γ. It follows that the differential $(d\varphi_t)_{\gamma(t)}$ as well as $\varphi_t(\gamma(t))$ depends continuously on t. Suppose now ψ_t, $0 \leqslant t \leqslant 1$, is another continuation of φ along γ and V_t the domain of definition of ψ_t. From the foregoing remarks it follows that the set of t for which $\varphi_t(\gamma(t)) = \psi_t(\gamma(t))$ and $(d\varphi_t)_{\gamma(t)} = (d\psi_t)_{\gamma(t)}$ is an open and closed subset of the unit interval and contains $t = 0$. Thus, for each t, φ_t and ψ_t coincide in the component of $\gamma(t)$ in $U_t \cap V_t$. Roughly speaking, we can therefore say in analogy with analytic continuation of holomorphic functions: the continuation of an isometry along a curve is unique whenever possible.

Proposition 11.3. *Let M and N be analytic and complete Riemannian manifolds, φ an isometry of a domain $U \subset M$ onto a domain in N. Let $\gamma(t)$, $0 \leqslant t \leqslant 1$, be a continuous curve in M such that $\gamma(0) \in U$. Then φ is extendable along γ.*

Proof. Let $p \in M$, $q \in N$, and suppose $B_\rho(p)$ and $B_\rho(q)$ are spherical normal neighborhoods of p and q, respectively. Suppose $r < \rho$ and suppose ψ is an isometry of $B_r(p)$ into N such that $\psi(p) = q$. Then ψ can be extended uniquely to an isometry ψ' of $B_\rho(p)$ onto $B_\rho(q)$. To see this, we note first that the expression of ψ in normal coordinates at p and $\psi(p)$ is a linear mapping which we can use to define ψ'. If g and h

denote the metric tensors on M and N, respectively, and if X and Y are analytic vector fields on $B_\rho(p)$, then $h(X^{\psi'}, Y^{\psi'}) \circ \psi' = g(X, Y)$ on $B_r(p)$ by the assumption; this relation also holds on $B_\rho(p)$ due to the analyticity of ψ' (see Lemma 4.3, Chapter VI). Hence ψ' is an isometry.

Without restriction of generality we can assume $\gamma(t)$ differentiable. Let s^* be the supremum of the parameter values s such that a continuation φ_t of φ exists along the curve $\gamma(t)$, $0 \leqslant t \leqslant s$. We put $p_t = \gamma(t)$ for $0 \leqslant t \leqslant 1$, and $q_t = \varphi_t(p_t)$ for $0 \leqslant t < s^*$. It is clear from the completeness assumption that the limit $q^* = \lim_{t \to s^*} \varphi_t(p_t)$ exists. We select $\rho > 0$ such that $B_{3\rho}(p_{s^*})$ and $B_{3\rho}(q^*)$ are convex normal balls. Select $s' < s^*$ such that

$$p_t \in B_\rho(p_{s^*}), \qquad q_t \in B_\rho(q^*) \qquad \text{for } s' \leqslant t < s^*.$$

Then $B_{2\rho}(p_{s'})$ and $B_{2\rho}(q_{s'})$ are normal neighborhoods of $p_{s'}$ and $q_{s'}$, respectively. As shown above, $\varphi_{s'}$ can be extended to an isometry of $B_{2\rho}(p_{s'})$ onto $B_{2\rho}(q_{s'})$. Since $p^* \in B_{2\rho}(p_{s'})$, this shows that $s^* = 1$ and that φ has a continuation along γ.

Proposition 11.4. *Let the assumptions be as in Prop. 11.3 and suppose $\delta(t)$, $0 \leqslant t \leqslant 1$, is a continuous curve in M, homotopic to γ. Let φ_t and ψ_t $(0 \leqslant t \leqslant 1)$ be continuations of φ along γ and δ, respectively. Then φ_1 and ψ_1 coincide in a neighborhood of $\gamma(1) = \delta(1)$.*

Proof. Since γ and δ are homotopic, there exists a continuous mapping α of the closed unit square into M such that

$$\alpha(0, t) = \gamma(t), \qquad 0 \leqslant t \leqslant 1,$$

$$\alpha(1, t) = \delta(t), \qquad 0 \leqslant t \leqslant 1;$$

$$\alpha(s, 0) = \gamma(0), \qquad \alpha(s, 1) = \gamma(1) \qquad \text{for all } 0 \leqslant s \leqslant 1.$$

For a fixed s, let α^s denote the continuous curve $t \to \alpha(s, t)$ $(0 \leqslant t \leqslant 1)$, and let φ_t^s $(0 \leqslant t \leqslant 1)$ denote the continuation of φ along α^s. Let σ denote the supremum of the values s^* such that for each s satisfying $0 \leqslant s \leqslant s^*$, φ_1^s coincides with φ_1 in a neighborhood of $\gamma(1)$. Consider now the continuous curve α^σ. The mapping $t \to \varphi_t^\sigma(\alpha^\sigma(t))$ is also a continuous curve. Hence there exists a number $r > 0$ such that for each t, $0 \leqslant t \leqslant 1$, the balls $B_{2r}(\alpha^\sigma(t))$ and $B_{2r}(\varphi_t^\sigma(\alpha^\sigma(t)))$ are normal neighborhoods of their centers. Now, there exists an $\epsilon > 0$ such that $d(\alpha^\sigma(t), \alpha^s(t)) < r$ for $0 \leqslant t \leqslant 1$, and $|\sigma - s| < \epsilon$. For such s, the family φ_t^σ gives a continuation of φ along α^s. Now as remarked before, the continuation of an isometry along a given curve is unique. It follows that if $|s - \sigma| < \epsilon$, the isometries φ_1^σ and φ_1^s coincide in a neighborhood of $\gamma(1)$. This shows firstly that $\sigma = 1$ and secondly that if $0 \leqslant s \leqslant 1$, then φ_1^1 and φ_1^s coincide in a neighborhood of $\gamma(1)$. This proves the proposition.

§ 12. Sectional Curvature

In this section we shall exhibit the classical geometric significance of the curvature tensor for a Riemannian manifold.

Let F be a Riemannian manifold of dimension 2 and let p be a point in F. Let $V_r(0)$ denote the open ball in the tangent plane F_p with center 0 and radius r. Suppose r is so small that Exp_p is a diffeomorphism of $V_r(0)$ onto the open ball $B_r(p)$. Let $A_0(r)$ and $A(r)$ denote the (two-dimensional) areas of $V_r(0)$ and $B_r(p)$, respectively.

Definition. The *curvature*† of F at p is defined as the limit

$$K = \lim_{r \to 0} 12 \frac{A_0(r) - A(r)}{r^2 A_0(r)}.$$

The existence of this limit is contained in the following lemma, which at the same time facilitates the computation of K.

Lemma 12.1. *Let f denote the "Radon-Nikodym derivative" of Exp_p on $V_r(0)$. Then*

$$K = -\frac{3}{2} [\Delta f](0)$$

where Δ is the Laplacian on the metric space F_p, that is, $\Delta = \partial^2/\partial x_1^2 + \partial^2/\partial x_2^2$, x_1 and x_2 being coordinates with respect to some orthonormal basis.

Proof. The definition of $f(X)$ is expressed in the formula $\operatorname{Exp}_p^*(dq) = f\,dX$ if dq and dX denote the surface elements (Chapter VIII, §2) of $B_r(p)$ and $V_r(0)$, respectively. Hence

$$A(r) = \int_{V_r(0)} f(X)\, dX.$$

The differentiable function $f(X)$ can be expanded in a finite Taylor series around the point $X = 0$. If this series is integrated over the disk $V_r(0)$ one finds

$$A(r) = A_0(r) \{f(0) + \tfrac{1}{8} r^2 [\Delta f](0) + O(r^3)\}.$$

The lemma now follows immediately, $f(0)$ being equal to 1.

† It is known from the differential geometry of surfaces that if F is a surface, then K is equal to the Gaussian curvature of F at p, but we shall not need this fact.

Let M be a Riemannian manifold and p a point in M. Let N_0 be a normal neighborhood of 0 in M_p and let $N_p = \operatorname{Exp} N_0$. Let S be a two-dimensional vector subspace of M_p. Then $\operatorname{Exp}(N_0 \cap S)$ is a connected submanifold of M of dimension 2 and has a Riemannian structure induced by that of M. The curvature of $\operatorname{Exp}(N_0 \cap S)$ at p is called the *sectional curvature* of M at p along the *plane section* S.

Theorem 12.2. *Let M be a Riemannian manifold with curvature tensor field R and Riemannian structure g. Let p be a point in M and S a two-dimensional vector subspace of the tangent space M_p. The sectional curvature of M at p along the section S is then*

$$K(S) = -\frac{g_p(R_p(Y, Z)\, Y, Z)}{|\, Y \vee Z\,|^2}.$$

Here Y and Z are any linearly independent vectors in S; $Y \vee Z$ denotes the parallelogram spanned by these vectors and $|\, Y \vee Z\,|$ the area.

Proof. We shall first assume that M and g are analytic in order to apply Theorem 6.5. We also assume temporarily that Y and Z are orthonormal vectors in S. Let $X_1, ..., X_m$ be an orthonormal basis of M_p such that $X_1 = Y$ and $X_2 = Z$. Then each $X \in S$ can be written $X = x_1X_1 + x_2X_2$, $x_1, x_2 \in \boldsymbol{R}$. The Laplacian Δ on S is

$$\Delta = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2}.$$

Let N_0 be a normal neighborhood of 0 in M_p and put $N_p = \operatorname{Exp} N_0$. A curve in the manifold $M_S = \operatorname{Exp}(S \cap N_0)$ has the same length regardless whether it is measured by means of the Riemannian structure on M or by means of the induced structure on M_S. If $q \in M_S$, the geodesic in N_p from p to q is the shortest curve in M_S joining p and q. It follows that the Exponential mappings at p for M and M_S, respectively, coincide on $S \cap N_0$. Let $X_1^*, ..., X_m^*, X^*$ denote the vector fields on N_p adapted to the tangent vectors $X_1, ..., X_m, X$. If $X \in S \cap N_0$, we put

$$v_1 = d\operatorname{Exp}_X(X_1), \qquad v_2 = d\operatorname{Exp}_X(X_2),$$

and define the functions $c^k{}_{ij}$ by

$$[X_i^*, X_j^*] = \sum_k c^k{}_{ij} X_k^*. \tag{1}$$

The mapping $\operatorname{Exp}(x_1X_1 + x_2X_2) \to (x_1, x_2)$ is a system of coordinates on the manifold $M_S = \operatorname{Exp}(S \cap N_0)$ and v_1 and v_2 are tangent vectors to M_S. If a and b are two vectors in a metric vector space, we denote by $a \vee b$ the parallelogram spanned by a and b and by $| a \vee b |$ the area. Let f denote the ratio of the surface elements in M_S and S (at the points Exp X and X). In other words,

$$f(X) = \frac{| v_1 \vee v_2 |}{| X_1 \vee X_2 |} = | v_1 \vee v_2 |.$$

The vectors v_1 and v_2 can be expressed

$$v_1 = \sum_{i=1}^{m} f_i X_i^*, \qquad v_2 = \sum_{j=1}^{m} g_j X_j^*,$$

where f_i and g_j are analytic functions of (x_1, x_2). These functions are determined by Theorem 6.5 for small (x_1, x_2). In fact, we have

$$v_1 = X_1^* - \tfrac{1}{2}[X^*, X_1^*] + \tfrac{1}{6}[X^*, [X^*, X_1^*]] - \dots \tag{2}$$

$$v_2 = X_2^* - \tfrac{1}{2}[X^*, X_2^*] + \tfrac{1}{6}[X^*, [X^*, X_2^*]] - \dots. \tag{2'}$$

The vectors $(X_i^*)_{\text{Exp}\, X}$ $(1 \leqslant i \leqslant m)$ form an orthonormal basis of $M_{\text{Exp}\, X}$ (Theorem 9.1 (ii)). The projection of $v_1 \vee v_2$ into the 2-plane spanned by the vectors $(X_i^*)_{\text{Exp}\, X}$ and $(X_j^*)_{\text{Exp}\, X}$ has area $| f_i g_j - g_i f_j |$. It follows that

$$| v_1 \vee v_2 |^2 = \sum_{i<j} (f_i g_j - f_j g_i)^2. \tag{3}$$

We denote this quantity by F. The relation $f = F^{1/2}$ implies

$$2f\Delta f = \Delta F - \frac{1}{2f^2}\left\{\left(\frac{\partial F}{\partial x_1}\right)^2 + \left(\frac{\partial F}{\partial x_2}\right)^2\right\}.$$

We have to evaluate this expression for $(x_1, x_2) = (0, 0)$. Since the torsion T vanishes, we have

$$[X_i^*, X_j^*] = \nabla_{X_i^*}(X_j^*) - \nabla_{X_j^*}(X_i^*)$$

and consequently the functions $c^k{}_{ij}$ vanish at p; in other words, the restrictions of $c^k{}_{ij}$ to M_S vanish for $(x_1, x_2) = (0, 0)$. From (2) and (2′) we obtain expansions for f_i, g_j:

$$f_i = \delta_{1i} - \frac{x_2}{2} c^i{}_{21} + \frac{x_1 x_2}{6}(X_1^* \cdot c^i{}_{21}) + \frac{x_2^2}{6}(X_2^* \cdot c^i{}_{21}) + \dots,$$

$$g_j = \delta_{2j} - \frac{x_1}{2} c^j{}_{12} + \frac{x_1 x_2}{6}(X_2^* \cdot c^j{}_{12}) + \frac{x_1^2}{6}(X_1^* \cdot c^j{}_{12}) + \dots,$$

where the terms which are not written vanish for $(x_1, x_2) = (0, 0)$ of higher than second order. It follows easily that $\partial F/\partial x_1$ and $\partial F/\partial x_2$ vanish for $(x_1, x_2) = (0, 0)$ and

$$2[\Delta f](0) = [\Delta F](0) = [\Delta (f_1 g_2)^2](0).$$

Omitting again terms of higher than 2nd order we have

$$(f_1 g_2)^2$$
$$= 1 - x_1 c^2{}_{12} - x_2 c^1{}_{21} + \frac{x_1^2}{3}(X_1^* c^2{}_{12}) + \frac{x_2^2}{3}(X_2^* c^1{}_{21}) + \frac{x_1 x_2}{3}(X_1^* c^1{}_{21} + X_2^* c^2{}_{12}).$$

Since

$$X_1 c^2{}_{12} = [X_1^* c^2{}_{12}](0) = \left[\frac{\partial}{\partial x_1} c^2{}_{12}\right](0), \quad \text{etc.},$$

we obtain

$$2[\Delta f](0) = -\frac{4}{3}(X_1 c^2{}_{12} + X_2 c^1{}_{21}),$$

and by Lemma 12.1

$$K(S) = X_1 \cdot g([X_1^*, X_2^*], X_2^*) + X_2 \cdot g([X_2^*, X_1^*], X_1^*). \tag{4}$$

On the other hand, using $[X_i^*, X_j^*]_p = 0$ and the formulas from Theorem 9.1, we get

$$\begin{aligned} -g_p(R_p(Y, Z)\,Y, Z) &= g_p(\nabla_{X_2^*}\nabla_{X_1^*} \cdot X_1^*, X_2^*) - g_p(\nabla_{X_1^*}\nabla_{X_2^*} \cdot X_1^*, X_2^*) \\ &= X_2 \cdot g(\nabla_{X_1^*} \cdot X_1^*, X_2^*) - X_1 \cdot g(\nabla_{X_2^*} \cdot X_1^*, X_2^*) \\ &\quad - g_p(\nabla_{X_1^*} \cdot X_1^*, \nabla_{X_2^*} \cdot X_2^*) + g_p(\nabla_{X_2^*} \cdot X_1^*, \nabla_{X_1^*} \cdot X_2^*). \end{aligned}$$

The two last terms vanish since in general $(\nabla_{X_i^*} \cdot X_j^*)_p = 0$. For the two other terms we use (2), § 9. Since $g(X_j^*, X_k^*)$ is constant we obtain

$$-g_p(R_p(Y, Z)\,Y, Z) = X_2 \cdot g([X_2^*, X_1^*], X_1^*) + X_1 \cdot g([X_1^*, X_2^*], X_2^*).$$

In view of (4) this proves Theorem 12.2 in the analytic case for Y, Z orthonormal. If Y, Z are linearly independent but not necessarily orthonormal, we can write $A = y_1 Y + z_1 Z$, $B = y_2 Y + z_2 Z$ where A, B are orthonormal vectors in S. Anticipating the first and third relations of Lemma 12.5 we find

$$\begin{aligned} K(S) &= -g_p(R_p(A, B)\,A, B) \\ &= -g_p(R_p(y_1 Y + z_1 Z, y_2 Y + z_2 Z) \cdot (y_1 Y + z_1 Z), y_2 Y + z_2 Z) \\ &= -(y_1 z_2 - y_2 z_1)^2 g_p(R_p(Y, Z)\,Y, Z) \\ &= -\frac{g_p(R_p(Y, Z)\,Y, Z)}{|Y \vee Z|^2}. \end{aligned}$$

Finally we drop the analyticity assumption. Let $\{x_1, \ldots, x_m\}$ be a system of coordinates valid on an open neighborhood U of p, such that $(\partial/\partial x_1)_p = Y$, $(\partial/\partial x_2)_p = Z$. Consider the function g_{ij} defined by

$$g_{ij} = g\left(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}\right).$$

Then for each $q \in U$, the matrix $(g_{ij})_q$ is symmetric and strictly positive definite. There exist analytic functions $\gamma_{ij} = \gamma_{ji}$ on an open set $V (p \in V \subset U)$ whose derivatives of order $0 \leqslant k \leqslant 3$ approximate those of g_{ij} as well as we please. For sufficiently good approximation the matrix $(\gamma_{ij})_q$ will be symmetric and strictly positive definite for each $q \in V$; we get an analytic Riemannian structure γ on V by requiring

$$\gamma\left(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}\right) = \gamma_{ij} \qquad (1 \leqslant i, j \leqslant m).$$

The sectional curvature and the curvature tensor field for γ approximate the corresponding sectional curvature and the curvature tensor field for g. Since Theorem 12.2 holds for the Riemannian manifold induced by the Riemannian structure γ on V, the theorem holds also for g.

The next proposition shows, that in a certain sense, the sectional curvature determines the curvature tensor.

Proposition 12.3. *Let M be a Riemannian manifold, p a point in M. Let g and g' be two Riemannian structures on M, R and R' the corresponding curvature tensors, and $K(S)$ and $K'(S)$ the corresponding sectional curvatures at p along a plane section $S \subset M_p$.*

Suppose that $g_p = g'_p$. If $K(S) = K'(S)$ for all plane sections $S \subset M_p$, then $R_p = R'_p$.

We first prove two simple lemmas.

Lemma 12.4. *Let A be a ring with identity element e such that $6a \neq 0$ for $a \neq 0$ in A. Let E be a module over A. Suppose a mapping $B : E \times E \times E \times E \to A$ is quadrilinear and satisfies the identities*

(a) $B(X, Y, Z, T) = -B(Y, X, Z, T)$.

(b) $B(X, Y, Z, T) = -B(X, Y, T, Z)$.

(c) $B(X, Y, Z, T) + B(Y, Z, X, T) + B(Z, X, Y, T) = 0$.

Then

(d) $B(X, Y, Z, T) = B(Z, T, X, Y)$.

If, in addition to (a), (b), and *(c), B satisfies*

(e) $B(X, Y, X, Y) = 0$ *for all* $X, Y \in E$,

then $B = 0$.

Proof. First we interchange T in (c) with X, Y, Z, respectively, and add the four obtained relations. Using (a) and (b) one obtains

$$(f) \qquad B(T, X, Y, Z) + B(T, Y, Z, X) + B(T, Z, X, Y) = 0.$$

From (c) and (a) it follows that

$$B(Z, X, T, Y) = + B(T, X, Z, Y) - B(T, Z, X, Y).$$

On substituting in (f), relation (d) follows. In particular, $B(X, Y, X, T)$ is symmetric in Y and T. Thus (e) implies $B(X, Y, X, T) = 0$. In view of (a) and (b) this implies that B is alternate; then (c) shows at once that $B = 0$.

Lemma 12.5. *Let M be a manifold with an affine connection ∇. Let R and T denote the curvature tensor and torsion tensor, respectively. Then R satisfies the following identities*

$$R(X, Y) = - R(Y, X).$$

If $T = 0$, then

$$R(X, Y) \cdot Z + R(Y, Z) \cdot X + R(Z, X) \cdot Y = 0 \quad (\textit{Bianchi's identity}).$$

If g is a pseudo-Riemannian structure on M and ∇ is the corresponding pseudo-Riemannian connection, then

$$g(R(X, Y)\, Z, V) = - g(R(X, Y)\, V, Z),$$
$$g(R(X, Y)\, Z, V) = \quad g(R(Z, V)\, X, Y).$$

Proof. The first identity $R(X, Y) = - R(Y, X)$ is obvious. For the second, we use $T = 0$ and obtain

$$(\nabla_X\nabla_Y - \nabla_Y\nabla_X)\, Z + (\nabla_Y\nabla_Z - \nabla_Z\nabla_Y)\, X + (\nabla_Z\nabla_X - \nabla_X\nabla_Z)\, Y$$
$$- \nabla_{[X,Y]}\, Z - \nabla_{[Y,Z]}\, X - \nabla_{[Z,X]}\, Y$$
$$= \nabla_X[Y, Z] - \nabla_{[Y,Z]}\, X + \nabla_Y[Z, X] - \nabla_{[Z,X]}\, Y + \nabla_Z[X, Y] - \nabla_{[X,Y]}\, Z$$
$$= [X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$$

by the Jacobi identity for vector fields (§2, No. 1). For the third identity we can assume the vector fields X, Y, Z, V are adapted to their values at some point p. From (2), §9, we find in this case

$$g(\nabla_X Z, Z) = 0.$$

For an arbitrary vector field W we have

$$Wg(X, Y) = g(\nabla_W X, Y) + g(X, \nabla_W Y)$$

from §9; also $(\nabla_W(Z))_p = 0$ since Z is adapted to Z_p. Hence

$$\begin{aligned} g_p(R_p(X, Y)Z, Z) &= g_p(\nabla_X \nabla_Y Z, Z) - g_p(\nabla_Y \nabla_X Z, Z) \\ &= X_p g(\nabla_Y Z, Z) - Y_p g(\nabla_X Z, Z) = 0. \end{aligned}$$

This proves the third identity above. The last follows from Lemma 12.4.

Returning now to Prop. 12.3, we use Lemma 12.4 on the quadrilinear function on $M_p \times M_p \times M_p \times M_p$ given by

$$B(X, Y, Z, T) = g_p(R_p(X, Y)Z, T) - g'_p(R'_p(X, Y)Z, T).$$

Since $g_p = g'_p$ the parallelogram $X \vee Y$ has the same area whether measured by means of g or g'. Now $K(S) = K'(S)$ implies $B(X, Y, X, Y) = 0$ for all $X, Y \in M_p$ so by Lemma 12.4, $B \equiv 0$. But g_p is nondegenerate so $R_p = R'_p$.

§ 13. Riemannian Manifolds of Negative Curvature

The local Riemannian geometry developed in §9 was mostly based on properties of the forms $(\text{Exp})^* \omega^i$, which are 1-forms in a normal neighborhood in the tangent space to the manifold at some fixed point. The forms ω^i are only defined locally and the same is therefore the case with the forms $(\text{Exp})^* \omega^i$. However, we shall now see that these last forms (in contrast to the ω^i) can be extended to the entire tangent space or at any rate to the part of the tangent space where Exp is defined and regular. No assumption will be made about the curvature for the time being.

Let M be a Riemannian manifold, p a point in M, M_p the tangent space at p. Let Exp stand for the mapping Exp_p. Let N_0 be any open subset of M_p star-shaped with respect to 0 such that Exp is a regular mapping of N_0 into M. Note that N_0 is not assumed to be a normal neighborhood of 0. Let $Y_1, ..., Y_m$ be any orthonormal basis of M_p. If $X \in N_0$, there exists an open neighborhood N_X of X in N_0 which Exp maps diffeomorphically onto an open neighborhood B of Exp X in M. For each $Y \in N_X$ let $(Y_i^*)_{\text{Exp}\, Y}$ denote the tangent vector at Exp Y which is obtained by parallel translating Y_i along the geodesic Exp tY $(0 \leqslant t \leqslant 1)$. Then $Y_1^*, ..., Y_m^*$ is a basis for the vector fields on B. Let $\Gamma_{ij}{}^k$, $R^k{}_{lij}$ be the corresponding functions on B as defined in §8 and let the 1-forms ω^i, $\omega^i{}_l$, $(1 \leqslant i, l \leqslant m)$ on B be determined by $\omega^i(Y_j^*) = \delta^i{}_j$, $\omega^i{}_l = \sum_k \Gamma_{kl}{}^i \omega^k$. Let V_X be the set of points $(t, a_1, ..., a_m) \in \boldsymbol{R} \times \boldsymbol{R}^m$ for which $ta_1Y_1 + ... + ta_mY_m \in N_X$. Consider the mapping $\Phi : V_X \to B$ given by $\Phi : (t, a_1, ..., a_m) \to \operatorname{Exp}(ta_1Y_1 + ... + ta_mY_m)$. Just as in §8 one proves that the 1-forms $\Phi^*(\omega^i)$ and $\Phi^*(\omega^i{}_l)$ are given by

$$\Phi^*(\omega^i) = a_i dt + \bar{\omega}^i, \qquad \Phi^*(\omega^i{}_l) = \bar{\omega}^i{}_l,$$

where $\bar{\omega}^i$ and $\bar{\omega}^i{}_l$ are 1-forms on V_X not containing dt. If X' is another point in N_0 we can as above construct 1-forms $\bar{\omega}^i$ and $\bar{\omega}^i{}_l$ on $V_{X'}$. It is clear from the definition that these forms agree on $V_X \cap V_{X'}$ with the ones previously constructed. Thus, if V denotes the set of points $(t, a_1, ..., a_m)$ in $\boldsymbol{R} \times \boldsymbol{R}^m$ such that $ta_1Y_1 + ... + ta_mY_m \in N_0$, it follows that the forms $\bar{\omega}^i$ and $\bar{\omega}^i{}_l$ can be defined on the entire set V. They satisfy the differential equations

$$\frac{\partial \bar{\omega}^i}{\partial t} = da_i + \sum_k a_k \bar{\omega}^i{}_k, \qquad \bar{\omega}^i(t, a_j; da_k)_{t=0} = 0, \tag{1}$$

$$\frac{\partial \bar{\omega}^i{}_l}{\partial t} = \sum_{j,k} R^i{}_{ljk} a_j \bar{\omega}^k, \qquad \bar{\omega}^i{}_l(t, a_j; da_k)_{t=0} = 0, \tag{2}$$

in each V_X where $R^i{}_{ljk}$ stands for $R^i{}_{ljk} \circ \Phi$; hence (1) and (2) hold in the entire set V. Combining (1) and (2) we obtain the system

$$\frac{\partial^2 \bar{\omega}^i}{\partial t^2} = \sum_{j,k,l} R^i{}_{ljk} a_l a_j \bar{\omega}^k, \qquad \bar{\omega}^i(t, a_j; da_k)_{t=0} = 0, \tag{3}$$

on V. This system is a generalization of the so-called Jacobi equation for surfaces.

Let S denote the unit sphere $\| X \| = 1$ in M_p and let U denote the set of all pairs $(t, X) \in \boldsymbol{R} \times S$ such that $tX \in N_0$. Consider the mapping $\Psi : (t, X) \to \operatorname{Exp} tX$ of U into M. If we identify M_p and $\boldsymbol{R}^m$ by means of the basis $Y_1, ..., Y_m$, the sphere S is identified with the submanifold $a_1^2 + ... + a_m^2 = 1$ of $\boldsymbol{R}^m$. Thus the forms $\bar{\omega}^i$ and $\bar{\omega}^i{}_l$ induce 1-forms on U which we denote by the same symbol. Using the fact that N_0 is star-shaped, one finds that Lemma 9.2, stating that

$$\Psi^* g = (dt)^2 + \sum_{i=1}^m (\bar{\omega}^i)^2, \qquad \text{on } U, \tag{4}$$

is still valid. If ξ denotes the mapping $(t, X) \to tX$ of U into N_0, then $\Psi = \operatorname{Exp} \circ \xi$. We can use (4) in the special case when $M = M_p$. Here $R = 0$, $\bar{\omega}^i = t da_i$ and $\Psi = \xi$.

Hence

$$\xi^*(g_p) = (dt)^2 + t^2 \sum_{i=1}^{m} (da_i)^2. \tag{5}$$

Let A be any fixed vector field on the sphere $a_1^2 + \ldots + a_m^2 = 1$ in $\boldsymbol{R}^m$ and put

$$\alpha_i = \bar{\omega}^i(A), \qquad \gamma_i = da_i(A) \qquad (1 \leqslant i \leqslant m).$$

The functions $\alpha_i = \alpha_i(t, a_1, \ldots, a_m)$ are defined on the set U and satisfy the equation

$$\sum_{i=1}^{m} \alpha_i \frac{\partial^2 \alpha_i}{\partial t^2} = \sum_{i,j,k,l} R^i{}_{ljk} a_l a_j \alpha_i \alpha_k, \qquad \alpha_i(0, a_1, \ldots, a_m) = 0. \tag{6}$$

We assume now that the sectional curvature of M is $\leqslant 0$ along each plane section at each point of M. For simplicity we express this condition by saying that *M has negative curvature.*

Consider now a *fixed* point $(t, a_1, \ldots, a_m) \in U$ and the corresponding numbers $\alpha_1, \ldots, \alpha_m$. Now the point $X = t(a_1 Y_1 + \ldots + a_m Y_m)$ lies in N_0 and the neighborhood N_X above is diffeomorphic to B. Let $q = \operatorname{Exp} X$, $Y = \sum_{i=1}^m a_i Y_i^*$, $Z = \sum_{i=1}^m \alpha_i Y_i^*$. Then $q \in B$ and Y and Z are vector fields on B. It is clear that

$$g_q(R_q(Y_q, Z_q)\, Y_q, Z_q) = \sum_{i,j,k,l} R^i{}_{ljk} a_l a_j \alpha_i \alpha_k$$

so (6) implies

$$\sum_i \alpha_i \frac{\partial^2 \alpha_i}{\partial t^2} \geqslant 0 \qquad \text{on } U, \tag{7}$$

due to the curvature assumption.

Again fix a point $(a_1, \ldots, a_m) \in S$. We put $h(t) = (\sum_i \alpha_i^2)^{1/2}$ for all $t \geqslant 0$ for which $(t, a_1, \ldots, a_m) \in U$. We assume temporarily that A does not vanish at the point $(a_1, \ldots, a_m)$. Then $h(0) = 0$ and $h(t) > 0$ for $t > 0$. Since

$$\left(\frac{\partial \bar{\omega}^i}{\partial t}\right)_{t=0} = da_i, \qquad \left(\frac{\partial \alpha_i}{\partial t}\right)_{t=0} = \gamma_i,$$

it follows that $h'(0) = (\sum_i \gamma_i^2)^{1/2}$. From (7) and the identity

$$h(t)^3\, h''(t) = h(t)^2 \sum_i \alpha_i \frac{\partial^2 \alpha_i}{\partial t^2} + \left(\sum_i \alpha_i^2 \sum_i \left(\frac{\partial \alpha_i}{\partial t}\right)^2 - \left(\sum_i \alpha_i \frac{\partial \alpha_i}{\partial t^2}\right)^2\right)$$

it follows that

$$h''(t) \geqslant 0, \qquad h'(t) \geqslant h'(0), \qquad \text{for } t > 0.$$

Consequently, $h(t) \geqslant th'(0)$, so

$$\sum_{i=1}^{m} (\bar{\omega}^i(A))^2 \geqslant t^2 \sum_{i=1}^{m} (da_i(A))^2 \tag{8}$$

at the point $(a_1, ..., a_m)$. If A vanishes at the point $(a_1, ..., a_m) \in S$, (8) holds trivially. Hence (8) holds for an arbitrary vector field A on S and for all points in U for which $t \geqslant 0$. Using now (4), (5), and the fact that $\Psi = \text{Exp} \circ \xi$, $\Psi^* = \xi^* \circ \text{Exp}^*$, we obtain

$$\| d\, \text{Exp}_X (Y) \| \geqslant \| Y \|,$$

if X is any point in N_0 and Y is any tangent vector to N_0 at X. This proves the following theorem.

Theorem 13.1. *Let M be a Riemannian manifold of negative curvature and p any point in M. Let N_0 be any open subset of M_p star-shaped with respect to 0 such that Exp (the Exponential mapping at p) is a regular mapping of N_0 into M. Then*

$$\| d\, \text{Exp}_X (Y) \| \geqslant \| Y \|,$$

if X is any point in N_0 and Y is any tangent vector to N_0 at X. In particular,

$$L(\text{Exp} \circ \Gamma) \geqslant L(\Gamma)$$

for any curve segment Γ in N_0, L denoting arc length.

Corollary 13.2. *Suppose M is a Riemannian manifold of negative curvature and V a minimizing convex normal ball in M. Let ABC be a triangle inside V whose angles are A, B, C and whose sides are geodesics of lengths a, b, and c. Then*

(i) $a^2 + b^2 - 2ab \cos C \leqslant c^2$;
(ii) $A + B + C \leqslant \pi$.

In fact, let us use Theorem 13.1 on Exp_C. Let Γ_a, Γ_b, and Γ_c denote the geodesics forming the sides of the triangle and let γ_a, γ_b, and γ_c denote the corresponding curve segments in the tangent space M_C, ($\text{Exp}_C (\gamma_a) = \Gamma_a$, etc.). Let γ_0 denote the straight line in M_C joining the end points of γ_c, and put $\Gamma_0 = \text{Exp}_C (\gamma_0)$.

Then

$$a = L(\gamma_a) = L(\Gamma_a),$$

$$b = L(\gamma_b) = L(\Gamma_b),$$

$$L(\gamma_0) \leqslant L(\gamma_c), \qquad L(\gamma_0)^2 = a^2 + b^2 - 2ab \cos C,$$

since the angle between γ_a and γ_b is C.

Suppose now that the sectional curvature is everywhere $\leqslant 0$. Then $L(\gamma_c) \leqslant L(\Gamma_c)$ and (i) follows. For (ii) we first observe that $c = d(A, B)$, etc., and consequently each length a, b, or c is majorized by the sum of the two others. We can therefore find an ordinary plane triangle with sides a, b, c. Denoting its angles by A', B', C' we have by (i): $A \leqslant A'$, $B \leqslant B'$, $C \leqslant C'$. Since $A' + B' + C' = \pi$, relation (ii) follows.

Theorem 13.3.

(i) *Let M be a complete Riemannian manifold of negative curvature and p any point in M. Then M contains no points conjugate to p.*

(ii) *Let M be a complete Riemannian manifold and suppose there exists a point $p \in M$ such that M contains no point conjugate to p. Then the pair (M_p, Exp_p) is a covering manifold of M. In particular, if M is simply connected, Exp_p is a diffeomorphism of M_p onto M.*

Proof. (i) As before, we write Exp instead of Exp_p and denote by $C(p)$ the (closed) set of points in M_p which are conjugate to p. If $C(p)$ were not empty, let X be a point in $C(p)$ at minimum distance from the origin. Then there exists a vector $Y \neq 0$ in M_p such that $d\,\text{Exp}_X(Y) = 0$. On the other hand, Theorem 13.1 implies that $\| d\,\text{Exp}_{tX}(Y) \| \geqslant \| Y \|$ for $0 \leqslant t < 1$. By continuity, $\| d\,\text{Exp}_X(Y) \| \geqslant \| Y \|$, which is a contradiction. Thus $C(p) = \emptyset$.

In order to prove the latter statement of the theorem we follow a suggestion of I. Singer and consider the tensor field $g^* = \text{Exp}^* g$ on M_p, g denoting the Riemannian structure on M. Owing to the regularity of Exp, g^* is a Riemannian structure on M_p. The space M_p with the Riemannian structure g^* is complete; in fact, the geodesics through the origin in M_p are straight lines. Thus M_p is complete at p (in the sense of the remark following Theorem 10.4), hence complete. Theorem 13.3 now follows from the next lemma.

Lemma 13.4.† *Let V and W be Riemannian manifolds, V complete, and φ a differentiable mapping of V onto W. Assume that $d\varphi_v$ is an isometry for each $v \in V$. Then (V, φ) is a covering space of W.*

† Ambrose [1], p. 360.

Proof (Simplification from Hicks [1]). Let $w \in W$, let N_0 be a normal neighborhood of 0 in W_w of the form $\|X\| < r$, and put $N_w = \text{Exp}_w N_0$. Let v be any point in $\varphi^{-1}(w)$. Let ψ denote the inverse of the mapping $\text{Exp}_w : N_0 \to N_w$. Since V is complete, the mapping $f = \text{Exp}_v \circ (d\varphi_v)^{-1} \circ \psi$ is a well-defined mapping of N_w onto a subset N_v of V and it is obvious that $\varphi \circ f$ is the identity mapping of N_w onto itself. Similarly, $f \circ \varphi$ is the identity mapping of N_v onto itself. Since $(d\varphi_v)^{-1}(N_0)$ is the ball $\|X\| < r$ in V_v, it is clear that N_v is contained in the open ball $B_r(v)$. On the other hand, $B_r(v) \subset N_v$ due to Theorem 10.4. Thus $N_v = B_r(v)$ and φ is a diffeomorphism of N_v onto N_w. Now suppose $v_1, v_2 \in \varphi^{-1}(w)$, $v_1 \neq v_2$. Then the balls $B_r(v_1)$ and $B_r(v_2)$ are disjoint because otherwise there would be a point $v^* \in B_r(v_2) \cap B_r(v_1)$ such that w and $\varphi(v^*)$ are joined by geodesics of different length lying inside N_w. Moreover,

$$\bigcup_{v \in \varphi^{-1}(w)} B_r(v) = \varphi^{-1}(N_w),$$

because each point in $\varphi^{-1}(N_w)$ can be joined to a point in $\varphi^{-1}(w)$ by means of a geodesic of length $< r$. This proves that (V, φ) is a covering manifold of W.

The next theorem, due to É. Cartan, has an important application to Lie groups (see Chapter VI); in fact, it leads to the only known proof of the conjugacy of maximal compact subgroups of a semisimple Lie group.

Theorem 13.5. *Let M be a complete simply connected Riemannian manifold of negative curvature. Let K be a compact Lie transformation group*† *of M whose elements are isometries of M. Then the members of K have a common fixed point.*

Proof. Let d denote the distance function on M and let dk denote the Haar measure on K, normalized by $\int_K dk = 1$. Select a point $p \in M$ and consider the real function J on M given by

$$J(q) = \int_K d^2(q, k \cdot p)\, dk.$$

Then J is a nonnegative continuous function on M. Since the orbit of p is compact, there exists a ball $B_r(p)$ such that $J(q) > J(p)$ for $q \notin B_r(p)$. The closure of $B_r(p)$ is compact, and contains therefore a minimum point q_0 for J. Then q_0 is also a minimum for J on M. It is

† See the definition in Chapter II, § 3.

clear that $J(k \cdot q_0) = J(q_0)$ for $k \in K$, so in order to prove the existence of the fixed point, it suffices to prove that

$$J(q) > J(q_0) \qquad \text{if } q \neq q_0. \tag{9}$$

Now, due to Theorem 13.3, any two points in M can be joined by a unique geodesic and its length is the distance between the points. Thus, due to Cor. 13.2, the cosine inequality

$$a^2 + b^2 - 2ab \cos C \leqslant c^2$$

is valid for an arbitrary geodesic triangle in M. Suppose now $q \neq q_0$ and let $t \to q_t$ $(0 \leqslant t \leqslant d(q_0, q))$ denote the geodesic joining q_0 to q. If $k \cdot p \neq q_t$, let $\alpha_t(k)$ denote the angle between the geodesics (q_t, q) and $(k \cdot p, q_t)$. In view of Lemma 13.6 we have

$$\frac{d}{dt} d^2(q_t, k \cdot p) = \begin{cases} 2d(q_t, k \cdot p) \cos \alpha_t(k), & \text{if } k \cdot p \neq q_t, \\ 0, & \text{if } k \cdot p = q_t. \end{cases} \tag{10}$$

We shall now prove that the function

$$F(t, k) = \frac{d}{dt} d^2(q_t, k \cdot p) \qquad (0 \leqslant t \leqslant d(q_0, q), k \in K),$$

is continuous at each point $(0, k)$, $k \in K$. Then F is clearly continuous everywhere. Let K_1 denote the (closed) set of elements k in K such that $k \cdot p = q_0$; let K_2 denote the complement $K - K_1$. Now, the mapping $k \to k \cdot p$ of K into M is differentiable; using Lemma 11.1 successively, it follows quickly that the function $(t, k) \to \cos \alpha_t(k)$ is continuous at $(0, k_0)$ if $k_0 \in K_2$. Next, let $k_0 \in K_1$ and suppose the sequence (t_n, k_n) converges to $(0, k_0)$. By (10) we have

$$| F(t, k) | \leqslant 2d(q_t, k \cdot p),$$

and since $d(q_{t_n}, k_n \cdot p) \to d(q_0, k_0 \cdot p) = 0$, it follows that $F(t_n, k_n) \to F(0, k_0)$. This proves the continuity of F. Thus the function $t \to J(q_t)$ is differentiable and its derivative can be obtained by differentiating under the integral sign. Since the minimum occurs for $t = 0$, we obtain from (10)

$$\int_{K_2} d(q_0, k \cdot p) \cos \alpha_0(k) \, dk = 0.$$

The cosine inequality above shows that

$$d^2(q, k \cdot p) \geqslant d^2(q_0, k \cdot p) + d^2(q_0, q) - 2d(q_0, q) \, d(q_0, k \cdot p) \cos (\pi - \alpha_0(k))$$

for $k \in K_2$. By integration we obtain

$$\int_{K_2} d^2(q, k \cdot p)\, dk \geqslant \int_{K_2} d^2(q_0, k \cdot p)\, dk + d^2(q_0, q) \int_{K_2} dk.$$

The similar inequality for K_1 is trivial; adding these inequalities we get

$$J(q) \geqslant J(q_0) + d^2(q_0, q),$$

which proves the theorem.

Lemma 13.6. *Let M be as in Theorem 13.5 and let $p \in M$. Let $\gamma : t \to q_t$ $(0 \leqslant t \leqslant L)$ be a curve segment not containing p, t being the arc parameter. Then*

$$\left[\frac{d}{dt}\, d(q_t, p)\right]_{t=0} = \cos \alpha,$$

where α denotes the angle between the geodesic γ and the geodesic (pq_0) at q_0 (see Fig. 1).

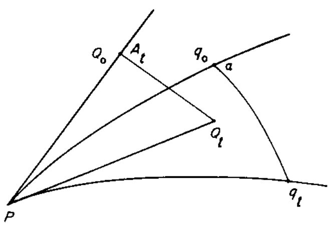

FIG. 1

Proof. Let Exp stand for Exp_p and determine $Q_t \in M_p$ such that $\mathrm{Exp}\, Q_t = q_t$. Let the distance in M_p also be denoted by d. Then

$$\lim_{t\to 0} \frac{1}{t} (d(q_t, p) - d(q_0, p)) = \lim_{t\to 0} \frac{1}{2d(q_0, p)\, t} (d^2(q_t, p) - d^2(q_0, p))$$

$$= \lim_{t\to 0} \frac{1}{2d(Q_0, p)\, t} (d^2(Q_t, p) - d^2(Q_0, p)).$$

Now

$$d^2(Q_t, p) - d^2(Q_0, p) = d^2(Q_0, Q_t) + 2d(Q_0, p)\, d(Q_0, Q_t) \cos A_t,$$

where A_t is the angle between the straight lines (pQ_0) and (Q_0Q_t). Since Exp is a diffeomorphism, the mapping $t \to Q_t$ $(0 \leqslant t \leqslant L)$ is a curve segment. Let Y denote its tangent vector at Q_0, L_t its arc length measured from Q_0 to Q_t, and A the angle between Y and (pQ_0). Then

$$\lim_{t\to 0} \frac{d(Q_0, Q_t)}{L_t} = 1, \qquad \lim_{t\to 0} \frac{L_t}{t \, || \, Y \, ||} = 1. \tag{11}$$

It follows at once that $(1/t)\,(d^2(Q_0, Q_t)) \to 0$ as $t \to 0$. Consequently,

$$\left[\frac{d}{dt}\, d(q_t, p)\right]_{t=0} = \lim_{t\to 0} \frac{1}{t}\, d(Q_0, Q_t) \cos A_t = ||\, Y \,|| \cos A.$$

Now $Y = Y_0 + Y_1$ where Y_0 has the direction of (pQ_0) and Y_1 is perpendicular to that direction. In view of Lemma 9.7, $d\,\mathrm{Exp}_{Q_0}(Y_1)$ is perpendicular to the geodesic (pq_0). Since $||\, d\,\mathrm{Exp}_{Q_0} Y_0 \,|| = ||\, Y_0 \,||$ we have

$$||\, Y \,|| \cos A = ||\, d\,\mathrm{Exp}_{Q_0}(Y) \,|| \cos \alpha = \cos \alpha,$$

and the lemma is proved.

§ 14. Totally Geodesic Submanifolds

Let M be a differentiable manifold, S a submanifold. Let $m = \dim M$, $s = \dim S$. A curve in S is of course a curve in M, but a curve in M contained in S is not necessarily a curve in S, because it may not even be continuous. However, we have:

Lemma 14.1. *Let φ be a differentiable mapping of a manifold V into the manifold M such that $\varphi(V)$ is contained in the submanifold S. If the mapping $\varphi : V \to S$ is continuous it is also differentiable.*

Let $p \in V$. In view of Prop. 3.2, there exists a coordinate system $\{x_1, \ldots, x_m\}$ valid on an open neighborhood N of $\varphi(p)$ in M such that the set

$$N_S = \{r \in N : x_j(r) = 0 \text{ for } s < j \leqslant m\}$$

together with the restrictions of $(x_1, \ldots, x_s)$ to N_S form a local chart on S containing $\varphi(p)$. By the continuity of φ there exists a local chart (W, ψ) around p such that $\varphi(W) \subset N_S$. The coordinates $x_j(\varphi(q))$ $(1 \leqslant j \leqslant m)$ depend differentiably on the coordinates of $q \in W$. In particular, this holds for the coordinates $x_j(\varphi(q))$ $(1 \leqslant j \leqslant s)$ so the mapping $\varphi : V \to S$ is differentiable.

As an immediate consequence of this lemma we have the following statement: Suppose that V and S are submanifolds of M and $V \subset S$. If S has the relative topology of M, then V is a submanifold of S.

In the remainder of this section we shall assume that M is a Riemannian manifold and S a connected submanifold. The Riemannian structure on M induces a Riemannian structure on S. Let d_M and d_S denote the distance functions in M and S, respectively. It is obvious that

$$d_M(p, q) \leqslant d_S(p, q)$$

for $p, q \in S$. In order to distinguish between geodesics in M and S we shall call them M-geodesics and S-geodesics, respectively.

Lemma 14.2. *Let γ be a curve in S, and suppose γ is an M-geodesic. Then γ is an S-geodesic.*

Proof. Let o and p be any points on γ, say $o = \gamma(r_0)$ and $p = \gamma(r)$; let N_o be a spherical normal neighborhood of o in M. If r is sufficiently close to r_0 the geodesic segment

$$\gamma_{op} : t \to \gamma(t), \qquad | t - r_0 | \leqslant | r - r_0 |,$$

is contained in N_o. In view of Lemma 9.3, the length of γ_{op} satisfies

$$L(\gamma_{op}) = d_M(o, p) \leqslant d_S(o, p) \leqslant L(\gamma_{op}).$$

Consequently $L(\gamma_{op}) = d_S(o, p)$; thus γ_{op} is a curve of shortest length in S joining o and p, hence an S-geodesic.

Definition. Let M be a Riemannian manifold and S a connected submanifold of M. Let $p \in S$. The submanifold S is said to be *geodesic at p* if each M-geodesic which is tangent to S at p is a curve in S. The submanifold S is called *totally geodesic* if it is geodesic at each of its points.

Lemma 14.3. *Suppose S is a submanifold of M, geodesic at a point $p \in S$. If γ is an S-geodesic through p, then γ is also an M-geodesic. If M is complete, then S is complete.*

Proof. Let Γ be the maximal M-geodesic tangent to γ at p. Then $\Gamma \subset S$ so by Lemma 14.2, Γ is an S-geodesic. Hence $\gamma \subset \Gamma$. Now suppose M is complete and let Exp_M and Exp_S denote the Exponential mapping at p for M and S, respectively. By assumption Exp_M is defined on the entire M_p. Since S is geodesic at p, Exp_S is the restriction of Exp_M to S_p, in particular, S is complete at p in the sense of the remark following Theorem 10.4. By that remark, S is complete.

Proposition 14.4. *Suppose S is a totally geodesic submanifold of M, and let I denote the identity mapping of S into M. For each $p \in S$ there exists an open neighborhood U_p of p in S on which I is distance preserving, that is,*

$$d_S(q_1, q_2) = d_M(q_1, q_2) \qquad \textit{for } q_1, q_2 \in U_p. \tag{1}$$

Proof. Let $B_\rho(p)$ be a minimizing convex normal ball around p in M. Since I is continuous, the intersection $B_\rho(p) \cap S$ is an open subset of S. Let U_p be a minimizing convex normal ball around p in S such that $U_p \subset B_\rho(p)$. Let q, r be arbitrary points in U_p and γ_{qr} the S-geodesic inside U_p joining q and r. Then $d_S(q, r) = L(\gamma_{qr})$. Consider the maximal M-geodesic Γ such that Γ and γ_{qr} have the same tangent vector at q. Since S is totally geodesic, $\Gamma \subset S$; from Lemma 14.2 follows that $\gamma_{qr} \subset \Gamma$, so γ_{qr} is an M-geodesic. Since $\gamma_{qr} \subset B_\rho(p)$ it follows that $d_M(q, r) = L(\gamma_{qr})$. This proves the proposition.

Remark. Examples are easily constructed (e.g., geodesics on a cone) which show that relation (1) does not in general hold for all $q_1, q_2 \in S$.

Theorem 14.5. *Let M be a Riemannian manifold and S a connected, complete submanifold of M. Then S is totally geodesic if and only if M-parallel translation along curves in S always transports tangents to S into tangents to S.*

Proof. Let $s = \dim S$, $m = \dim M$ and let o be an arbitrary point in S. In view of Prop. 3.2 there exists an open neighborhood N of o in M on which a coordinate system $\{x_1, ..., x_m\}$ is valid such that the set

$$U = \{q \in N : x_j(q) = 0 \quad \text{for} \quad s+1 \leqslant j \leqslant m\}$$

is a normal neighborhood of o in S and such that the restrictions of $(x_1, ..., x_s)$ to U form a coordinate system on U.

Let $p \in U$ and let $\gamma : t \to \gamma(t)$ be a curve in U such that $p = \gamma(0)$. Let $Y(t)$ be a family of tangent vectors to M which is M-parallel along the curve γ and such that $Y(0) \in S_p$. Writing $Y(t) = \sum_{a=1}^m Y^a(t)\, \partial/\partial x_a$ the coefficients $Y^a(t)$ satisfy the equations

$$\dot{Y}^a(t) + \sum_{b,c=1}^m \Gamma_{bc}{}^a\, \dot{x}_b(t)\, Y^c(t) = 0, \qquad 1 \leqslant a, b, c \leqslant m, \tag{2}$$

$$Y^a(0) = 0, \qquad x_a(t) \equiv 0, \qquad s < a \leqslant m.$$

Let $\pi : t \to \pi(t)$ be an M-geodesic tangent to S at p, t being the arc parameter measured from p. If t is sufficiently small and we write $x_a(t)$ for $x_a(\pi(t))$,

$$\ddot{x}_a(t) + \sum_{b,c=1}^m \Gamma_{bc}{}^a\, \dot{x}_b(t)\, \dot{x}_c(t) = 0, \qquad 1 \leqslant a, b, c, \leqslant m, \tag{3}$$

$$\dot{x}_a(0) = 0, \qquad s < a \leqslant m.$$

In the computations below we adopt the following range of indices:

$$1 \leqslant i, j, k \leqslant s,$$
$$s + 1 \leqslant \alpha, \beta, \gamma \leqslant m.$$

Suppose now S is totally geodesic. Then the M-geodesic π above is a curve in S, hence an S-geodesic. For small t, $\pi(t)$ lies in U so $x_\alpha(t) \equiv 0$. Since every S-geodesic is now an M-geodesic, (3) implies

$$\Gamma_{jk}{}^{\alpha}(p) = 0, \qquad p \in U. \tag{4}$$

For the curve γ above we have $\dot{x}_\alpha(t) \equiv 0$. In view of (4) we obtain

$$\dot{Y}^{\alpha}(t) + \sum_{j,\beta} \Gamma_{j\beta}{}^{\alpha}\, \dot{x}_j(t)\, Y^{\beta}(t) = 0 \qquad \text{on } \gamma. \tag{5}$$

Now, $Y(0) \in S_p$ so $Y^{\beta}(0) = 0$. Owing to the uniqueness theorem for the system (5) of linear differential equations we have $Y^{\beta}(t) \equiv 0$. Consequently, the family $Y(t)$ is tangent to S. Finally, let $\beta : t \to \beta(t)$ $(t \in J)$ be an arbitrary curve in S and let $Z(t)$ be an M-parallel family along β such that $Z(t_0) \in S_{\beta(t_0)}$ for some $t_0 \in J$. The set of $t \in J$ such that $Z(t) \in S_{\beta(t)}$ is clearly closed in J. The argument above shows that this set is open in J. Thus $Z(t) \in S_{\beta(t)}$ for all $t \in J$ and the first half of the theorem is proved.

To prove the converse, suppose that for each curve as above, the relation $Y(0) \in S_p$ implies $Y(t) \in S_{\gamma(t)}$ for each t. In (2) we have therefore

$$Y^{\alpha}(t) \equiv 0 \qquad \dot{x}_\alpha(t) \equiv 0$$

and (4) follows. Now substitute (4) into (3). Since $\Gamma_{bc}{}^{a} = \Gamma_{cb}{}^{a}$ (torsion is 0), we obtain

$$\ddot{x}_\alpha(t) + 2 \sum_{j,\beta} \Gamma_{j\beta}{}^{\alpha}\, \dot{x}_j(t)\, \dot{x}_\beta(t) + \sum_{\beta,\gamma} \Gamma_{\beta\gamma}{}^{\alpha}\, \dot{x}_\beta(t)\, \dot{x}_\gamma(t) = 0. \tag{6}$$

Since $\dot{x}_\alpha(0) = 0$ we conclude from the uniqueness theorem for the non-linear system (6) that $x_\alpha(t)$ is constant, that is, $x_\alpha(t) = 0$ for all t in a certain interval around 0. The functions $x_j(t)$ are differentiable; consequently, a piece π' of π containing p is a curve in U, hence a curve in S. Let π^* be the maximal S-geodesic tangent to π at p, parametrized by the arc length t^* measured from p. Since S is complete, t^* runs from $-\infty$ to ∞. Now $\pi^*(t) = \pi(t)$ if t is sufficiently small; moreover, the set of t-values for which $\pi(t) = \pi^*(t)$ is open and closed. Thus $\pi \subset \pi^*$, π is a curve in S and the theorem is proved.

Theorem 14.6. *Let M be a simply connected, complete Riemannian manifold of negative curvature. Let S be a closed totally geodesic submanifold of M. For each $p \in S$, the geodesics in M which are perpendicular to S at p make up a submanifold $S^{\perp}(p)$ of M and M is the disjoint union:*

$$M = \bigcup_{p \in S} S^{\perp}(p).$$

Proof. Let Exp_p denote the Exponential mapping of M_p into M and let T_p denote the orthogonal complement of the tangent space S_p in M_p. Since S is complete and since S-geodesics are M-geodesics, it follows that

$$S = \mathrm{Exp}_p\,(S_p).$$

Moreover, we have by definition

$$S^{\perp}(p) = \mathrm{Exp}_p\,(T_p).$$

Since Exp_p is a diffeomorphism (Theorem 13.3), $S^{\perp}(p)$ is a submanifold of M.

Now, let q be an arbitrary point in M lying outside S; S being closed there exists a point $p_0 \in S$ at shortest distance from q. The unique geodesic Γ from q to p_0 is perpendicular to S. In fact, if $\sigma : t \to s_t$ is a curve in S, the derivative $(d/dt)\, d(s_t, q)$ equals $\cos \alpha_t$ where α_t is the angle between σ and the geodesic from q to s_t (Lemma 13.6). On the other hand, if a geodesic connecting q to S is perpendicular to S, then this geodesic must coincide with Γ since the sum of the angles in a geodesic triangle in M is $\leqslant \pi$. This shows that each point $q \in M$ lies in exactly one of the manifolds $S^{\perp}(p)$ and the theorem is proved.

EXERCISES

A. Manifolds

1. Let M be a paracompact manifold, A and B disjoint closed subsets of M. Then there exists a function $f \in C^{\infty}(M)$ such that $f \equiv 1$ on A, $f \equiv 0$ on B.

2. Let M be a connected manifold and p, q two points in M. Then there exists a diffeomorphism Φ of M onto itself such that $\Phi(p) = q$.

3. Let M be a Hausdorff space and let δ and δ' be two differentiable structures on M. Let $\mathfrak{F}$ and $\mathfrak{F}'$ denote the corresponding sets of C^{∞} functions. Then $\delta = \delta'$ if and only if $\mathfrak{F} = \mathfrak{F}'$.

Deduce that the real line $\boldsymbol{R}$ with its ordinary topology has infinitely many different differentiable structures.

4. Let Φ be a differentiable mapping of a manifold M onto a manifold N. A vector field X on M is called *projectable* (Koszul [1]) if there exists a vector field Y on N such that $d\Phi \cdot X = Y$.

(i) Show that X is projectable if and only if $X\mathfrak{F}_0 \subset \mathfrak{F}_0$ where $\mathfrak{F}_0 = \{f \circ \Phi : f \in C^\infty(N)\}$.

(ii) A necessary condition for X to be projectable is that

$$d\Phi_p(X_p) = d\Phi_q(X_q) \tag{1}$$

whenever $\Phi(p) = \Phi(q)$. If, in addition, $d\Phi_p(M_p) = N_{\Phi(p)}$ for each $p \in M$, this condition is also sufficient.

(iii) Let $M = \boldsymbol{R}$ with the usual differentiable structure and let N be the topological space $\boldsymbol{R}$ with the differentiable structure obtained by requiring the homeomorphism $\psi : x \to x^{1/3}$ of M onto N to be a diffeomorphism. In this case the identity mapping $\Phi : x \to x$ is a differentiable mapping of M onto N. The vector field $X = \partial/\partial x$ on M is not projectable although (1) is satisfied.

B. The Lie Derivative

1. Let M be a manifold, X a vector field on M. The Lie derivative $\theta(X) : Y \to [X, Y]$ which maps $\mathfrak{D}^1(M)$ into itself can be extended uniquely to a mapping of $\mathfrak{D}(M)$ into itself such that:

(i) $\theta(X)f = Xf$ for $f \in C^\infty(M)$.

(ii) $\theta(X)$ is a derivation of $\mathfrak{D}(M)$ preserving type of tensors.

(iii) $\theta(X)$ commutes with contractions.

2. Let Φ be a diffeomorphism of a manifold M onto itself. Then Φ induces a unique type-preserving automorphism $T \to \Phi \cdot T$ of the tensor algebra $\mathfrak{D}(M)$ such that:

(i) The automorphism commutes with contractions.

(ii) $\Phi \cdot X = X^\Phi$, $(X \in \mathfrak{D}^1(M))$, $\Phi \cdot f = f^\Phi$, $(f \in C^\infty(M))$.

Prove that $\Phi \cdot \omega = (\Phi^{-1})^* \omega$ for $\omega \in \mathfrak{D}_*(M)$.

3. Let g_t be a one-parameter Lie transformation group of M and denote by X the vector field on M induced by g_t (Chapter II, §3). Then

$$\theta(X) = \lim_{t \to 0} \frac{1}{t}(T - g_t \cdot T)$$

for each tensor field T in M ($g_t \cdot T$ is defined in Exercise 2).

4. The Lie derivative $\theta(X)$ on a manifold M has the following properties:

(i) $\theta([X, Y]) = \theta(X)\,\theta(Y) - \theta(Y)\,\theta(X), \qquad X, Y \in \mathfrak{D}^1(M)$.

(ii) $\theta(X)$ commutes with the alternation $A : \mathfrak{D}_*(M) \to \mathfrak{A}_*(M)$ and therefore induces a derivation of the Grassmann algebra of M.

(iii) $\theta(X)\, d = d\theta(X)$, that is, $\theta(X)$ commutes with exterior differentiation.

C. Affine Connections

1. Let M be a connected manifold with a countable basis. Using partition of unity show that M has a Riemannian structure. On the other hand, a Riemannian manifold has a countable basis (Prop. 9.6).

2. Let ∇ be the affine connection on $\boldsymbol{R}^n$ determined by $\nabla_X(Y) = 0$ for $X = \partial/\partial x_i$, $Y = \partial/\partial x_j$, $1 \leqslant i, j \leqslant n$. Find the corresponding affine transformations.

3. Let M be a manifold with a torsion-free affine connection ∇. Suppose $X_1, ..., X_m$ is a basis for the vector fields on an open subset U of M. Let the forms $\omega^1, ..., \omega^m$ on U be determined by $\omega^i(X_j) = \delta^i{}_j$. Prove the formula

$$d\theta = \sum_{i=1}^{m} \omega_i \wedge \nabla_{X_i}(\theta)$$

for each differential form θ on U.

4. In Prop. 10.7 it was proved that a complete noncompact Riemannian manifold M always contains a ray. Does M always contain a straight line, that is, a geodesic $\gamma(t)$ $(-\infty < t < \infty)$ which realizes the shortest distance between any two of its points?

5. Let M and N be analytic, complete, simply connected Riemannian manifolds. Suppose that an open subset of M is isometric to an open subset of N. Using results from §11 show that M and N are isometric (Myers-Rinow).

D. Submanifolds

1. Let M and N be differentiable manifolds and Φ a differentiable mapping of M into N. Consider the mapping $\varphi : m \to (m, \Phi(m))$ $(m \in M)$ and the graph

$$G_\Phi = \{(m, \Phi(m)) : m \in M\}$$

of Φ with the topology induced by the product space $M \times N$. Then φ is a homeomorphism of M onto G_Φ and if the differentiable structure of M is transferred to G_Φ by φ, the graph G_Φ becomes a closed submanifold of $M \times N$.

2. Let N be a manifold and M a topological space, $M \subset N$. Show that there exists at most one differentiable structure on the topological space M such that M is a submanifold of N.

3. Using the figure 8 as a subset of R^2 show that

(i) A closed connected submanifold of a connected manifold does not necessarily carry the relative topology.

(ii) A subset M of a connected manifold N may have two different topologies and differentiable structures such that in both cases M is a submanifold of N.

4. Let M be a Riemannian manifold and S a connected, complete submanifold of M. Show that S is totally geodesic if and only if M-parallel translation of tangent vectors to S along curves in S always coincides with the S-parallel translation (see (2), Chapter V, § 6).

E. Surfaces

1. Let M be a Riemannian manifold of dimension 2, p a point in M, $r(q)$ the distance $d(p, q)$. Show that the curvature K of M at p satisfies

$$K = -3 \lim_{r \to 0} \Delta (\log r)$$

where Δ is the Laplace-Beltrami operator on M (Chapter X, § 2).

2. (Minding) Let S be an orientable surface in $\mathbf{R}^3$, oriented by means of a continuous family of unit normal vectors $\xi_s (s \in S)$. Let $t \to \gamma_S(t)$ be a curve in S, t being the arc-parameter. The triple vector product $(\xi \times \dot{\gamma}_S \cdot \ddot{\gamma}_S)(s)$ is called the *geodesic curvature* of γ_S at s. Show that the geodesic curvature can be expressed in terms of $\dot{\gamma}_S$, $\ddot{\gamma}_S$, the Riemannian structure of S, and its derivatives with respect to local coordinates. Deduce that the geodesic curvature is invariant under orientation-preserving isometries.

3. (Levi-Civita) Suppose a surface S in $\mathbf{R}^3$ rolls without slipping on a plane π. Let the point of contact run through a curve γ_S on S and a curve γ_π on π. Using the result of Exercise E. 2 show that the Euclidean parallelism along γ_π corresponds to the parallelism along γ_S in the sense of the Riemannian connection on S.

NOTES

§1-§3. The treatment of differentiable manifolds given here is closest to that of Chevalley [2].

§4-§5. For the classical literature on affine connections, see Struik [1] and Schouten [1]. The definition of an affine connection adopted here is due to J. L.

Koszul, see Nomizu [2]. It is equivalent to the customary definition in terms of the bundle of frames over the manifold, see Nomizu [4], Chapter III, §4 (where the term linear connection is used). Useful as the frame bundle definition is in global differential geometry, we have nevertheless preferred the vector field definition because the spaces with which we are mainly concerned are coset spaces G/K and for these the natural bundle to consider is the group G itself; the frame bundle would only be extra baggage.

§6. Theorem 6.2 is due to Whitehead [2]. The proof in the text is a slight simplification of Whitehead's proof using stronger differentiability hypotheses. Theorem 6.5 was proved by the author [4].

§9-§10. The treatment of local Riemannian geometry given here is partly based on É. Cartan [22], Chapter X. The equivalence of the completeness conditions (Theorem 10.3) for a two-dimensional Riemannian manifold is due to Hopf-Rinow [1]. A simplified proof was given by de Rham [1]. See also Myers [1] and Whitehead [3].

§11. Theorem 11.1 is due to Myers-Steenrod [1]. Their proof was simplified by Palais [1]. The remaining results in §11 which deal with continuations of local isometries and are useful in Chapter IV are based on de Rham's paper [1]. See also Rinow [1].

§12. The definition of sectional curvature goes back to Riemann's original lecture [1] and the formula of Theorem 12.2 can be found in most books on Riemannian geometry. Lemma 12.1 and the ensuing proof of Theorem 12.2 are from the author's paper [4].

§13-§14. The treatment of Riemannian manifolds of negative curvature is based on É. Cartan's book [22], Note III. In the two-dimensional case, Theorem 13.3 is due to Hadamard [1]. The concept of a totally geodesic submanifold is due to Hadamard [2]. Theorem 14.5 is proved in Cartan [22], p. 115. Cartan also proves, [22], p. 232, that if every submanifold of dimension $s \geqslant 2$ which is geodesic at a point is also totally geodesic then the manifold has constant sectional curvature. Theorem 14.6 can be regarded as a generalization of a decomposition theorem for a semisimple Lie group (Theorem 1.4, Chapter VI) due to Mostow.

CHAPTER II

LIE GROUPS AND LIE ALGEBRAS

A Lie group is, roughly speaking, an analytic manifold with a group structure such that the group operations are analytic. Lie groups arise in a natural way as transformation groups of geometric objects. For example, the group of all affine transformations of a connected manifold with an affine connection and the group of all isometries of a pseudo-Riemannian manifold are known to be Lie groups in the compact open topology. However, the group of all diffeomorphisms of a manifold is too big to form a Lie group in any reasonable topology.

The tangent space $\mathfrak{g}$ at the identity element of a Lie group G has a rule of composition $(X,Y) \to [X,Y]$ derived from the bracket operation on the left invariant vector fields on G. The vector space $\mathfrak{g}$ with this rule of composition is called the Lie algebra of G. The structures of $\mathfrak{g}$ and G are related by the exponential mapping exp: $\mathfrak{g} \to$ G which sends straight lines through the origin in $\mathfrak{g}$ onto one-paramater subgroups of G. Several properties of this mapping are developed already in §1 because they can be derived as special cases of properties of the Exponential mapping for a suitable affine connection on G. Although the structure of $\mathfrak{g}$ is determined by an arbitrary neighborhood of the identity element of G, the exponential mapping sets up a far-reaching relationship between $\mathfrak{g}$ and the group G in the large. We shall for example see in Chapter VII that the center of a compact simply connected Lie group G is explicitly determined by the Lie algebra $\mathfrak{g}$. In §2 the correspondence (induced by exp) between subalgebras and subgroups is developed. This correspondence is of basic importance in the theory in spite of its weakness that the subalgebra does not in general decide whether the corresponding subgroup will be closed or not, an important distinction when coset spaces are considered.

In §4 we investigate the relationship between homogeneous spaces and coset spaces. It is shown that if a manifold M has a separable transitive Lie transformation group G acting on it, then M can be identified with a coset space G/H (H closed) and therefore falls inside the realm of Lie group theory. Thus, one can, for example, conclude that if H is compact, then M has a G-invariant Riemannian structure.

Let G be a connected Lie group with Lie algebra $\mathfrak{g}$. If $\sigma \in G$, the inner automorphism $g \to \sigma g \sigma^{-1}$ induces an automorphism Ad (σ) of $\mathfrak{g}$ and the mapping $\sigma \to$ Ad (σ) is an analytic homomorphism of G onto an analytic subgroup Ad (G) of $GL(\mathfrak{g})$, the adjoint group. The group Ad (G) can be defined by $\mathfrak{g}$ alone and since its Lie algebra is isomorphic to $\mathfrak{g}/\mathfrak{z}$ ($\mathfrak{z}$ = center of $\mathfrak{g}$), one can, for example, conclude that a semisimple Lie algebra over $\boldsymbol{R}$ is isomorphic to the Lie algebra of a Lie group. This fact holds for arbitrary Lie algebras over $\boldsymbol{R}$ but will not be needed in this book in that generality.

Section 6 deals with some preliminary results about semisimple Lie groups. The main result is Weyl's theorem stating that the universal covering group of a compact semisimple Lie group is compact.

§ 1. The Exponential Mapping

1. The Lie Algebra of a Lie Group

Definition. A *Lie group* is a group G which is also an analytic manifold such that the mapping $(\sigma, \tau) \rightarrow \sigma\tau^{-1}$ of the product manifold $G \times G$ into G is analytic.

Examples. 1. Let G be the group of all isometries of the Euclidean plane $\boldsymbol{R}^2$ which preserve the orientation. If $\sigma \in G$, let $(x(\sigma), y(\sigma))$ denote the coordinates of the point $\sigma \cdot 0$ ($0 =$ origin of $\boldsymbol{R}^2$) and let $\theta(\sigma)$ denote the angle between the x-axis and the line from 0 to $\sigma \cdot 0$. Then the mapping $\varphi : \sigma \rightarrow (x(\sigma), y(\sigma), \theta(\sigma))$ maps G in a one-to-one fashion onto the product manifold $\boldsymbol{R}^2 \times \boldsymbol{S}^1$ ($\boldsymbol{S}^1 = \boldsymbol{R} \bmod 2\pi$). We can turn G into an analytic manifold by requiring φ to be an analytic diffeomorphism. An elementary computation shows that for $\sigma, \tau \in G$

$$\begin{aligned} x(\sigma\tau^{-1}) &= x(\sigma) - x(\tau)\cos(\theta(\sigma) - \theta(\tau)) + y(\tau)\sin(\theta(\sigma) - \theta(\tau)); \\ y(\sigma\tau^{-1}) &= y(\sigma) - x(\tau)\sin(\theta(\sigma) - \theta(\tau)) - y(\tau)\cos(\theta(\sigma) - \theta(\tau)); \\ \theta(\sigma\tau^{-1}) &= \theta(\sigma) - \theta(\tau) \pmod{2\pi}. \end{aligned}$$

Since the functions sin and cos are analytic, it follows that G is a Lie group.

2. Let $\tilde{G}$ be the group of all isometries of $\boldsymbol{R}^2$. If s is the symmetry of $\boldsymbol{R}^2$ with respect to a line, then $\tilde{G} = G \cup sG$ (disjoint union). We can turn sG into an analytic manifold by requiring the mapping $\sigma \rightarrow s\sigma$ ($\sigma \in G$) to be an analytic diffeomorphism of G onto sG. This makes $\tilde{G}$ a Lie group.

On the other hand, if G_1 and G_2 are two components of a Lie group G and $x_1 \in G_1$, $x_2 \in G_2$, then the mapping $g \rightarrow x_2x_1^{-1}g$ is an analytic diffeomorphism of G_1 onto G_2.

Remark. A Lie group is always paracompact. In fact, let G be a topological group which is *locally Euclidean*, that is, has a neighborhood of the identity e, homeomorphic to a Euclidean space. Let G_0 denote the *identity component* of G (that is, the component of G containing e). Then G_0 is a connected topological group and as such it is generated by any neighborhood of the identity. It follows that G_0 has a countable base, in particular G_0 is paracompact. The same statement follows for G by the definition of paracompactness.

Let G be a connected topological group. A *covering group* of G is a pair $(\tilde{G}, \pi)$ where $\tilde{G}$ is a topological group and π is a homomorphism of $\tilde{G}$ into G such that $(\tilde{G}, \pi)$ is a covering space of G. In the case when

G is a Lie group, then $\tilde{G}$ has clearly an analytic structure such that $\tilde{G}$ is a Lie group, π analytic and $(\tilde{G}, \pi)$ a covering manifold of G.

Definition. A homomorphism of a Lie group into another which is also an analytic mapping is called an *analytic homomorphism*. An isomorphism of one Lie group onto another which is also an analytic diffeomorphism is called an *analytic isomorphism*.

Let G be a Lie group. If $\rho \in G$, the left translation $L_\rho : g \rightarrow \rho g$ of G onto itself is an analytic diffeomorphism. A vector field Z on G is called left invariant if $dL_\rho Z = Z$ for all $\rho \in G$. Given a tangent vector $X \in G_e$ there exists exactly one left invariant vector field $\tilde{X}$ on G such that $\tilde{X}_e = X$ and this $\tilde{X}$ is analytic. In fact, $\tilde{X}$ can be defined by

$$[\tilde{X}f](\rho) = Xf^{L_{\rho^{-1}}} = \left\{\frac{d}{dt} f(\rho\gamma(t))\right\}_{t=0}$$

if $f \in C^\infty(G)$, $\rho \in G$, and $\gamma(t)$ is any curve in G with tangent vector X for $t = 0$. If $X, Y \in G_e$, then the vector field $[\tilde{X}, \tilde{Y}]$ is left invariant due to Prop. 3.3, Chapter I. The tangent vector $[\tilde{X}, \tilde{Y}]_e$ is denoted by $[X, Y]$. The vector space G_e, with the rule of composition $(X, Y) \rightarrow [X, Y]$ we denote by $\mathfrak{g}$ (or $\mathfrak{L}(G)$) and call the *Lie algebra of* G.

More generally, let $\mathfrak{a}$ be a vector space over a field K (of characteristic 0). The set $\mathfrak{a}$ is called a *Lie algebra over* K if there is given a rule of composition $(X, Y) \rightarrow [X, Y]$ in $\mathfrak{a}$ which is bilinear and satisfies (a) $[X, X] = 0$ for all $X \in \mathfrak{a}$; (b) $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$ for all $X, Y, Z \in \mathfrak{a}$. The identity (b) is called the *Jacobi identity*. The Lie algebra of G above is clearly a Lie algebra over $\boldsymbol{R}$.

If $\mathfrak{a}$ is a Lie algebra over K and $X \in \mathfrak{a}$, the linear transformation $Y \rightarrow [X, Y]$ of $\mathfrak{a}$ is denoted by $\operatorname{ad} X$ (or $\operatorname{ad}_{\mathfrak{a}} X$ when a confusion would otherwise be possible). Let $\mathfrak{b}$ and $\mathfrak{c}$ be two vector subspaces of $\mathfrak{a}$. Then $[\mathfrak{b}, \mathfrak{c}]$ denotes the vector subspace of $\mathfrak{a}$ generated by the set of elements $[X, Y]$ where $X \in \mathfrak{b}$, $Y \in \mathfrak{c}$. A vector subspace $\mathfrak{b}$ of $\mathfrak{a}$ is called a *subalgebra* of $\mathfrak{a}$ if $[\mathfrak{b}, \mathfrak{b}] \subset \mathfrak{b}$ and an *ideal* in $\mathfrak{a}$ if $[\mathfrak{b}, \mathfrak{a}] \subset \mathfrak{b}$. If $\mathfrak{b}$ is an ideal in $\mathfrak{a}$ then the factor space $\mathfrak{a}/\mathfrak{b}$ is a Lie algebra with the bracket operation inherited from $\mathfrak{a}$. Let $\mathfrak{a}$ and $\mathfrak{b}$ be two Lie algebras over the same field K and σ a linear mapping of $\mathfrak{a}$ into $\mathfrak{b}$. The mapping σ is called a *homomorphism* if $\sigma([X, Y]) = [\sigma X, \sigma Y]$ for all $X, Y \in \mathfrak{a}$. If σ is a homomorphism then $\sigma(\mathfrak{a})$ is a subalgebra of $\mathfrak{b}$ and the kernel $\sigma^{-1}\{0\}$ is an ideal in $\mathfrak{a}$. If $\sigma^{-1}\{0\} = \{0\}$, then σ is called an *isomorphism* of $\mathfrak{a}$ into $\mathfrak{b}$. An isomorphism of a Lie algebra onto itself is called an *automorphism*.

Let V be a vector space over a field K and let $\mathfrak{gl}(V)$ denote the vector space of all endomorphisms of V with the bracket operation $[A, B] = AB - BA$. Then $\mathfrak{gl}(V)$ is a Lie algebra over K. Let $\mathfrak{a}$ be a Lie algebra over K. A homomorphism of $\mathfrak{a}$ into $\mathfrak{gl}(V)$ is called a *representation* of $\mathfrak{a}$ on V. In particular, since $\operatorname{ad}([X, Y]) = \operatorname{ad} X \operatorname{ad} Y - \operatorname{ad} Y \operatorname{ad} X$, the linear mapping $X \to \operatorname{ad} X$ ($X \in \mathfrak{a}$) is a representation of $\mathfrak{a}$ on $\mathfrak{a}$. It is called the *adjoint representation* of $\mathfrak{a}$ and is denoted ad (or $\operatorname{ad}_{\mathfrak{a}}$ when a confusion would otherwise be possible). The kernel of $\operatorname{ad}_{\mathfrak{a}}$ is called the *center* of $\mathfrak{a}$. If the center of $\mathfrak{a}$ equals $\mathfrak{a}$, $\mathfrak{a}$ is said to be *abelian*. Thus $\mathfrak{a}$ is abelian if and only if $[\mathfrak{a}, \mathfrak{a}] = \{0\}$.

Let $\mathfrak{a}$ and $\mathfrak{b}$ be two Lie algebras over the same field K. The vector space $\mathfrak{a} \times \mathfrak{b}$ becomes a Lie algebra over K if we define

$$[(X, Y), (X', Y')] = ([X, X'], [Y, Y']).$$

This Lie algebra is called the *Lie algebra product* of $\mathfrak{a}$ and $\mathfrak{b}$. The sets $\{(X, 0) : X \in \mathfrak{a}\}$, $\{(0, Y) : Y \in \mathfrak{b}\}$ are ideals in $\mathfrak{a} \times \mathfrak{b}$ and $\mathfrak{a} \times \mathfrak{b}$ is the direct sum of these ideals.

In the following a Lie algebra shall always mean a finite-dimensional Lie algebra unless the contrary is stated.

2. The Universal Enveloping Algebra

Let $\mathfrak{a}$ be a Lie algebra over a field K. The rule of composition $(X, Y) \to [X, Y]$ is rarely associative; we shall now assign to $\mathfrak{a}$ an associative algebra with unit, the *universal enveloping algebra* of $\mathfrak{a}$, denoted $U(\mathfrak{a})$. This algebra is defined as the factor algebra $T(\mathfrak{a})/J$ where $T(\mathfrak{a})$ is the tensor algebra over $\mathfrak{a}$ (considered as a vector space) and J is the two-sided ideal in $T(\mathfrak{a})$ generated by the set of all elements of the form $X \otimes Y - Y \otimes X - [X, Y]$ where $X, Y \in \mathfrak{a}$. If $X \in \mathfrak{a}$, let X^* denote the image of X under the canonical mapping π of $T(\mathfrak{a})$ onto $U(\mathfrak{a})$. The identity element in $U(\mathfrak{a})$ will be denoted by 1. Then $1 \neq 0$ if $\mathfrak{a} \neq \{0\}$.

Proposition 1.1. *Let V be a vector space over K. There is a natural one-to-one correspondence between the set of all representations of $\mathfrak{a}$ on V and the set of all representations of $U(\mathfrak{a})$ on V. If ρ is a representation of $\mathfrak{a}$ on V and ρ^* is the corresponding representation of $U(\mathfrak{a})$ on V, then*

$$\rho(X) = \rho^*(X^*) \qquad \text{for } X \in \mathfrak{a}. \tag{1}$$

Proof. Let ρ be a representation of $\mathfrak{a}$ on V. Then there exists a unique representation $\tilde{\rho}$ of $T(\mathfrak{a})$ on V satisfying $\tilde{\rho}(X) = \rho(X)$ for all $X \in \mathfrak{a}$. The mapping $\tilde{\rho}$ vanishes on the ideal J because

$$\tilde{\rho}(X \otimes Y - Y \otimes X - [X, Y]) = \rho(X)\rho(Y) - \rho(Y)\rho(X) - \rho([X, Y]) = 0.$$

Thus we can define a representation ρ^* of $U(\mathfrak{a})$ on V by the condition $\rho^* \circ \pi = \tilde{\rho}$. Then (1) is satisfied and determines ρ^* uniquely. On the other hand, suppose σ is a representation of $U(\mathfrak{a})$ on V. If $X \in \mathfrak{a}$ we put $\rho(X) = \sigma(X^*)$. Then the mapping $X \to \rho(X)$ is linear and in fact a representation of $\mathfrak{a}$ on V, because

$$\begin{aligned}\rho([X, Y]) &= \sigma([X, Y]^*) = \sigma(\pi(X \otimes Y - Y \otimes X)) \\ &= \sigma(X^*Y^* - Y^*X^*) = \rho(X)\,\rho(Y) - \rho(Y)\,\rho(X)\end{aligned}$$

for $X, Y \in \mathfrak{a}$. This proves the proposition.

Let $X_1, ..., X_n$ be a basis of $\mathfrak{a}$ and put $X^*(t) = \sum_{i=1}^n t_i X_i^*$ $(t_i \in K)$. Let $M = (m_1, ..., m_n)$ be an ordered set of integers $m_i \geqslant 0$. We shall call M a *positive integral n-tuple.* We put $|M| = m_1 + ... + m_n$, $t^M = t_1^{m_1} ... t_n^{m_n}$. Considering $t_1, ..., t_n$ as indeterminates the various t^M are linearly independent over K and for $|M| > 0$ we can define $X^*(M) \in U(\mathfrak{a})$ as the coefficient of t^M in the expansion of $(|M|!)^{-1}(X^*(t))^{|M|}$. Put $X^*(M) = 1$ if $|M| = 0$.

Proposition 1.2. *The smallest vector subspace of* $U(\mathfrak{a})$ *containing all the elements* $X^*(M)$ *(where* M *is a positive integral n-tuple) is* $U(\mathfrak{a})$ *itself.*

Proof. It suffices to prove that each element $X_{i_1}^* X_{i_2}^* ... X_{i_p}^*$ $(1 \leqslant i_1, ..., i_p \leqslant n)$ can be expressed as a finite sum $\sum_{|M| \leqslant p} a_M X^*(M)$ where $a_M \in K$. Consider the element

$$u_p = \frac{1}{p!} \sum_{\sigma} X_{i_{\sigma(1)}}^* ... X_{i_{\sigma(p)}}^*,$$

where σ runs over all permutations of the set $\{1, 2, ..., p\}$. It is clear that $u_p = cX^*(M)$, where $c \in K$ and M is a suitable positive integral n-tuple. Using the relation $X_j^* X_k^* - X_k^* X_j^* = [X_j, X_k]^*$ we see that

$$X_{i_1}^* ... X_{i_p}^* - X_{i_{\sigma(1)}}^* ... X_{i_{\sigma(p)}}^*$$

is a linear combination (with coefficients in K) of elements of the form $X_{j_1}^* ... X_{j_{p-1}}^*$ $(1 \leqslant j_1 ... j_{p-1} \leqslant n)$ where each X_{j_q} $(1 \leqslant q \leqslant p - 1)$ belongs to the subalgebra of $\mathfrak{a}$ generated by $X_{i_1}, ..., X_{i_p}$. The formula

$$X_{i_1}^* X_{i_2}^* ... X_{i_p}^* = \sum_{|M| \leqslant p} a_M X^*(M)$$

now follows by induction on p.

Corollary 1.3. *Let* $\mathfrak{b}$ *be a subalgebra of* $\mathfrak{a}$. *Suppose* $\mathfrak{b}$ *has dimension* $n - r$ *and let the basis* $X_1, ..., X_n$ *of* $\mathfrak{a}$ *be chosen in such a way that the*

$n - r$ last elements lie in $\mathfrak{b}$. Let $\mathfrak{B}$ denote the vector subspace of $U(\mathfrak{a})$ spanned by all elements $X^*(M)$ where M varies over all positive integral n-tuples of the form $(0, ..., 0, m_{r+1}, ..., m_n)$. Then $\mathfrak{B}$ is a subalgebra of $U(\mathfrak{a})$.

In fact, the proof above shows that the product $X^*_{i_1} ... X^*_{i_p}$ $(r < i_1, ..., i_p \leqslant n)$ can be written as a linear combination of elements $X^*(M)$ for which $m_1 = ... = m_r = 0$.

3. Left Invariant Affine Connections

Let G be a Lie group, and ∇ an affine connection on G; ∇ is said to be *left invariant* if each L_σ $(\sigma \in G)$ is an affine transformation of G. Let $X_1, ..., X_n$ be a basis of the Lie algebra $\mathfrak{g}$ of G and let $\tilde{X}_1, ..., \tilde{X}_n$ denote the corresponding left invariant vector fields on G. Then if ∇ is left invariant, the vector fields $\nabla_{\tilde{X}_i}(\tilde{X}_j)$ $(1 \leqslant i, j \leqslant n)$ are obviously left invariant. On the other hand suppose ∇ is an affine connection on G such that the vector fields $\nabla_{\tilde{X}_i}(\tilde{X}_j)$ $(1 \leqslant i, j \leqslant n)$ are left invariant. Let Z, Z' be arbitrary vector fields in $\mathfrak{D}^1$. Then $Z = \Sigma_i f_i \tilde{X}_i$, $Z' = \Sigma_j g_j \tilde{X}_j$ where $f_i, g_j \in C^\infty(G)$. Using the axioms ∇_1 and ∇_2 and Prop. 3.3 in Chapter I we find easily that $\nabla_{dL_\sigma Z}(dL_\sigma Z') = dL_\sigma \nabla_Z(Z')$ for each $\sigma \in G$ so ∇ is left invariant.

Proposition 1.4. *There is a one-to-one correspondence between the set of left invariant affine connections ∇ on G and the set of bilinear functions α on $\mathfrak{g} \times \mathfrak{g}$ with values in $\mathfrak{g}$ given by*

$$\alpha(X, Y) = (\nabla_{\tilde{X}}(\tilde{Y}))_e.$$

Let $X \in \mathfrak{g}$. The following statements are then equivalent:

(i) $\alpha(X, X) = 0$;

(ii) *The geodesic $t \to \gamma_X(t)$ is an analytic homomorphism of $\mathbf{R}$ into G.*

Proof. Given a bilinear mapping $\alpha : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, we define the affine connection ∇ by the requirement

$$\nabla_{\tilde{X}_i}(\tilde{X}_j) = \alpha(X_i, X_j)^\sim \qquad (1 \leqslant i, j \leqslant n).$$

By the remark above, ∇ is left invariant, and the correspondence follows. Also, ∇ is analytic.

Next let $X \in \mathfrak{g}$ and let $\tilde{X}$ be the corresponding left invariant vector field on G. Locally there exist integral curves to the vector field $\tilde{X}$ (Chapter I, §7). In other words, there exists a number $\epsilon > 0$ and a curve segment $\Gamma : t \to \Gamma(t)$ $(0 \leqslant t \leqslant \epsilon)$ in G such that

$$\Gamma(0) = e, \qquad \dot{\Gamma}(s) = \tilde{X}_{\Gamma(s)} \tag{2}$$

for $0 \leqslant s \leqslant \epsilon$. Using induction we define $\Gamma(t)$ for all $t \geqslant 0$ by the requirement

$$\Gamma(t) = \Gamma(n\epsilon)\,\Gamma(t - n\epsilon), \qquad \text{if} \qquad n\epsilon \leqslant t \leqslant (n+1)\,\epsilon,$$

n being a nonnegative integer. On the interval $n\epsilon \leqslant t \leqslant (n+1)\,\epsilon$ we have $\Gamma \circ L_{-n\varepsilon} = L_{\Gamma(n\varepsilon)^{-1}} \circ \Gamma$. We use both sides of this equation on the tangent vector $(d/dt)_t$ $(n\epsilon \leqslant t \leqslant (n+1)\,\epsilon)$. From (2) we obtain

$$\begin{aligned} \dot{\Gamma}(t) = d\Gamma\left(\frac{d}{dt}\right)_t &= dL_{\Gamma(n\varepsilon)} \circ d\Gamma \circ dL_{-n\varepsilon}\left(\frac{d}{dt}\right)_t \\ &= dL_{\Gamma(n\varepsilon)} \cdot \tilde{X}_{\Gamma(t-n\varepsilon)} \\ &= \tilde{X}_{\Gamma(t)}. \end{aligned}$$

Thus (2) holds for all $s \geqslant 0$ (including the points $n\epsilon$).

Assume now $\alpha(X, X) = 0$. Then, due to the left invariance of the corresponding affine connection ∇, we have $\nabla_{\tilde{X}}(\tilde{X}) = 0$. Hence the curve segment $\Gamma(t)$ $(t \geqslant 0)$ is a geodesic segment, and by the uniqueness of such, we have $\Gamma(t) = \gamma_X(t)$ for all $t \geqslant 0$. For any affine connection, $\gamma_{-X}(t) = \gamma_X(-\,t)$. Since $\alpha(-\,X, -\,X) = 0$, it follows that $\gamma_X(t)$ is defined for all $t \in \boldsymbol{R}$. Now let $s \geqslant 0$. Then the curves $t \to \gamma_X(s + t)$ and $t \to \gamma_X(s)\,\gamma_X(t)$ are both geodesics in G (since ∇ is left invariant) passing through $\gamma_X(s)$. These geodesics have tangent vectors $\dot{\gamma}_X(s)$ and $dL_{\gamma_X(s)} \cdot X$, respectively, at the point $\gamma_X(s)$. These are equal since (2) holds for all $s \geqslant 0$. We conclude that

$$\gamma_X(s + t) = \gamma_X(s)\,\gamma_X(t) \tag{3}$$

for $s \geqslant 0$ and all t. Using again $\gamma_{-X}(t) = \gamma_X(-\,t)$, we see that (3) holds for all s and t. This proves that (i) $\Rightarrow$ (ii).

Suppose now θ is any analytic homomorphism of $\boldsymbol{R}$ into G such that $\dot{\theta}(0) = X$. Then from $\theta(s + t) = \theta(s)\,\theta(t)$, $(t, s \in \boldsymbol{R})$, follows that

$$\theta(0) = e, \qquad \dot{\theta}(s) = \tilde{X}_{\theta(s)} \qquad \text{for all } s \in \boldsymbol{R}. \tag{4}$$

In particular, if γ_X is an analytic homomorphism, we have $\nabla_{\tilde{X}}(\tilde{X}) = 0$ on the curve γ_X; hence $\alpha(X, X) = (\nabla_{\tilde{X}}(\tilde{X}))_e = 0$.

Corollary 1.5. *Let $X \in \mathfrak{g}$. There exists a unique analytic homomorphism θ of $\boldsymbol{R}$ into G such that $\dot{\theta}(0) = X$.*

Proof. Let ∇ be any affine connection on G for which $\alpha(X, X) = 0$. Then $\theta = \gamma_X$ is a homomorphism with the required properties. For the uniqueness we observe that (4), in connection with $\alpha(X, X) = 0$, shows, that any homomorphism θ with the required properties must be a geodesic; by the uniqueness of geodesics (Prop. 5.3, Chapter I), $\theta = \gamma_X$.

Definition. For each $X \in \mathfrak{g}$, we put $\exp X = \theta(1)$ if θ is the homomorphism of Cor. 1.5. The mapping $X \to \exp X$ of $\mathfrak{g}$ into G is called the *exponential mapping.*

We have the formula

$$\exp(t+s)X = \exp tX \exp sX$$

for all $s, t \in \boldsymbol{R}$ and all $X \in \mathfrak{g}$. This follows immediately from the fact that if $\alpha(X, X) = 0$, then $\theta(t) = \gamma_X(t) = \gamma_{tX}(1) = \exp tX$.

Definition. A *one-parameter subgroup* of a Lie group G is an analytic homomorphism of $\boldsymbol{R}$ into G.

We have seen above that the one-parameter subgroups are the mappings $t \to \exp tX$ where X is an element of the Lie algebra.

We see from Prop. 1.4 and the corollary that the exponential mapping agrees with the mapping Exp_e (from Chapter I) for all left invariant affine connections on G satisfying $\alpha(X, X) = 0$ for all $X \in \mathfrak{g}$. The classical examples (Cartan and Schouten [1]) are $\alpha \equiv 0$ (the $(-)$-connection), $\alpha(X, Y) = \frac{1}{2}[X, Y]$ (the (0)-connection) and $\alpha(X, Y) = [X, Y]$ (the $(+)$-connection).

From Theorem 6.1, Chapter I, we deduce the following statement.

Proposition 1.6. *There exists an open neighborhood N_0 of 0 in $\mathfrak{g}$ and an open neighborhood N_e of e in G such that exp is an analytic diffeomorphism of N_0 onto N_e.*

Let $X_1, \ldots, X_n$ be a basis of $\mathfrak{g}$. The mapping

$$\exp(x_1X_1 + \ldots + x_nX_n) \to (x_1, \ldots, x_n)$$

of N_e onto N_0 is a coordinate system on N_e, called a system of *canonical coordinates* with respect to $X_1, \ldots, X_n$. The set N_e is called a *canonical coordinate neighborhood.* Note that N_0 is not required to be star-shaped.

4. Taylor's Formula and Applications

Let G be a Lie group with Lie algebra $\mathfrak{g}$. Let $X \in \mathfrak{g}$, $g \in G$, and $f \in C^\infty(G)$. Since the homomorphism $\theta(t) = \exp tX$ satisfies $\dot{\theta}(0) = X$ we obtain

$$\tilde{X}_g f = X(f \circ L_g) = \left\{\frac{d}{dt} f(g \exp tX)\right\}_{t=0}, \tag{5}$$

It follows that the value of $\tilde{X}f$ at $g \exp uX$ is

$$[\tilde{X}f](g \exp uX) = \left\{\frac{d}{dt} f(g \exp uX \exp tX)\right\}_{t=0} = \frac{d}{du} f(g \exp uX)$$

and by induction

$$[\tilde{X}^n f]\,(g \exp uX) = \frac{d^n}{du^n} f(g \exp uX).$$

Suppose now that f is analytic at g. Then there exists a star-shaped neighborhood N_0 of 0 in $\mathfrak{g}$ such that

$$f(g \exp X) = P(x_1, ..., x_n) \qquad (X \in N_0),$$

where P denotes an absolutely convergent power series and $(x_1, ..., x_n)$ are the coordinates of X with respect to a fixed basis of $\mathfrak{g}$. Then we have for a fixed $X \in N_0$

$$f(g \exp tX) = P(tx_1, ..., tx_n) = \sum_0^\infty \frac{1}{m!} a_m t^m \qquad (a_m \in \boldsymbol{R}),$$

for $0 \leqslant t \leqslant 1$. It follows that each coefficient a_m equals the mth derivative of $f(g \exp tX)$ for $t = 0$; consequently

$$a_m = [\tilde{X}^m f]\,(g).$$

This proves the "Taylor formula";

$$f(g \exp X) = \sum_0^\infty \frac{1}{n!} [\tilde{X}^n f]\,(g) \tag{6}$$

for $X \in N_0$.

Theorem 1.7. *Let G be a Lie group with Lie algebra $\mathfrak{g}$. The exponential mapping of the manifold $\mathfrak{g}$ into G has the differential*

$$d \exp_X = d(L_{\exp X})_e \circ \frac{1 - e^{-\operatorname{ad} X}}{\operatorname{ad} X} \qquad (X \in \mathfrak{g}).$$

As usual, $\mathfrak{g}$ is here identified with the tangent space $\mathfrak{g}_X$.

Proof. We consider the left invariant affine connection on G given by $\alpha(X, Y) = 0$ for all $X, Y \in \mathfrak{g}$ (Prop. 1.4). Then the left invariant vector field $\nabla_{\tilde{X}}(\tilde{Y})$ vanishes identically on G. It follows that the vector field X^* adapted to X (Chapter I, §6) coincides with $\tilde{X}$ in a neighborhood of e in G. From Theorem 6.5, Chapter I we conclude that the equation

$$d(L_{\exp-tX}) \circ d \exp_{tX}(Y) = \frac{1 - e^{-\operatorname{ad} tX}}{\operatorname{ad} tX}(Y)$$

holds for all t in some interval $|t| < \delta$. Both sides of this equation are analytic functions on $\boldsymbol{R}$ with values in $\mathfrak{g}$. Since they agree for $|t| < \delta$, they must agree for all $t \in \boldsymbol{R}$; this proves the theorem.

The following result will be needed later.

Lemma 1.8. *Let G be a Lie group with Lie algebra $\mathfrak{g}$, and let exp be the exponential mapping of $\mathfrak{g}$ into G. Then, if $X, Y \in \mathfrak{g}$,*

(i) $\exp tX \exp tY = \exp \{t(X + Y) + \frac{t^2}{2}[X, Y] + O(t^3)\}$,

(ii) $\exp(-tX)\exp(-tY)\exp tX \exp tY = \exp\{t^2[X, Y] + O(t^3)\}$,

(iii) $\exp tX \exp tY \exp(-tX) = \exp\{tY + t^2[X, Y] + O(t^3)\}$.

In each case $O(t^3)$ denotes a vector in $\mathfrak{g}$ with the property: there exists an $\epsilon > 0$ such that $(1/t^3)\, O(t^3)$ is bounded and analytic for $|t| < \epsilon$.

We first prove (i). Let f be analytic at e. Then using the formula

$$[\tilde{X}^n f](g \exp tX) = \frac{d^n}{dt^n} f(g \exp tX)$$

twice we obtain

$$[\tilde{X}^n \tilde{Y}^m f](e) = \left[\frac{d^n}{dt^n}\frac{d^m}{ds^m} f(\exp tX \exp sY)\right]_{s=0,\, t=0}.$$

Therefore, the Taylor series for $f(\exp tX \exp sY)$ is

$$f(\exp tX \exp sY) = \sum_{m,n \geqslant 0} \frac{t^n}{n!}\frac{s^m}{m!}[\tilde{X}^n \tilde{Y}^m f](e) \tag{7}$$

for sufficiently small t and s. On the other hand,

$$\exp tX \exp tY = \exp Z(t)$$

for sufficiently small t where $Z(t)$ is a function with values in $\mathfrak{g}$, analytic at $t = 0$. We have $Z(t) = tZ_1 + t^2 Z_2 + O(t^3)$ where Z_1 and Z_2 are fixed vectors in $\mathfrak{g}$. Then if f is any of the canonical coordinate functions $\exp(x_1X_1 + \ldots + x_nX_n) \to x_i$ we have

$$f(\exp Z(t)) = f(\exp(tZ_1 + t^2Z_2)) + O(t^3)$$

$$= \sum_0^\infty \frac{1}{n!} [(t\tilde{Z}_1 + t^2\tilde{Z}_2)^n f](e) + O(t^3). \tag{8}$$

If we compare (7) for $t = s$ and (8) we find $Z_1 = X + Y$, $\frac{1}{2}\tilde{Z}_1^2 + \tilde{Z}_2 = \frac{1}{2}\tilde{X}^2 + \tilde{X}\tilde{Y} + \frac{1}{2}\tilde{Y}^2$. Consequently

$$Z_1 = X + Y, \qquad Z_2 = \tfrac{1}{2}[X, Y],$$

which proves (i). The relation (ii) is obtained by applying (i) twice. To prove (iii), let again f be analytic at e; then for small t

$$f(\exp tX \exp tY \exp(-tX)) = \sum_{m,n,p\geq 0} \frac{t^m}{m!}\frac{t^n}{n!}\frac{t^p}{p!} [\tilde{X}^m\tilde{Y}^n(-\tilde{X})^p f](e) \tag{9}$$

and

$$\exp tX \exp tY \exp(-tX) = \exp S(t)$$

where $S(t) = tS_1 + t^2S_2 + O(t^3)$ and $S_1, S_2 \in \mathfrak{g}$. If f is any canonical coordinate function, then

$$f(\exp S(t)) = f(\exp(tS_1 + t^2S_2)) + O(t^3)$$

$$= \sum_0^\infty \frac{1}{n!} [(t\tilde{S}_1 + t^2\tilde{S}_2)^n f](e) + O(t^3), \tag{10}$$

and we find by comparing coefficients in (9) and (10), $S_1 = Y$, $S_2 = [X, Y]$, which proves (iii).

Remark. The relation (ii) gives a geometric interpretation of the bracket $[X, Y]$; in fact, it shows that $[X, Y]$ is the tangent vector at e to the curve segment

$$s \to \exp(-\sqrt{s}\,X)\exp(-\sqrt{s}\,Y)\exp\sqrt{s}\,X\exp\sqrt{s}\,Y \qquad (s \geqslant 0).$$

Let $D(G)$ denote the algebra of operators on $C^\infty(G)$ generated by all the left invariant vector fields† on G and I (the identity operator on

† It will be shown in Chapter X that $D(G)$ is the algebra of all left invariant differential operators on G.

$C^\infty(G)$). If $X \in \mathfrak{g}$ we shall also denote the corresponding left invariant vector field on G by X. Similarly the operator $\tilde{X}_1 \cdot \tilde{X}_2 \ldots \tilde{X}_k$ ($X_i \in \mathfrak{g}$) will be denoted by $X_1 \cdot X_2 \ldots X_k$ for simplicity. Let $X_1, \ldots, X_n$ be any basis of $\mathfrak{g}$ and put $X(t) = \sum_{i=1}^n t_i X_i$. Let $M = (m_1, \ldots, m_n)$ be a positive integral n-tuple, let $t^M = t_1^{m_1} \cdots t_n^{m_n}$ and let $X(M)$ denote the coefficient of t^M in the expansion of $(|M|!)^{-1}(X(t))^{|M|}$. If $|M| = 0$ put $X(M) = I$. It is clear that $X(M) \in \boldsymbol{D}(G)$.

Proposition 1.9.

(a) *As M varies through all positive integral n-tuples the elements $X(M)$ form a basis of $\boldsymbol{D}(G)$ (considered as a vector space over $\boldsymbol{R}$).*

(b) *The universal enveloping algebra $U(\mathfrak{g})$ is isomorphic to $\boldsymbol{D}(G)$.*

Proof. Let f be an analytic function at $g \in G$; we have by (6)

$$f(g \exp X(t)) = \sum_M t^M [X(M)f](g), \tag{11}$$

if the t_i are sufficiently small. If we compare this formula with the ordinary Taylor formula for the function F defined by $F(t_1, \ldots, t_n) = f(g \exp X(t))$, we obtain

$$[X(M)f](g) = \frac{1}{m_1! \ldots m_n!} \left\{ \frac{\partial^{|M|}}{\partial t_1^{m_1} \ldots \partial t_n^{m_n}} f(g \exp X(t)) \right\}_{t_1 = \ldots = t_n = 0}. \tag{12}$$

It follows immediately that the various $X(M)$ are linearly independent. The Lie algebra $\mathfrak{g}$ has a representation ρ on $C^\infty(G)$ if we associate to each $X \in \mathfrak{g}$ the corresponding left invariant vector field. The representation ρ^* from Prop. 1.1 gives a homomorphism of $U(\mathfrak{g})$ into $\boldsymbol{D}(G)$ such that $\rho^*(X^*) = \rho(X)$ for $X \in \mathfrak{g}$. The mapping ρ^* sends the element $X_{i_1}^* \ldots X_{i_p}^* \in U(\mathfrak{g})$ into $X_{i_1} \ldots X_{i_p} \in \boldsymbol{D}(G)$; thus $\rho^*(U(\mathfrak{g})) = \boldsymbol{D}(G)$. Moreover, ρ^* sends the element $X^*(M) \in U(\mathfrak{g})$ into $X(M) \in \boldsymbol{D}(G)$. Since the elements $X(M)$ are linearly independent, the proposition follows from Prop. 1.2.

Corollary 1.10. *With the notation above, the elements $X_1^{e_1} \ldots X_n^{e_n}$ ($e_i \geqslant 0$) form a basis of $\boldsymbol{D}(G)$.*

Since $X_i X_j - X_j X_i = [X_i, X_j]$ it is clear that each $X(M)$ can be written as a real linear combination of elements $X_1^{e_1} \ldots X_n^{e_n}$ where $e_1 + \ldots + e_n \leqslant |M|$. On the other hand, as noted in the proof of Prop. 1.2, each $X_1^{e_1} \ldots X_n^{e_n}$ can be written as a real linear combination of elements $X(M)$ for which $|M| \leqslant e_1 + \ldots + e_n$. Since the number of elements $X(M)$, $|M| \leqslant e_1 + \ldots + e_n$ equals the number of elements

$X_1^{f_1} \dots X_n^{f_n}$ $(f_1 + \dots + f_n \leqslant e_1 + \dots + e_n)$, the corollary follows from Prop. 1.9.

This corollary shows quickly that $D(G)$ has no divisors of 0.

Definition. Let G and G' be two Lie groups with identity elements e and e'. These groups are said to be *isomorphic* if there exists an analytic isomorphism of G onto G'. The groups G and G' are said to be *locally isomorphic* if there exist open neighborhoods U and U' of e and e', respectively, and an analytic diffeomorphism f of U onto U' satisfying:

(a) If $x, y, xy \in U$, then $f(xy) = f(x) f(y)$.

(b) If $x', y', x'y' \in U'$, then $f^{-1}(x'y') = f^{-1}(x') f^{-1}(y')$.

Theorem 1.11. *Two Lie groups are locally isomorphic if and only if their Lie algebras are isomorphic.*

Proof. Let G be a Lie group with Lie algebra $\mathfrak{g}$. Let $X_1, \dots, X_n$ be a basis of $\mathfrak{g}$. Owing to Prop. 1.9 we can legitimately write $X(M)$ instead of $X^*(M)$; there exist uniquely determined constants $C^P{}_{MN} \in \boldsymbol{R}$ such that

$$X(M) X(N) = \sum_P C^P{}_{MN} X(P),$$

M, N, and P denoting positive integral n-tuples. Owing to Prop. 1.9, the constants $C^P{}_{MN}$ depend only on the Lie algebra $\mathfrak{g}$. If N_e is a canonical coordinate neighborhood of $e \in G$ and $g \in N_e$ let $g_1, \dots, g_n$ denote the canonical coordinates of g. Then if $x, y, xy \in N_e$, we have

$$x = \exp(x_1 X_1 + \dots + x_n X_n), \qquad y = \exp(y_1 X_1 + \dots + y_n X_n),$$
$$xy = \exp((xy)_1 X_1 + \dots + (xy)_n X_n).$$

We also put

$$x^M = x_1^{m_1} \dots x_n^{m_n}, \qquad y^M = y_1^{m_1} \dots y_n^{m_n}.$$

Using (7) on the function $f: x \to x_k$ we find for sufficiently small x_i, y_j

$$(xy)_k = \sum_{M,N} x^M y^N [X(M) X(N) x_k](e). \tag{13}$$

From (12) it follows that

$$[X(P) x_k](e) = \begin{cases} 1 \text{ if } P = (\delta_{k1}, \delta_{k2}, \dots, \delta_{kn}), \\ 0 \text{ otherwise.} \end{cases}$$

Putting $[k] = (\delta_{1k}, \dots, \delta_{nk})$, relation (13) becomes

$$(xy)_k = \sum_{M,N} C^{[k]}{}_{MN}\, x^M y^N, \tag{14}$$

if x_i, y_j $(1 \leqslant i, j \leqslant n)$ are sufficiently small. This last formula shows that the group law is determined in a neighborhood of e by the Lie algebra. In particular, Lie groups with isomorphic Lie algebras are locally isomorphic. Before proving the converse of Theorem 1.11 we prove a general lemma about homomorphisms.

Lemma 1.12. *Let H and K be Lie groups with Lie algebras $\mathfrak{h}$ and $\mathfrak{k}$, respectively. Let φ be an analytic homomorphism of H into K. Then $d\varphi_e$ is a homomorphism of $\mathfrak{h}$ into $\mathfrak{k}$ and*

$$\varphi(\exp X) = \exp d\varphi_e(X) \qquad (X \in \mathfrak{h}). \tag{15}$$

Proof. Let $X \in \mathfrak{h}$. The mapping $t \to \varphi(\exp tX)$ is an analytic homomorphism of $\boldsymbol{R}$ into K. If we put $X' = d\varphi_e(X)$, Cor. 1.5 implies that $\varphi(\exp tX) = \exp tX'$ for all $t \in \boldsymbol{R}$. Since φ is a homomorphism, we have $\varphi \circ L_\sigma = L_{\varphi(\sigma)} \circ \varphi$ for $\sigma \in H$. It follows that

$$(d\varphi_\sigma) \circ dL_\sigma \cdot X = dL_{\varphi(\sigma)} \cdot X'.$$

This means that the left invariant vector fields $\tilde{X}$ and $\tilde{X}'$ are φ-related. Hence, by Prop. 3.3, Chapter I, $d\varphi_e$ is a homomorphism and the lemma is proved.

To finish the proof of Theorem 1.11, we suppose now that the Lie groups G and G' are locally isomorphic. Let $\mathfrak{g}$ and $\mathfrak{g}'$ denote their respective Lie algebras. There is no restriction of generality in assuming G and G' connected. Let $(\tilde{G}, \pi)$ and $(\tilde{G}', \pi')$ be the universal covering groups of G and G', respectively. It follows from Lemma 1.12 that the mappings $d\pi_e$ and $d\pi'_e$ are Lie algebra isomorphisms. From the first part of the proof it follows that π and π' are local isomorphisms. The given local isomorphism between G and G' therefore induces a local isomorphism θ between $\tilde{G}$ and $\tilde{G}'$. Now $\tilde{G}$ is simply connected. Owing to a well-known theorem on topological groups, see e.g. Chevalley [2], p. 49, there exists a continuous homomorphism $\bar{\theta}$ of $\tilde{G}$ into $\tilde{G}'$ which coincides with θ in a neighborhood of the identity; in particular, $\bar{\theta}$ is analytic. On interchanging $\tilde{G}$ and $\tilde{G}'$ we see that $\bar{\theta}$ is an isomorphism of $\tilde{G}$ onto $\tilde{G}'$. From Lemma 1.12 it follows that $d\bar{\theta}_e$ is an isomorphism between the corresponding Lie algebras. This in turn gives the desired isomorphism of the Lie algebras $\mathfrak{g}$ and $\mathfrak{g}'$.

Example. Let $\boldsymbol{GL}(n, \boldsymbol{R})$ denote the group of all real nonsingular $n \times n$ matrices and let $\mathfrak{gl}(n, \boldsymbol{R})$ denote the Lie algebra of all real $n \times n$ matrices, the bracket being $[A, B] = AB - BA$, $A, B \in \mathfrak{gl}(n, \boldsymbol{R})$. If we consider the matrix $\sigma = (x_{ij}(\sigma)) \in \boldsymbol{GL}(n, \boldsymbol{R})$ as the set of coordinates of a point in $\boldsymbol{R}^{n^2}$ then $\boldsymbol{GL}(n, \boldsymbol{R})$ can be regarded as an open submanifold of $\boldsymbol{R}^{n^2}$. With this analytic structure $\boldsymbol{GL}(n, \boldsymbol{R})$ is a Lie group;

this is obvious by considering the expression of $x_{ij}(\sigma\tau^{-1})$ $(\tau, \sigma \in \boldsymbol{GL}(n, \boldsymbol{R}))$ in terms of $x_{kl}(\sigma)$, $x_{pq}(\tau)$, given by matrix multiplication.

Let X be an element of $\mathfrak{L}(\boldsymbol{GL}(n, \boldsymbol{R}))$ and let $\tilde{X}$ denote the left invariant vector field on $\boldsymbol{GL}(n, \boldsymbol{R})$ such that $\tilde{X}_e = X$. Let $(a_{ij}(X))$ denote the matrix $(\tilde{X}_e x_{ij})$. We shall prove that the mapping $\varphi : X \rightarrow (a_{ij}(X))$ is an isomorphism of $\mathfrak{L}(\boldsymbol{GL}(n, \boldsymbol{R}))$ onto $\mathfrak{gl}(n, \boldsymbol{R})$. The mapping φ is linear and one-to-one; in fact, the relation $(a_{ij}(X)) = 0$ implies $\tilde{X}_e f = 0$ for all differentiable functions f, hence $\tilde{X} = 0$. Considering the dimensions of the Lie algebras we see that the range of φ is $\mathfrak{gl}(n, \boldsymbol{R})$. Next we consider $[\tilde{X} x_{ij}](\sigma) = (dL_\sigma X)\, x_{ij} = X(x_{ij} \circ L_\sigma)$. If $\tau \in \boldsymbol{GL}(n, \boldsymbol{R})$, then

$$(x_{ij} \circ L_\sigma)(\tau) = x_{ij}(\sigma\tau) = \sum_{k=1}^{n} x_{ik}(\sigma)\, x_{kj}(\tau). \tag{16}$$

Hence

$$[\tilde{X} x_{ij}](\sigma) = \sum_{k=1}^{u} x_{ik}(\sigma)\, a_{kj}(X). \tag{17}$$

It follows that

$$[(\tilde{X}\tilde{Y} - \tilde{Y}\tilde{X})\, x_{ij}](e) = \sum_{k=1}^{n} a_{ik}(X)\, a_{kj}(Y) - a_{ik}(Y)\, a_{kj}(X)$$

$$= [\varphi(X), \varphi(Y)]_{ij}.$$

Consequently, the Lie algebra of $\boldsymbol{GL}(n, \boldsymbol{R})$ can be identified with $\mathfrak{gl}(n, \boldsymbol{R})$ so we now write X_{ij} instead of $a_{ij}(X)$ above. Using the general formula

$$[\tilde{X} f](\exp tX) = \frac{d}{dt} f(\exp tX)$$

for a differentiable function f, we obtain from (16) and (17)

$$\frac{d}{dt} x_{ij}(\exp tX) = \sum_{k=1}^{n} x_{ik}(\exp tX)\, X_{kj}.$$

Thus the matrix function $Y(t) = \exp tX$ satisfies the differential equation

$$\frac{dY(t)}{dt} = Y(t)\, X, \qquad Y(0) = I.$$

Since this equation is also satisfied by the matrix exponential function

$$Y(t) = e^{tX} = I + tX + \frac{t^2 X^2}{2!} + \dots,$$

we conclude that $\exp X = e^X$ for all $X \in \mathfrak{gl}(n, \boldsymbol{R})$.

Let V be an n-dimensional vector space over $\boldsymbol{R}$. Let $\mathfrak{gl}(V)$ be the Lie algebra of all endomorphisms of V and let $GL(V)$ be the group of invertible endomorphisms of V. Fix a basis $e_1, ..., e_n$ of V. To each $\sigma \in \mathfrak{gl}(V)$ we associate the matrix $(x_{ij}(\sigma))$ given by

$$\sigma e_j = \sum_{i=1}^{n} x_{ij}(\sigma)\, e_i.$$

The mapping $J_e : \sigma \to (x_{ij}(\sigma))$ is an isomorphism of $\mathfrak{gl}(V)$ onto $\mathfrak{gl}(n, \boldsymbol{R})$ whose restriction to $GL(V)$ is an isomorphism of $GL(V)$ onto $GL(n, \boldsymbol{R})$. This isomorphism turns $GL(V)$ into a Lie group with Lie algebra isomorphic to $\mathfrak{gl}(V)$. If $f_1, ..., f_n$ is another basis of V, we get another isomorphism $J_f : \mathfrak{gl}(V) \to \mathfrak{gl}(n, \boldsymbol{R})$. If $A \in GL(V)$ is determined by $Ae_i = f_i$ $(1 \leqslant i \leqslant n)$, then J_f and J_e are connected by the equation $J_e(\sigma) = J_f(A) J_f(\sigma) J_f(A^{-1})$. Since the mapping $g \to J_f(A) g J_f(A^{-1})$ is an analytic isomorphism of $GL(n, \boldsymbol{R})$ onto itself, we conclude: (1) The analytic structure of $GL(V)$ is independent of the choice of basis. (2) There is an isomorphism of $\mathfrak{L}(GL(V))$ onto $\mathfrak{gl}(V)$ (namely, $J_e^{-1} \circ dJ_e$) which is independent of the choice of basis of V.

§ 2. Lie Subgroups and Subalgebras

Definition. Let G be a Lie group. A submanifold H of G is called a *Lie subgroup* if

(i) H is a subgroup of the (abstract) group G;

(ii) H is a topological group.

A Lie subgroup is itself a Lie group; in order to see this, consider the analytic mapping $\alpha : (x, y) \to xy^{-1}$ of $G \times G$ into G. Let α_H denote the restriction of α to $H \times H$. Then the mapping $\alpha_H : H \times H \to G$ is analytic, and by (ii) the mapping $\alpha_H : H \times H \to H$ is continuous. In view of Lemma 14.1, Chapter I, the mapping α_H is an analytic mapping of $H \times H$ into H so H is a Lie group.

A connected Lie subgroup is often called an *analytic subgroup.*

Theorem 2.1. *Let G be a Lie group. If H is a Lie subgroup of G, then the Lie algebra $\mathfrak{h}$ of H is a subalgebra of $\mathfrak{g}$, the Lie algebra of G. Each subalgebra of $\mathfrak{g}$ is the Lie algebra of exactly one connected Lie subgroup of G.*

Proof. If I denotes the identity mapping of H into G, then by Lemma 1.12 dI_e is a homomorphism of $\mathfrak{h}$ into $\mathfrak{g}$. Since H is a submanifold of G, dI_e is one-to-one. Thus $\mathfrak{h}$ can be regarded as a subalgebra of $\mathfrak{g}$.

Let $\exp_{\mathfrak{h}}$ and $\exp_{\mathfrak{g}}$, respectively, denote the exponential mappings of $\mathfrak{h}$ into H and of $\mathfrak{g}$ into G. From Cor. 1.5 we get immediately

$$\exp_{\mathfrak{h}}(X) = \exp_{\mathfrak{g}}(X), \qquad X \in \mathfrak{h}. \tag{1}$$

We can therefore drop the subscripts and write exp instead of $\exp_{\mathfrak{h}}$ and $\exp_{\mathfrak{g}}$. If $X \in \mathfrak{h}$, then the mapping $t \to \exp tX$ $(t \in \boldsymbol{R})$ is a curve in H. On the other hand, suppose $X \in \mathfrak{g}$ such that the mapping $t \to \exp tX$ is a path in H, that is, a continuous curve in H. By Lemma 14.1, Chapter I, the mapping $t \to \exp tX$ is an analytic mapping of $\boldsymbol{R}$ into H. Thus $X \in \mathfrak{h}$, so we have

$$\mathfrak{h} = \{X \in \mathfrak{g} : \text{the map } t \to \exp tX \text{ is a path in } H\}. \tag{2}$$

To prove the second statement of Theorem 2.1, suppose $\mathfrak{h}$ is any subalgebra of $\mathfrak{g}$. Let H be the smallest subgroup of G containing $\exp \mathfrak{h}$. Let $(X_1, ..., X_n)$ be a basis of $\mathfrak{g}$ such that (X_i) $(r < i \leqslant n)$ is a basis of $\mathfrak{h}$. Then we know from Cor. 1.3 (and Prop. 1.9) that all real linear combinations of elements $X(M)$, where the n-tuple M has the form $(0, ..., 0, m_{r+1}, ..., m_n)$, actually form a subalgebra of $U(\mathfrak{g})$. Let $|X| = (x_1^2 + ... + x_n^2)^{1/2}$ if $X = x_1X_1 + ... + x_nX_n$ $(x_i \in \boldsymbol{R})$. Choose $\delta > 0$ such that exp is a diffeomorphism of the open ball $B_\delta = \{X : |X| < \delta\}$ onto an open neighborhood N_e of e in G and such that (14), §1, holds for $x, y, xy \in N_e$. Denote the subset $\exp(\mathfrak{h} \cap B_\delta)$ of N_e by V. The mapping

$$\exp(x_{r+1}X_{r+1} + ... + x_nX_n) \to (x_{r+1}, ..., x_n)$$

is a coordinate system on V with respect to which V is a connected manifold. Since $\mathfrak{h} \cap B_\delta$ is a submanifold of B_δ, V is a submanifold of N_e; hence V is a submanifold of G. Now suppose $x, y \in V$, $xy \in N_e$, and consider the canonical coordinates of xy as given by (14), §1. Since $x_k = y_k = 0$, if $1 \leqslant k \leqslant r$, we find (using the remark above about $X(M)$) that $(xy)_k = 0$ if $1 \leqslant k \leqslant r$. Thus we have

$$VV \cap N_e \subset V. \tag{3}$$

Let $\mathscr{V}$ denote the family of subsets of H containing a neighborhood of e in V. Let us verify that $\mathscr{V}$ satisfies the following six axioms for a topological group (Chevalley [2], Chapter II, § II).

I. The intersection of any two sets of $\mathscr{V}$ lies in $\mathscr{V}$.
II. The intersection of all sets of $\mathscr{V}$ is $\{e\}$.
III. Any subset of H containing a set in $\mathscr{V}$ lies in $\mathscr{V}$.
IV. If $U \in \mathscr{V}$, there exists a set $U_1 \in \mathscr{V}$ such that $U_1U_1 \subset U$.
V. If $U \in \mathscr{V}$, then $U^{-1} \in \mathscr{V}$.
VI. If $U \in \mathscr{V}$ and $h \in H$, then $hUh^{-1} \in \mathscr{V}$.

Of these axioms, I, II, III, and V are obvious and IV follows from (3). For VI, let $U \in \mathscr{V}$ and $h \in H$. Let log denote the inverse of the mapping $\exp : B_\delta \to N_e$. Then log maps V onto $\mathfrak{h} \cap B_\delta$. If $X \in \mathfrak{g}$, there exists a unique vector $X' \in \mathfrak{g}$ such that $h \exp tX\, h^{-1} = \exp tX'$ for all $t \in \boldsymbol{R}$. The mapping $X \to X'$ is an automorphism of $\mathfrak{g}$ (Lemma 1.12); it maps $\mathfrak{h}$ into itself as is easily seen from (3) by using a decomposition $h = \exp Z_1 \ldots \exp Z_p$ where each Z_i belongs to $B_\delta \cap \mathfrak{h}$. Consequently, we can select δ_1 $(0 < \delta_1 < \delta)$ such that the open ball B_{δ_1} satisfies

$$h \exp (B_{\delta_1} \cap \mathfrak{h})\, h^{-1} \subset V,$$

$$h (\exp B_{\delta_1})\, h^{-1} \subset N_e.$$

The mapping $X \to \log (h \exp X h^{-1})$ of $B_{\delta_1} \cap \mathfrak{h}$ into $B_\delta \cap \mathfrak{h}$ is regular so the image of $B_{\delta_1} \cap \mathfrak{h}$ is a neighborhood of 0 in $\mathfrak{h}$. Applying the mapping exp we see that $h \exp (B_{\delta_1} \cap \mathfrak{h})\, h^{-1}$ is a neighborhood of e in V. This shows that $hUh^{-1} \in \mathscr{V}$. Axioms I-VI are therefore satisfied. Hence there exists a topology on H such that H is a topological group and such that $\mathscr{V}$ is the family of neighborhoods of e in H. In particular V is a neighborhood of e in H.

For each $z \in G$, consider the mapping

$$\Phi_z : z \exp (x_1 X_1 + \ldots + x_n X_n) \to (x_1, \ldots, x_n),$$

which maps zN_e onto B_δ. Let φ_z denote the restriction of Φ_z to zV. If $z \in H$ then φ_z maps the neighborhood zV (of z in H) onto the open subset $B_\delta \cap \mathfrak{h}$ in Euclidean space $\boldsymbol{R}^{n-r}$. Moreover, if $z_1, z_2 \in H$ the mapping $\varphi_{z_1} \circ \varphi_{z_2}^{-1}$ is the restriction of $\Phi_{z_1} \circ \Phi_{z_2}^{-1}$ to an open subset of $\mathfrak{h}$, hence analytic. The space H with the collection of maps φ_z, $z \in H$, is therefore an analytic manifold.

Now V is a submanifold of G. Since left translations are diffeomorphisms of H it follows that H is a submanifold of G. Hence H is a Lie subgroup of G.

We know that $\dim H = \dim \mathfrak{h}$. For $i > r$ the mapping $t \to \exp tX_i$ is a curve in H so by (2), $X_i \in \mathfrak{h}$. This proves that H has Lie algebra $\mathfrak{h}$. Moreover, H is connected since it is generated by $\exp \mathfrak{h}$ which is a connected neighborhood of e in H.

Finally, in order to prove uniqueness, suppose H_1 is any connected Lie subgroup of G with $(H_1)_e = \mathfrak{h}$. From (1) we see that $H = H_1$ (set theoretically). Since exp is an analytic diffeomorphism of a neighborhood of 0 in $\mathfrak{h}$ onto a neighborhood of e in H and H_1, it is clear that the Lie groups H and H_1 coincide.

Corollary 2.2. *Suppose H_1 and H_2 are two Lie subgroups of a Lie group G such that $H_1 = H_2$ (as topological groups). Then $H_1 = H_2$ (as Lie groups).*

Relation (2) shows, in fact, that H_1 and H_2 have the same Lie algebra. By Theorem 2.1, their identity components coincide as Lie groups. Since left translations on H_1 and H_2 are analytic, it follows at once that the Lie groups H_1 and H_2 coincide.

Theorem 2.3. *Let G be a Lie group with Lie algebra $\mathfrak{g}$ and H an (abstract) subgroup of G. Suppose H is a closed subset of G. Then there exists a unique analytic structure on H such that H is a topological Lie subgroup of G.*

We begin by proving a simple lemma.

Lemma 2.4. *Suppose $\mathfrak{g}$ is a direct sum $\mathfrak{g} = \mathfrak{m} + \mathfrak{n}$ where $\mathfrak{m}$ and $\mathfrak{n}$ are two vector subspaces of $\mathfrak{g}$. Then there exist bounded, open, connected neighborhoods $U_{\mathfrak{m}}$ and $U_{\mathfrak{n}}$ of 0 in $\mathfrak{m}$ and $\mathfrak{n}$, respectively, such that the mapping $\Phi : (A, B) \to \exp A \exp B$ is a diffeomorphism of $U_{\mathfrak{m}} \times U_{\mathfrak{n}}$ onto an open neighborhood of e in G.*

Proof. Let $X_1, ..., X_n$ be a basis of $\mathfrak{g}$ such that $X_i \in \mathfrak{m}$ for $1 \leqslant i \leqslant r$, $X_j \in \mathfrak{n}$ for $r < j \leqslant n$. Let $\{t_1, ..., t_n\}$ denote the canonical coordinates of the element $\exp(x_1X_1 + ... + x_rX_r) \exp(x_{r+1}X_{r+1} + ... + x_nX_n)$ with respect to this basis. Then $t_j = \varphi_j(x_1, ..., x_n)$, $1 \leqslant j \leqslant n$, where the functions φ_j are analytic at $(0, ..., 0)$. If $x_i = \delta_{ij}s$, then $t_i = \delta_{ij}s$ and the Jacobian determinant $\partial(\varphi_1, ..., \varphi_n)/\partial(x_1, ..., x_n)$ equals 1 for $x_1 = ... = x_n = 0$. This proves the lemma (Prop. 3.1, Chapter I).

Remark. The lemma generalizes immediately to an arbitrary direct decomposition $\mathfrak{g} = \mathfrak{m}_1 + ... + \mathfrak{m}_s$ of $\mathfrak{g}$ into subspaces.

Turning now to the proof of Theorem 2.3, let $\mathfrak{h}$ denote the subset of $\mathfrak{g}$ given by

$$\mathfrak{h} = \{X : \exp tX \in H \text{ for all } t \in \boldsymbol{R}\}.$$

We shall prove that $\mathfrak{h}$ is a subalgebra of $\mathfrak{g}$. First we note that $X \in \mathfrak{h}$, $s \in \boldsymbol{R}$ implies $sX \in \mathfrak{h}$. Next, suppose $X, Y \in \mathfrak{h}$. By Lemma 1.8 we have for a given $t \in \boldsymbol{R}$,

$$\left(\exp \frac{t}{n} X \exp \frac{t}{n} Y\right)^n = \exp\left\{t(X + Y) + \frac{t^2}{2n}[X, Y] + O\left(\frac{1}{n^2}\right)\right\},$$

$$\left(\exp\left(-\frac{t}{n} X\right) \exp\left(-\frac{t}{n} Y\right) \exp \frac{t}{n} X \exp \frac{t}{n} Y\right)^{n^2} = \exp\left\{t^2[X, Y] + O\left(\frac{1}{n}\right)\right\}.$$

The left-hand sides of these equations belong to H; since H is closed, the limit as $n \to \infty$ also belongs to H. Thus $t(X + Y) \in \mathfrak{h}$ and $t^2[X, Y] \in \mathfrak{h}$ as desired.

Let H^* denote the connected Lie subgroup of G with Lie algebra $\mathfrak{h}$. Then $H^* \subset H$. We shall now prove, that if H is given the relative topology of G, and H_0 is the identity component of H, then $H^* = H_0$ (as topological groups). It suffices to prove that if N is a neighborhood of e in H^*, then N is a neighborhood of e in H. If N were not a neighborhood of e in H, there would exist a sequence $(c_k) \subset H - N$ such that $c_k \to e$ (in the topology of G). Using Lemma 2.4 for $\mathfrak{h} = \mathfrak{n}$ and $\mathfrak{m}$ any complementary subspace we can assume that $c_k = \exp A_k \exp B_k$ where $A_k \in U_{\mathfrak{m}}$, $B_k \in U_{\mathfrak{n}}$, and $\exp B_k \in N$. Then

$$A_k \neq 0, \qquad \lim A_k = 0.$$

Since $A_k \neq 0$, there exists an integer $r_k > 0$ such that

$$r_k A_k \in U_{\mathfrak{m}}, \qquad (r_k + 1) A_k \notin U_{\mathfrak{m}}.$$

Now, $U_{\mathfrak{m}}$ is bounded, so we can assume, passing to a subsequence if necessary, that the sequence $(r_k A_k)$ converges to a limit $A \in \mathfrak{m}$. Since $(r_k + 1) A_k \notin U_{\mathfrak{m}}$ and $A_k \to 0$, we see that A lies on the boundary of $U_{\mathfrak{m}}$; in particular $A \neq 0$.

Let p, q be any integers ($q > 0$). Then we can write $pr_k = qs_k + t_k$ where s_k, t_k are integers and $0 \leqslant t_k < q$. Then $\lim (t_k/q) A_k = 0$, so

$$\exp \frac{p}{q} A = \lim_k \exp \frac{pr_k}{q} A_k = \lim_k (\exp A_k)^{s_k},$$

which belongs to H. By continuity, $\exp tA \in H$ for each $t \in \boldsymbol{R}$, so $A \in \mathfrak{h}$. This contradicts the fact that $A \neq 0$ and $A \in \mathfrak{m}$.

We have therefore proved: (1) H_0 is open in H (taking $N = H^*$); (2) H_0 (and therefore H) has an analytic structure compatible with the relative topology of G in which it is a submanifold of G, hence a Lie subgroup of G. The uniqueness statement of Theorem 2.3 is immediate from Cor. 2.2.

Remark. The subgroup H above is discrete if and only if $\mathfrak{h} = 0$.

Lemma 2.5. *Let G be a Lie group and H a Lie subgroup. Let $\mathfrak{g}$ and $\mathfrak{h}$ denote the corresponding Lie algebras. Suppose H is a topological subgroup of G. Then there exists an open neighborhood V of 0 in $\mathfrak{g}$ such that:*

(i) *exp maps V diffeomorphically onto an open neighborhood of e in G.*

(ii) *$\exp (V \cap \mathfrak{h}) = (\exp V) \cap H$.*

Proof. First select a neighborhood W_0 of 0 in $\mathfrak{g}$ such that exp is one-to-one on W_0. Then select an open neighborhood N_0 of 0 in $\mathfrak{h}$ such that $N_0 \subset W_0$ and such that exp is a diffeomorphism of N_0 onto an open neighborhood N_e of e in H. Now, since H is a topological subspace of G there exists a neighborhood U_e of e in G such that $U_e \cap H = N_e$. Finally select an open neighborhood V of 0 in $\mathfrak{g}$ such that $V \subset W_0$, $V \cap \mathfrak{h} \subset N_0$ and such that exp is a diffeomorphism of V onto an open subset of G contained in U_e. Then V satisfies (i). Condition (ii) is also satisfied. In fact, let $X \in V$ such that $\exp X \in H$. Since $\exp X \in U_e \cap H = N_e$ there exists a vector $X_{\mathfrak{h}} \in N_0$ such that $\exp X_{\mathfrak{h}} = \exp X$. Since $X, X_{\mathfrak{h}} \in W_0$ we have $X = X_{\mathfrak{h}}$ so $\exp X \in \exp (V \cap \mathfrak{h})$. This proves $(\exp V) \cap H \subset \exp (V \cap \mathfrak{h})$. The converse inclusion being obvious the lemma is proved.

Theorem 2.6. *Let G and H be Lie groups and φ a continuous homomorphism of G into H. Then φ is analytic.*

Proof. Let the Lie algebras of G and H be denoted by $\mathfrak{g}$ and $\mathfrak{h}$, respectively. The product manifold $G \times H$ is a Lie group whose Lie algebra is the product $\mathfrak{g} \times \mathfrak{h}$ as defined in §1, No. 1. The graph of φ is the subset of $G \times H$ given by $K = \{(g, \varphi(g)) : g \in G\}$. It is obvious that K is a closed subgroup of $G \times H$. As a result of Theorem 2.3, K has a unique analytic structure under which it is a topological Lie subgroup of $G \times H$. Its Lie algebra is given by

$$\mathfrak{k} = \{(X, Y) \in \mathfrak{g} \times \mathfrak{h} : (\exp tX, \exp tY) \in K \text{ for } t \in \boldsymbol{R}\}. \tag{4}$$

Let N_0 be an open neighborhood of 0 in $\mathfrak{h}$ such that exp maps N_0 diffeomorphically onto an open neighborhood N_e of e in H. Let M_0 and M_e be chosen similarly for G. We may assume that $\varphi(M_e) \subset N_e$. In view of Lemma 2.5 we can also assume that exp is a diffeomorphism of $(M_0 \times N_0) \cap \mathfrak{k}$ onto $(M_e \times N_e) \cap K$. We shall now show that for a given $X \in \mathfrak{g}$ there exists a unique $Y \in \mathfrak{h}$ such that $(X, Y) \in \mathfrak{k}$. The uniqueness is obvious from (4); in fact, if (X, Y_1) and (X, Y_2) belong to $\mathfrak{k}$, then $(0, Y_1 - Y_2) \in \mathfrak{k}$ so by (4), $(e, \exp t(Y_1 - Y_2)) \in K$ for all $t \in \boldsymbol{R}$. By the definition of K, $\exp t(Y_1 - Y_2) = \varphi(e) = e$ for $t \in \boldsymbol{R}$ so $Y_1 - Y_2 = 0$. In order to prove the existence of Y, select an integer $r > 0$ such that the vector $X_r = (1/r)\, X$ lies in M_0. Since $\varphi(\exp X_r) \in N_e$, there exists a unique vector $Y_r \in N_0$ such that $\exp Y_r = \varphi(\exp X_r)$ and a unique $Z_r \in (M_0 \times N_0) \cap \mathfrak{k}$ such that

$$\exp Z_r = (\exp X_r, \exp Y_r).$$

Now exp is one-to-one on $M_0 \times N_0$ so this relation implies $Z_r = (X_r, Y_r)$

and we can put $Y = rY_r$. The mapping $\psi : X \to Y$ thus obtained is clearly a homomorphism of $\mathfrak{g}$ into $\mathfrak{h}$. Relation (4) shows that

$$\varphi(\exp tX) = \exp t\psi(X), \qquad X \in \mathfrak{g}. \tag{5}$$

Let $X_1, \ldots, X_n$ be a basis of $\mathfrak{g}$. Then by (5)

$$\varphi((\exp t_1X_1)(\exp t_2X_2) \ldots (\exp t_nX_n)) = (\exp t_1\psi(X_1))(\exp t_2\psi(X_2)) \ldots (\exp t_n\psi(X_n)). \tag{6}$$

The remark following Lemma 2.4 shows that the mapping $(\exp t_1X_1) \ldots (\exp t_nX_n) \to (t_1, \ldots, t_n)$ is a coordinate system on a neighborhood of e in G. But then by (6), φ is analytic at e, hence everywhere on G.

We shall now see that a simple countability assumption makes it possible to sharpen relation (2) and Cor. 2.2 substantially.

Proposition 2.7. *Let G be a Lie group and H a Lie subgroup. Let $\mathfrak{g}$ and $\mathfrak{h}$ denote the corresponding Lie algebras. Assume that the Lie group H has at most countably many components. Then*

$$\mathfrak{h} = \{X \in \mathfrak{g} : \exp tX \in H \text{ for all } t \in \boldsymbol{R}\}.$$

Proof. We use Lemma 2.4 for $\mathfrak{n} = \mathfrak{h}$ and $\mathfrak{m}$ any complementary subspace to $\mathfrak{h}$ in $\mathfrak{g}$. Let V denote the set $\exp U_{\mathfrak{m}} \exp U_{\mathfrak{h}}$ (from Lemma 2.4) with the relative topology of G and put

$$\mathfrak{a} = \{A \in U_{\mathfrak{m}} : \exp A \in H\}.$$

Then

$$H \cap V = \bigcup_{A \in \mathfrak{a}} \exp A \exp U_{\mathfrak{h}}$$

and this is a disjoint union due to Lemma 2.4. Each member of this union is a neighborhood in H. Since H has a countable basis the set $\mathfrak{a}$ must be countable. Consider now the mapping π of V onto $U_{\mathfrak{m}}$ given by $\pi(\exp X \exp Y) = X$ ($X \in U_{\mathfrak{m}}$, $Y \in U_{\mathfrak{h}}$). This mapping is continuous and maps $H \cap V$ onto $\mathfrak{a}$. The component of e in $H \cap V$ (in the topology of V) is mapped by π onto a connected countable subset of $U_{\mathfrak{m}}$, hence the single point 0. Since $\pi^{-1}(0) = \exp U_{\mathfrak{h}}$ we conclude that $\exp U_{\mathfrak{h}}$ is the component of e in $H \cap V$ (in the topology of V).

Now let $X \in \mathfrak{g}$ such that $\exp tX \in H$ for all $t \in \boldsymbol{R}$. The mapping $\varphi : t \to \exp tX$ of $\boldsymbol{R}$ into G is continuous. Hence there exists a connected neighborhood U of 0 in $\boldsymbol{R}$ such that $\varphi(U) \subset V$. Then $\varphi(U) \subset H \cap V$ and since $\varphi(U)$ is connected, $\varphi(U) \subset \exp U_{\mathfrak{h}}$. But $\exp U_{\mathfrak{h}}$ is an arbitrarily small neighborhood of e in H so the mapping φ is a continuous mapping of $\boldsymbol{R}$ into H. By (2) we have $X \in \mathfrak{h}$ and the proposition is proved.

Remark. The countability assumption is essential in Prop. 2.7 as is easily seen by considering a Lie subgroup with the discrete topology.

Corollary 2.8. *Let G be a Lie group and let H_1 and H_2 be two Lie subgroups each having countably many components. Suppose that $H_1 = H_2$ (set theoretically). Then $H_1 = H_2$ (as Lie groups).*

In fact, Prop. 2.7 shows that H_1 and H_2 have the same Lie algebra.

Corollary 2.9. *Let G be a Lie group and let K and H be two analytic subgroups of G. Assume $K \subset H$. Then the Lie group K is an analytic subgroup of the Lie group H.*

Let $\mathfrak{k}$ and $\mathfrak{h}$ denote the Lie algebras of K and H. Then $\mathfrak{k} \subset \mathfrak{h}$ by Prop. 2.7. Let K^* denote the analytic subgroup of H with Lie algebra $\mathfrak{k}$. Then the analytic subgroups K and K^* of G have the same Lie algebra. By Theorem 2.1 the Lie groups K and K^* coincide.

Let $\boldsymbol{S}^1$ denote the unit circle and $\boldsymbol{T}$ the group $\boldsymbol{S}^1 \times \boldsymbol{S}^1$. Let $t \to \gamma(t)$ ($t \in \boldsymbol{R}$) be a continuous one-to-one homomorphism of $\boldsymbol{R}$ into $\boldsymbol{T}$. If we carry the analytic structure of $\boldsymbol{R}$ over by the homomorphism we obtain a Lie subgroup $\Gamma = \gamma(\boldsymbol{R})$ of $\boldsymbol{T}$. This Lie subgroup is neither closed in $\boldsymbol{T}$ nor a topological subgroup of $\boldsymbol{T}$. We shall now see that these anomalies go together.

Theorem 2.10. *Let G be a Lie group and H a Lie subgroup of G.*

(i) *If H is a topological subgroup of G then H is closed in G.*

(ii) *If H has at most countably many components and is closed in G then H is a topological subgroup of G.*

Proof. (i) Although no countability assumptions are imposed it suffices to prove that if a sequence $(h_n) \subset H$ converges (in G) to an element $g \in G$ then $g \in H$. Let $\mathfrak{h}$ and $\mathfrak{g}$ denote the Lie algebras of H and G, respectively. Let V be a neighborhood of 0 in $\mathfrak{g}$ with the properties of Lemma 2.5. We may take V bounded. Let U be an open neighborhood of 0 in $\mathfrak{g}$ contained in V such that $\exp(-U) \exp U \subset \exp V$. Then there exists an integer N such that $h_n \in g \exp U$ for $n \geqslant N$. It follows that $h_N^{-1}h_n \in (\exp V) \cap H$ for $n \geqslant N$. By Lemma 2.5 there exists an element $X_n \in V \cap \mathfrak{h}$ such that

$$h_N^{-1}h_n = \exp X_n \qquad (n \geqslant N).$$

Since V is bounded we may, passing to a subsequence if necessary, assume that the sequence (X_n) converges to an element $Z \in \mathfrak{g}$. Then $Z \in \mathfrak{h}$ and (h_n) converges to $h_N \exp Z$ so $g \in H$.

(ii) H being closed in G it has by Theorem 2.3 an analytic structure in which it is a topological Lie subgroup of G. Let H' denote this Lie subgroup. Then the identity mapping $I : H \to H'$ is continuous. Each component of H lies in a component of H'. Since H has countably many components the same holds for H'. Now (ii) follows from Cor. 2.8.

§ 3. Lie Transformation Groups

Let M be a Hausdorf space and G a topological group such that to each $g \in G$ is associated a homeomorphism $p \to g \cdot p$ of M onto itself such that

(1) $g_1 g_2 \cdot p = g_1 \cdot (g_2 \cdot p)$ for $p \in M$, $g_1, g_2 \in G$;

(2) the mapping $(g, p) \to g \cdot p$ is a continuous mapping of the product space $G \times M$ onto M.

The group G is then called a *topological transformation group* of M. From (1) follows that $e \cdot p = p$ for all $p \in M$. If e is the only element of G which leaves each $p \in M$ fixed then G is called *effective* and is said to act effectively on M.

Example. Suppose A is a topological group and F a closed subgroup of A. The system of left cosets aF, $a \in A$ is denoted A/F; let π denote the natural mapping of A onto A/F. The set A/F can be given a topology, the *natural topology*, which is uniquely determined by the condition that π is a continuous and open mapping. This makes A/F a Hausdorf space and it is not difficult to see that if to each $a \in A$ we assign the mapping $\tau(a) : bF \to abF$, then A is a topological transformation group of A/F. The group A is effective if and only if F contains no normal subgroup of A. The coset space A/F is a *homogeneous space*, that is, has a transitive group of homeomorphisms, namely $\tau(A)$. Theorem 3.2 below deals with the converse question, namely that of representing a homogeneous space by means of a coset space.

Lemma 3.1 (the category theorem). *If a locally compact space M is a countable union*

$$M = \bigcup_{n=1}^{\infty} M_n,$$

where each M_n is a closed subset, then at least one M_n contains an open subset of M.

Proof. Suppose no M_n contains an open subset of M. Let U_1 be an open subset of M whose closure $\bar{U}_1$ is compact. Select successively

$a_1 \in U_1 - M_1$ and a neighborhood U_2 of a_1 such that $\bar{U}_2 \subset U_1$ and $\bar{U}_2 \cap M_1 = \emptyset$;

$a_2 \in U_2 - M_2$ and a neighborhood U_3 of a_2 such that $\bar{U}_3 \subset U_2$ and $\bar{U}_3 \cap M_2 = \emptyset$, etc.

Then $\bar{U}_1, \bar{U}_2, \ldots$ is a decreasing sequence of compact sets $\neq \emptyset$. Thus there is a point $b \in M$ in common to all $\bar{U}_n$. But this implies $b \notin M_n$ for each n which is a contradiction.

Theorem 3.2. *Let G be a locally compact group with a countable base. Suppose G is a transitive topological transformation group of a locally compact Hausdorf space M. Let p be any point in M and H the subgroup of G which leaves p fixed. Then H is closed and the mapping*

$$gH \to g \cdot p$$

is a homeomorphism of G/H onto M.

Proof. Since the mapping $\varphi : g \to g \cdot p$ of G onto M is continuous, it follows that $H = \varphi^{-1}(p)$ is closed in G. The natural mapping $\pi : G \to G/H$ is open and continuous. Thus, in order to prove Theorem 3.2, it suffices to prove that φ is open. Let V be an open subset of G and g a point in V. Select a compact neighborhood U of e in G such that $U = U^{-1}$, $gU^2 \subset V$. There exists a sequence $(g_n) \subset G$ such that $G = \bigcup_n g_n U$. The group G being transitive, this implies $M = \bigcup_n g_n U \cdot p$. Each summand is compact, hence a closed subset of M. By the lemma above, some summand, and therefore $U \cdot p$, contains an inner point $u \cdot p$. Then p is an inner point of $u^{-1}U \cdot p \subset U^2 \cdot p$ and consequently $g \cdot p$ is an inner point of $V \cdot p$. This shows that the mapping φ is open.

Definition. The group H is called the *isotropy* group at p (or the isotropy subgroup of G at p).

Corollary 3.3. *Let G and X be two locally compact groups. Assume G has a countable base. Then every continuous homomorphism ψ of G onto X is open.*

In fact, if we associate to each $g \in G$ the homeomorphism $x \to \psi(g)\, x$ of X onto itself, then G becomes a transitive topological transformation group of X. If f denotes the identity element of X, the proof above shows that the mapping $g \to \psi(g) f$ of G onto X is open.

Let G be a Lie group and M a differentiable manifold. Suppose G is a topological transformation group of M; G is said to be a *Lie transformation group* of M if the mapping $(g, p) \to g \cdot p$ is a differentiable mapping of $G \times M$ onto M. It follows that for each $g \in G$ the mapping $p \to g \cdot p$ is a diffeomorphism of M onto itself.

Let G be a Lie transformation group of M. To each X in $\mathfrak{g}$, the Lie algebra of G, we can associate a vector field X^+ on M by the formula

$$[X^+f](p) = \lim_{t\to 0} \frac{f(\exp tX \cdot p) - f(p)}{t}$$

for $f \in C^\infty(M)$, $p \in M$. The existence of X^+ follows from the fact that the mapping $(g, p) \to g \cdot p$ is a differentiable mapping of $G \times M$ onto M. It is also easy to check that X^+ is a derivation of $C^\infty(M)$. It is called the vector field on M *induced* by the one-parameter subgroup $\exp tX$, $t \in \boldsymbol{R}$.

Theorem 3.4. *Let G be a Lie transformation group of M. Let X, Y be in $\mathfrak{g}$, the Lie algebra of G, and let X^+, Y^+ be the vector fields on M induced by $\exp tX$ and $\exp tY$, $(t \in \boldsymbol{R})$. Then*

$$[X^+, Y^+] = \lim_{t\to 0} \frac{1}{t}(Y^+ - dg_t Y^+),$$

where $g_t = \exp tX$.

Proof. Let $f \in C^\infty(M)$ and $q \in M$. The function $F(t, q) = f(\exp tX \cdot q)$ is differentiable with respect to t so

$$F(t, q) - F(0, q) = t \int_0^1 \left[\frac{\partial F}{\partial t}\right](st, q)\, ds = h(t, q)\, t,$$

where $h \in C^\infty(\boldsymbol{R} \times M)$ and $h(0, q) = [X^+f](q)$. Then

$$(dg_t \cdot Y^+)_p f = [Y^+(f \circ g_t)](g_t^{-1} \cdot p) = [Y^+f](g_t^{-1} \cdot p) + t[Y^+h](t, g_t^{-1} \cdot p),$$

so

$$\lim_{t\to 0} \frac{1}{t}(Y^+ - dg_t Y^+)_p f = \lim_{t\to 0} \frac{1}{t}\{[Y^+f](p) - [Y^+f](g_t^{-1} \cdot p)\}$$

$$+ \lim_{t\to 0} \frac{1}{t}\{[Y^+f](g_t^{-1} \cdot p) - [Y^+(f \circ g_t)](g_t^{-1} \cdot p)\}$$

$$= -[-X^+Y^+f](p) - \lim_{t\to 0}[Y^+h](t, g_t^{-1} \cdot p) = (X^+Y^+ - Y^+X^+)_p f.$$

§ 4. Coset Spaces and Homogeneous Spaces

Let G be a Lie group and H a closed subgroup. The group H will always be given the analytic structure from Theorem 2.3. Let $\mathfrak{g}$ and $\mathfrak{h}$ denote the Lie algebras of G and H, respectively, and let $\mathfrak{m}$ denote some vector subspace of $\mathfrak{g}$ such that $\mathfrak{g} = \mathfrak{m} + \mathfrak{h}$ (direct sum). Let π be the natural mapping of G onto the space G/H of left cosets gH, $g \in G$. As usual we give G/H the natural topology determined by the requirement that π should be continuous and open. We put $p_0 = \pi(e)$ and let ψ denote the restriction of exp to $\mathfrak{m}$.

Lemma 4.1. *There exists a neighborhood U of 0 in $\mathfrak{m}$ which is mapped homeomorphically under ψ and such that π maps $\psi(U)$ homeomorphically onto a neighborhood of p_0 in G/H.*

Proof. Let $U_{\mathfrak{m}}$, $U_{\mathfrak{h}}$ have the property described in Lemma 2.4 for $\mathfrak{h} = \mathfrak{n}$. Then since H has the relative topology of G, we can select a neighborhood V of e in G such that $V \cap H = \exp U_{\mathfrak{h}}$. Let U be a compact neighborhood of 0 in $U_{\mathfrak{m}}$ such that $\exp(-U) \exp U \subset V$. Then ψ is a homeomorphism of U onto $\psi(U)$. Moreover, π is one-to-one on $\psi(U)$ because if $X', X'' \in U$ satisfy $\pi(\exp X') = \pi(\exp X'')$, then $\exp(-X'') \exp X' \subset V \cap H$ so $\exp X' = \exp X'' \exp Z$ where $Z \in U_{\mathfrak{h}}$. From Lemma 2.4 we can conclude that $X' = X''$, $Z = 0$; consequently, π is one-to-one on $\psi(U)$, hence a homeomorphism. Finally, $U \times U_{\mathfrak{h}}$ is a neighborhood of $(0, 0)$ in $U_{\mathfrak{m}} \times U_{\mathfrak{h}}$; hence $\exp U \exp U_{\mathfrak{h}}$ is a neighborhood of e in G and since π is an open mapping, the set $\pi(\exp U \exp U_{\mathfrak{h}}) = \pi(\psi(U))$ is a neighborhood of p_0 in G/H. This proves the lemma. The set $\psi(U)$ will be referred to as a *local cross section.*

Let N_0 denote the interior of the set $\pi(\psi(U))$ and let $X_1, \ldots, X_r$ be a basis of $\mathfrak{m}$. If $g \in G$, then the mapping

$$\pi(g \exp(x_1 X_1 + \ldots + x_r X_r)) \to (x_1, \ldots, x_r)$$

is a homeomorphism of the open set $g \cdot N_0$ onto an open subset of $\boldsymbol{R}^r$. It is easy to see that with these charts, G/H is an analytic manifold. Moreover, if $x \in G$, the mapping $\tau(x) : yH \to xyH$ is an analytic diffeomorphism of G/H.

Theorem 4.2. *Let G be a Lie group, H a closed subgroup of G, G/H the space of left cosets gH with the natural topology. Then G/H has a unique analytic structure with the property that G is a Lie transformation group of G/H.*

We use the notation above and let $B = \psi(\mathring{U})$ where $\mathring{U}$ is the interior of U. Remembering that the mapping Φ in Lemma 2.4 is a diffeomorphism, the set B is a submanifold of G. The mappings in the diagram

$$\begin{tikzcd} G \times B \arrow[r, "\Phi"] \arrow[d, "I \times \pi"'] & G \arrow[d, "\pi"] \\ G \times N_0 & G/H \end{tikzcd}$$

are:

$$I \times \pi : (g, x) \to (g, xH), \qquad g \in G, x \in B;$$
$$\Phi : (g, x) \to gx, \qquad g \in G, x \in B.$$

Then the mapping $(g, xH) \to gxH$ of $G \times N_0$ onto G/H can be written $\pi \circ \Phi \circ (I \times \pi)^{-1}$ which is analytic. Thus G is a Lie transformation group of G/H. The uniqueness results from the following proposition which should be compared with Theorem 3.2.

Proposition 4.3. *Let G be a transitive Lie transformation group of a C^∞ manifold M. Let p_0 be a point in M and let G_{p_0} denote the subgroup of G that leaves p_0 fixed. Then G_{p_0} is closed. Let α denote the mapping $gG_{p_0} \to g \cdot p_0$ of G/G_{p_0} onto M.*

(a) *If α is a homeomorphism, then it is a diffeomorphism (G/G_{p_0} having the analytic structure defined above).*

(b) *Suppose α is a homeomorphism and that M is connected. Then G_0, the identity component of G, is transitive on M.*

Proof. (a) We put $H = G_{p_0}$ and use Lemma 4.1. Let B and N_0 have the same meaning as above. Then B is a submanifold of G, diffeomorphic to N_0 under π. Let i denote the identity mapping of B into G and let β denote the mapping $g \to g \cdot p_0$ of G onto M. By assumption, α_{N_0}, the restriction of α to N_0, is a homeomorphism of N_0 onto an open subset of M. The mapping α is differentiable since $\alpha_{N_0} = \beta \circ i \circ \pi^{-1}$. To show that α^{-1} is differentiable, we begin by showing that the Jacobian of β at $g = e$ has rank r_β equal to dim M.

$$\begin{tikzcd} B \arrow[r, "i"] \arrow[d, "\pi"'] & G \arrow[d, "\beta"] \\ N_0 \arrow[r, "\alpha"] & M \end{tikzcd}$$

The mapping $d\beta_e$ is a linear mapping of $\mathfrak{g}$ into M_{p_0}. Suppose X is in the kernel of $d\beta_e$. Then if $f \in C^\infty(M)$, we have

$$0 = (d\beta_e X)f = X(f \circ \beta) = \left\{\frac{d}{dt} f(\exp tX \cdot p_0)\right\}_{t=0}. \tag{1}$$

Let $s \in \boldsymbol{R}$; we use (1) on the function $f^*(q) = f(\exp sX \cdot q)$, $q \in M$. Then

$$0 = \left\{\frac{d}{dt} f^*(\exp tX \cdot p_0)\right\}_{t=0} = \left\{\frac{d}{dt} f(\exp tX \cdot p_0)\right\}_{t=s},$$

which shows that $f(\exp sX \cdot p_0)$ is constant in s. Since f is arbitrary, we have $\exp sX \cdot p_0 = p_0$ for all s so $X \in \mathfrak{h}$. On the other hand, it is obvious that $d\beta_e$ vanishes on $\mathfrak{h}$ so $\mathfrak{h} = \text{kernel } (d\beta_e)$. Hence $r_\beta = \dim \mathfrak{g} - \dim \mathfrak{h}$. But since α is a homeomorphism, it follows from the topological invariance of dimension, that $\dim G/H = \dim M$. Thus $d\beta_e$ maps $\mathfrak{g}$ onto M_{p_0} and $r_\beta = \dim M$. This proves (a).

(b) If α is a homeomorphism, β above is an open mapping. There exists a subset $\{x_\gamma : \gamma \in I\}$ of G such that $G = \bigcup_{\gamma \in I} G_0 x_\gamma$. Each orbit $G_0 x_\gamma \cdot p_0$ is an open subset of M; two orbits $G_0 x_\gamma \cdot p_0$ and $G_0 x_{\gamma'} \cdot p_0$ are either disjoint or equal. Therefore, since M is connected, all orbits must coincide and (b) follows.

Definition. In the sequel the coset space G/H (G a Lie group, H a closed subgroup) will always be taken with the analytic structure described in Theorem 4.2. If $x \in G$, the diffeomorphism $yH \to xyH$ of G/H onto itself will be denoted by $\tau(x)$. The group H is called the *isotropy group*. The group H^* of linear transformations $(d\tau(h))_{\pi(e)}$, $(h \in H)$, is called the *linear isotropy group*.

Let N be a Lie subgroup of G. Then the subset $N \cap H$ is closed in N and the coset space $N/N \cap H$ is in one-to-one correspondence with the orbit of $\pi(e)$ in G/H under N. If $\mathfrak{n}$ and $\mathfrak{h}_1$ denote the Lie algebras of N and $N \cap H$, respectively, then $\mathfrak{h}_1 = \mathfrak{h} \cap \mathfrak{n}$ by (2), §2, and Theorem 2.3.

Proposition 4.4.

(a) *The orbit $N/N \cap H$ is a submanifold of G/H.*

(b) *If N is a topological subgroup of G, and if H is compact, then the submanifold $N/N \cap H$ is a closed topological subspace of G/H.*

Proof. (a) We consider the commutative diagram

$$\begin{tikzcd} N \arrow[r, "i"] \arrow[d, "\pi_1"'] & G \arrow[d, "\pi"] \\ N/N \cap H \arrow[r, "I"] & G/H \end{tikzcd}$$

where π_1 and π are the natural mappings of N onto $N/N \cap H$ and of G onto G/H, respectively. The identity mapping of N into G is denoted by i and I denotes the mapping $n(N \cap H) \to i(n)H$ of $N/N \cap H$ into G/H. Let $\mathfrak{n}_1$ be a complementary subspace of $\mathfrak{h}_1$ in $\mathfrak{n}$ and $\mathfrak{g}_1$ a complementary subspace of $\mathfrak{h} + \mathfrak{n}_1$ in $\mathfrak{g}$. We use Lemma 4.1 on the decompositions $\mathfrak{n} = \mathfrak{h}_1 + \mathfrak{n}_1$ and $\mathfrak{g} = \mathfrak{h} + (\mathfrak{n}_1 + \mathfrak{g}_1)$. We can then get submanifolds $B_N \subset N$, $B_G \subset G$ through e which π_1 and π map diffeomorphically onto open neighborhoods of $\pi_1(e)$ in $N/N \cap H$ and of $\pi(e)$ in G/H, respectively. Here we can take B_N as a submanifold of B_G. Put $V_1 = \pi_1(B_N)$, $V = \pi(B_G)$. The restriction of I to V_1, say I_{V_1}, can be written $I_{V_1} = \pi \circ i \circ \pi_1^{-1}$. The Jacobian of I_{V_1} at $\pi_1(e)$ therefore has rank equal to dim $(N/N \cap H)$. Hence I is regular, proving (a).

(b) Since N is now a topological subgroup of G, the diagram above shows that I is a homeomorphism of $N/N \cap H$ into G/H. In order to show that $N/N \cap H$ is closed, let (p_k) be a sequence in $N/N \cap H$ converging to a point $q \in G/H$. Select $g \in G$ such that $\pi(g) = q$. We may assume that all p_k belong to the neighborhood $g \cdot V$ of q. Hence there is a unique element $g_k \in gB_G$ such that $\pi(g_k) = p_k$. Since π is a homeomorphism of gB_G onto $g \cdot V$, we have $\lim g_k = g$.

On the other hand, for each index k there exists an element $n_k \in N$ such that $\pi_1(n_k) = p_k$. Thus there exists an element $h_k \in H$ such that $g_k = n_k h_k$. Since H is compact we may assume that (h_k) is a convergent sequence. It follows that the sequence (n_k) is convergent; the limit n^* lies in N since N is closed in G. Consequently, $\pi_1(n^*) = q$ so the orbit $N/N \cap H$ is closed.

Remark. Prop. 4.4 holds in greater generality; see Exercise B.3 following this chapter and Exercise 5, Chapter IV.

§ 5. The Adjoint Group

Let $\mathfrak{a}$ be a Lie algebra over $\boldsymbol{R}$. Let $\boldsymbol{GL}(\mathfrak{a})$ as usual denote the group of all nonsingular endomorphisms of $\mathfrak{a}$. We recall that an endomorphism of a vector space V (in particular of a Lie algebra) simply means a linear mapping of V into itself. The Lie algebra $\mathfrak{gl}(\mathfrak{a})$ of $\boldsymbol{GL}(\mathfrak{a})$ consists of the vector space of all endomorphisms of $\mathfrak{a}$ with the bracket operation $[A, B] = AB - BA$. The mapping $X \to \operatorname{ad} X$, $X \in \mathfrak{a}$ is a homomorphism of $\mathfrak{a}$ onto a subalgebra ad $(\mathfrak{a})$ of $\mathfrak{gl}(\mathfrak{a})$. Let Int $(\mathfrak{a})$ denote the analytic subgroup of $\boldsymbol{GL}(\mathfrak{a})$ whose Lie algebra is ad $(\mathfrak{a})$; Int $(\mathfrak{a})$ is called the *adjoint group* of $\mathfrak{a}$.

The group Aut $(\mathfrak{a})$ of all automorphisms of $\mathfrak{a}$ is a closed subgroup of $\boldsymbol{GL}(\mathfrak{a})$. Thus Aut $(\mathfrak{a})$ has a unique analytic structure in which it is a topological Lie subgroup of $\boldsymbol{GL}(\mathfrak{a})$. Let $\partial(\mathfrak{a})$ denote the Lie algebra of Aut $(\mathfrak{a})$. From §2 we know that $\partial(\mathfrak{a})$ consists of all endomorphisms D of $\mathfrak{a}$ such that $e^{tD} \in$ Aut $(\mathfrak{a})$ for each $t \in \boldsymbol{R}$. Let $X, Y \in \mathfrak{a}$. The relation $e^{tD}[X, Y] = [e^{tD}X, e^{tD}Y]$ for all $t \in \boldsymbol{R}$ implies

$$D[X, Y] = [DX, Y] + [X, DY]. \tag{1}$$

An endomorphism D of $\mathfrak{a}$ satisfying (1) for all $X, Y \in \mathfrak{a}$ is called a *derivation* of $\mathfrak{a}$. By induction we get from (1)

$$D^k[X, Y] = \sum_{i+j=k} \frac{k!}{i!j!} [D^iX, D^jY], \qquad i \geqslant 0, j \geqslant 0, \tag{2}$$

where D^0 means the identity mapping of $\mathfrak{a}$. From (2) follows that $e^{tD}[X, Y] = [e^{tD}X, e^{tD}Y]$ and thus $\partial(\mathfrak{a})$ *consists of all derivations of* $\mathfrak{a}$. Using the Jacobi identity we see that $\operatorname{ad}(\mathfrak{a}) \subset \partial(\mathfrak{a})$ and therefore $\operatorname{Int}(\mathfrak{a}) \subset \operatorname{Aut}(\mathfrak{a})$. The elements of $\operatorname{ad}(\mathfrak{a})$ and $\operatorname{Int}(\mathfrak{a})$, respectively, are called the *inner derivations* and *inner automorphisms* of $\mathfrak{a}$. Since $\operatorname{Aut}(\mathfrak{a})$ is a topological subgroup of $\boldsymbol{GL}(\mathfrak{a})$ the identity mapping of $\operatorname{Int}(\mathfrak{a})$ into $\operatorname{Aut}(\mathfrak{a})$ is continuous. In view of Lemma 14.1, Chapter I, $\operatorname{Int}(\mathfrak{a})$ is a Lie subgroup of $\operatorname{Aut}(\mathfrak{a})$. We shall now prove that $\operatorname{Int}(\mathfrak{a})$ is a normal subgroup of $\operatorname{Aut}(\mathfrak{a})$. Let $s \in \operatorname{Aut}(\mathfrak{a})$. Then the mapping $\sigma : g \to sgs^{-1}$ is an automorphism of $\operatorname{Aut}(\mathfrak{a})$, and $(d\sigma)_e$ is an automorphism of $\partial(\mathfrak{a})$. If A, B are endomorphisms of a vector space and A^{-1} exists, then $Ae^BA^{-1} = e^{ABA^{-1}}$. Considering Lemma 1.12 we have

$$(d\sigma)_e D = sDs^{-1} \qquad \text{for } D \in \partial(\mathfrak{a}).$$

If $X \in \mathfrak{a}$, we have $s \operatorname{ad} X s^{-1} = \operatorname{ad}(s \cdot X)$, so

$$(d\sigma)_e \operatorname{ad} X = \operatorname{ad}(s \cdot X),$$

and consequently

$$\sigma \cdot e^{\operatorname{ad} X} = e^{\operatorname{ad}(s \cdot X)} \qquad (X \in \mathfrak{a}).$$

Now, the group $\operatorname{Int}(\mathfrak{a})$ is connected, so it is generated by the elements $e^{\operatorname{ad} X}$, $X \in \mathfrak{a}$. It follows that $\operatorname{Int}(\mathfrak{a})$ is a normal subgroup of $\operatorname{Aut}(\mathfrak{a})$ and the automorphism s of $\mathfrak{a}$ induces the analytic isomorphism $g \to sgs^{-1}$ of $\operatorname{Int}(\mathfrak{a})$ onto itself.

More generally, if s is an isomorphism of a Lie algebra $\mathfrak{a}$ onto a Lie algebra $\mathfrak{b}$ (both Lie algebras over $\boldsymbol{R}$) then the mapping $g \to sgs^{-1}$ is an isomorphism of $\operatorname{Aut}(\mathfrak{a})$ onto $\operatorname{Aut}(\mathfrak{b})$ which maps $\operatorname{Int}(\mathfrak{a})$ onto $\operatorname{Int}(\mathfrak{b})$.

Let G be a Lie group. If $\sigma \in G$, the mapping $I(\sigma) : g \to \sigma g \sigma^{-1}$ is an analytic isomorphism of G onto itself. We put $\operatorname{Ad}(\sigma) = dI(\sigma)_e$. Sometimes we write $\operatorname{Ad}_G(\sigma)$ instead of $\operatorname{Ad}(\sigma)$ when a misunderstanding might otherwise arise. The mapping $\operatorname{Ad}(\sigma)$ is an automorphism of $\mathfrak{g}$, the Lie algebra of G. We have by Lemma 1.12

$$\exp \operatorname{Ad}(\sigma) X = \sigma \exp X \sigma^{-1} \qquad \text{for } \sigma \in G,\ X \in \mathfrak{g}. \tag{3}$$

The mapping $\sigma \to \operatorname{Ad}(\sigma)$ is a homomorphism of G into $\boldsymbol{GL}(\mathfrak{g})$. This homomorphism is called the *adjoint representation* of G. Let us prove that this homomorphism is analytic. For this it suffices to prove that for each $X \in \mathfrak{g}$ and each linear function ω on $\mathfrak{g}$ the function $\sigma \to \omega(\operatorname{Ad}(\sigma) X)$, $(\sigma \in G)$, is analytic at $\sigma = e$. Select $f \in C^\infty(G)$ such that f is analytic at $\sigma = e$ and such that $Yf = \omega(Y)$ for all $Y \in \mathfrak{g}$.

Then, using (3), we obtain

$$\omega(\operatorname{Ad}(\sigma)X) = (\operatorname{Ad}(\sigma)X)f = \left\{\frac{d}{dt}f(\sigma \exp tX\sigma^{-1})\right\}_{t=0},$$

which is clearly analytic at $\sigma = e$.

Next, let X and Y be arbitrary vectors in $\mathfrak{g}$. From Lemma 1.8 (iii) we have

$$\exp(\operatorname{Ad}(\exp tX)\,tY) = \exp(tY + t^2[X, Y] + O(t^3)).$$

It follows that

$$\operatorname{Ad}(\exp tX)\,Y = Y + t[X, Y] + O(t^2). \tag{4}$$

The differential $d\operatorname{Ad}_e$ is a homomorphism of $\mathfrak{g}$ into $\mathfrak{gl}(\mathfrak{g})$ and due to (4) we have

$$d\operatorname{Ad}_e(X) = \operatorname{ad} X, \qquad X \in \mathfrak{g}.$$

Applying the exponential mapping on both sides we obtain (Lemma 1.12)

$$\operatorname{Ad}(\exp X) = e^{\operatorname{ad} X}, \qquad X \in \mathfrak{g}. \tag{5}$$

Let G be a connected Lie group and H an analytic subgroup. Let $\mathfrak{g}$ and $\mathfrak{h}$ denote the corresponding Lie algebras. Relations (3) and (5) show that H is a normal subgroup of G if and only if $\mathfrak{h}$ is an ideal in $\mathfrak{g}$.

Lemma 5.1. *Let G be a connected Lie group with Lie algebra $\mathfrak{g}$ and let φ be an analytic homomorphism of G into a Lie group X with Lie algebra $\mathfrak{x}$. Then:*

(i) *The kernel $\varphi^{-1}(e)$ is a topological Lie subgroup of G. Its Lie algebra is the kernel of $d\varphi$ $(=d\varphi_e)$.*

(ii) *The image $\varphi(G)$ is a Lie subgroup of X with Lie algebra $d\varphi(\mathfrak{g}) \subset \mathfrak{x}$.*

(iii) *The factor group $G/\varphi^{-1}(e)$ with its natural analytic structure is a Lie group and the mapping $g\varphi^{-1}(e) \rightarrow \varphi(g)$ is an analytic isomorphism of $G/\varphi^{-1}(e)$ onto $\varphi(G)$. In particular, the mapping $\varphi : G \rightarrow \varphi(G)$ is analytic.*

Proof. (i) According to Theorem 2.3, $\varphi^{-1}(e)$ has a unique analytic structure with which it is a topological Lie subgroup of G. Moreover, its Lie algebra contains a vector $Z \in \mathfrak{g}$ if and only if $\varphi(\exp tZ) = e$ for all $t \in \boldsymbol{R}$. Since $\varphi(\exp tZ) = \exp td\varphi(Z)$, the condition is equivalent to $d\varphi(Z) = 0$.

(ii) Let X_1 denote the analytic subgroup of X with Lie algebra $d\varphi(\mathfrak{g})$. The group $\varphi(G)$ is generated by the elements $\varphi(\exp Z)$, $Z \in \mathfrak{g}$. The group X_1 is generated by the elements $\exp(d\varphi(Z))$, $Z \in \mathfrak{g}$. Since $\varphi(\exp Z) = \exp d\varphi(Z)$ it follows that $\varphi(G) = X_1$.

(iii) Let H be any closed normal subgroup of G. Then H is a topological Lie subgroup and the factor group G/H has a unique analytic structure such that the mapping $(g, xH) \to g\,xH$ is an analytic mapping of $G \times G/H$ onto G/H. In order to see that G/H is a Lie group in this analytic structure we use the local cross section $\psi(U)$ from Lemma 4.1. Let $B = \psi(\mathring{U})$ where $\mathring{U}$ is the interior of U. In the commutative diagram

$$\begin{tikzcd} G \times G/H \arrow[rr, "\Phi"] \arrow[dr, "\pi \times I"'] & & G/H \\ & G/H \times G/H \arrow[ur, "\alpha"'] & \end{tikzcd}$$

the symbols Φ, $\pi \times I$, and α denote the mappings:

$$\begin{aligned} \Phi &: (g, xH) \to g^{-1}xH, && x, g \in G; \\ \pi \times I &: (g, xH) \to (gH, xH), && x, g \in G; \\ \alpha &: (gH, xH) \to g^{-1}xH, && x, g \in G. \end{aligned}$$

The mapping α is well defined since H is a normal subgroup of G. Let g_0, x_0 be arbitrary two points in G. The restriction of $\pi \times I$ to $(g_0B) \times (G/H)$ is an analytic diffeomorphism of $(g_0B) \times (G/H)$ onto a neighborhood N of (g_0H, x_0H) in $G/H \times G/H$. On N we have $\alpha = \Phi \circ (\pi \times I)^{-1}$ which shows that α is analytic. Hence G/H is a Lie group.

Now choose for H the group $\varphi^{-1}(e)$ and let $\mathfrak{h}$ denote the Lie algebra of H. Then $\mathfrak{h} = d\varphi^{-1}(0)$ so $\mathfrak{h}$ is an ideal in $\mathfrak{g}$. By (ii) the Lie algebra of G/H is $d\pi(\mathfrak{g})$ which is isomorphic to the algebra $\mathfrak{g}/\mathfrak{h}$. On the other hand, the mapping $Z + \mathfrak{h} \to d\varphi(Z)$ is an isomorphism of $\mathfrak{g}/\mathfrak{h}$ onto $d\varphi(\mathfrak{g})$. The corresponding local isomorphism between G/H and $\varphi(G)$ coincides with the (algebraic) isomorphism $gH \to \varphi(g)$ on some neighborhood of the identity. It follows that this last isomorphism is analytic at e, hence everywhere.

Corollary 5.2. *Let G be a connected Lie group with Lie algebra $\mathfrak{g}$. Let Z denote the center of G. Then:*

(i) *Ad_G is an analytic homomorphism of G onto Int $(\mathfrak{g})$ with kernel Z.*

(ii) *The mapping $gZ \to Ad_G(g)$ is an analytic isomorphism of G/Z onto Int $(\mathfrak{g})$.*

In fact $\mathrm{Ad}_G(G) = \mathrm{Int}(\mathfrak{g})$ due to (5) and $\mathrm{Ad}_G^{-1}(e) = Z$ due to (3). The remaining statements are contained in Lemma 5.1.

Corollary 5.3. *Let $\mathfrak{g}$ be a Lie algebra over $\boldsymbol{R}$ with center $\{0\}$. Then the center of Int $(\mathfrak{g})$ consists of the identity element alone.*

In fact, let $G' = \text{Int}(\mathfrak{g})$ and let Z denote the center of G'. Let ad denote the adjoint representation of $\mathfrak{g}$ and let Ad′ and ad′ denote the adjoint representation of G' and ad $(\mathfrak{g})$, respectively. The mapping

$$\theta : gZ \to \text{Ad}'(g), \qquad g \in G',$$

is an isomorphism of G'/Z onto Int (ad $(\mathfrak{g})$). On the other hand, the mapping $s : X \to \text{ad}\, X$ $(X \in \mathfrak{g})$ is an isomorphism of $\mathfrak{g}$ onto ad $(\mathfrak{g})$ and consequently the mapping $S : g \to s \circ g \circ s^{-1}$ $(g \in G')$ is an isomorphism of G' onto Int (ad $(\mathfrak{g})$). Moreover, if $X \in \mathfrak{g}$, we obtain from (5)

$$S(e^{\text{ad}\, X}) = s \circ e^{\text{ad}\, X} \circ s^{-1} = e^{(\text{ad}'(\text{ad}\, X))} = \text{Ad}'(e^{\text{ad}\, X}),$$

ad $(\mathfrak{g})$ being the Lie algebra of G'. It follows that $S^{-1} \circ \theta$ is an isomorphism of G'/Z onto G', mapping gZ onto g $(g \in G')$. Obviously Z must consist of the identity element alone.

Remark. The conclusion of Cor. 5.3 does not hold in general, if $\mathfrak{g}$ has nontrivial center. Let, for example, $\mathfrak{g}$ be the three-dimensional Lie algebra $\mathfrak{g} = \boldsymbol{R}X_1 + \boldsymbol{R}X_2 + \boldsymbol{R}X_3$ with the bracket defined by: $[X_1, X_2] = X_3$, $[X_1, X_3] = [X_2, X_3] = 0$. Here $\mathfrak{g}$ is nonabelian, whereas Int $(\mathfrak{g})$ is abelian and has dimension 2.

Definition. Let $\mathfrak{g}$ be a Lie algebra over $\boldsymbol{R}$. Let $\mathfrak{k}$ be a subalgebra of $\mathfrak{g}$ and K^* the analytic subgroup of Int $(\mathfrak{g})$ which corresponds to the subalgebra $\text{ad}_{\mathfrak{g}}(\mathfrak{k})$ of $\text{ad}_{\mathfrak{g}}(\mathfrak{g})$. The subalgebra $\mathfrak{k}$ is called a *compactly imbedded subalgebra of* $\mathfrak{g}$ if K^* is compact. The Lie algebra $\mathfrak{g}$ is said to be *compact* if it is compactly imbedded in itself or equivalently if Int $(\mathfrak{g})$ is compact.

It should be observed that the topology of K^* might *a priori* differ from the relative topology of the group Int $(\mathfrak{g})$ which again might differ from the relative topology of $\boldsymbol{GL}(\mathfrak{g})$. The next proposition clarifies this situation.

Proposition 5.4. *Let $\tilde{K}$ denote the abstract group K^* with the relative topology of $\boldsymbol{GL}(\mathfrak{g})$. Then K^* is compact if and only if $\tilde{K}$ is compact.*

The identity mapping of K^* into $\boldsymbol{GL}(\mathfrak{g})$ is analytic, in particular, continuous. Thus $\tilde{K}$ is compact if K^* is compact. On the other hand, if $\tilde{K}$ is compact, then it is closed in $\boldsymbol{GL}(\mathfrak{g})$; by Theorem 2.10, K^* and $\tilde{K}$ are homeomorphic.

Remark. Suppose G is any connected Lie group with Lie algebra $\mathfrak{g}$. Let K be the analytic subgroup of G with Lie algebra $\mathfrak{k}$. Then the group K^* above coincides with $\text{Ad}_G(K)$; in fact, both groups are generated by $\text{Ad}_G(\exp X)$, $X \in \mathfrak{k}$.

§ 6. Semisimple Lie Groups

Let $\mathfrak{g}$ be a Lie algebra over a field of characteristic 0. Denoting by Tr the trace of a vector space endomorphism we consider the bilinear form $B(X, Y) = \operatorname{Tr}(\operatorname{ad} X \operatorname{ad} Y)$ on $\mathfrak{g} \times \mathfrak{g}$. The form B is called the *Killing form* of $\mathfrak{g}$.

Let σ be an automorphism of $\mathfrak{g}$. Then $\operatorname{ad} \sigma X = \sigma \circ \operatorname{ad} X \circ \sigma^{-1}$ so $\operatorname{Tr}(\operatorname{ad} \sigma X \operatorname{ad} \sigma Y) = \operatorname{Tr}(\sigma \operatorname{ad} X \operatorname{ad} Y \sigma^{-1}) = \operatorname{Tr}(\operatorname{ad} X \operatorname{ad} Y)$. Since $\operatorname{Tr}(AB) = \operatorname{Tr}(BA)$ for any endomorphisms A and B, we have $\operatorname{Tr}(\operatorname{ad}[Z, X] \operatorname{ad} Y) = \operatorname{Tr}((\operatorname{ad} Z \operatorname{ad} X - \operatorname{ad} X \operatorname{ad} Z) \operatorname{ad} Y) = \operatorname{Tr}((\operatorname{ad} Y \operatorname{ad} Z - \operatorname{ad} Z \operatorname{ad} Y) \operatorname{ad} X) = \operatorname{Tr}(\operatorname{ad}[Y, Z] \operatorname{ad} X)$. Thus we have the formulas

$$B(\sigma X, \sigma Y) = B(X, Y), \qquad \sigma \in \operatorname{Aut}(\mathfrak{g}),$$

$$B(X, [Y, Z]) = B(Y, [Z, X]) = B(Z, [X, Y]), \qquad X, Y, Z \in \mathfrak{g}. \tag{1}$$

Suppose $\mathfrak{a}$ is an ideal in $\mathfrak{g}$. Then it is easily verified that the Killing form of $\mathfrak{a}$ coincides with the restriction of B to $\mathfrak{a} \times \mathfrak{a}$.

Definition. A Lie algebra $\mathfrak{g}$ over a field of characteristic 0 is called *semisimple* if the Killing B of $\mathfrak{g}$ is nondegenerate. We shall call a Lie algebra $\mathfrak{g} \neq \{0\}$ *simple*† if it is semisimple and has no ideals except $\{0\}$ and $\mathfrak{g}$. A Lie group is called semisimple (simple) if its Lie algebra is semisimple (simple).

Proposition 6.1. *Let* $\mathfrak{g}$ *be a semisimple Lie algebra,* $\mathfrak{a}$ *an ideal in* $\mathfrak{g}$. *Let* $\mathfrak{a}^\perp$ *denote the set of elements* $X \in \mathfrak{g}$ *which are orthogonal to* $\mathfrak{a}$ *with respect to* B. *Then* $\mathfrak{a}$ *is semisimple,* $\mathfrak{a}^\perp$ *is an ideal and*

$$\mathfrak{g} = \mathfrak{a} + \mathfrak{a}^\perp \quad (\textit{direct sum}).$$

Proof. The fact that $\mathfrak{a}^\perp$ is an ideal is obvious from (1). Since B is nondegenerate, we have $\dim \mathfrak{a} + \dim \mathfrak{a}^\perp = \dim \mathfrak{g}$. If $Z \in \mathfrak{g}$ and $X, Y \in \mathfrak{a} \cap \mathfrak{a}^\perp$, we have $B(Z, [X, Y]) = B([Z, X], Y) = 0$ so $[X, Y] = 0$. Hence $\mathfrak{a} \cap \mathfrak{a}^\perp$ is an abelian ideal in $\mathfrak{g}$. Let $\mathfrak{b}$ be any subspace of $\mathfrak{g}$ complementary to $\mathfrak{a} \cap \mathfrak{a}^\perp$. If $Z \in \mathfrak{g}$ and $T \in \mathfrak{a} \cap \mathfrak{a}^\perp$, then the endomorphism $\operatorname{ad} T \operatorname{ad} Z$ maps $\mathfrak{a} \cap \mathfrak{a}^\perp$ into $\{0\}$, and $\mathfrak{b}$ into $\mathfrak{a} \cap \mathfrak{a}^\perp$. In particular, $\operatorname{Tr}(\operatorname{ad} T \operatorname{ad} Z) = 0$. It follows that $\mathfrak{a} \cap \mathfrak{a}^\perp = \{0\}$ and we get the direct decomposition $\mathfrak{g} = \mathfrak{a} + \mathfrak{a}^\perp$. Since the Killing form of $\mathfrak{a}$ is the restriction of B to $\mathfrak{a} \times \mathfrak{a}$, the semisimplicity of $\mathfrak{a}$ is obvious.

† This definition of a simple Lie algebra is convenient for our purposes but is formally different from the usual one: A Lie algebra $\mathfrak{g}$ is simple if it is nonabelian and has no ideals except $\{0\}$ and $\mathfrak{g}$. However, the two definitions are equivalent in view of Cartan's criterion: A Lie algebra is semisimple if and only if it contains no abelian ideal $\neq \{0\}$.

Corollary 6.2. *A semisimple Lie algebra has center* $\{0\}$.

Corollary 6.3. *A semisimple Lie algebra* $\mathfrak{g}$ *is the direct sum*

$$\mathfrak{g} = \mathfrak{g}_1 + \dots + \mathfrak{g}_r,$$

where $\mathfrak{g}_i$ $(1 \leqslant i \leqslant r)$ *are all the simple ideals in* $\mathfrak{g}$. *Each ideal* $\mathfrak{a}$ *of* $\mathfrak{g}$ *is the direct sum of certain* $\mathfrak{g}_i$.

In fact, Prop. 6.1 implies that $\mathfrak{g}$ can be written as a direct sum of simple ideals $\mathfrak{g}_i$ $(1 \leqslant i \leqslant s)$ such that $\mathfrak{a}$ is the direct sum of certain of these $\mathfrak{g}_i$. If $\mathfrak{b}$ were a simple ideal which does not occur among the ideals $\mathfrak{g}_i$ $(1 \leqslant i \leqslant s)$, then $[\mathfrak{g}_i, \mathfrak{b}] \subset \mathfrak{g}_i \cap \mathfrak{b} = \{0\}$ for $1 \leqslant i \leqslant s$. This contradicts Cor. 6.2.

Proposition 6.4. *If* $\mathfrak{g}$ *is semisimple, then* $ad\,(\mathfrak{g}) = \partial(\mathfrak{g})$, *that is, every derivation is an inner derivation.*

Proof. The algebra $\operatorname{ad}(\mathfrak{g})$ is isomorphic to $\mathfrak{g}$, hence semisimple. If D is a derivation of $\mathfrak{g}$ then $\operatorname{ad}(DX) = [D, \operatorname{ad} X]$ for $X \in \mathfrak{g}$, hence $\operatorname{ad}(\mathfrak{g})$ is an ideal in $\partial(\mathfrak{g})$. Its orthogonal complement, say $\mathfrak{a}$, is also an ideal in $\partial(\mathfrak{g})$. Then $\mathfrak{a} \cap \operatorname{ad}(\mathfrak{g})$ is orthogonal to $\operatorname{ad}(\mathfrak{g})$ also with respect to the Killing form of $\operatorname{ad}(\mathfrak{g})$, hence $\mathfrak{a} \cap \operatorname{ad}(\mathfrak{g}) = \{0\}$. Consequently $D \in \mathfrak{a}$ implies $[D, \operatorname{ad} X] \in \mathfrak{a} \cap \operatorname{ad}(\mathfrak{g}) = \{0\}$. Thus $\operatorname{ad}(DX) = 0$ for each $X \in \mathfrak{g}$, hence $D = 0$. This shows that $\mathfrak{a} = \{0\}$ so $\operatorname{ad}(\mathfrak{g}) = \partial(\mathfrak{g})$.

Corollary 6.5. *For a semisimple Lie algebra* $\mathfrak{g}$ *over* $\boldsymbol{R}$, *the adjoint group Int* $(\mathfrak{g})$ *is the identity component of Aut* $(\mathfrak{g})$. *In particular, Int* $(\mathfrak{g})$ *is a closed topological subgroup of Aut* $(\mathfrak{g})$.

Remark. If $\mathfrak{g}$ is not semisimple, the group $\operatorname{Int}(\mathfrak{g})$ is not necessarily closed in $\operatorname{Aut}(\mathfrak{g})$ (see Exercise C.3 for this chapter).

Proposition 6.6.

(i) *Let* $\mathfrak{g}$ *be a semisimple Lie algebra over* $\boldsymbol{R}$. *Then* $\mathfrak{g}$ *is compact if and only if the Killing form of* $\mathfrak{g}$ *is strictly negative definite.*

(ii) *Every compact Lie algebra* $\mathfrak{g}$ *is the direct sum* $\mathfrak{g} = \mathfrak{z} + [\mathfrak{g}, \mathfrak{g}]$ *where* $\mathfrak{z}$ *is the center of* $\mathfrak{g}$ *and the ideal* $[\mathfrak{g}, \mathfrak{g}]$ *is semisimple and compact.*

Proof. Suppose $\mathfrak{g}$ is a Lie algebra over $\boldsymbol{R}$ whose Killing form is strictly negative definite. Let $\boldsymbol{O}(B)$ denote the group of all linear transformations of $\mathfrak{g}$ which leave B invariant. Then $\boldsymbol{O}(B)$ is compact in the relative topology of $\boldsymbol{GL}(\mathfrak{g})$. We have $\operatorname{Aut}(\mathfrak{g}) \subset \boldsymbol{O}(B)$, so by Cor. 6.5, $\operatorname{Int}(\mathfrak{g})$ is compact.

Suppose now $\mathfrak{g}$ is an arbitrary compact Lie algebra. The Lie subgroup $\operatorname{Int}(\mathfrak{g})$ of $\boldsymbol{GL}(\mathfrak{g})$ is compact; hence it carries the relative topology of $\boldsymbol{GL}(\mathfrak{g})$. There exists a strictly positive definite quadratic form Q on $\mathfrak{g}$

invariant under the action of the compact linear group Int $(\mathfrak{g})$. There exists a basis $X_1, \dots, X_n$ of $\mathfrak{g}$ such that $Q(X) = \sum_{i=1}^n x_i^2$ if $X = \sum_{i=1}^n x_i X_i$. By means of this basis, each $\sigma \in$ Int $(\mathfrak{g})$ is represented by an orthogonal matrix and each ad X, $(X \in \mathfrak{g})$, by a skew symmetric matrix, say $(a_{ij}(X))$. Now the center $\mathfrak{z}$ of $\mathfrak{g}$ is invariant under Int $(\mathfrak{g})$, that is $\sigma \cdot \mathfrak{z} \subset \mathfrak{z}$ for each $\sigma \in$ Int $(\mathfrak{g})$. The orthogonal complement $\mathfrak{g}'$ of $\mathfrak{z}$ in $\mathfrak{g}$ with respect to Q is also invariant under Int $(\mathfrak{g})$ and under ad $(\mathfrak{g})$. Hence $\mathfrak{g}'$ is an ideal in $\mathfrak{g}$. This being so, the Killing form B' of $\mathfrak{g}'$ is the restriction to $\mathfrak{g}' \times \mathfrak{g}'$ of the Killing form B of $\mathfrak{g}$. Now, if $X \in \mathfrak{g}$

$$B(X, X) = \operatorname{Tr}(\operatorname{ad} X \operatorname{ad} X) = \sum_{i,j} a_{ij}(X)\, a_{ji}(X) = -\sum_{i,j} a_{ij}(X)^2 \leqslant 0.$$

The equality sign holds if and only if ad $X = 0$, that is, if and only if $X \in \mathfrak{z}$. This proves that $\mathfrak{g}'$ is semisimple and compact. The decomposition in Cor. 6.3 shows that $[\mathfrak{g}', \mathfrak{g}'] = \mathfrak{g}'$. Hence $\mathfrak{g}' = [\mathfrak{g}, \mathfrak{g}]$ and the proposition is proved.

Corollary 6.7. *A Lie algebra $\mathfrak{g}$ over $\boldsymbol{R}$ is compact if and only if there exists a compact Lie group G with Lie algebra isomorphic to $\mathfrak{g}$.*

For this corollary one just has to remark that every abelian Lie algebra is isomorphic to the Lie algebra of a torus $\boldsymbol{S}^1 \times \dots \times \boldsymbol{S}^1$.

Proposition 6.8. *Let $\mathfrak{g}$ be a Lie algebra over $\boldsymbol{R}$ and let $\mathfrak{z}$ denote the center of $\mathfrak{g}$. Suppose $\mathfrak{k}$ is a compactly imbedded subalgebra of $\mathfrak{g}$. If $\mathfrak{k} \cap \mathfrak{z} = \{0\}$ then the Killing form of $\mathfrak{g}$ is strictly negative definite on $\mathfrak{k}$.*

Proof. Let B denote the Killing form of $\mathfrak{g}$, and let K denote the analytic subgroup of the adjoint group Int $(\mathfrak{g})$ with Lie algebra $\operatorname{ad}_{\mathfrak{g}}(\mathfrak{k})$. Owing to our assumptions, K is a compact Lie subgroup of $\boldsymbol{GL}(\mathfrak{g})$; hence it carries the relative topology of $\boldsymbol{GL}(\mathfrak{g})$. There exists a strictly positive definite quadratic form Q on $\mathfrak{g}$ invariant under K. There exists a basis of $\mathfrak{g}$ such that each endomorphism $\operatorname{ad}_{\mathfrak{g}}(T)$, $(T \in \mathfrak{k})$, is expressed by means of a skew symmetric matrix, say $(a_{ij}(T))$. Then

$$B(T, T) = \sum_{i,j} a_{ij}(T)\, a_{ji}(T) = -\sum_{i,j} a_{ij}(T)^2 \leqslant 0,$$

and equality sign holds only if $T \in \mathfrak{z} \cap \mathfrak{k} = \{0\}$.

Theorem 6.9. *Let G be a compact, connected semisimple Lie group. Then the universal covering group G^* of G is compact.*

Proof. Let $\mathfrak{g}$ denote the Lie algebra of G (and G^*), and let B be the

Killing form of $\mathfrak{g}$. There exists unique left invariant Riemannian structures Q and Q^* on G and G^*, respectively, such that $Q_e = Q^*_{e^*} = -B$. Here e and e^* denote the identity elements in G and G^*, respectively. Since

$$B(\mathrm{Ad}\,(g)\,X, \mathrm{Ad}(g)\,Y) = B(X, Y), \qquad X, Y \in \mathfrak{g}, g \in G,$$

it follows that Q and Q^* are also invariant under right translations on G and G^*. Let π denote the covering mapping of G^* onto G. Then $Q^* = \pi^* Q$ and since G is complete, the covering manifold G^* is also complete (Prop. 10.6, Chapter I). From (1) we have

$$Q^*([\tilde{Z}, \tilde{X}], \tilde{Y}) + Q^*(\tilde{X}, [\tilde{Z}, \tilde{Y}]) = 0,$$

where $\tilde{X}$, $\tilde{Y}$, and $\tilde{Z}$ are the left invariant vector fields on G^* which have values X, Y, Z at e. Let ∇ denote the Riemannian connection on G^* induced by Q^* (Theorem 9.1, Chapter I). Then we see from (2), §9, Chapter I, that $\nabla_{\tilde{X}}(\tilde{X}) = 0$ for all $X \in \mathfrak{g}$. From Prop. 1.4 we deduce that the geodesics in G^* through e^* are the one-parameter subgroups. This implies again that G^* is complete.

Suppose now the theorem were false for G. Then, due to Prop. 10.7, Chapter I, G^* contains a ray emanating from e^*. Let γ be the one-parameter subgroup containing this ray. Then γ is a "straight line" in G^*, that is, it realizes the shortest distance in G^* between any two of its points. In fact, any pair of points on γ can be moved by a left translation on a pair of points on the ray. We parametrize γ by arc length t measured from the point $e^* = \gamma(0)$. The set $\pi(\gamma)$ is a one-parameter subgroup of G; its closure in G is a compact, abelian, connected subgroup, hence a torus. By the classical theorem of Kronecker, there exists a sequence $(t_n) \subset \boldsymbol{R}$ such that $t_n \to \infty$ and $\pi(\gamma(t_n)) \to e$. We can assume that all $\pi(\gamma(t_n))$ lie in a minimizing convex normal ball $B_r(e)$ and that each component C of $\pi^{-1}(B_r(e))$ is diffeomorphic to $B_r(e)$ under π. Then the mapping $\pi : C \to B_r(e)$ is distance-preserving; hence there exists an element $z_n \in G^*$ such that

$$\pi(z_n) = e, \tag{2}$$
$$d(z_n, \gamma(t_n)) = d(e, \pi(\gamma(t_n))).$$

Here d denotes the distance in G as well as in G^*. Since (G^*, π) is a covering group of G, the kernel of π is contained in the center Z of G^*. Hence by (2), we have $z_n \in Z$. We intend to show $\gamma \subset Z$.

Now for a given element $a \in G^*$, consider the one-parameter subgroup $\delta : t \to a\gamma(t)a^{-1}$, $(t \in \boldsymbol{R})$. Since left and right translations on G^* are isometries, δ is a "straight line" and $|t|$ is the arc parameter measured from e^*. Since $z_n \in Z$ we have $d(\delta(t_n), z_n) = d(\gamma(t_n), z_n)$ and this shows that

$$\lim_{n\to\infty} d(\gamma(t_n), \delta(t_n)) = 0. \tag{3}$$

Suppose now $\delta(t) \neq \gamma(t)$ for some $t \neq 0$. Then the angle between the vectors $\dot\gamma(0)$ and $\dot\delta(0)$ is different from 0 (possibly 180°). In any case, we have from Lemma 9.8, Chapter I and subsequent remark

$$d(\gamma(-1), \delta(+1)) < d(e^*, \gamma(-1)) + d(e^*, \delta(1)) = 2.$$

From (3) we can determine an integer N such that

$$t_N > 1, \quad d(\gamma(t_N), \delta(t_N)) < 2 - d(\gamma(-1), \delta(+1)).$$

We consider now the following broken geodesic ζ: from $\gamma(-1)$ to $\delta(+1)$ along a shortest geodesic, from $\delta(+1)$ to $\delta(t_N)$ on δ, from $\delta(t_N)$ to $\gamma(t_N)$ along a shortest geodesic. The curve ζ joins $\gamma(-1)$ to $\gamma(t_N)$ and has length

$$d(\gamma(-1), \delta(+1)) + (t_N - 1) + d(\delta(t_N), \gamma(t_N)),$$

which is strictly smaller than $t_N + 1 = d(\gamma(-1), \gamma(t_N))$. This contradicts the property of γ being a straight line.

It follows that $\delta(t) = \gamma(t)$ for all $t \in \boldsymbol{R}$. Since $a \in G^*$ was arbitrary it follows that $\gamma \subset Z$. But then $\mathfrak{z}$, the Lie algebra of Z, is $\neq \{0\}$, and this contradicts the semisimplicity of $\mathfrak{g}$.

EXERCISES

A. On the Geometry of Lie Groups

1. Let $\gamma(t)$ $(t \in \boldsymbol{R})$ be a one-parameter subgroup of a Lie group. Assume that γ intersects itself. Then γ is a "closed" one-parameter subgroup, that is, there exists a number $L > 0$ such that $\gamma(t + L) = \gamma(t)$ for all $t \in \boldsymbol{R}$.

2. Let G be a connected Lie group, H a closed subgroup. Prove that the manifold G/H is complete in any G-invariant Riemannian structure.

3. Let G be a connected Lie group with Lie algebra $\mathfrak{g}$. Let B be a nondegenerate symmetric bilinear form on $\mathfrak{g} \times \mathfrak{g}$. Then there exists a unique left invariant pseudo-Riemannian structure Q on G such that $Q_e = B$. Show, using Prop. 1.4 and (2), §9, Chapter I, that the following conditions are equivalent:

(i) The geodesics through e are the one-parameter subgroups.
(ii) $B(X, [X, Y]) = 0, \qquad$ for all $X, Y \in \mathfrak{g}$.

(iii) $B(X, [Y, Z]) = B([X, Y], Z)$ for all $X, Y, Z \in \mathfrak{g}$.

(iv) Q is invariant under all right translations on G.

(v) Q is invariant under the mapping $g \to g^{-1}$ of G onto itself.

4. Let G be a connected Lie group with Lie algebra $\mathfrak{g}$ and let Q be a Riemannian structure on G which is left and right invariant. Let $X, Y \in \mathfrak{g}$. The parallel translate of X along the geodesic $\exp tY$ $(0 \leqslant t \leqslant 1)$ is given by

$$dL\,(\exp \tfrac{1}{2}\,Y)\,dR\,(\exp \tfrac{1}{2}\,Y) \cdot X$$

if $L(x)$ and $R(x)$ denote the left and right translation by the group element x.

5. Let $\boldsymbol{SL}(2, \boldsymbol{R})$ denote the group of all real 2×2 matrices with determinant 1. Its Lie algebra $\mathfrak{sl}(2, \boldsymbol{R})$ consists of all real 2×2 matrices of trace 0.

(i) Let $X \in \mathfrak{sl}(2,\boldsymbol{R})$, $I =$ unit matrix. Show that

$$e^X = \cosh\,(-\det X)^{1/2}\,I + \frac{\sinh\,(-\det X)^{1/2}}{(-\det X)^{1/2}}\,X \qquad \text{if } \det X < 0$$

$$e^X = \cos\,(\det X)^{1/2}\,I + \frac{\sin\,(\det X)^{1/2}}{(\det X)^{1/2}}\,X \qquad \text{if } \det X > 0.$$

(ii) Let us consider one-parameter subgroups the same if they have proportional tangent vectors at e. Then the matrix

$$\begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix} \in \boldsymbol{SL}(2, \boldsymbol{R}) \qquad (\lambda \neq 1)$$

lies on exactly one one-parameter subgroup if $\lambda > 0$, on infinitely many one-parameter subgroups if $\lambda = -1$ and on no one-parameter subgroup if $\lambda < 0$, $\lambda \neq -1$.

6. Show that the group $\boldsymbol{SL}(2, \boldsymbol{R})$ admits a bi-invariant pseudo-Riemannian structure. This pseudo-Riemannian manifold is *complete* in the sense that the geodesics are indefinitely extendable (Exercise 3). Show that:

(i) Two points in $\boldsymbol{SL}(2, \boldsymbol{R})$ can not in general be joined by a geodesic.

(ii) Two points in $\boldsymbol{SL}(2, \boldsymbol{R})$ can always be joined by a singly broken geodesic.

(iii) On any connected manifold M with an affine connection an arbitrary pair of points can always be joined by a finitely broken geodesic. The number of breaks necessary may be unbounded even if M is complete (Hicks [2]).

B. Subgroups and Transformation Groups

1. Verify the description of the Lie algebras of the various subgroups of $GL(n, C)$ listed in Chapter IX, §4.

2. Show that a commutative connected Lie group is isomorphic to a product group of the form $R^n \times T^m$ where T^m is an m-dimensional torus.

3. Let G be a Lie transformation group of a manifold M. Then each orbit $G \cdot p$ is a submanifold of M, diffeomorphic to G/G_p. (Proceed as in the proof of Prop. 4.3.)

4. Derive the formula ((4), §5)

$$\operatorname{ad} X = \lim_{t \to 0} \frac{1}{t} (\operatorname{Ad}(\exp tX) - I)$$

as a special case of Theorem 7.1, Chapter I, and as a special case of Theorem 3.4, Chapter II.

5. Let G be a locally connected topological group. Suppose the identity component G_0 has an analytic structure compatible with the topology in which it is a Lie group. Show that G has the same property. (Hint: Use Theorem 2.6).

This shows that the definition of a Lie group adopted here is equivalent to that of Chevalley [2].

C. Closed Subgroups

1. Let Γ be a discrete subgroup of R^2 such that R^2/Γ is compact. Show that an analytic subgroup of R^2 is always closed but that its image R^2/Γ under the natural mapping is not necessarily closed.

2. Let $H \subset G$ be connected Lie groups. Suppose the identity mapping $I : H \to G$ is a continuous (algebraic) isomorphism. Then H is a Lie subgroup of G.

3. Let G denote the five-dimensional manifold $C \times C \times R$ with multiplication defined as follows (van Est [1], Hochschild [1]):

$$(c_1, c_2, r)(c_1', c_2', r') = (c_1 + e^{2\pi i r} c_1', c_2 + e^{2\pi i h r} c_2', r + r'),$$

where h is a fixed irrational number and $c_1, c_2, c_1', c_2' \in C$, $r, r' \in R$. Then G is a Lie group.

(i) Let $s, t \in R$ and define the mapping $\alpha_{s,t} : G \to G$ by $\alpha_{s,t}(c_1, c_2, r) = (e^{2\pi i s} c_1, e^{2\pi i t} c_2, r)$. Show that $\alpha_{s,t}$ is an analytic isomorphism.

(ii) If $t = hs + hn$ where n is an integer, then $\alpha_{s,t}$ coincides with the inner automorphism

$$(c_1, c_2, r) \to (0, 0, s + n)(c_1, c_2, r)(0, 0, s + n)^{-1}$$

(iii) Let $\mathfrak{g}$ denote the Lie algebra of G and let $A_{s,t}$ denote the automorphism $d\alpha_{s,t}$ of $\mathfrak{g}$. If $s_n \to s_0$, $t_n \to t_0$ then $A_{s_n,t_n} \to A_{s_0,t_0}$ in Aut ($\mathfrak{g}$).

(iv) Show that $A_{0,1/3} \notin$ Int ($\mathfrak{g}$). Deduce from (iii) that Int ($\mathfrak{g}$) is not closed in Aut ($\mathfrak{g}$).

4*. Let G be a connected Lie group and H an analytic subgroup. Let $\mathfrak{g}$ and $\mathfrak{h}$ denote the corresponding Lie algebras.

(i) Assume G simply connected. If $\mathfrak{h}$ is an ideal in $\mathfrak{g}$ then H is closed in G (Chevalley [2], p. 127).

(ii) Assume G simply connected. If $\mathfrak{h}$ is semisimple then H is closed in G (Mostow [2], p. 615).

(iii) Assume G compact. If $\mathfrak{h}$ is semisimple then H is closed in G (Mostow [2], p. 615).

(iv) Assume $G = GL(n, C)$. If $\mathfrak{h}$ is semisimple then H is closed in G (Goto [1]).

(v) Suppose H is not closed in G. Then there exists a one-parameter subgroup γ of H whose closure (in G) is not contained in H (Goto [1]).

NOTES

In the early days of Lie group theory, the late nineteenth century, the notion of a Lie group had, in the hands of S. Lie, W. Killing, and É. Cartan, a primarily local character. Global Lie groups were not emphasized until during the 1920's through the work of H. Weyl, É. Cartan, and O. Schreier. These two viewpoints, the infinitesimal method and the integral method, were not completely coordinated until É. Cartan proved in 1936 ([20]) that every Lie algebra over $\boldsymbol{R}$ is the Lie algebra of a Lie group. The book of Chevalley [2] gives a systematic exposition of Lie group theory from the global point of view.

§1-§3. No generality is gained by replacing the analyticity requirement in the definition of a Lie group by differentiability (F. Schmidt (1891). Pontrjagin [1], § 56). Even the differentiability assumption is not essential; a locally Euclidean topological group has an analytic structure (actually unique) in which the group operations are analytic. This theorem, proved by A. Gleason, D. Montgomery, and L. Zippin, constitutes an affirmative solution to a problem posed by Hilbert in 1900 (see Montgomery-Zippin [1]).

The universal enveloping algebra was defined by G. Birkhoff [1] and Witt [1]; it plays an important role in Harish-Chandra's work on the representations of Lie groups. The isomorphism $U(\mathfrak{g}) \approx D(G)$ in Prop. 1.9 is proved in Harish-

Chandra [8] and is attributed to L. Schwartz in Godement [3]. Various invariant affine connections on a Lie group were introduced by Cartan-Schouten [1]; Prop. 1.4 is proved by Nomizu [2]. Theorem 1.7 which we have derived as a special case of Theorem 6.5, Chapter I, is (by duality) equivalent to Cartan's formula [16], p. 21 (see also Chevalley [2], p. 157), giving the Maurer-Cartan forms in canonical coordinates. The treatment of Lie subgroups and subalgebras in §2 is primarily based on Chevalley [2]. Several simplifications have been possible since the exponential mapping is available. In particular, the proofs of the basic Theorems 1.11, 2.1 and 2.3 are from Bruhat [2]. The proof of Theorem 2.5 also occurs there (although oversimplified) and in Freudenthal [4]. The use of the graph of a homomorphism also occurs in another context in Chevalley [2], p. 112. Theorem 2.3 on closed subgroups was originally proved by von Neumann [1] for matrix groups and generalized by É. Cartan [16] to arbitrary Lie groups. Theorem 3.2 was proved by Arens [1], but Cor. 3.3 is older (see Pontrjagin [1]). The interpretation of the bracket (Theorem 3.4) is classical (see Pontrjagin [1], §60 and Nomizu [4]).

§5-§6. The adjoint group goes back to Lie and É. Cartan. The existence of a positive definite quadratic form invariant under a given compact linear group is one of H. Weyl's important applications of the invariant measure. The fundamental Theorem 6.9 is also due to Weyl [1], Kap IV, Satz 2. For a fuller exposition of Weyl's proof see Pontrjagin [1], §64. The proof given in the text is due to Samelson [1].

CHAPTER III

STRUCTURE OF SEMISIMPLE LIE ALGEBRAS

As will become clear in Chapter V, the study of symmetric spaces leads quickly and rather surprisingly to semisimple Lie algebras. This chapter is devoted to a preliminary study of these Lie algebras. The central result is Theorem 6.3, asserting that any complex semisimple Lie algebra has a compact real form. The proof is based on the root space decomposition of a complex semisimple Lie algebra with respect to a Cartan subalgebra. However, the existence of a Cartan subalgebra is not a trivial matter. It is based on Lie's theorem on solvable Lie algebras, proved in §2. Section 1 gives a brief review of some well-known facts related to the Jordan canonical form for linear transformations of a vector space.

§ 1. Preliminaries

Let K be a field and V a finite-dimensional vector space over K. We shall recall some facts concerning endomorphisms of V. Let $\operatorname{Hom}(V, V)$ denote the ring of all endomorphisms of V. Let $e_1, \ldots, e_n$ be a basis of V. To each $A \in \operatorname{Hom}(V, V)$ we associate the matrix

$$(\alpha_{ij}) = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \cdots & & \cdots & \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn} \end{pmatrix}$$

which is determined by the condition $Ae_j = \sum_{i=1}^{n} \alpha_{ij} e_i$ $(1 \leqslant j \leqslant n)$. We shall call the matrix (α_{ij}) the *expression* of A in terms of the basis $e_1, \ldots, e_n$. The mapping $A \to (\alpha_{ij})$ is an isomorphism of $\operatorname{Hom}(V, V)$ onto the ring $M_n(K)$ of all $n \times n$ matrices with entries in K. If $f_1, \ldots, f_n$ is a basis dual to $e_1, \ldots, e_n$, then the endomorphism ${}^tA : V^* \to V^*$ has matrix expression (α_{ji}), the transpose of (α_{ij}). A matrix (α_{ij}) for which $\alpha_{ij} = 0$ if $i > j$ is called an *upper triangular* matrix, a matrix (β_{ij}) for which $\beta_{ij} = 0$ if $i < j$ is called a *lower triangular* matrix. A matrix which is both upper and lower triangular is called a *diagonal matrix.*

If $\lambda \in K$, let V_λ denote the set of elements $e \in V$ such that $Ae = \lambda e$. If $V_\lambda \neq \{0\}$, then λ is called an *eigenvalue* of A and V_λ is called the *eigenspace* of A for the eigenvalue λ. A vector $v \neq 0$ in V which belongs to some eigenspace of A is called an *eigenvector* of A. The equation in λ,

$$\det(\lambda I - A) = 0 \qquad (I = \text{identity endomorphism})$$

is called the *characteristic equation* of A. Those solutions of this equation which lie in K coincide with the eigenvalues of A. The left-hand side of the equation is called the *characteristic polynomial* of A.

An endomorphism $N \in \text{Hom}(V, V)$ is called *nilpotent* if $N^k = 0$ for some integer $k > 0$. If $V \neq \{0\}$ and if $N \in \text{Hom}(V, V)$ is nilpotent, then N has exactly one eigenvalue, namely 0. Let $e_1 \neq 0$ be a vector in V such that $Ne_1 = 0$. If E_1 denotes the one-dimensional subspace of V spanned by e_1, N induces an endomorphism N_1 of the factor space V/E_1. This endomorphism N_1 is again nilpotent and if $\dim V/E_1 \neq 0$, we can select $e_2 \neq 0$ in V such that the vector $(e_2 + E_1) \in V/E_1$ is an eigenvector of N_1. By a continuation of this process we obtain a basis $e_1, ..., e_n$ of V such that

$$Ne_1 = 0,\ Ne_p = 0 \mod (e_1, ..., e_{p-1}),\ 2 \leqslant p \leqslant n.$$

Here $(e_1, ..., e_{p-1})$ denotes the subspace of V spanned by the vectors $e_1, ..., e_{p-1}$. The matrix (n_{ij}) expressing N in terms of the basis $e_1, ..., e_n$ has 0 on and below the diagonal. On the other hand, let (n_{ij}) be an $n \times n$ matrix with 0 on and below the diagonal. If $f_1, ..., f_n$ is any basis of V, the endomorphism N given by $Nf_j = \sum_{i=1}^n n_{ij}f_i$ is nilpotent.

Thus N is nilpotent if and only if it has a matrix expression with 0 *on and below the diagonal.*

Consider now a subset $\mathfrak{S} \subset \text{Hom}(V, V)$. A subspace W of V is called *invariant* (under $\mathfrak{S}$) if $SW \subset W$ for each $S \in \mathfrak{S}$. The space V is called *irreducible* if its only invariant subspaces are $\{0\}$ and V. The set $\mathfrak{S}$ is called *semisimple* if each invariant subspace (under $\mathfrak{S}$) has a complementary invariant subspace. In this case V can be written as a direct sum $V = \sum_i V_i$ where the spaces V_i are invariant and irreducible (under $\mathfrak{S}$). If $\mathfrak{S}$ is a commutative family of endomorphisms, then $\mathfrak{S}$ is semisimple if and only if each $S \in \mathfrak{S}$ is semisimple.

Each $A \in \text{Hom}(V, V)$ can be uniquely decomposed:

$$A = S + N,\ S \text{ semisimple},\ N \text{ nilpotent},\ SN = NS. \qquad (1)$$

In this decomposition, S and N are polynomials in A, and are called the *semisimple part* and *nilpotent part* of A respectively. Suppose λ is an eigenvalue of A. Select a vector $e \neq 0$ such that $Ae = \lambda e$. Since N is a polynomial in A, but has 0 as the only eigenvalue, it follows that $Ne = 0$ and $Se = \lambda e$. On the other hand, let λ be an eigenvalue of S

and let V_λ be the eigenspace of S for the eigenvalue λ. Since $AS = SA$, we have $AV_\lambda \subset V_\lambda$ and the restriction of $A - \lambda I$ to V_λ is nilpotent. In particular, λ is an eigenvalue of A. Thus, A and S have the same eigenvalues, say $\lambda_1, \cdots, \lambda_r$.

Suppose now that K is algebraically closed. Let $\mathfrak{S}$ be a semisimple commutative family of endomorphisms of V. Then each nonzero invariant irreducible subspace (under $\mathfrak{S}$) is one-dimensional, so V has a basis in terms of which all $S \in \mathfrak{S}$ are expressed by diagonal matrices. In particular, this is the case when $\mathfrak{S}$ consists of the single endomorphism S above. Thus V can be written as a direct sum

$$V = \sum_{i=1}^{r} V_i,$$

where V_i is the eigenspace of S for the eigenvalue λ_i. Let $v \neq 0$ be a vector in V and $\lambda \in K$ such that

$$(A - \lambda I)^k v = 0$$

for some integer k. Taking k as small as possible, we see that λ is an eigenvalue of A, let us say $\lambda = \lambda_1$. We can write $v = \sum_i v_i$ where $v_i \in V_i$, and since V_i is invariant under A, we obtain

$$(A - \lambda_1 I)^k v_i = 0 \qquad \text{for } 1 \leqslant i \leqslant r.$$

Now, the endomorphism $N_i = A - \lambda_i I$ is nilpotent on V_i, whereas the equation

$$(N_i + (\lambda_i - \lambda_1) I)^k v_i = 0$$

shows that if $v_i \neq 0$, N_i has the eigenvalue $\lambda_1 - \lambda_i$ on V_i. It follows that $v_i = 0$ if $i > 1$; hence $v \in V_1$. Summarizing, we get:

Proposition 1.1. *Let V be a finite-dimensional vector space over an algebraically closed field K. Let $A \in Hom(V, V)$, and let $\lambda_1, \ldots, \lambda_r \in K$ be the different eigenvalues of A. Put*

$$V_i = \{v \in V : (A - \lambda_i I)^k v = 0 \text{ for } k \text{ sufficiently large}\}.$$

Then

(i) $$V = \sum_{i=1}^{r} V_i \qquad (\textit{direct sum}).$$

(ii) *Each V_i is invariant under A.*

(iii) *The semisimple part of* A *is given by*

$$S\left(\sum_{i=1}^{r} v_i\right) = \sum_{i=1}^{r} \lambda_i v_i \qquad (v_i \in V_i).$$

(iv) *The characteristic polynomial of* A *is*

$$\det(\lambda I - A) = (\lambda - \lambda_1)^{d_1} \dots (\lambda - \lambda_r)^{d_r},$$

where $d_i = \dim V_i$ $(1 \leqslant i \leqslant r)$.

§ 2. Theorems of Lie and Engel

Throughout this section, K denotes a field of characteristic 0 and $\tilde{K}$ its algebraic closure. Let $\mathfrak{g}$ be a Lie algebra over K. The vector space spanned by all elements $[X, Y]$, $X, Y \in \mathfrak{g}$, is an ideal in $\mathfrak{g}$, called the *derived* algebra of $\mathfrak{g}$. The derived algebra will be denoted $\mathfrak{D}\mathfrak{g}$ and the nth derived algebra $\mathfrak{D}^n\mathfrak{g}$ of $\mathfrak{g}$ is defined inductively by $\mathfrak{D}^0\mathfrak{g} = \mathfrak{g}$ and $\mathfrak{D}^n\mathfrak{g} = \mathfrak{D}(\mathfrak{D}^{n-1}\mathfrak{g})$. Each $\mathfrak{D}^n\mathfrak{g}$ is an ideal in $\mathfrak{g}$.

Definition. The Lie algebra $\mathfrak{g}$ is called *solvable* if there exists an integer $n \geqslant 0$ such that $\mathfrak{D}^n\mathfrak{g} = \{0\}$. A Lie group is called solvable if its Lie algebra is solvable.

Let $\mathfrak{g}$ be a solvable Lie algebra $\neq \{0\}$ and let n be the smallest integer for which $\mathfrak{D}^n\mathfrak{g} = \{0\}$. Then $\mathfrak{D}^{n-1}\mathfrak{g}$ is a nonzero abelian ideal in $\mathfrak{g}$. Hence $\mathfrak{g}$ is not semisimple (Prop. 6.1, Chapter II).

Definition. A Lie algebra $\mathfrak{g}$ is said to satisfy the *chain condition* if for each ideal $\mathfrak{h} \neq \{0\}$ in $\mathfrak{g}$ there exists an ideal $\mathfrak{h}_1$ of $\mathfrak{h}$ of codimension 1.

Lemma 2.1. *A Lie algebra* $\mathfrak{g}$ *is solvable if and only if it satisfies the chain condition.*

Proof. If $\mathfrak{g}$ is solvable and $\neq \{0\}$, then $\mathfrak{D}\mathfrak{g} \neq \mathfrak{g}$. Hence there exists a subspace $\mathfrak{h}$ of $\mathfrak{g}$ of codimension 1, containing $\mathfrak{D}\mathfrak{g}$. Then $\mathfrak{h}$ is an ideal in $\mathfrak{g}$. Each ideal (even each subalgebra) of a solvable Lie algebra is solvable; consequently, $\mathfrak{g}$ satisfies the chain condition. On the other hand, suppose $\mathfrak{g}$ is a Lie algebra satisfying the chain condition. Then there exists a sequence

$$\mathfrak{g} = \mathfrak{g}_0 \supset \mathfrak{g}_1 \supset \cdots \supset \mathfrak{g}_{n-1} \supset \mathfrak{g}_n = \{0\},$$

where $\mathfrak{g}_r$ is an ideal in $\mathfrak{g}_{r-1}$ of codimension 1 $(1 \leqslant r \leqslant n)$, and thus $\mathfrak{D}(\mathfrak{g}_{r-1}) \subset \mathfrak{g}_r$. By induction, $\mathfrak{g}$ is solvable.

Theorem 2.2. (Lie) *Let* $\mathfrak{g}$ *be a solvable Lie algebra over* K. *Let* $V \neq \{0\}$ *be a finite-dimensional vector space over* $\tilde{K}$, *the algebraic closure of* K. *Let* π *be a homomorphism of* $\mathfrak{g}$ *into* $\mathfrak{gl}(V)$. *Then there exists a vector* $v \neq 0$ *in* V *which is an eigenvector of all the members of* $\pi(\mathfrak{g})$.

Proof. We shall prove the theorem by induction on $\dim \mathfrak{g}$. If $\dim \mathfrak{g} = 1$, the theorem is a consequence of Prop. 1.1; we assume now that the theorem holds for all solvable Lie algebras over K of dimension $< \dim \mathfrak{g}$. Let $\mathfrak{h}$ be an ideal in $\mathfrak{g}$ of codimension 1. Then $\mathfrak{h}$ is solvable, so by assumption there exists a vector $e_0 \neq 0$ in V and a linear function $\lambda : \mathfrak{h} \to \tilde{K}$ such that

$$\pi(H)\, e_0 = \lambda(H)\, e_0 \qquad \text{for all } H \in \mathfrak{h}.$$

Select $X \in \mathfrak{g}$ such that $X \notin \mathfrak{h}$ and put

$$e_{-1} = 0, \qquad e_p = \pi(X)^p\, e_0, \qquad p = 1, 2, \ldots .$$

The subspace W of V spanned by all e_p $(p \geqslant 0)$ is invariant under $\pi(X)$. We shall now prove by induction that

$$\pi(H)\, e_p \equiv \lambda(H)\, e_p \qquad \mod (e_0, \ldots, e_{p-1}) \qquad \text{for all } H \in \mathfrak{h},\ p \geqslant 0. \tag{1}$$

In fact, (1) holds for $p = 0$, and, assuming it for p, we have

$$\pi(H)\, e_{p+1} = \pi(H)\, \pi(X)\, e_p = \pi([H, X])\, e_p + \pi(X)\, \pi(H)\, e_p$$

$$\equiv \lambda([H, X])\, e_p + \pi(X)\, \lambda(H)\, e_p \qquad \mod (e_0, \ldots, e_{p-1}, \pi(X)\, e_0, \ldots, \pi(X)\, e_{p-1})$$

so

$$\pi(H)\, e_{p+1} \equiv \lambda(H)\, e_{p+1} \qquad \mod (e_0, e_1 \ldots e_p) \qquad \text{for } H \in \mathfrak{h}.$$

It follows that W is invariant under $\pi(\mathfrak{g})$ and $\mathrm{Tr}_W\, \pi(H) = \lambda(H) \dim W$. Now $\pi([H, X]) = \pi(H)\, \pi(X) - \pi(X)\, \pi(H)$ so $\mathrm{Tr}_W\, \pi([H, X]) = 0$. Since $\dim W > 0$ we obtain $\lambda([H, X]) = 0$. Now, we have

$$\pi(H)\, e_{p+1} = \pi([H, X])\, e_p + \pi(X)\, \pi(H)\, e_p$$

and the relation

$$\pi(H)\, e_p = \lambda(H)\, e_p \qquad (H \in \mathfrak{h}, \text{ all } p \geqslant 0)$$

follows by induction on p. This shows that for $H \in \mathfrak{h}$, $\pi(H) = \lambda(H)\, I$ on W. Since $\pi(X)$ leaves W invariant it has an eigenvector $v \neq 0$ in W. This vector has the properties stated in the theorem.

Corollary 2.3. *Let* $\mathfrak{g}$ *be a solvable Lie algebra over a field* K *and* π *a representation of* $\mathfrak{g}$ *on a finite-dimensional vector space* $V \neq \{0\}$ *over* $\tilde{K}$, *the algebraic closure of K. Then there exists a basis $e_1, \ldots, e_n$ of V, in terms of which all the endomorphisms $\pi(X)$, $X \in \mathfrak{g}$, are expressed by upper triangular matrices.*

In fact, we can apply Theorem 2.2. Let $e_1 \neq 0$ be a common eigenvector of all $\pi(X)$, $X \in \mathfrak{g}$, and consider the subspace E_1 of V spanned by e_1. The representation π induces a representation π_1 of $\mathfrak{g}$ on the factor space V/E_1, so if dim $V/E_1 \neq 0$ we can select $e_2 \in V$ such that the vector $(e_2 + E_1) \in V/E_1$ is an eigenvector for all $\pi_1(X)$. Continuing in this manner we find a basis $e_1, \ldots, e_n$ of V such that for each $X \in \mathfrak{g}$

$$\pi(X)\, e_i \equiv 0 \qquad \mod (e_1, e_2 \ldots e_i).$$

This means that the matrix representing $\pi(X)$ has zeros below the main diagonal.

Definition. A Lie algebra $\mathfrak{g}$ over K is said to be *nilpotent* if for each $Z \in \mathfrak{g}$, $\mathrm{ad}_{\mathfrak{g}}\, Z$ is a nilpotent endomorphism of $\mathfrak{g}$. A Lie group is called nilpotent if its Lie algebra is nilpotent.

Theorem 2.4. (Engel) *Let V be a nonzero finite-dimensional vector space over K, and let $\mathfrak{g}$ be a subalgebra of $\mathfrak{gl}(V)$ consisting of nilpotent elements. Then*

(i) $\mathfrak{g}$ *is nilpotent.*

(ii) *There exists a vector $v \neq 0$ in V such that $Zv = 0$ for all $Z \in \mathfrak{g}$.*

(iii) *There exists a basis $e_1, \ldots, e_n$ of V in terms of which all the endomorphisms $X \in \mathfrak{g}$ are expressed by matrices with zeros on and below the diagonal.*

Proof. (i) For $Z \in \mathfrak{gl}(V)$ consider the endomorphisms L_Z and R_Z on $\mathfrak{gl}(V)$ given by $L_Z X = ZX$, $R_Z X = XZ$ $(X \in \mathfrak{gl}(V))$. Then L_Z and R_Z commute and if ad denotes the adjoint representation of $\mathfrak{gl}(V)$, we have $\mathrm{ad}\, Z = L_Z - R_Z$. It follows that for $X \in \mathfrak{g}$ and any integer $p \geqslant 0$

$$(\mathrm{ad}\, Z)^p\,(X) = \sum_{i=0}^{p} (-1)^i \binom{p}{i} Z^{p-i}\, X Z^i. \tag{2}$$

Suppose $Z \in \mathfrak{g}$. Then Z is nilpotent and by relation (2) ad Z is nilpotent on $\mathfrak{gl}(V)$. Since $\mathrm{ad}_{\mathfrak{g}}\, Z$ is the restriction of ad Z to $\mathfrak{g}$, it follows that $\mathrm{ad}_{\mathfrak{g}}\, Z$ is nilpotent.

For the second part of the theorem let $r = \dim \mathfrak{g}$. We shall use induction on r. If $r = 1$, (ii) is trivial. Assume now that (ii) holds for algebras of dimension $< r$. Let $\mathfrak{h}$ be a proper subalgebra of $\mathfrak{g}$ of maximum dimension. If $H \in \mathfrak{h}$, then by (i), $\mathrm{ad}_{\mathfrak{g}}\, H$ is a nilpotent endomorphism of $\mathfrak{g}$

and maps $\mathfrak{h}$ into itself, hence $\mathrm{ad}_{\mathfrak{g}}\, H$ induces a nilpotent endomorphism H^* on the vector space $\mathfrak{g}/\mathfrak{h}$. The set $\{H^* : H \in \mathfrak{h}\}$ is a subalgebra of $\mathfrak{gl}(\mathfrak{g}/\mathfrak{h})$ having dimension $< r$ and consisting of nilpotent elements. Using the induction hypothesis we conclude that there exists an element $X \in \mathfrak{g}$, $X \notin \mathfrak{h}$, such that $\mathrm{ad}_{\mathfrak{g}}\, H(X) \in \mathfrak{h}$ for all $H \in \mathfrak{h}$. The subspace $\mathfrak{h} + KX$ of $\mathfrak{g}$ is therefore a subalgebra of $\mathfrak{g}$ which, due to the maximality of $\mathfrak{h}$, must coincide with $\mathfrak{g}$. Thus $\mathfrak{h}$ is an ideal in $\mathfrak{g}$.

Now let W denote the subspace of V given by

$$W = \{e \in V : He = 0 \text{ for all } H \in \mathfrak{h}\}.$$

Owing to the induction hypothesis, $W \neq \{0\}$. Moreover, if $e \in W$ we have

$$HXe = [H, X]\, e + XHe = 0$$

so $X \cdot W \subset W$. The restriction of X to W is nilpotent and there exists a vector $v \neq 0$ in W such that $Xv = 0$. This vector v has the property required in (ii).

To prove (iii) let e_1 be any vector in V such that $e_1 \neq 0$ and $Ze_1 = 0$ for all $Z \in \mathfrak{g}$. Let E_1 be the subspace of V spanned by e_1. Then each $Z \in \mathfrak{g}$ induces a nilpotent endomorphism Z^* of the vector space V/E_1. If $V/E_1 \neq \{0\}$ we can select $e_2 \in V$, $e_2 \notin E_1$ such that $e_2 + E_1 \in V/E_1$ is annihilated by all Z^*, $(Z \in \mathfrak{g})$. Continuing in this manner we find a basis $e_1, \ldots, e_n$ of V such that for each $Z \in \mathfrak{g}$

$$Ze_1 = 0, \qquad Ze_i \equiv 0 \qquad \mathrm{mod}\, (e_1, \ldots, e_{-i1}), \qquad 2 \leqslant i \leqslant n. \tag{3}$$

The matrix expressing Z in terms of the basis $e_1, \ldots, e_n$ has zeros on and below the diagonal.

Corollary 2.5. *In the notation of Theorem* 2.4 *we have*

$$X_1 X_2 \ldots X_s = 0$$

if $s \geqslant \dim V$ *and* $X_i \in \mathfrak{g}$ $(1 \leqslant i \leqslant s)$.

In fact, this is an immediate consequence of (3).

Corollary 2.6. *A nilpotent Lie algebra* $\mathfrak{g}$ *is solvable.*

In fact, the algebra $\mathrm{ad}_{\mathfrak{g}}(\mathfrak{g})$ is a subalgebra of $\mathfrak{gl}(\mathfrak{g})$ and consists of nilpotent endomorphisms of $\mathfrak{g}$. The product of s such endomorphisms is 0 if $s \geqslant \dim \mathfrak{g}$ (Cor. 2.5). In particular, $\mathfrak{g}$ is solvable.

Definition. For a Lie algebra $\mathfrak{l}$, we define

$$\mathscr{C}^0\mathfrak{l} = \mathfrak{l}, \qquad \mathscr{C}^{p+1}\mathfrak{l} = [\mathfrak{l}, \mathscr{C}^p\mathfrak{l}], \qquad p = 0, 1, \ldots.$$

The series

$$\mathscr{C}^0\mathfrak{l} \supset \mathscr{C}^1\mathfrak{l} \supset \mathscr{C}^2\mathfrak{l} \supset \dots$$

is called the *central decending series* of $\mathfrak{l}$.

Corollary 2.7. *A Lie algebra $\mathfrak{l}$ over K is nilpotent if and only if $\mathscr{C}^m\mathfrak{l} = \{0\}$ for $m \geqslant \dim \mathfrak{l}$.*

In fact, if $\mathfrak{l}$ is nilpotent we can use Cor. 2.5 on the Lie algebra $\mathfrak{g} = \text{ad}(\mathfrak{l})$ and deduce $\mathscr{C}^m\mathfrak{l} = \{0\}$ if $m \geqslant \dim \mathfrak{l}$. The converse is trivial.

Corollary 2.8. *A nilpotent Lie algebra $\mathfrak{l} \neq \{0\}$ has nonzero center.*

In fact, if $\mathscr{C}^m\mathfrak{l} = \{0\}$ then $\mathscr{C}^{m-1}\mathfrak{l}$ lies in the center of $\mathfrak{l}$.

§ 3. Cartan Subalgebras

In this section $\mathfrak{g}$ denotes an arbitrary fixed semisimple Lie algebra $\mathfrak{g}$ over the complex numbers C. The adjoint representation of $\mathfrak{g}$ will be denoted ad.

Definition. A *Cartan subalgebra* of $\mathfrak{g}$ is a subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ satisfying the following conditions:

(i) $\mathfrak{h}$ is a maximal abelian subalgebra of $\mathfrak{g}$.

(ii) For each $H \in \mathfrak{h}$, the endomorphism ad H of $\mathfrak{g}$ is semisimple.

In this section we shall prove that every semisimple Lie algebra $\mathfrak{g}$ over C has a Cartan subalgebra. Later on we shall see that for any two Cartan subalgebras $\mathfrak{h}_1$ and $\mathfrak{h}_2$ of $\mathfrak{g}$, there exists an automorphism σ of $\mathfrak{g}$ such that $\sigma\cdot\mathfrak{h}_1 = \mathfrak{h}_2$.

Let H be any element in $\mathfrak{g}$ and let $0 = \lambda_0, \lambda_1, \dots, \lambda_r$ be the different eigenvalues of ad H. For each $\lambda \in C$ we consider the subspace

$$\mathfrak{g}(H, \lambda) = \{X \in \mathfrak{g} : (\text{ad } H - \lambda I)^k X = 0 \text{ for some } k\}.$$

Then, according to Prop. 1.1 we have $\mathfrak{g}(H, \lambda) = 0$ unless $\lambda = \lambda_i$ for some i; also

$$\mathfrak{g} = \sum_{i=0}^{r} \mathfrak{g}(H, \lambda_i) \qquad \text{(direct sum)}.$$

Definition. The element $H \in \mathfrak{g}$ is called *regular* if

$$\dim \mathfrak{g}(H, 0) = \min_{X \in \mathfrak{g}} (\dim \mathfrak{g}(X, 0)).$$

We shall now prove the following theorem, which ensures the existence of Cartan subalgebras.

Theorem 3.1. *Let H_0 be a regular element in $\mathfrak{g}$. Then $\mathfrak{g}(H_0, 0)$ is a Cartan subalgebra of $\mathfrak{g}$.*

This will be proved by means of the theorems of Lie and Engel. We put $\mathfrak{h} = \mathfrak{g}(H_0, 0)$.

Lemma 3.2. *If $Z \in \mathfrak{g}$, then $[\mathfrak{g}(Z, \lambda), \mathfrak{g}(Z, \mu)] \subset \mathfrak{g}(Z, \lambda + \mu)$. In particular, $\mathfrak{h}$ is a subalgebra of $\mathfrak{g}$.*

Let $X_\lambda \in \mathfrak{g}(Z, \lambda)$, $X_\mu \in \mathfrak{g}(Z, \mu)$. Then

$$(\operatorname{ad} Z - (\lambda + \mu) I)\, [X_\lambda, X_\mu] = [(\operatorname{ad} Z - \lambda I)\, X_\lambda, X_\mu] + [X_\lambda, (\operatorname{ad} Z - \mu I)\, X_\mu],$$

and by induction

$$(\operatorname{ad} Z - (\lambda + \mu) I)^n\, [X_\lambda, X_\mu] = \sum_{i=0}^{n} \binom{n}{i} [(\operatorname{ad} Z - \lambda I)^i X_\lambda, (\operatorname{ad} Z - \mu I)^{n-i} X_\mu].$$

The lemma follows immediately.

Lemma 3.3. *The algebra $\mathfrak{h}$ is nilpotent.*

Let $0 = \lambda_0, \lambda_1, \ldots, \lambda_r$ be the different eigenvalues of $\operatorname{ad} H_0$ and let $\mathfrak{g}'$ denote the subspace $\sum_{i=1}^{r} \mathfrak{g}(H_0, \lambda_i)$ of $\mathfrak{g}$. Then $[\mathfrak{h}, \mathfrak{g}'] \subset \mathfrak{g}'$ due to Lemma 3.2. For each $H \in \mathfrak{h}$ let H' denote the restriction of $\operatorname{ad} H$ to $\mathfrak{g}'$. Let $d(H) = \det H'$. Then the function $H \to d(H)$ is a polynomial function on $\mathfrak{h}$ and since H_0' has only nonzero eigenvalues we have $d(H_0) \neq 0$. Now if a polynomial function vanishes on an open set it must vanish identically. We conclude that the subset S of $\mathfrak{h}$ consisting of all points $H \in \mathfrak{h}$ for which $d(H) \neq 0$ is a dense subset of $\mathfrak{h}$. If H is any element in S, the endomorphism H' of $\mathfrak{g}'$ has all its eigenvalues $\neq 0$; it follows that $\mathfrak{g}(H, 0) \subset \mathfrak{h}$. Since H_0 is regular we conclude that $\mathfrak{g}(H, 0) = \mathfrak{h}$. This means that the restriction of $\operatorname{ad} H$ to $\mathfrak{h}$ is nilpotent. This restriction is $\operatorname{ad}_{\mathfrak{h}} H$, so if $l = \dim \mathfrak{h}$ we have

$$(\operatorname{ad}_{\mathfrak{h}} H)^l = 0 \qquad \text{for each } H \in S. \tag{1}$$

Since S is dense in $\mathfrak{h}$, relation (1) follows by continuity for all $H \in \mathfrak{h}$; thus $\mathfrak{h}$ is nilpotent.

The definition of a regular element and the proof of Lemma 3.3 holds for any complex Lie algebra $\mathfrak{g}$. In the next lemma, however, we make use of the semisimplicity of $\mathfrak{g}$.

Lemma 3.4. *The algebra $\mathfrak{h}$ is abelian and in fact a maximal abelian subalgebra of $\mathfrak{g}$.*

As usual, let B denote the Killing form of $\mathfrak{g}$. If $X \in \mathfrak{g}(H_0, \lambda)$ and

$H \in \mathfrak{h}$, then ad X ad H maps the subspace $\mathfrak{g}(H_0, \mu)$ into $\mathfrak{g}(H_0, \lambda + \mu)$; choosing a basis of $\mathfrak{g}$ composed of bases of the spaces $\mathfrak{g}(H_0, \lambda_i)$ we see that

$$\mathrm{Tr}\,(\mathrm{ad}\,X\,\mathrm{ad}\,H) = B(X, H) = 0 \qquad \text{if } H \in \mathfrak{h},\ X \in \mathfrak{g}'. \tag{2}$$

Now $\mathfrak{h}$, being nilpotent, is solvable (Cor. 2.6) so, by Cor. 2.3, there exists a basis of $\mathfrak{g}$ with respect to which all the endomorphisms ad H ($H \in \mathfrak{h}$) are expressed by upper triangular matrices. If A, B, C are upper triangular matrices, then ABC and BAC have the same diagonal elements; hence $\mathrm{Tr}\,(ABC) = \mathrm{Tr}\,(BAC)$. In particular we have

$$\mathrm{Tr}\,(\mathrm{ad}\,[H_1, H_2]\,\mathrm{ad}\,H) = 0 \qquad \text{if } H_1, H_2, H \in \mathfrak{h}.$$

Combining this with (2) we see that $[H_1, H_2]$ is orthogonal to $\mathfrak{g}$ (with respect to B). Making now use of the semisimplicity of $\mathfrak{g}$ we conclude that $\mathfrak{h}$ is abelian. The maximality is immediate from the definition of $\mathfrak{h}$.

Passing now to the proof of Theorem 3.1, let λ be one of the nonzero eigenvalues $\lambda_1, \ldots, \lambda_r$ of ad H_0. Each endomorphism ad H ($H \in \mathfrak{h}$) leaves $\mathfrak{g}(H_0, \lambda)$ invariant, so if $\mathrm{ad}_\lambda H$ denotes the restriction of ad H to $\mathfrak{g}(H_0, \lambda)$, the mapping $\mathrm{ad}_\lambda : H \to \mathrm{ad}_\lambda H$ is a representation of $\mathfrak{h}$ on $\mathfrak{g}(H_0, \lambda)$. Due to Cor. 2.3 there exists a basis $e_1, \ldots, e_s$ of $\mathfrak{g}(H_0, \lambda)$ with respect to which all the endomorphisms $\mathrm{ad}_\lambda H$ ($H \in \mathfrak{h}$) are expressed by upper triangular matrices. The diagonal elements $\alpha_1(H), \ldots, \alpha_s(H)$ are linear functions on $\mathfrak{h}$ and $\alpha_1(H_0) = \ldots = \alpha_s(H_0) = \lambda$. Let β be any linear function on $\mathfrak{h}$ such that $\beta(H_0) = \lambda$. Let V_β be the subspace of $\mathfrak{g}(H_0, \lambda)$ spanned by the basis vectors e_j for which $\alpha_j(H) \equiv \beta(H)$ for all $H \in \mathfrak{h}$. Then V_β is the set of vectors $X \in \mathfrak{g}(H_0, \lambda)$ such that

$$(\mathrm{ad}\,H - \beta(H)\,I)^k X = 0 \tag{3}$$

for all $H \in \mathfrak{h}$ and some fixed k (independent of H). On the other hand, since $\beta(H_0) = \lambda$, no vector $X \in \mathfrak{g}$ which does not lie in $\mathfrak{g}(H_0, \lambda)$ could satisfy (3). It follows that

$$V_\beta = \{X \in \mathfrak{g} : (\mathrm{ad}\,H - \beta(H)\,I)^k X = 0 \text{ for all } H \in \mathfrak{h} \text{ and some fixed } k\}. \tag{4}$$

Given any linear function β on $\mathfrak{h}$ we can define a subset V_β of $\mathfrak{g}$ by (4). Then V_β is a subspace of $\mathfrak{g}$ and $V_0 = \mathfrak{h}$. Moreover, the relation

$$[V_\alpha, V_\beta] \subset V_{\alpha+\beta} \tag{5}$$

is proved just as Lemma 3.2.

It has been shown above that $\mathfrak{g}$ is a direct sum of certain of the spaces V_β, say

$$\mathfrak{g} = \sum_i V_{\beta_i}.$$

Then, if $H, H' \in \mathfrak{h}$ we have

$$B(H, H') = \operatorname{Tr}(\operatorname{ad} H \operatorname{ad} H') = \sum_i \beta_i(H)\,\beta_i(H') \dim V_{\beta_i}. \tag{6}$$

The endomorphism $\operatorname{ad} H$ can be decomposed

$$\operatorname{ad} H = S + N,$$

where S is semisimple, N is nilpotent, and $SN = NS$. Since S is a polynomial in $\operatorname{ad} H$, S leaves each V_β invariant and (4) shows that $SX = \beta(H)\,X$ for all $X \in V_\beta$. From (5) it now follows quickly that S is a derivation of $\mathfrak{g}$. From Prop. 6.4, Chapter II we know that every derivation of $\mathfrak{g}$ is inner; in other words, there exists an element $Z \in \mathfrak{g}$ such that $S = \operatorname{ad} Z$. Now S, being a polynomial in $\operatorname{ad} H$, commutes with all elements of $\operatorname{ad} \mathfrak{h}$. Since $\mathfrak{h}$ is maximal abelian in $\mathfrak{g}$, this implies $Z \in \mathfrak{h}$. Now, $\operatorname{ad} Z(X) = \beta(H)\,X$ for $X \in V_\beta$, so by the definition of V_β we conclude that $\beta(H) = \beta(Z)$ for every linear function β on $\mathfrak{h}$ for which $V_\beta \neq \{0\}$. But then (6) and (2) show that $Z - H$ is orthogonal to $\mathfrak{g}$ (with respect to B). Consequently, $Z = H$ so $\operatorname{ad} H$ is semisimple. This concludes the proof of Theorem 3.1.

§ 4. Root Space Decomposition

The structure theory of semisimple Lie algebras is based on the following theorem of which a proof was given in §3.

Theorem 4.1. *Every semisimple Lie algebra over C contains a Cartan subalgebra.*

In §4 and §5 let $\mathfrak{g}$ be an arbitrary semisimple Lie algebra over C and let $\mathfrak{h}$ denote an arbitrary fixed Cartan subalgebra.

Let α be a linear function on the complex vector space $\mathfrak{h}$. Let $\mathfrak{g}^\alpha$ denote the linear subspace of $\mathfrak{g}$ given by

$$\mathfrak{g}^\alpha = \{X \in \mathfrak{g} : [H, X] = \alpha(H)\,X \text{ for all } H \in \mathfrak{h}\}.$$

The linear function α is called a *root* (of $\mathfrak{g}$ with respect to $\mathfrak{h}$) if $\mathfrak{g}^\alpha \neq \{0\}$. In that case, $\mathfrak{g}^\alpha$ is called a *root subspace*. Since $\mathfrak{h}$ is a maximal abelian subalgebra of $\mathfrak{g}$ we have $\mathfrak{g}^0 = \mathfrak{h}$. From the Jacobi identity we obtain easily the relation

$$[\mathfrak{g}^\alpha, \mathfrak{g}^\beta] \subset \mathfrak{g}^{\alpha+\beta} \tag{1}$$

for any pair α, β of C-linear functions on $\mathfrak{h}$.

Let Δ denote the set of all nonzero roots and as before let B denote the Killing form on $\mathfrak{g} \times \mathfrak{g}$.

Theorem 4.2.

(i) $\mathfrak{g} = \mathfrak{h} + \sum_{\alpha \in \Delta} \mathfrak{g}^{\alpha}$ *(direct sum).*

(ii) *dim* $\mathfrak{g}^{\alpha} = 1$ *for each* $\alpha \in \Delta$.

(iii) *Let* α, β *be two roots such that* $\alpha + \beta \neq 0$. *Then* $\mathfrak{g}^{\alpha}$ *and* $\mathfrak{g}^{\beta}$ *are orthogonal under* B.

(iv) *The restriction of* B *to* $\mathfrak{h} \times \mathfrak{h}$ *is nondegenerate. For each root* α *there exists a unique element* $H_{\alpha} \in \mathfrak{h}$ *such that*

$$B(H, H_{\alpha}) = \alpha(H) \qquad \text{for all } H \in \mathfrak{h}.$$

(v) *If* $\alpha \in \Delta$, *then* $-\alpha \in \Delta$ *and*

$$[\mathfrak{g}^{\alpha}, \mathfrak{g}^{-\alpha}] = CH_{\alpha}, \qquad \alpha(H_{\alpha}) \neq 0.$$

Proof. (i) We prove first that the sum is direct. If this were not so we would have a relation

$$H^* + \sum_i X_{\alpha_i} = 0,$$

where $H^* \in \mathfrak{h}$, $X_{\alpha_i} \neq 0$ in $\mathfrak{g}^{\alpha_i}$ where the roots α_i are different and not 0. Then we can select $H \in \mathfrak{h}$ such that the numbers $\alpha_i(H)$ are all different and not zero. In fact, the subset N of $\mathfrak{h}$ where all α_i are different and nonzero is the complement of the union of finitely many hyperplanes. In particular N is not empty. Then H^* and the vectors X_{α_i} lie in different eigenspaces of ad H and are therefore linearly independent. On the other hand, since the set $\mathrm{ad}_{\mathfrak{g}}(\mathfrak{h})$ is semisimple, we can write $\mathfrak{g} = \sum_i \mathfrak{g}_i$ (direct sum), where each $\mathfrak{g}_i$ is a one-dimensional subspace of $\mathfrak{g}$, invariant under $\mathrm{ad}_{\mathfrak{g}}(\mathfrak{h})$. This means that $\mathfrak{g}_i \subset \mathfrak{g}^{\alpha}$ for a suitable root α, and (i) follows. As a consequence of (i) we have

$$\text{If} \qquad \alpha(H_0) = 0 \qquad \text{for each } \alpha \in \Delta, \text{ then } H_0 = 0. \tag{2}$$

In fact, (i) shows that $[H_0, X] = 0$ for all $X \in \mathfrak{g}$; hence $H_0 = 0$, the center of $\mathfrak{g}$ being $\{0\}$. In order to prove (iii), select any $X \in \mathfrak{g}^{\alpha}$, $Y \in \mathfrak{g}^{\beta}$. Then ad X ad Y maps $\mathfrak{g}^{\gamma}$ into $\mathfrak{g}^{\gamma+\alpha+\beta}$; since $\alpha + \beta \neq 0$, it is clear that $\mathfrak{g}^{\gamma} \cap \mathfrak{g}^{\gamma+\alpha+\beta} = \{0\}$. Therefore, if the endomorphism ad X ad Y is expressed by means of a basis, each of whose elements lies in a root subspace $\mathfrak{g}^{\gamma}$, it is obvious that Tr (ad X ad Y) $= 0$. We next prove (iv). If $H_0 \in \mathfrak{h}$ satisfies $B(H_0, H) = 0$ for all $H \in \mathfrak{h}$, then by (iii), $B(H_0, X) = 0$ for all $X \in \mathfrak{g}$. Thus $H_0 = 0$. The latter part of (iv) is a consequence of the first. To prove (v) let $\alpha \in \Delta$. If $\mathfrak{g}^{-\alpha}$ were $\{0\}$, then by (iii) each $X_{\alpha} \in \mathfrak{g}^{\alpha}$

would satisfy $B(X_\alpha, X) = 0$ for all $X \in \mathfrak{g}$ which is impossible. Let H, X_α, $X_{-\alpha}$ be arbitrary elements in $\mathfrak{h}$, $\mathfrak{g}^\alpha$, $\mathfrak{g}^{-\alpha}$, respectively. Then

$$B([X_\alpha, X_{-\alpha}], H) = B(X_\alpha, [X_{-\alpha}, H]) = B(X_\alpha, X_{-\alpha})\, B(H_\alpha, H).$$

Using (iv) we obtain therefore

$$[X_\alpha, X_{-\alpha}] = B(X_\alpha, X_{-\alpha})\, H_\alpha. \tag{3}$$

Next consider any element $E_{-\alpha} \neq 0$ in $\mathfrak{g}^{-\alpha}$. Since B induces a nondegenerate bilinear form on $\mathfrak{g}^\alpha \times \mathfrak{g}^{-\alpha}$, there exists a vector $E_\alpha \in \mathfrak{g}^\alpha$ such that

$$B(E_\alpha, E_{-\alpha}) = +1.$$

Let β be any root and put $\mathfrak{g}^* = \sum_{n \in N} \mathfrak{g}^{\beta+n\alpha}$ where N is the set of all integers n for which $\beta + n\alpha$ is a root. Owing to (1), the subspace $\mathfrak{g}^*$ is invariant under $\operatorname{ad} E_{-\alpha}$, $\operatorname{ad} E_\alpha$, $\operatorname{ad} H_\alpha$. We compute the trace of $\operatorname{ad} H_\alpha$ in two ways. Since $[E_\alpha, E_{-\alpha}] = H_\alpha$ it follows that

$$\operatorname{Tr}_{\mathfrak{g}^*} \operatorname{ad} H_\alpha = -\operatorname{Tr}_{\mathfrak{g}^*} \operatorname{ad} E_{-\alpha} \operatorname{ad} E_\alpha + \operatorname{Tr}_{\mathfrak{g}^*} \operatorname{ad} E_\alpha \operatorname{ad} E_{-\alpha} = 0.$$

On the other hand, $\operatorname{ad} H_\alpha$ leaves each space $\mathfrak{g}^{\beta+n\alpha}$ invariant, so

$$\operatorname{Tr}_{\mathfrak{g}^*} \operatorname{ad} H_\alpha = \sum_{n \in N} (\beta + n\alpha)(H_\alpha) \dim \mathfrak{g}^{\beta+n\alpha}.$$

Thus we have the relation

$$\beta(H_\alpha) \sum_{n \in N} \dim \mathfrak{g}^{\beta+n\alpha} = -\alpha(H_\alpha) \sum_{n \in N} n \dim \mathfrak{g}^{\beta+n\alpha} \tag{4}$$

for each $\alpha \in \Delta$ and each root β.

Since $\dim \mathfrak{g}^\beta > 0$ we deduce from (2) and (4) that $\alpha(H_\alpha) \neq 0$. This proves (v). To prove (ii), suppose $\dim \mathfrak{g}^\alpha$ were > 1. Then (E_α and $E_{-\alpha}$ being as above), there exists a vector $D_\alpha \neq 0$ in $\mathfrak{g}^\alpha$ such that

$$B(D_\alpha, E_{-\alpha}) = 0.$$

We put $D_{-1} = 0$, $D_n = (\operatorname{ad} E_\alpha)^n D_\alpha$, $n = 0, 1, 2, \dots$. Then

$$[E_{-\alpha}, D_n] = -\frac{n(n+1)}{2}\, \alpha(H_\alpha)\, D_{n-1}, \qquad n = 0, 1, 2 \dots \tag{5}$$

For $n = 0$ this is clear from (3). Assuming (5) true for n, we have

$$[E_{-\alpha}, D_{n+1}] = [E_{-\alpha}, [E_\alpha, D_n]] = -[E_\alpha, [D_n, E_{-\alpha}]] - [D_n, [E_{-\alpha}, E_\alpha]]$$

$$= -\frac{n(n+1)}{2}\alpha(H_\alpha)[E_\alpha, D_{n-1}] + [D_n, H_\alpha]$$

$$= -\left(\frac{n(n+1)}{2} + (n+1)\right)\alpha(H_\alpha) D_n = -\frac{(n+1)(n+2)}{2}\alpha(H_\alpha) D_n,$$

where we have used the fact that $D_n \in \mathfrak{g}^{(n+1)\alpha}$. Since $D_0 \neq 0$, (5) shows that all $D_n \neq 0$, $n = 0, 1, 2 \ldots$ which is impossible. This proves (ii) so Theorem 4.2 is proved.

Let $\alpha \in \Delta$ and let β be any root. The α-*series* containing β is by definition the set of all roots of the form $\beta + n\alpha$ where n is an integer.

Theorem 4.3. *Let β be a root, and $\alpha \in \Delta$.*

(i) *The α-series containing β has the form $\beta + n\alpha$ $(p \leqslant n \leqslant q)$ (the α-series is an uninterrupted string). Also*

$$-2\frac{\beta(H_\alpha)}{\alpha(H_\alpha)} = p + q.$$

(ii) *Let $X_\alpha \in \mathfrak{g}^\alpha$, $X_{-\alpha} \in \mathfrak{g}^{-\alpha}$, $X_\beta \in \mathfrak{g}^\beta$ where $\beta \neq 0$. Then*

$$[X_{-\alpha}, [X_\alpha, X_\beta]] = \frac{q(1-p)}{2}\alpha(H_\alpha) B(X_\alpha, X_{-\alpha}) X_\beta.$$

(iii) *The only roots proportional to α are $-\alpha, 0, \alpha$.*

(iv) *Suppose $\alpha + \beta \neq 0$. Then $[\mathfrak{g}^\alpha, \mathfrak{g}^\beta] = \mathfrak{g}^{\alpha+\beta}$.*

Proof. Let $E_{-\alpha}$, E_α be any elements in $\mathfrak{g}^{-\alpha}$ and $\mathfrak{g}^\alpha$ respectively satisfying

$$B(E_\alpha, E_{-\alpha}) = 1.$$

To prove (i), let r, s be two integers such that $\beta + n\alpha$ is a root for $r \leqslant n \leqslant s$ but neither for $n = r - 1$ nor $n = s + 1$. Such a set $\beta + n\alpha$, $(r \leqslant n \leqslant s)$, we shall call a *maximal string*. The subspace

$$\mathfrak{g}^* = \sum_{n=r}^{s} \mathfrak{g}^{\beta+n\alpha}$$

is invariant under ad $E_{-\alpha}$, ad E_α, ad H_α. Since $H_\alpha = [E_\alpha, E_{-\alpha}]$ we have

$$\mathrm{Tr}_{\mathfrak{g}^*}(\mathrm{ad}\, H_\alpha) = 0.$$

On the other hand,

$$\mathrm{Tr}_{\mathfrak{g}^*}(\mathrm{ad}\, H_\alpha) = \sum_{n=r}^{s} (\beta + n\alpha)(H_\alpha)$$

and since $\alpha(H_\alpha) \neq 0$ we obtain

$$-2\frac{\beta(H_\alpha)}{\alpha(H_\alpha)} = r + s. \tag{6}$$

The α-series containing β is of course a union of maximal strings $\beta + n\alpha$, $r_i \leqslant n \leqslant s_i$. Since (6) holds for each maximal string, the α-series can only consist of one such string. This proves (i). For (iii) let $n\alpha$ $(p \leqslant n \leqslant q)$ be the α-series containing 0. The subspace

$$\mathfrak{s} = \sum_{n=-1}^{q} \mathfrak{g}^{n\alpha}$$

is then invariant under ad $E_{-\alpha}$, ad E_α, and ad H_α. We find as before

$$0 = \mathrm{Tr}_{\mathfrak{s}}\, \mathrm{ad}\, H_\alpha = \sum_{n=-1}^{q} n\alpha(H_\alpha),$$

which implies $q = 1$. Using Theorem 4.2 (v), we see that $p = -1$. Now suppose there were a complex number c which is not an integer such that $c\alpha$ is a root. Using (i) on the root $\beta = c\alpha$ we see that $c = n + \frac{1}{2}$ where n is an integer. The α-series containing β will also contain $-(n + \frac{1}{2})\,\alpha$, and since this series consists of just one string it must contain $-\frac{1}{2}\alpha$ and $\frac{1}{2}\alpha$. But since $\alpha = 2(\frac{1}{2}\alpha)$, this contradicts the first part of the proof. We next prove (iv). We have $[\mathfrak{g}^\alpha, \mathfrak{g}^\beta] \subset \mathfrak{g}^{\alpha+\beta}$ so (iv) is obvious if $\alpha + \beta$ is not a root. Suppose $\alpha + \beta$ is a root and $\beta + n\alpha$, $p \leqslant n \leqslant q$, is the α-series containing β. Then $q \geqslant 1$. If $[\mathfrak{g}^\alpha, \mathfrak{g}^\beta]$ were $\{0\}$, then the subspace

$$\mathfrak{t} = \sum_{n=p}^{0} \mathfrak{g}^{\beta+n\alpha}$$

would be invariant under ad E_α, ad $E_{-\alpha}$, ad H_α. We find as before

$$0 = \mathrm{Tr}_{\mathfrak{t}}\, \mathrm{ad}\, H_\alpha = \sum_{n=p}^{0} (\beta + n\alpha)(H_\alpha)$$

or $-2\beta(H_\alpha) = \alpha(H_\alpha)\,p$. Then (i) implies that $q = 0$ which is a contradiction. Finally we prove (ii). First observe that $\beta + p\alpha \neq 0$. Select a vector $E_p \neq 0$ in $\mathfrak{g}^{\beta+p\alpha}$, and put $E_n = (\operatorname{ad} E_\alpha)^{n-p} E_p$ for $n \geqslant p$. Then $E_n = 0$ for $n > q$; if $p \leqslant n \leqslant q$, then $E_n \neq 0$ as a consequence of (iv) and Theorem 4.2 (v). We shall now prove

$$[E_{-\alpha}, [E_\alpha, E_n]] = \frac{(q-n)(1-p+n)}{2} \alpha(H_\alpha) E_n \qquad (n \geqslant p). \tag{7}$$

Since X_β is a scalar multiple of E_0, (7) would imply (ii). We prove (7) by induction and consider first the case $n = p$. By the Jacobi identity

$$\begin{aligned}[E_{-\alpha}, [E_\alpha, E_p]] &= -[E_\alpha, [E_p, E_{-\alpha}]] - [E_p, [E_{-\alpha}, E_\alpha]] \\ &= 0 + [E_p, H_\alpha] = -(\beta + p\alpha)(H_\alpha) E_p,\end{aligned}$$

which by (i) equals $\frac{1}{2}(q-p)\,\alpha(H_\alpha) E_p$. Now assume (7); we have

$$[E_{-\alpha}, [E_\alpha, E_{n+1}]] = -[E_\alpha, [E_{n+1}, E_{-\alpha}]] - [E_{n+1}, [E_{-\alpha}, E_\alpha]]. \tag{8}$$

The first term on the right is $[E_\alpha, [E_{-\alpha}, [E_\alpha, E_n]]$ which by induction hypothesis equals

$$\frac{(q-n)(1-p+n)}{2} \alpha(H_\alpha) E_{n+1}.$$

The last term on the right-hand side of (8) is $-(\beta + (n+1)\alpha)(H_\alpha) E_{n+1}$ since $E_{n+1} \in \mathfrak{g}^{\beta+(n+1)\alpha}$. Using (i) we find that these two terms add up to

$$\begin{aligned}&\tfrac{1}{2}\alpha(H_\alpha) E_{n+1}\{(q-n)(1-p+n) + p + q - 2n - 2\} \\ &\qquad = \frac{(q-n-1)(n+2-p)}{2} \alpha(H_\alpha) E_{n+1}.\end{aligned}$$

This proves (7), and Theorem 4.3 is proved.

Theorem 4.4. *Let* $\mathfrak{h}^* = \sum_{\alpha\in\Delta} RH_\alpha$. *Then*
(i) *B is real and strictly positive definite on* $\mathfrak{h}^* \times \mathfrak{h}^*$.
(ii) $\mathfrak{h} = \mathfrak{h}^* + i\mathfrak{h}^*$ (*direct sum*).

Proof. We have for $H, H' \in \mathfrak{h}$

$$B(H, H') = \operatorname{Tr}(\operatorname{ad} H \operatorname{ad} H') = \sum_{\beta\in\Delta} \beta(H)\,\beta(H'). \tag{9}$$

From Theorem 4.3 (i) we know that

$$2\beta(H_\alpha) = -\alpha(H_\alpha)(p_{\beta,\alpha} + q_{\beta,\alpha}), \qquad p_{\beta,\alpha},\ q_{\beta,\alpha} \text{ integers},$$

so

$$\alpha(H_\alpha) = B(H_\alpha, H_\alpha) = \tfrac{1}{4}\alpha(H_\alpha)^2 \sum_{\beta\in\Delta} (p_{\beta,\alpha} + q_{\beta,\alpha})^2.$$

Since $\alpha(H_\alpha) \neq 0$ this shows that $\alpha(H_\alpha)$ is real (and positive) and $\beta(H)$ is real for each $H \in \mathfrak{h}^*$. Using (2) and (9), part (i) follows. Moreover, (i) shows that $\mathfrak{h}^* \cap i\mathfrak{h}^* = \{0\}$. Finally, the spaces $\mathfrak{h}$ and $\sum_{\alpha\in\Delta} CH_\alpha$ must coincide; in fact, suppose the contrary were the case. Then there exists a linear function λ on $\mathfrak{h}$ which is not identically 0 but vanishes on the subspace $\sum_{\alpha\in\Delta} CH_\alpha$. There exists a unique element $H_\lambda \in \mathfrak{h}$ such that $B(H, H_\lambda) = \lambda(H)$ for all $H \in \mathfrak{h}$. In particular, $\alpha(H_\lambda) = 0$ for all $\alpha \in \Delta$, so by (2), $H_\lambda = 0$ and $\lambda \equiv 0$. This contradiction proves (ii).

§ 5. Significance of the Root Pattern

To a semisimple Lie algebra $\mathfrak{g}$ and a Cartan subalgebra $\mathfrak{h}$ we have associated a system of vectors H_α, $\alpha \in \Delta$. We shall now see that this system determines $\mathfrak{g}$ up to isomorphism.

Let S be any subset of Δ. The *hull* of S, denoted $\bar{S}$, is by definition the set of all roots in Δ of the form $\pm\alpha$, $\pm(\alpha+\beta)$ where α, β run through S. For each pair $\alpha, -\alpha \in \bar{S}$ we select vectors $E_\alpha \in \mathfrak{g}^\alpha$, $E_{-\alpha} \in \mathfrak{g}^{-\alpha}$ such that

$$B(E_\alpha, E_{-\alpha}) = 1 \qquad \text{for } \alpha, -\alpha \in \bar{S}. \tag{1}$$

Let α, β be any elements in $\bar{S}$ such that $\alpha + \beta \neq 0$ and such that either $\alpha + \beta \in \bar{S}$ or $\alpha + \beta \notin \Delta$. We define the number $N_{\alpha,\beta}$ by

$$[E_\alpha, E_\beta] = N_{\alpha,\beta} E_{\alpha+\beta} \qquad \text{if } \alpha+\beta \in \bar{S},$$
$$N_{\alpha,\beta} = 0 \qquad \text{if } \alpha+\beta \notin \Delta.$$

Thus $N_{\alpha,\beta}$ is defined under the conditions:

(a) $\alpha, \beta \in \bar{S}$; (b) $\alpha + \beta \neq 0$; (c) $\alpha+\beta \in \bar{S}$, or $\alpha + \beta \notin \Delta$.

We have obviously $N_{\alpha,\beta} = -N_{\beta,\alpha}$.

Lemma 5.1. *Suppose $\alpha, \beta, \gamma \in \bar{S}$ and $\alpha + \beta + \gamma = 0$. Then*

$$N_{\alpha,\beta} = N_{\beta,\gamma} = N_{\gamma,\alpha} \qquad (\textit{all } N \textit{ are defined}).$$

Proof. We use the Jacobi identity on the vectors E_α, E_β, E_γ. In view of (3), §4, we obtain

$$N_{\beta,\gamma}H_\alpha + N_{\gamma,\alpha}H_\beta + N_{\alpha,\beta}H_\gamma = 0.$$

On the other hand, $-H_\gamma = H_\alpha + H_\beta$ and β is not proportional to α. The lemma follows.

Lemma 5.2. *Suppose* $\alpha, \beta \in S$, $\alpha + \beta \in \Delta$. *Let* $\beta + n\alpha$ $(p \leqslant n \leqslant q)$ *denote the* α*-series containing* β. *Then*

$$N_{\alpha,\beta}N_{-\alpha,-\beta} = -\frac{q(1-p)}{2}\alpha(H_\alpha).$$

Proof. Since $-\alpha + (\alpha + \beta) = \beta$, $N_{-\alpha,\alpha+\beta}$ is defined; using Theorem 4.3 (ii), we have

$$N_{\alpha,\beta}N_{-\alpha;\alpha+\beta} = \frac{q(1-p)}{2}\alpha(H_\alpha).$$

From Lemma 5.1 we have

$$N_{-\alpha,\alpha+\beta} = N_{\alpha+\beta,-\beta} = N_{-\beta,-\alpha}$$

and Lemma 5.2 follows.

Lemma 5.3. *Suppose* $\alpha, \beta, \gamma, \delta$ *are four roots in* S (*not necessarily distinct*), *no two of which have sum* 0. *If*

$$\alpha + \beta + \gamma + \delta = 0,$$

then

$$N_{\alpha,\beta}N_{\gamma,\delta} + N_{\beta,\gamma}N_{\alpha,\delta} + N_{\gamma,\alpha}N_{\beta,\delta} = 0 \qquad (\textit{all } N \textit{ are defined}).$$

Proof. Suppose first that $\beta + \gamma$ is a root. Then $\beta + \gamma \in S$ and since $\alpha + (\beta + \gamma) = -\delta$, $N_{\alpha,\beta+\gamma}$ is defined. In this case, we have

$$[E_\alpha, [E_\beta, E_\gamma]] = N_{\beta,\gamma}N_{\alpha,\beta+\gamma}E_{-\delta}.$$

Applying Lemma 5.1 to α, $\beta + \gamma$, δ, we have $N_{\alpha,\beta+\gamma} = N_{\delta,\alpha}$ and therefore

$$[E_\alpha, [E_\beta, E_\gamma]] = -N_{\beta,\gamma}N_{\alpha,\delta}E_{-\delta}. \tag{2}$$

This relation holds also if $\beta + \gamma$ is not a root because then both sides are 0. In (2) we permute the letters α, β, γ cyclically and use Jacobi's identity. Since $E_{-\delta} \neq 0$, Lemma 5.3 follows.

A set M is said to be *ordered* (or totally ordered) by means of a relation $<$ if to any two elements a, b in the set exactly one of the following relations holds: $a < b$, $b < a$ or $a = b$. Moreover, it is assumed that whenever $a < b$ and $b < c$, then $a < c$. The relation $a > b$ is to mean the same as $b < a$. If the relation $a < b$ (or $a > b$) is defined only for certain pairs in M, the set M is called *partially ordered.*

Let V be a finite-dimensional vector space over $\boldsymbol{R}$; V is said to be an *ordered vector space* if it is an ordered set and the ordering relation $<$ satisfies the conditions: (1) $X > 0$, $Y > 0$ implies $X + Y > 0$.

(2) If $X > 0$ and a is a positive real number, then $aX > 0$. Note that (1) implies: $X > 0$ if and only if $-X < 0$. If $X_1, ..., X_n$ is a basis of V then V can be turned into an ordered vector space as follows: Let $X, Y \in V$. We say $X > Y$ if $X - Y = \sum_{i=1}^{n} a_i X_i$ and the first nonzero number in the sequence $a_1, a_2, ..., a_n$ is > 0. It is readily verified that V is an ordered vector space with this ordering, which is called the *lexicographic ordering* of V with respect to the basis $X_1, ..., X_n$. Let $V^\wedge$ denote the dual space of V. Let $\lambda, \mu \in V^\wedge$. We say $\lambda > \mu$ if the first nonzero number in the sequence $\lambda(X_1) - \mu(X_1), ..., \lambda(X_n) - \mu(X_n)$ is positive. This ordering of $V^\wedge$ is called the lexicographic ordering with respect to the basis $X_1, ..., X_n$ of $V(= (V^\wedge)^\wedge)$. The element $\lambda \in V^\wedge$ is called positive if $\lambda > 0$.

Let V and W be two vector spaces over $\boldsymbol{R}$ and $V^\wedge$ and $W^\wedge$ their duals. If φ is a linear mapping of V into W then ${}^t\varphi$ shall denote the dual mapping of $W^\wedge$ into $V^\wedge$ which is determined by

$$({}^t\varphi)(F)(v) = F(\varphi(v)) \qquad \text{if } v \in V, F \in W^\wedge.$$

The following important theorem shows that a semisimple Lie algebra over $\boldsymbol{C}$ is determined (up to isomorphism) by means of a Cartan subalgebra and the corresponding pattern of roots.

Theorem 5.4. *Let $\mathfrak{g}$ and $\mathfrak{g}'$ be two semisimple Lie algebras, $\mathfrak{h}$ and $\mathfrak{h}'$ Cartan subalgebras of $\mathfrak{g}$ and $\mathfrak{g}'$, respectively. Let Δ and Δ' denote the corresponding sets of nonzero roots and as usual let*

$$\mathfrak{h}^* = \sum_{\alpha \in \Delta} \boldsymbol{R} H_\alpha, \qquad (\mathfrak{h}')^* = \sum_{\alpha' \in \Delta'} \boldsymbol{R} H_{\alpha'}.$$

Then Δ and Δ' can be considered as subsets of the dual space of $\mathfrak{h}^*$ and $(\mathfrak{h}')^*$, respectively, since each $\beta \in \Delta$ ($\beta' \in \Delta'$) is real on $\mathfrak{h}^*$ ($(\mathfrak{h}')^*$).

Suppose φ is a one-to-one $\boldsymbol{R}$-linear mapping of $\mathfrak{h}^*$ onto $(\mathfrak{h}')^*$ such that ${}^t\varphi$ maps Δ' into Δ. Then φ can be extended to an isomorphism $\tilde{\varphi}$ of $\mathfrak{g}$ onto $\mathfrak{g}'$.

Proof. The Killing forms B and B' induce positive definite metrics on $\mathfrak{h}^*$ and $(\mathfrak{h}')^*$. We shall first prove that φ is an isometry, that is,

$$B(H_1, H_2) = B'(\varphi H_1, \varphi H_2) \qquad \text{for } H_1, H_2 \in \mathfrak{h}^*.$$

It suffices to prove that

$$B(H_\alpha, H_\beta) = B'(H_{\alpha'}, H_{\beta'})$$

for all $\alpha, \beta \in \Delta$, where $\alpha = {}^t\varphi \cdot \alpha'$, $\beta = {}^t\varphi \cdot \beta'$. Using Theorem 4.3 (i), we obtain

$$\frac{\beta(H_\alpha)}{\alpha(H_\alpha)} = \frac{\beta'(H_{\alpha'})}{\alpha'(H_{\alpha'})}$$

for all $\alpha, \beta \in \Delta$. Interchanging α and β we find that

$$B(H_\alpha, H_\beta) = cB(H_{\alpha'}, H_{\beta'}), \qquad \alpha, \beta \in \Delta,$$

where c is a constant $\neq 0$, independent of α and β. Since

$$B(H_\alpha, H_\beta) = \sum_{\gamma \in \Delta} \gamma(H_\alpha)\,\gamma(H_\beta) = c^2 \sum_{\gamma' \in \Delta'} \gamma'(H_{\alpha'})\,\gamma'(H_{\beta'}) = c^2 B(H_{\alpha'}, H_{\beta'})$$

we see that $c^2 = c$ so $c = 1$ as stated.

Now given any basis in $\mathfrak{h}^*$ we can introduce a lexicographic ordering in the dual space of $\mathfrak{h}^*$. Thus Δ becomes an ordered set and we shall prove Theorem 5.4 by induction with respect to this ordering. We select for each $\alpha \in \Delta$ an element $E_\alpha \in \mathfrak{g}^\alpha$ such that (1) holds for all $\alpha \in \Delta$. The numbers $N_{\alpha,\beta}$ are then defined for all $\alpha, \beta \in \Delta$ for which $\alpha + \beta \neq 0$.

We shall show that to each $\alpha' \in \Delta'$ there exists an element $E_{\alpha'} \in \mathfrak{g}'^{(\alpha')}$ such that

$$B'(E_{\alpha'}, E_{-\alpha'}) = 1, \qquad \alpha \in \Delta, \tag{3}$$

and such that

$$[E_{\alpha'}, E_{\beta'}] = N_{\alpha,\beta} E_{\alpha'+\beta'} \qquad \text{if } \alpha, \beta, \alpha + \beta \in \Delta. \tag{4}$$

Let $\rho \in \Delta$ be a positive root and let Δ_ρ denote the set of roots $\alpha \in \Delta$ satisfying $-\rho < \alpha < \rho$. If there exists a root greater than ρ, then ρ^* will denote the smallest such root.

The induction hypothesis is that for each $\alpha \in \Delta_\rho$, the vector $E_{\alpha'}$ can be chosen in $\mathfrak{g}'^{(\alpha')}$ in such a way that (3) is satisfied for all $\alpha \in \Delta_\rho$ and (4) is satisfied if $\alpha, \beta, \alpha + \beta \in \Delta_\rho$. We shall then define $E_{\rho'}$ and $E_{-\rho'}$ in such a way that (3) is satisfied for all $\alpha \in \Delta_{\rho^*}$ and (4) is satisfied if $\alpha, \beta, \alpha + \beta \in \Delta_{\rho^*}$. If ρ^* does not exist, Δ_{ρ^*} is to mean Δ itself.

If ρ has no decomposition $\rho = \alpha + \beta$ with $\alpha, \beta \in \Delta_\rho$ then we just have to take for $E_{\rho'}$ an arbitrary nonzero element in $\mathfrak{g}'^{(\rho')}$ and then fix $E_{-\rho'}$ by the relation

$$B'(E_{\rho'}, E_{-\rho'}) = 1. \tag{5}$$

If ρ has a decomposition $\rho = \alpha + \beta$, $\alpha, \beta \in \Delta_\rho$ we select the particular decomposition for which α is as small as possible and define $E_{\rho'}$ by means of the equation

$$[E_{\alpha'}, E_{\beta'}] = N_{\alpha,\beta} E_{\rho'}. \tag{6}$$

Then $E_{\rho'} \neq 0$ and we can again define $E_{-\rho'}$ by (5). In order to prove that (4) holds for all $\alpha, \beta, \alpha + \beta \in \Delta_{\rho^*}$, we define the numbers $M_{\gamma,\delta}$ by means of the relation

$$[E_{\gamma'}, E_{\delta'}] = M_{\gamma,\delta} E_{\gamma'+\delta'} \qquad \text{if } \gamma, \delta, \gamma + \delta \in \Delta_{\rho^*}. \tag{7}$$

We also put $M_{\gamma,\delta} = 0$ if γ, δ satisfy the conditions $\gamma, \delta \in \Delta_{\rho^*}, \gamma + \delta \neq 0$, $\gamma + \delta \notin \Delta$. We shall prove that $N_{\gamma,\delta} = M_{\gamma,\delta}$ whenever $\gamma, \delta, \gamma + \delta \in \Delta_{\rho^*}$. We have to consider various possibilities:

1. $\gamma, \delta, \gamma + \delta \in \Delta_\rho$. Then $N_{\gamma,\delta} = M_{\gamma,\delta}$ by the induction hypothesis.

2. $\gamma + \delta = \rho$. Then $\gamma, \delta \in \Delta_\rho$. We can assume that the decomposition $\rho = \gamma + \delta$ differs from the decomposition $\rho = \alpha + \beta$. Then the roots $\alpha, \beta, -\gamma, -\delta$ satisfy the relation $\alpha + \beta + (-\gamma) + (-\delta) = 0$ and no two of these roots have sum 0. We can apply Lemmas 5.2 and 5.3 to $\mathfrak{g}$ for $S = \Delta$. We obtain

$$N_{\alpha,\beta} N_{-\gamma,-\delta} = -N_{\beta,-\gamma} N_{\alpha,-\delta} - N_{-\gamma,\alpha} N_{\beta,-\delta}, \tag{8}$$

$$N_{\gamma,\delta} N_{-\gamma,-\delta} = -\frac{l(1-k)}{2} \gamma(H_\gamma), \tag{8'}$$

if $\delta + n\gamma$ $(k \leqslant n \leqslant l)$ is the γ-series containing δ. We can also apply Lemmas 5.2 and 5.3 to $\mathfrak{g}'$ by taking for S the set of roots $\alpha', \beta', -\gamma', -\delta'$. To see that the lemmas can be applied, we note that ${}^t\varphi(\bar{S}) \subset \Delta_{\rho^*}$ so $E_{\mu'}$ is defined for each $\mu' \in \bar{S}$; also $M_{\mu,\nu}$ is defined under the required conditions (a) $\mu', \nu' \in \bar{S}$, (b) $\mu' + \nu' \neq 0$, (c) $\mu' + \nu' \in \bar{S}$ or $\mu' + \nu' \notin \Delta'$. We obtain

$$M_{\alpha,\beta} M_{-\gamma,-\delta} = -M_{\beta,-\gamma} M_{\alpha,-\delta} - M_{-\gamma,\alpha} M_{\beta,-\delta}, \tag{9}$$

$$M_{-\gamma,-\delta} M_{\gamma,\delta} = -\frac{l(1-k)}{2} (-\gamma'(H_{-\gamma'})), \tag{9'}$$

because $-\delta + n(-\gamma)$, $(k \leqslant n \leqslant l)$, is the $(-\gamma)$-series containing $-\delta$. From **1** we know that the right-hand sides of (9) and (8) are the same. From (6) and (7) we have $M_{\alpha,\beta} = N_{\alpha,\beta} \neq 0$. It follows that $N_{-\gamma,-\delta} = M_{-\gamma,-\delta}$. By the first part of the proof, $\gamma(H_\gamma) = \gamma'(H_{\gamma'})$ so by (8′) and (9′), $N_{\gamma,\delta} = M_{\gamma,\delta}$ as we wished to prove.

3. $\gamma + \delta = -\rho$. Then $(-\gamma) + (-\delta) = \rho$, $(-\gamma), (-\delta) \in \Delta_\rho$. By **2**, we have $N_{\gamma,\delta} = M_{\gamma,\delta}$.

4. One of the roots γ, δ equals $\pm \rho$. Suppose, for example, $\gamma = -\rho$. Then $\delta \neq \pm \rho$ and $\rho = \delta + (-\gamma - \delta)$ where $\delta, -\gamma - \delta \in \Delta_\rho$. Using **1** we have $N_{\delta,-\gamma-\delta} = M_{\delta,-\gamma-\delta}$. We apply Lemma 5.1 to $\mathfrak{g}$ for $S = \Delta$, and get

$$N_{\delta,-\gamma-\delta} = N_{-\gamma-\delta,\gamma} = N_{\gamma,\delta}.$$

Applying Lemma 5.1 to $\mathfrak{g}'$ for $S = \{\delta', -\gamma' - \delta', \gamma'\}$ we obtain (since ${}^t\varphi(\bar{S}) \subset \Delta_{\rho*}$)

$$M_{\delta,-\gamma-\delta} = M_{-\gamma-\delta,\gamma} = M_{\gamma,\delta}$$

and $N_{\gamma,\delta} = M_{\gamma,\delta}$ follows. This proves relations (3) and (4).

Consider now the linear mapping $\tilde{\varphi} : \mathfrak{g} \to \mathfrak{g}'$ determined by

$$\tilde{\varphi}(H_\alpha) = \varphi(H_\alpha) = H_{\alpha'},$$
$$\tilde{\varphi}(E_\alpha) = E_{\alpha'}, \qquad \alpha \in \Delta.$$

Relations (3) and (4) then show that $\tilde{\varphi}$ is an isomorphism of $\mathfrak{g}$ onto $\mathfrak{g}'$.

Remark 1. The extension $\tilde{\varphi}$ is not in general unique; in fact, if $H \in \mathfrak{h}$, then $e^{\operatorname{ad} H}$ is an automorphism of $\mathfrak{g}$ whose restriction to $\mathfrak{h}$ is the identity mapping.

Remark 2. In the proof above the ordering of Δ was introduced in order to carry out an induction process. However, the ordering of Δ plays an important role in other contexts which will come up later. Therefore, it should be borne in mind that "ordering of Δ" shall always mean the ordering of the set Δ induced by some vector space ordering of the dual space of $\mathfrak{h}^*$.

Theorem 5.5. *For each $\alpha \in \Delta$ a vector $X_\alpha \in \mathfrak{g}^\alpha$ can be chosen such that for all $\alpha, \beta \in \Delta$*

$$[X_\alpha, X_{-\alpha}] = H_\alpha, \quad [H, X_\alpha] = \alpha(H)\, X_\alpha \quad \textit{for } H \in \mathfrak{h};$$
$$[X_\alpha, X_\beta] = 0 \qquad \textit{if } \alpha + \beta \neq 0 \textit{ and } \alpha + \beta \notin \Delta;$$
$$[X_\alpha, X_\beta] = N_{\alpha,\beta} X_{\alpha+\beta} \qquad \textit{if } \alpha + \beta \in \Delta,$$

where the constants $N_{\alpha,\beta}$ satisfy

$$N_{\alpha,\beta} = - N_{-\alpha,-\beta}.$$

For any such choice

$$N^2_{\alpha,\beta} = \frac{q(1-p)}{2}\, \alpha(H_\alpha),$$

where $\beta + n\alpha$ ($p \leqslant n \leqslant q$) is the α-series containing β.

Proof. Let φ denote the mapping $H \to -H$ of $\mathfrak{h}^*$ onto itself. Then $({}^t\varphi)(\lambda) = -\lambda$ for each real linear form λ on $\mathfrak{h}^*$. In particular, ${}^t\varphi$ maps the set Δ onto itself so from Theorem 5.4 we know that φ can be extended to an automorphism $\tilde{\varphi}$ of $\mathfrak{g}$. For each $\alpha \in \Delta$ we select $E_\alpha \in \mathfrak{g}^\alpha$ such that

$$B(E_\alpha, E_{-\alpha}) = 1 \qquad \text{for all } \alpha \in \Delta.$$

Since $\tilde{\varphi}(E_\alpha) \in \mathfrak{g}^{-\alpha}$ and $\dim \mathfrak{g}^{-\alpha} = 1$ there exists a complex number $c_{-\alpha}$ such that $\tilde{\varphi}(E_\alpha) = c_{-\alpha}E_{-\alpha}$. Since B is invariant under $\tilde{\varphi}$ we have $c_\alpha c_{-\alpha} = 1$. For each $\alpha \in \Delta$ one can select a number a_α such that

$$a_\alpha^2 = -c_\alpha, \qquad a_\alpha a_{-\alpha} = +1 \qquad \text{for } \alpha \in \Delta.$$

We put now

$$X_\alpha = a_\alpha E_\alpha, \qquad \alpha \in \Delta.$$

By (3), §4,

$$[X_\alpha, X_{-\alpha}] = B(X_\alpha, X_{-\alpha})\, H_\alpha = a_\alpha a_{-\alpha} B(E_\alpha, E_{-\alpha})\, H_\alpha = H_\alpha.$$

Also,

$$\tilde{\varphi}(X_\alpha) = a_\alpha \tilde{\varphi}(E_\alpha) = a_\alpha c_{-\alpha} E_{-\alpha} = -a_{-\alpha}E_{-\alpha} = -X_{-\alpha}.$$

If $\alpha, \beta, \alpha + \beta \in \Delta$ we define $N_{\alpha,\beta}$ by $[X_\alpha, X_\beta] = N_{\alpha,\beta}X_{\alpha+\beta}$. Then

$$-N_{\alpha,\beta}X_{-\alpha-\beta} = \tilde{\varphi}(N_{\alpha,\beta}X_{\alpha+\beta}) = \tilde{\varphi}[X_\alpha, X_\beta] = [-X_{-\alpha}, -X_{-\beta}] = N_{-\alpha,-\beta}X_{-\alpha-\beta},$$

so $N_{\alpha,\beta} = -N_{-\alpha,-\beta}$. The last relation of the theorem now follows from Lemma 5.2.

§ 6. Real Forms

Let V be a vector space over $\boldsymbol{R}$ of finite dimension. A *complex structure* on V is an $\boldsymbol{R}$-linear endomorphism J of V such that $J^2 = -I$, where I is the identity mapping of V. A vector space V over $\boldsymbol{R}$ with a complex structure J can be turned into a vector space $\tilde{V}$ over $\boldsymbol{C}$ by putting

$$(a + ib)\, X = aX + bJX, \qquad X \in V,\ a, b \in \boldsymbol{R}.$$

In fact, $J^2 = -I$ implies $\alpha(\beta X) = (\alpha\beta)\, X$ for $\alpha, \beta \in \boldsymbol{C}$ and $X \in V$. We have clearly $\dim_{\boldsymbol{C}} \tilde{V} = \frac{1}{2} \dim_{\boldsymbol{R}} V$ and consequently V must be even-dimensional. We call $\tilde{V}$ the *complex vector space associated to* V. Note that V and $\tilde{V}$ agree set theoretically.

On the other hand, if E is a vector space over $\boldsymbol{C}$ we can consider E as a vector space $E^{\boldsymbol{R}}$ over $\boldsymbol{R}$. The multiplication by i on E then becomes a complex structure J on $E^{\boldsymbol{R}}$ and it is clear that $E = (E^{\boldsymbol{R}})^{\sim}$.

A Lie algebra $\mathfrak{v}$ over $\boldsymbol{R}$ is said to have a complex structure J if J is a complex structure on the vector space $\mathfrak{v}$ and in addition

$$[X, JY] = J[X, Y], \qquad \text{for } X, Y \in \mathfrak{v}. \tag{1}$$

Condition (1) means $(\operatorname{ad} X) \circ J = J \circ \operatorname{ad} X$ for all $X \in \mathfrak{v}$, or equivalently, $\operatorname{ad}(JX) = J \circ \operatorname{ad} X$ for all $X \in \mathfrak{v}$. It follows from (1) that

$$[JX, JY] = -[X, Y].$$

The complex vector space $\tilde{\mathfrak{v}}$ becomes a Lie algebra over $\boldsymbol{C}$ with the bracket operation inherited from $\mathfrak{v}$. In fact

$$\begin{aligned}[(a+ib)X, (c+id)Y] &= [aX + bJX, cY + dJY] \\ &= ac[X, Y] + bcJ[X, Y] + adJ[X, Y] - bd[X, Y]\end{aligned}$$

so

$$[(a+ib)X, (c+id)Y] = (a+ib)(c+id)[X, Y].$$

On the other hand, suppose $\mathfrak{e}$ is a Lie algebra over $\boldsymbol{C}$. The vector space $\mathfrak{e}^R$ has a complex structure J given by multiplication by i on $\mathfrak{e}$. With the bracket operation inherited from $\mathfrak{e}$, $\mathfrak{e}^R$ becomes a Lie algebra over $\boldsymbol{R}$ with the complex structure J.

Now suppose W is an arbitrary finite-dimensional vector space over $\boldsymbol{R}$. The product $W \times W$ is again a vector space over $\boldsymbol{R}$ and the endomorphism $J: (X, Y) \to (-Y, X)$ is a complex structure on $W \times W$. The complex vector space $(W \times W)^\sim$ is called the *complexification* of W and will be denoted W^C. We have of course $\dim_{\mathbf{C}} W^C = \dim_R W$. The elements of W^C are the pairs (X, Y) where $X, Y \in W$ and since $(X, Y) = (X, 0) + i(Y, 0)$ we write $X + iY$ instead of (X, Y). Then since

$$(a + bJ)(X, Y) = a(X, Y) + b(-Y, X) = (aX - bY, aY + bX)$$

we have

$$(a + ib)(X + iY) = aX - bY + i(aY + bX).$$

On the other hand, each finite-dimensional vector space E over $\boldsymbol{C}$ is isomorphic to W^C for a suitable vector space W over $\boldsymbol{R}$; in fact, if (Z_i) is any basis of E, one can take W as the set of all vectors of the form $\sum_i a_i Z_i$, $a_i \in \boldsymbol{R}$.

Let $\mathfrak{l}_0$ be a Lie algebra over $\boldsymbol{R}$; owing to the conventions above, the complex vector space $\mathfrak{l} = (\mathfrak{l}_0)^C$ consists of all symbols $X + iY$, where $X, Y \in \mathfrak{l}_0$. We define the bracket operation in $\mathfrak{l}$ by

$$[X + iY, Z + iT] = [X, Z] - [Y, T] + i([Y, Z] + [X, T]),$$

and this bracket operation is bilinear over C. It is clear that $\mathfrak{l} = (\mathfrak{l}_0)^C$ is a Lie algebra over C; it is called the *complexification of the Lie algebra* $\mathfrak{l}_0$. The Lie algebra $\mathfrak{l}^R$ is a Lie algebra over R with a complex structure J derived from multiplication by i on $\mathfrak{l}$.

Lemma 6.1. *Let K_0, K, and K^R denote the Killing forms of the Lie algebras $\mathfrak{l}_0$, $\mathfrak{l}$, and $\mathfrak{l}^R$. Then*

$$K_0(X, Y) = K(X, Y) \qquad \textit{for } X, Y \in \mathfrak{l}_0,$$

$$K^R(X, Y) = 2 \operatorname{Re}(K(X, Y)) \qquad \textit{for } X, Y \in \mathfrak{l}^R \qquad (\operatorname{Re} = \textit{real part}).$$

The first relation is obvious. For the second, suppose X_i $(1 \leqslant i \leqslant n)$ is any basis of $\mathfrak{l}$; let $B + iC$ denote the matrix of ad X ad Y with respect to this basis, B and C being real. Then $X_1, \ldots, X_n, JX_1, \ldots, JX_n$ is a basis of $\mathfrak{l}^R$ and since the linear transformation ad X ad Y of $\mathfrak{l}^R$ commutes with J, it has the matrix expression

$$\begin{pmatrix} B & -C \\ C & B \end{pmatrix}$$

and the second relation above follows.

As a consequence of Lemma 6.1 we note that the algebras $\mathfrak{l}_0$, $\mathfrak{l}$, and $\mathfrak{l}^R$ are all semisimple if and only if one of them is.

Definition. Let $\mathfrak{g}$ be a Lie algebra over C. A *real form* of $\mathfrak{g}$ is a subalgebra $\mathfrak{g}_0$ of the real Lie algebra $\mathfrak{g}^R$ such that

$$\mathfrak{g}^R = \mathfrak{g}_0 + J\mathfrak{g}_0 \qquad \text{(direct sum of vector spaces).}$$

In this case, each $Z \in \mathfrak{g}$ can be uniquely written

$$Z = X + iY, \qquad X, Y \in \mathfrak{g}_0.$$

Thus $\mathfrak{g}$ is isomorphic to the complexification of $\mathfrak{g}_0$. The mapping σ of $\mathfrak{g}$ onto itself given by $\sigma : X + iY \to X - iY$ $(X, Y \in \mathfrak{g}_0)$ is called the *conjugation* of $\mathfrak{g}$ with respect to $\mathfrak{g}_0$. The mapping σ has the properties

$$\sigma(\sigma(X)) = X, \qquad \sigma(X + Y) = \sigma(X) + \sigma(Y),$$

$$\sigma(\alpha X) = \bar{\alpha} X, \qquad \sigma[X, Y] = [\sigma X, \sigma Y],$$

for $X, Y \in \mathfrak{g}$, $\alpha \in C$. Thus σ is not an automorphism of $\mathfrak{g}$, but it is an automorphism of the real algebra $\mathfrak{g}^R$. On the other hand, let σ be a mapping of $\mathfrak{g}$ onto itself with the properties above. Then the set $\mathfrak{g}_0$ of fixed points of σ is a real form of $\mathfrak{g}$ and σ is the conjugation of $\mathfrak{g}$ with respect to $\mathfrak{g}_0$. In fact, $J\mathfrak{g}_0$ is the eigenspace of σ for the eigenvalue -1 and consequently $\mathfrak{g}^R = \mathfrak{g}_0 + J\mathfrak{g}_0$. If B is the Killing form on $\mathfrak{g} \times \mathfrak{g}$, it is easy to see from Lemma 6.1 that $B(\sigma X, \sigma Y)$ is the complex conjugate of $B(X, Y)$. Another useful remark in this connection is the following: Let $\mathfrak{g}_1$ and $\mathfrak{g}_2$ be two real forms of $\mathfrak{g}$ and σ_1 and σ_2 the corresponding conjugations. Then σ_1 leaves $\mathfrak{g}_2$ invariant if and only if σ_1 and σ_2 commute; in this case we have the direct decompositions

$$\mathfrak{g}_1 = \mathfrak{g}_1 \cap \mathfrak{g}_2 + \mathfrak{g}_1 \cap (i\mathfrak{g}_2),$$
$$\mathfrak{g}_2 = \mathfrak{g}_1 \cap \mathfrak{g}_2 + \mathfrak{g}_2 \cap (i\mathfrak{g}_1).$$

Lemma 6.2. *Suppose $\mathfrak{g}$ is a semisimple Lie algebra over C, $\mathfrak{g}_0$ a real form of $\mathfrak{g}$, and σ the conjugation of $\mathfrak{g}$ with respect to $\mathfrak{g}_0$. Let ad denote the adjoint representation of $\mathfrak{g}^R$ and Int $(\mathfrak{g}^R)$ the adjoint group of $\mathfrak{g}^R$. If G_0 denotes the analytic subgroup of Int $(\mathfrak{g}^R)$ whose Lie algebra is ad $(\mathfrak{g}_0)$, then G_0 is a closed subgroup of Int $(\mathfrak{g}^R)$ and analytically isomorphic to Int $(\mathfrak{g}_0)$.*

Proof. Every automorphism s of $\mathfrak{g}^R$ gives rise to an automorphism $\tilde{s}$ of Int $(\mathfrak{g}^R)$ satisfying $\tilde{s}(e^{\mathrm{ad}X}) = e^{\mathrm{ad}(s \cdot X)}$, $(X \in \mathfrak{g}^R)$. In particular there exists an automorphism $\tilde{\sigma}$ of Int $(\mathfrak{g}^R)$ such that $(d\tilde{\sigma})_e$ (ad X) = ad $(\sigma \cdot X)$ for $X \in \mathfrak{g}^R$. Since ad is an isomorphism, this proves that ad $(\mathfrak{g}_0)$ is the set of fixed points of $(d\tilde{\sigma})_e$; thus G_0 is the identity component of the set of fixed points of $\tilde{\sigma}$. Now, let ad_0 denote the adjoint representation of $\mathfrak{g}_0$ and for each endomorphism A of $\mathfrak{g}^R$ leaving $\mathfrak{g}_0$ invariant, let A_0 denote its restriction to $\mathfrak{g}_0$. Then if $X \in \mathfrak{g}_0$, we have $(\mathrm{ad}\, X)_0 = \mathrm{ad}_0\, X$ and the mapping $A \to A_0$ maps G_0 onto Int $(\mathfrak{g}_0)$. This mapping is an isomorphism of G_0 onto Int $(\mathfrak{g}_0)$. In fact, suppose $A \in G_0$ such that A_0 is the identity. Since A commutes with the complex structure J, it follows that A is the identity. Finally since the isomorphism is regular at the identity it is an analytic isomorphism.

The following theorem is of fundamental importance in the theory of semisimple Lie algebras and symmetric spaces.

Theorem 6.3. *Every semisimple Lie algebra $\mathfrak{g}$ over C has a real form which is compact.*

Proof. As always, let B denote the Killing form on $\mathfrak{g} \times \mathfrak{g}$. Let $\mathfrak{h}$ be a Cartan subalgebra of $\mathfrak{g}$, and Δ the corresponding set of nonzero roots. For each $\alpha \in \Delta$ we select $X_\alpha \in \mathfrak{g}^\alpha$ with the properties of Theorem 5.5. The first relation $[X_\alpha, X_{-\alpha}] = H_\alpha$ implies $B(X_\alpha, X_{-\alpha}) = 1$ by (3), § 4, and consequently

$$B(X_\alpha - X_{-\alpha}, X_\alpha - X_{-\alpha}) = -2,$$
$$B(i(X_\alpha + X_{-\alpha}), i(X_\alpha + X_{-\alpha})) = -2,$$
$$B(X_\alpha - X_{-\alpha}, i(X_\alpha + X_{-\alpha})) = 0,$$
$$B(iH_\alpha, iH_\alpha) = -\alpha(H_\alpha) < 0,$$

the last relation following from Theorem 4.4. Since $B(X_\alpha, X_\beta) = 0$ if $\alpha + \beta \neq 0$, it follows that B is strictly negative definite on the $\boldsymbol{R}$-linear subspace

$$\mathfrak{g}_k = \sum_{\alpha\in\Delta} \boldsymbol{R}(iH_\alpha) + \sum_{\alpha\in\Delta} \boldsymbol{R}(X_\alpha - X_{-\alpha}) + \sum_{\alpha\in\Delta} \boldsymbol{R}(i(X_\alpha + X_{-\alpha})). \tag{2}$$

Moreover, $\mathfrak{g} = \mathfrak{g}_k + i\mathfrak{g}_k$ (direct vector space sum). Using now (for the first time) the relation $N_{\alpha,\beta} = -N_{-\alpha,-\beta}$ (which in view of Lemma 5.2 implies that each $N_{\alpha,\beta}$ is real), we see that $X, Y \in \mathfrak{g}_k$ implies $[X, Y] \in \mathfrak{g}_k$. Thus $\mathfrak{g}_k$ is a real form of $\mathfrak{g}$. The Killing form of $\mathfrak{g}_k$ is strictly negative definite, being the restriction of B to $\mathfrak{g}_k \times \mathfrak{g}_k$. Thus $\mathfrak{g}_k$ is compact and the theorem is proved.

§ 7. Cartan Decompositions

Theorem 7.1. *Let $\mathfrak{g}_0$ be a semisimple Lie algebra over $\boldsymbol{R}$, $\mathfrak{g}$ its complexification, and $\mathfrak{u}$ any compact real form of $\mathfrak{g}$. Let σ and τ denote the conjugations of $\mathfrak{g}$ with respect to $\mathfrak{g}_0$ and $\mathfrak{u}$, respectively. Then there exists an automorphism φ of $\mathfrak{g}$ such that the compact real form $\varphi \cdot \mathfrak{u}$ is invariant under σ.*

Proof. The Hermitian form B_τ on $\mathfrak{g} \times \mathfrak{g}$ given by

$$B_\tau(X, Y) = -B(X, \tau Y), \qquad X, Y \in \mathfrak{g},$$

is strictly positive definite since $\mathfrak{u}$ is compact. The linear transformation $N = \sigma\tau$ is an automorphism of the complex algebra $\mathfrak{g}$ and hence leaves the Killing form invariant. Using $\sigma^2 = \tau^2 = I$, we obtain

$$B(NX, \tau Y) = B(X, N^{-1}\tau Y) = B(X, \tau NY)$$

or

$$B_\tau(NX, Y) = B_\tau(X, NY).$$

This shows that N is self-adjoint with respect to B_τ. Let $X_1, ..., X_n$ be a basis of $\mathfrak{g}$ with respect to which N is represented by a diagonal matrix. Then the endomorphism $P = N^2$ is represented by a diagonal matrix with positive diagonal elements $\lambda_1, ..., \lambda_n$. For each $t \in \boldsymbol{R}$, let P^t denote the linear transformation of $\mathfrak{g}$ represented by the diagonal matrix with diagonal elements $(\lambda_i)^t > 0$. Then each P^t commutes with N. Let $c^k{}_{ij}$ denote the constants determined by

$$[X_i, X_j] = \sum_{k=1}^{n} c^k{}_{ij} X_k$$

for $1 \leqslant i, j \leqslant n$. Since P is an automorphism, we have

$$\lambda_i \lambda_j c^k{}_{ij} = \lambda_k c^k{}_{ij} \qquad (1 \leqslant i, j, k \leqslant n).$$

This equation implies

$$(\lambda_i)^t (\lambda_j)^t c^k{}_{ij} = (\lambda_k)^t c^k{}_{ij} \qquad (t \in \boldsymbol{R}),$$

which shows that each P^t is an automorphism of $\mathfrak{g}$.

Consider now the mapping $\tau_1 = P^t \tau P^{-t}$ of $\mathfrak{g}$ into itself. The subspace $P^t \mathfrak{u}$ is a compact real form of $\mathfrak{g}$ and τ_1 is the conjugation of $\mathfrak{g}$ with respect to this form. Moreover we have $\tau N \tau^{-1} = N^{-1}$ so $\tau P \tau^{-1} = P^{-1}$. By a simple matrix computation the relation $\tau P = P^{-1}\tau$ implies $\tau P^t = P^{-t}\tau$ for all $t \in \boldsymbol{R}$. Consequently,

$$\sigma\tau_1 = \sigma P^t \tau P^{-t} = \sigma\tau P^{-2t} = NP^{-2t},$$
$$\tau_1\sigma = (\sigma\tau_1)^{-1} = P^{2t} N^{-1} = N^{-1} P^{2t}.$$

If $t = \frac{1}{4}$, then $\sigma\tau_1 = \tau_1\sigma$. Thus the automorphism $\varphi = P^{1/4}$ has the desired properties.

Remark. The proof has shown that the automorphism φ can be chosen as $P^{1/4}$ where P^t $(t \in \boldsymbol{R})$ is a one-parameter group of semisimple, positive definite (for B_τ) automorphisms of $\mathfrak{g}$ satisfying $P^1 = (\sigma\tau)^2$.

Definition. Let $\mathfrak{g}_0$ be a semisimple Lie algebra over $\boldsymbol{R}$, $\mathfrak{g}$ its complexification, σ the conjugation of $\mathfrak{g}$ with respect to $\mathfrak{g}_0$. A direct decomposition $\mathfrak{g}_0 = \mathfrak{k}_0 + \mathfrak{p}_0$ of $\mathfrak{g}_0$ into a subalgebra $\mathfrak{k}_0$ and a vector subspace $\mathfrak{p}_0$ is called a *Cartan decomposition* if there exists a compact real form $\mathfrak{g}_k$ of $\mathfrak{g}$ such that

$$\sigma \cdot \mathfrak{g}_k \subset \mathfrak{g}_k, \qquad \mathfrak{k}_0 = \mathfrak{g}_0 \cap \mathfrak{g}_k, \qquad \mathfrak{p}_0 = \mathfrak{g}_0 \cap (i\mathfrak{g}_k). \tag{1}$$

It is an immediate consequence of Theorems 6.3 and 7.1 that each semisimple Lie algebra $\mathfrak{g}_0$ over $\boldsymbol{R}$ has a Cartan decomposition. The following theorem shows that any two Cartan decompositions are conjugate under an inner automorphism.

Theorem 7.2. *Let $\mathfrak{g}_0$ be a semisimple Lie algebra over $\boldsymbol{R}$. Suppose*

$$\mathfrak{g}_0 = \mathfrak{k}_1 + \mathfrak{p}_1, \qquad \mathfrak{g}_0 = \mathfrak{k}_2 + \mathfrak{p}_2$$

are two Cartan decompositions of $\mathfrak{g}_0$. Then there exists an element $\psi \in Int\,(\mathfrak{g}_0)$ such that

$$\psi \cdot \mathfrak{k}_1 = \mathfrak{k}_2, \qquad \psi \cdot \mathfrak{p}_1 = \mathfrak{p}_2.$$

Proof. Let $\mathfrak{g}$ be the complexification of $\mathfrak{g}_0$, σ the conjugation of $\mathfrak{g}$ with respect to $\mathfrak{g}_0$. Then there exist compact real forms $\mathfrak{u}_1$ and $\mathfrak{u}_2$ of $\mathfrak{g}$ such that

$$\sigma \cdot \mathfrak{u}_j \subset \mathfrak{u}_j, \qquad \mathfrak{k}_j = \mathfrak{g}_0 \cap \mathfrak{u}_j, \qquad \mathfrak{p}_j = \mathfrak{g}_0 \cap (i\mathfrak{u}_j) \qquad (j = 1, 2).$$

Let τ_1 and τ_2 denote the conjugations of $\mathfrak{g}$ with respect to $\mathfrak{u}_1$ and $\mathfrak{u}_2$, respectively. Then from Theorem 7.1 and the subsequent remark there exists a one-parameter group P^t of positive definite semisimple automorphisms of $\mathfrak{g}$ such that $P^1 = (\tau_1\tau_2)^2$ and $P^{1/4} \cdot \mathfrak{u}_2$ is a compact real form of $\mathfrak{g}$ invariant under τ_1. It follows that $P^{1/4} \cdot \mathfrak{u}_2$ is the direct sum of its intersections with $\mathfrak{u}_1$ and $i\mathfrak{u}_1$. Now B, the Killing form of $\mathfrak{g}$, is strictly positive definite on the subspace $i\mathfrak{u}_1$, and strictly negative definite on the compact form $P^{1/4}\mathfrak{u}_2$. Consequently, the intersection $P^{1/4}\mathfrak{u}_2 \cap i\mathfrak{u}_1$ reduces to $\{0\}$, so

$$\mathfrak{u}_1 = P^{1/4} \cdot \mathfrak{u}_2.$$

Since $\sigma\mathfrak{u}_j \subset \mathfrak{u}_j$, we have $\sigma\tau_j = \tau_j\sigma$ $(j = 1, 2)$ and consequently σ commutes with P^1. Now P^t is the unique positive definite tth power of P^1; hence σ commutes with each P^t so P^t leaves $\mathfrak{g}_0$ invariant. The restriction of the linear transformation P^t to $\mathfrak{g}_0$ gives rise to a one-parameter subgroup $\{\exp tX\}$ of Aut $(\mathfrak{g}_0)$. As a result of Cor. 6.5, Chapter II, we have $\{\exp tX\} \subset \text{Int}\,(\mathfrak{g}_0)$. The theorem follows if we take $\psi = \exp \frac{1}{4} X$.

The proof has the following

Corollary 7.3. *If $\mathfrak{u}_1$ and $\mathfrak{u}_2$ are any compact real forms of a semisimple Lie algebra $\mathfrak{g}$ over $\boldsymbol{C}$ then there exists a one-parameter subgroup ψ^t $(t \in \boldsymbol{R})$ of automorphisms of $\mathfrak{g}$ such that $\psi^1\mathfrak{u}_1 = \mathfrak{u}_2$.*

Proposition 7.4. *Let $\mathfrak{g}_0$ be a semisimple Lie algebra over $\boldsymbol{R}$ which is the direct sum $\mathfrak{g}_0 = \mathfrak{k}_0 + \mathfrak{p}_0$ where $\mathfrak{k}_0$ is a subalgebra and $\mathfrak{p}_0$ a vector subspace. The following conditions are equivalent.*

(i) *$\mathfrak{g}_0 = \mathfrak{k}_0 + \mathfrak{p}_0$ is a Cartan decomposition of $\mathfrak{g}_0$.*

(ii) *$B(T, T) < 0$ for $T \neq 0$ in $\mathfrak{k}_0$,*
$B(X, X) > 0$ for $X \neq 0$ in $\mathfrak{p}_0$

and the mapping

$$s_0 : T + X \to T - X, \qquad T \in \mathfrak{k}_0,\ X \in \mathfrak{p}_0,$$

is an automorphism of $\mathfrak{g}_0$.

If these conditions are satisfied, $\mathfrak{k}_0$ is a maximal compactly imbedded subalgebra of $\mathfrak{g}_0$.

Proof. (ii) $\Rightarrow$ (i). Let $\mathfrak{g}$ denote the complexification of $\mathfrak{g}_0$, and let $\mathfrak{g}^R$ denote the Lie algebra $\mathfrak{g}$ when considered as a Lie algebra over $\boldsymbol{R}$. Since s_0 is an automorphism, we have $B(\mathfrak{k}_0, \mathfrak{p}_0) = 0$, $[\mathfrak{k}_0, \mathfrak{p}_0] \subset \mathfrak{p}_0$, and $[\mathfrak{p}_0, \mathfrak{p}_0] \subset \mathfrak{k}_0$. It follows that the subspace $\mathfrak{g}_k = \mathfrak{k}_0 + i\mathfrak{p}_0$ of $\mathfrak{g}^R$ is a subalgebra and in fact a compact real form of $\mathfrak{g}$, satisfying the relations (1). On the other hand, relations (1) show that (i) $\Rightarrow$ (ii). We know from Lemma 6.2 that the groups $\mathrm{Int}(\mathfrak{g}_0)$ and $\mathrm{Int}(\mathfrak{g}_k)$ can be regarded as closed subgroups of $\mathrm{Int}(\mathfrak{g}^R)$. Now $\mathrm{Int}(\mathfrak{g}_k)$ is compact and the same is true of $\mathrm{Int}(\mathfrak{g}_0) \cap \mathrm{Int}(\mathfrak{g}_k)$. This last group is a Lie subgroup of $\mathrm{Int}(\mathfrak{g}_0)$ and has Lie algebra $\mathfrak{g}_0 \cap \mathfrak{g}_k = \mathfrak{k}_0$. Thus $\mathfrak{k}_0$ is compactly imbedded in $\mathfrak{g}_0$. If $\mathfrak{k}_0$ were not maximal let $\mathfrak{k}_1$ be a compactly imbedded subalgebra of $\mathfrak{g}_0$, properly containing $\mathfrak{k}_0$. Then there exists an element $X \neq 0$ in $\mathfrak{k}_1 \cap \mathfrak{p}_0$. Let η denote the conjugation of $\mathfrak{g}$ with respect to $\mathfrak{g}_k$. Then $\eta\mathfrak{g}_0 \subset \mathfrak{g}_0$ and the bilinear form B_η on $\mathfrak{g}_0 \times \mathfrak{g}_0$ defined by

$$B_\eta(Y, Z) = -B(Y, \eta Z), \qquad Y, Z \in \mathfrak{g}_0,$$

is symmetric and strictly positive definite. Since

$$B([X, Y], \eta Z) = -B(Y, [X, \eta Z]) = B(Y, [\eta X, \eta Z])$$

we have

$$B_\eta(\mathrm{ad}\, X(Y), Z = B_\eta(Y, \mathrm{ad}\, X(Z)).$$

Thus $\mathrm{ad}\, X$ has all its eigenvalues real, and not all zero. But then the powers $e^{n\, \mathrm{ad}\, X}$ can not lie in a compact matrix group. This contradicts the fact that $\mathfrak{k}_1$ is a compactly imbedded subalgebra of $\mathfrak{g}_0$.

Corollary 7.5. *Let* $\mathfrak{g}$ *be a semisimple Lie algebra over* $\boldsymbol{C}$ *and let* $\mathfrak{u}$ *be any compact real form of* $\mathfrak{g}$. *Let* $\mathfrak{g}^R$ *denote the Lie algebra* $\mathfrak{g}$ *considered as a real Lie algebra and let* J *denote the complex structure of* $\mathfrak{g}^R$ *which corresponds to the multiplication by* i *on* $\mathfrak{g}$. *Then*

$$\mathfrak{g}^R = \mathfrak{u} + J\mathfrak{u}$$

is a Cartan decomposition of $\mathfrak{g}^R$.

In fact, let B^R, B, and B^C denote the Killing forms on $\mathfrak{g}^R$, $\mathfrak{g}$, and $(\mathfrak{g}^R)^C$. Then by Lemma 3.1 $B^C = B^R = 2\,\mathrm{Re}\, B$ on $\mathfrak{g}^R \times \mathfrak{g}^R$. Since B is strictly negative definite on $\mathfrak{u} \times \mathfrak{u}$ and strictly positive definite on $J\mathfrak{u} \times J\mathfrak{u}$, the same holds for B^C. Using Prop. 7.4 the corollary follows.

EXERCISES

A. Solvable and Nilpotent Lie Algebras

1. A solvable Lie algebra has no semisimple subalgebra $\neq \{0\}$.

2. Let $\mathfrak{t}(n)$ denote the subalgebra of $\mathfrak{gl}(n, \boldsymbol{R})$ formed by all upper triangular $n \times n$ matrices and let $\mathfrak{n}(n)$ denote the subalgebra of matrices in $\mathfrak{t}(n)$ having all diagonal elements 0. Prove that:

(i) $\mathfrak{t}(n)$ is solvable, $\mathfrak{n}(n)$ is nilpotent and coincides with the derived algebra of $\mathfrak{t}(n)$.

(ii) The Lie algebras $\mathfrak{t}(n)$ and $\mathfrak{n}(n)$ both have centers of dimension 1.

(iii) Let B denote the Killing form of $\mathfrak{t}(n)$. Then $B(\mathfrak{t}(n), \mathfrak{n}(n)) = 0$.

3*. A Lie algebra is solvable if and only if its derived algebra is nilpotent (É. Cartan [1], p. 47, Bourbaki [3], §5).

4. Let G denote the group of all mappings $T_{a,b} : x \to ax + b$ $(x \in \boldsymbol{R})$, where a and b are real numbers, $a > 0$. Let G have the analytic structure determined by the condition that the mapping $(a, b) \to T_{a,b}$ is an analytic diffeomorphism. Show that G is a solvable Lie group and that its Lie algebra is the only noncommutative Lie algebra over $\boldsymbol{R}$ of dimension 2 (up to isomorphism).

5. Let $\mathfrak{g}$ be a solvable Lie algebra over $\boldsymbol{C}$ and let $\mathfrak{a}$ be a minimal proper ideal in $\mathfrak{g}$. Then $\mathfrak{a}$ has dimension 1. Formulate and prove an analogous result for solvable Lie algebras over $\boldsymbol{R}$.

B. Semisimple Lie Algebras

1. Let $\mathfrak{g}$ be a semisimple Lie algebra over $\boldsymbol{C}$, $\mathfrak{h}$ a Cartan subalgebra. Let Δ denote the set of nonzero roots of $\mathfrak{g}$ with respect to $\mathfrak{h}$. Let Γ be a subset of Δ satisfying the conditions: If $\gamma \in \Gamma$, then $-\gamma \in \Gamma$; if $\gamma, \delta \in \Gamma$ and $\gamma + \delta \in \Delta$, then $\gamma + \delta \in \Gamma$. Let $\mathfrak{g}_\Gamma$ be the smallest subalgebra of $\mathfrak{g}$ containing all the root subspaces $\mathfrak{g}^\gamma$, $\gamma \in \Gamma$. Then $\mathfrak{g}_\Gamma$ is semisimple and $\mathfrak{h} \cap \mathfrak{g}_\Gamma$ is a Cartan subalgebra of $\mathfrak{g}_\Gamma$.

2. Let $\mathfrak{g}$ be a semisimple Lie algebra over $\boldsymbol{C}$ and $\mathfrak{h}$ a Cartan subalgebra. Let Δ denote the system of nonzero roots and put $\mathfrak{h}^* = \sum_{\alpha \in \Delta} \boldsymbol{R} H_\alpha$. The system $\{H_\alpha : \alpha \in \Delta\}$ is invariant under the reflections in the hyperplanes $\alpha(H) = 0$ of $\mathfrak{h}^*$ (Theorem 4.3 (i)).

3. Consider the Lie algebras $\mathfrak{sl}(n, C)$, $\mathfrak{so}(n, C)$, and $\mathfrak{sp}(n, C)$ from Chapter IX, §4. Let E_{ij} denote the matrix $(\delta_{ai}\delta_{bj})_{1 \leq a,b \leq n}$. Show that:

(i) $\mathfrak{sl}(n, C)$ has Cartan subalgebra consisting of all diagonal matrices of trace 0. Its Killing form is

$$B(X, Y) = 2n \operatorname{Tr}(XY), \qquad X, Y \in \mathfrak{sl}(n, C).$$

(ii) $\mathfrak{so}(n, \boldsymbol{C})$ has Cartan subalgebra spanned by the vectors $E_{2i-1\,2i} - E_{2i\,2i-1}$ $(1 \leqslant i \leqslant [\frac{1}{2}n])$. Its Killing form is

$$B(X, Y) = (n-2)\operatorname{Tr}(XY), \qquad X, Y \in \mathfrak{so}(n, \boldsymbol{C}).$$

(iii) $\mathfrak{sp}(n, \boldsymbol{C})$ has Cartan subalgebra spanned by the vectors $E_{ii} - E_{n+i\,n+i}$ $(1 \leqslant i \leqslant n)$. Its Killing form is

$$B(X, Y) = (2n+2)\operatorname{Tr}(XY), \qquad X, Y \in \mathfrak{sp}(n, \boldsymbol{C}).$$

4. A representation ρ of a Lie algebra $\mathfrak{g}$ on a finite-dimensional vector space V is called ***semisimple*** if each subspace of V invariant under $\rho(\mathfrak{g})$ has a complementary invariant subspace.

(i) Any finite-dimensional representation of a compact Lie algebra is semisimple.

(ii) (Weyl's unitary trick.) Using a compact real form, prove that any finite-dimensional representation of a semisimple Lie algebra over $\boldsymbol{R}$ is semisimple.

NOTES

§2. Lie's theorem is proved in Lie [1], III, p. 678. The name "solvable" ("integrable" in Lie's terminology) is derived from the analogous concept for finite groups. In the theory of Picard-Vessiot extensions of differential fields solvable Lie groups play a role analogous to the role of solvable finite groups in Galois theory (Kolchin [1]). Concerning Engel's theorem, see É. Cartan [1], p. 46.

§3-§5. The root space decomposition was used in É. Cartan's classification of simple Lie algebras over $\boldsymbol{C}$, [1]. The roots are there introduced as zeros of the characteristic polynomial $\det(\lambda I - \operatorname{ad} X)$. In Weyl [1, 2] the Cartan subalgebra is brought more to the foreground and many simplifications and methods were introduced. Among these is the ordering of the roots, a tool whose power is convincingly illustrated in the proof of Theorem 5.4, due to Weyl. Using Cartan's classification, Chevalley has in [5], Théorème 1, sharpened Theorem 5.5. Chevalley shows that if H_α, X_α $(\alpha \in \Delta)$ are replaced by proportional vectors H'_α, X'_α such that $\alpha(H'_\alpha) = 2$ and the commutation relations still satisfied, then the new coefficients $N_{\alpha,\beta}$ have absolute value $1 - p$. Hence it follows from Theorem 4.3 (i) that $\mathfrak{g}$ has a basis Z_i such that the structural constants $c^i{}_{jk}$ determined by $[Z_j, Z_k] = \Sigma_i\, c^i{}_{jk}\, Z_i$ are all integers. Thus $\mathfrak{g}$ gives rise to a Lie algebra over any field of any characteristic.

§6. The real forms of all simple Lie algebras over $\boldsymbol{C}$ were classified by É. Cartan [2], thereby proving Theorem 6.3 for the first time. The proof in the text, which is independent of the classification is due to Weyl [1], Kap. III, Satz 6. No proof seems to be known which does not use the Cartan subalgebra (see interesting indications by É. Cartan in [12], p. 23). Such a proof would have the existence of Cartan subalgebras as a corollary without appeal to the theorems of Lie and Engel.

§7. Theorem 7.1 is proved by É. Cartan [12] and simplified by Mostow [1]. The proof in the text is modelled after Samelson. Theorem 7.2 is also proved in Cartan [12]; the proof in the text is simpler, but possibly less instructive.

CHAPTER IV

SYMMETRIC SPACES

In this chapter we return to Riemannian geometry and begin a study of the Riemannian locally symmetric spaces. These are defined as Riemannian manifolds for which the curvature tensor is invariant under all parallel translations. É. Cartan set himself the problem of giving a complete classification of these spaces. In an ingeneous manner he gave the problem two different group-theoretic formulations [6]. One of these is particularly effective and strikingly enough reduces the problem to the classification of simple Lie algebras over $\boldsymbol{R}$, a problem which Cartan himself had solved already in 1914.

Cartan's first method was based on the so-called *holonomy group*. If o is a point in a Riemannian manifold M, then the holonomy group of M is the group of all linear transformations of the tangent space M_o obtained by parallel translation along closed curves starting at o. It is readily seen that different points of M give isomorphic holonomy groups. Of course each element of the holonomy group leaves the Riemannian structure g_o invariant; if M is locally symmetric the curvature tensor R_o is also left invariant. Hence it follows from the structural equations (6) and (7), §8, Chapter I, that each element of the holonomy group induces an isometry of a neighborhood of o in M onto itself leaving o fixed. This leads to algebraic relations between the Lie algebra $\mathfrak{k}$ of the identity component of the holonomy group and the tensors g_o and R_o, namely,

$$\begin{aligned} g_o(AX,Y) + g_o(X,AY) &= 0, && A \in \mathfrak{k},\ X,Y \in M_o; \\ [A,R_o(X,Y)] &= R_o(AX,Y) + R_o(X,AY), && A \in \mathfrak{k},\ X,Y \in M_o; \\ R_o(X,Y) &\in \mathfrak{k}, && X,Y \in M_o, \end{aligned}$$

proved here in §5. Cartan now showed ([6], p. 225) that if for a given $\mathfrak{k}$ a tensor R_o of type (1,3) satisfies these formulas and in addition fulfills the general symmetry conditions for a curvature tensor (Lemma 12.5, Chapter I), then there exists a locally symmetric space for which it is the curvature tensor at o. In our treatment this proof is used in §5 to construct a globally symmetric space from a locally symmetric one. Since a symmetric space is determined locally from R_o and g_o (Lemma 1.2), the problem is reduced to two others: 1. Classify all possible holonomy groups of symmetric spaces. 2. Determine (up to a constant factor) the curvature tensor of a symmetric space by means of the holonomy group. For the second problem see Theorem 3.6, Chapter IX, and Cartan [6], pp. 221-224. For the first problem Cartan used his earlier work on finite-dimensional representations of Lie algebras, but the necessarily extensive calculations were not carried out in all details in [6] since a simpler method became available.

This second method of Cartan [7] (and the one which we shall follow) brings the problem of classifying locally symmetric spaces more directly into the realm of group theory. It is based on the fact that the invariance of the curvature tensor under parallelism is equivalent to the condition that the geodesic symmetry with respect to each point be a local isometry. (This explains the term "locally symmetric.") The consequences of this fact are more conveniently expressed for Riemannian globally symmetric spaces for which the geodesic symmetry by definition always extends to a global isometry. Such spaces have a transitive group of isometries and can be represented as coset spaces G/K where G is a connected Lie group with an involutive automorphism σ whose fixed point set is (essentially) K. The group G becomes semisimple after dividing out a direct factor which is a motion group of a Euclidean space. In this way the problem is reduced to the study of certain involutive automorphisms of semisimple Lie algebras.

In §1 the two definitions of locally symmetric spaces are considered and shown equivalent. The group of isometries of a Riemannian manifold is studied in §2; in §3 the results are applied to Riemannian globally symmetric spaces which then are represented as coset spaces of a special kind. The curvature tensor of a Riemannian globally symmetric space is computed in §4. The result is used in §5 for constructing a Riemannian globally symmetric space, a piece of which is isometric to a piece of a given locally symmetric space. In the last section it is shown how totally geodesic subspaces of a symmetric space are connected with Lie triple systems of the Lie algebra of the group of isometries.

§ 1. Affine Locally Symmetric Spaces

Let M be a C^∞ manifold with an affine connection ∇. Let p be a point in M and let N_0 be a normal neighborhood of the origin 0 in M_p, symmetric with respect to 0. As usual, put $N_p = \operatorname{Exp}_p N_0$. For each $q \in N_p$, consider the geodesic $t \to \gamma(t)$ within N_p passing through p and q such that $\gamma(0) = p$, $\gamma(1) = q$. We put $q' = \gamma(-1)$. The mapping $q \to q'$ of N_p onto itself is called the *geodesic symmetry* with respect to p and will be denoted by s_p. In normal coordinates $\{x_1, ..., x_m\}$ at p, s_p has the expression $(x_1, ..., x_m) \to (-x_1, ..., -x_m)$. In particular, s_p is a diffeomorphism of N_p onto itself and $(ds_p)_p = -I$ where I denotes the identity mapping.

Definition. Let M be a manifold with an affine connection ∇ which has torsion tensor T and curvature tensor R; M is called *affine locally symmetric* if each point $m \in M$ has an open neighborhood N_m on which the geodesic symmetry s_m is an affine transformation.

Theorem 1.1. *A manifold M is affine locally symmetric if and only if $T = 0$ and $\nabla_Z R = 0$ for all $Z \in \mathfrak{D}^1(M)$.*

Proof. To begin with let M be any manifold with an affine connection ∇ and let φ be a diffeomorphism of M onto itself. We define the connection ∇' on M by

$$\nabla'_X(Y) = (\nabla_{X^\varphi}(Y^\varphi))^{\varphi^{-1}} \qquad \text{for } X, Y \in \mathfrak{D}^1(M). \tag{1}$$

If we denote the mapping $X \to X^{\varphi}$ ($X \in \mathfrak{D}^1$) by φ, we have $\nabla'_X = \varphi^{-1} \circ \nabla_{\varphi X} \circ \varphi$ on $\mathfrak{D}^1$. Let T' and R' denote the corresponding torsion and curvature tensors. It is trivial to verify

$$T'(X, Y) = \varphi^{-1}(T(\varphi X, \varphi Y)), \tag{2}$$

$$R'(X, Y) = \varphi^{-1} \circ R(\varphi X, \varphi Y) \circ \varphi \tag{3}$$

for $X, Y \in \mathfrak{D}^1$. We shall now prove the relation

$$(\nabla'_Z R')(X, Y) = \varphi^{-1} \circ ((\nabla_{\varphi Z} R)(\varphi X, \varphi Y)) \circ \varphi \tag{4}$$

for $X, Y, Z \in \mathfrak{D}^1$. From relation (7) in Chapter I, §7, we have

$$[\nabla_Z, R(X, Y)] = \nabla_Z(R(X, Y)). \tag{5}$$

We now apply ∇_Z to the tensor field $X \otimes Y \otimes R$. Using (5) and the fact that ∇_Z commutes with contractions we obtain

$$(\nabla_Z R)(X, Y) = [\nabla_Z, R(X, Y)] - R(\nabla_Z X, Y) - R(X, \nabla_Z Y).$$

If we combine this with the similar formula for $(\nabla'_Z R')(X, Y)$, relation (4) follows easily.

Let m be an arbitrary point in M and let N_m be a normal neighborhood of m invariant under s_m. Suppose first that M is affine locally symmetric. Let $Z \in \mathfrak{D}^1(M)$ and γ a geodesic in N_m, passing through m with tangent vector Z_m at m. Let p and q be two points on γ, symmetric with respect to m. Let τ and τ_m denote the parallel translation (along γ) from p to q and m, respectively. Consider a vector $L \in M_p$. The vectors L and $\tau_m L$ are parallel with respect to the geodesic (mp). Since s_m is an affine transformation, the vectors $ds_m L$ and $ds_m \tau_m L$ are parallel with respect to the geodesic (mq). Since $ds_m \tau_m L = -\tau_m L$ it follows that

$$ds_m(L) = -\tau L. \tag{6}$$

Using (2) and (3) for $s_m = \varphi$, $R' = R$, $T' = T$, we obtain (§7, Chapter I)

$$\tau R_p = R_q, \qquad \tau T_p = -T_q. \tag{7}$$

Using the definition of $(\nabla_Z R)_m$ (§7, Chapter I) we deduce from (7) that $(\nabla_Z R)_m = 0$. Moreover, putting $p = q = m$, it follows from (7) that $T_m = 0$.

Next we prove the converse. The diffeomorphism s_m of N_m defines a new connection on N_m by (1). Since T and $\nabla_Z R$ vanish we find from (2) and (4) that $T' = 0$ and $\nabla'_Z R' = 0$. Finally we have $R'_m = R_m$

as a consequence of $(ds_m)_m = -I$. The theorem will thus follow from:

Lemma 1.2. *Let M and M' be two manifolds with affine connections ∇ and ∇'. Assume that*

$$\nabla_Z T = 0, \qquad \nabla_Z R = 0 \qquad \textit{for all } Z \in \mathfrak{D}^1(M),$$

$$\nabla'_{Z'} T' = 0, \qquad \nabla'_{Z'} R' = 0 \qquad \textit{for all } Z' \in \mathfrak{D}^1(M').$$

Let $p \in M$, $p' \in M'$ and let A be a linear one-to-one mapping of M_p onto $M'_{p'}$. Let $\tilde{A}$ denote the unique type-preserving isomorphism of the mixed tensor algebra $\mathfrak{D}(p)$ onto $\mathfrak{D}(p')$ extending A such that $\tilde{A}$ coincides with $({}^tA)^{-1}$ on the dual $(M_p)^\wedge$. Assume now that $\tilde{A} \cdot R_p = R'_{p'}$, $\tilde{A} \cdot T_p = T'_{p'}$. Then there exists an open neighborhood U_p of p in M and an affine transformation φ of U_p onto an open neighborhood $U_{p'}$ of p' in M' such that $\varphi(p) = p'$ and $d\varphi_p = A$.

Proof. Let N_0 and N'_0 be normal neighborhoods of the origin in M_p and $M'_{p'}$, respectively, and put $N_p = \operatorname{Exp} N_0$, $N'_{p'} = \operatorname{Exp} N'_0$. Let $Y_1, \ldots, Y_m$ be a basis of M_p and let $T^i{}_{jk}$ and $R^i{}_{ljk}$ be the coefficients of T and R in terms of the adapted vector field basis $Y_1^*, \ldots, Y_m^*$. These coefficients are then constants in N_p, due to the assumption. Now we express T' and R' in terms of the vector field basis on $N'_{p'}$, adapted to the basis $AY_1, \ldots, AY_m$ of $M'_{p'}$. Then the coefficients $T'^i{}_{jk}$ and $R'^i{}_{ljk}$ are *the same constants* as $T^i{}_{jk}$ and $R^i{}_{ljk}$.

Let V (respectively, V') denote the set of points $(t, a_1, \ldots, a_m) \in \boldsymbol{R} \times \boldsymbol{R}^m$ for which $ta_1Y_1 + \ldots + ta_mY_m \in N_0$, $(ta_1(AY_1) + \ldots + ta_m(AY_m) \in N'_0)$. The differential equations (6) and (7) in Chapter I, §8, are exactly the same for both connections ∇ and ∇'. For ∇ they hold on V and for ∇' they hold on V'. Since the equations are linear, the uniqueness of their solutions holds globally. Consequently, the solutions of the two sets of equations agree on $V \cap V'$. Let

$$U_p = \{\operatorname{Exp}_p (t(a_1Y_1 + \ldots + a_mY_m)) : (t, a_1, \ldots, a_m) \in V \cap V'\};$$

$$U_{p'} = \{\operatorname{Exp}_{p'} (t(a_1AY_1 + \ldots + a_mAY_m)) : (t, a_1, \ldots, a_m) \in V \cap V'\}.$$

Then U_p and $U_{p'}$ are normal neighborhoods of $p \in M$ and $p' \in M'$, respectively. If $q \in U_p$ and $q' \in U_{p'}$ have the same normal coordinates with respect to (Y_i) and (AY_i), respectively, the mapping $\varphi : q \to q'$ is an affine transformation of U_p onto $U_{p'}$ such that $\varphi(p) = p'$ and $d\varphi_p = A$.

Definition. A Riemannian manifold M is called a *Riemannian locally symmetric space* if for each $p \in M$ there exists a normal neighborhood of p on which the geodesic symmetry with respect to p is an isometry.

Theorem 1.3. *Let M be a Riemannian manifold. Then M is a Riemannian locally symmetric space if and only if the sectional curvature is invariant under all parallel translations.*

If M is locally symmetric, then the Riemannian structure g and the curvature tensor R are both invariant under all parallel translations. The invariance of the sectional curvature follows from Theorem 12.2, Chapter I. On the other hand, suppose the sectional curvatures invariant under all parallel translations. Let $p, q \in M$, γ a curve segment joining p to q, and τ the parallel translation from p to q along γ. Then if X, $Y \in M_p$, we have

$$g_p(R_p(X, Y)\, X, Y) = g_q(R_q(\tau X, \tau Y)\, \tau X, \tau Y),$$
$$g_p(R_p(X, Y)\, X, Y) = g_q(\tau(R_p(X, Y)\, X), \tau Y).$$

The quadrilinear form B given by

$$B(X, Y, Z, T) = g_q(R_q(\tau X, \tau Y)\, \tau Z, \tau T) - g_q(\tau(R_p(X, Y)\, Z), \tau T)$$

for $X, Y, Z, T \in M_p$, satisfies the conditions of Lemma 12.4, Chapter I. Consequently, $B \equiv 0$ so

$$\tau(R_p(X, Y)\, Z) = R_q(\tau X, \tau Y)\, \tau Z, \quad \text{that is,} \quad \tau \cdot R_p = R_q,$$

which shows that $\nabla_U R = 0$ for each $U \in \mathfrak{D}^1$. Since the geodesic symmetry s_m induces an isometry of M_m, the theorem now follows from Theorem 1.1 and:

Lemma 1.4. *Let φ be an affine transformation of a pseudo-Riemannian manifold M. Suppose that for some point $q \in M$, the mapping $d\varphi_q : M_q \to M_{\varphi(q)}$ is an isometry. Then φ is an isometry of M onto itself.*

Proof. Let $p \in M$ and $X, Y \in M_p$. We join p to q by a curve γ and let τ denote the parallel translation from p to q along γ. Then

$$g_p(X, Y) = g_q(\tau X, \tau Y) = g_{\varphi(q)}(d\varphi_q \tau X, d\varphi_q \tau Y).$$

This last quantity equals $g_{\varphi(p)}\,(d\varphi_p X, d\varphi_p Y)$ because φ, being an affine transformation, transforms vectors that are parallel along γ into vectors that are parallel along $\varphi \cdot \gamma$. This proves the lemma.

§ 2. Groups of Isometries

Let M be a Riemannian manifold and $I(M)$ the set of all isometries of M. Let $g_1, g_2 \in I(M)$. The composite mapping $g_1 \circ g_2$ is again an isometry. If we put $g_1 g_2 = g_1 \circ g_2$, $I(M)$ becomes a group. We shall always consider $I(M)$ with the *compact open topology*. This is defined as follows: Let C and U, respectively, be a compact and an open subset of M, and put

$$W(C, U) = \{g \in I(M) : g \cdot C \subset U\}.$$

The compact open topology on $I(M)$ is defined as the smallest topology on $I(M)$ for which all the sets $W(C, U)$ are open. It is obvious that $I(M)$ is a Hausdorf space. The identity component of $I(M)$ will be denoted $I_0(M)$.

Lemma 2.1. *The space $I(M)$ has a countable base.*

Proof. Since M is a separable metric space (Chapter I, §9) there exists a countable basis $O_1, \ldots, O_i, \ldots$ for the open subsets of M. Since M is locally compact, we can assume that the closure $\bar{O}_i$ is compact for each i. Let $C \subset M$ be compact, $U \subset M$ open, and f any element of $W(C, U)$. For each $p \in C$ there exists an index i and an index j such that $p \in O_i$, $f(O_i) \subset O_j \subset U$. We can find coverings $O_{i_1} \ldots O_{i_N}$ of C and $O_{j_1} \ldots O_{j_N}$ of $f(C)$ such that

$$f(\bar{O}_{i_k}) \subset O_{j_k} \subset U \qquad (1 \leqslant k \leqslant N).$$

It follows that

$$f \in \bigcap_{k=1}^{N} W(\bar{O}_{i_k}, O_{j_k}) \subset W(C, U).$$

This shows that the set Ω of all finite intersections of sets of the form $W(\bar{O}_i, O_j)$ forms a basis of the open sets of $I(M)$. Since Ω is countable, the lemma follows.

Theorem 2.2. *Let M be a Riemannian manifold and (f_n) a sequence in $I(M)$. Suppose there exists a point $o \in M$ such that the sequence $(f_n \cdot o)$ is convergent. Then there exists an element $f \in I(M)$ and a subsequence (f_{n_ν}) of (f_n) which converges to f in the compact open topology.*

We first prove a lemma.

Lemma 2.3. *Assume that a sequence (f_n) in $I(M)$ converges pointwise on a set $A \subset M$. Then (f_n) also converges pointwise on $\bar{A}$ (the closure of A).*

Proof. Let $p \in \bar{A}$ and choose $r > 0$ such that the open ball $B_r(p)$ has compact closure. Let ϵ be given, $0 < \epsilon < r$. We first select a point $p_1 \in A$ such that $d(p, p_1) < \epsilon/3$, then an integer N such that

$$d(f_n \cdot p_1, f_m \cdot p_1) < \epsilon/3 \qquad \text{for } n, m \geqslant N.$$

It follows that

$$d(f_n \cdot p, f_m \cdot p) \leqslant d(f_n \cdot p, f_n \cdot p_1) + d(f_n \cdot p_1, f_m \cdot p_1) + d(f_m \cdot p_1, f_m \cdot p) < \epsilon \tag{1}$$

if $m, n \geqslant N$. Therefore, all $f_n \cdot p$ ($n \geqslant N$) lie inside the ball $B_\epsilon(f_N \cdot p)$ which has compact closure as well as $B_r(p)$. We can thus select a subsequence of $(f_n \cdot p)$ which converges to a limit, say p^*. From (1) we conclude that

$$d(f_n \cdot p, p^*) \leqslant \epsilon$$

for $n \geqslant N$ and the lemma is proved.

To prove Theorem 2.2, let S denote the set of points $q \in M$ for which the sequence $(f_n \cdot q)$ has compact closure. If (p_i) is a sequence in S and $(f_n{}^*)$ a subsequence of (f_n) we can, using a diagonal process, find a subsequence of $(f_n{}^*)$ which converges at each p_i. By Lemma 2.3 S is closed. We shall now prove that S is open. Since $o \in S$ and M is connected, this will prove that $S = M$.

Let $p^* \in S$ and choose $r > 0$ such that the ball $B_r(p^*)$ has compact closure. Let $p \in B_{r/4}(p^*)$ and let $(f_n{}^* \cdot p)$ be any subsequence of the sequence $(f_n \cdot p)$. There exists a subsequence (f_{n_μ}) of $(f_n{}^*)$ such that the sequence $(f_{n_\mu} \cdot p^*)$ converges to a limit q^* and such that the entire sequence $(f_{n_\mu} \cdot p^*)$ is contained in the ball $B_{r/4}(q^*)$. Then the sequence $(f_{n_\mu} \cdot p)$ is contained in the ball $B_{r/2}(q^*)$ which has compact closure since it is contained in each of the balls $f_{n_\mu} \cdot B_r(p^*)$. Hence $(f_n{}^* \cdot p)$ has a convergent subsequence. This proves that $S = M$.

Now M has a dense sequence of points, and using a diagonal process we can find a subsequence (φ_ν) of (f_n) which converges at all these points and by Lemma 2.3 on the entire M. The following lemma completes the proof of the theorem.

Lemma 2.4. *Let (φ_ν) be a sequence in $I(M)$ which converges pointwise on M to a mapping $f : M \to M$. Then $f \in I(M)$ and $\lim \varphi_\nu = f$ in the compact open topology.*

Proof. Let C be a compact subset of M and $\epsilon > 0$. We select points $p_1, \ldots, p_n$ such that each $p \in C$ has distance less than $\epsilon/3$ from some p_i. We can choose an integer N such that

$$d(\varphi_\nu \cdot p_i, f \cdot p_i) < \epsilon/3 \qquad \text{for } 1 \leqslant i \leqslant n, \qquad \nu > N.$$

If $p \in C$, we select p_j such that $d(p, p_j) < \epsilon/3$. Then, since f preserves distances, we have for $\nu > N$

$$d(\varphi_\nu \cdot p, f \cdot p) \leqslant d(\varphi_\nu \cdot p, \varphi_\nu \cdot p_j) + d(\varphi_\nu \cdot p_j, f \cdot p_j) + d(f \cdot p_j, f \cdot p) < \epsilon.$$

We finally prove that $f \cdot M = M$. Let $q \in f \cdot M$ and determine $q' \in M$ by $f(q') = q$. Then

$$0 = \lim d(q, \varphi_\nu \cdot q') = \lim d(\varphi_\nu^{-1} \cdot q, q').$$

Thus the sequence $\varphi_\nu^{-1} \cdot q$ converges to q'. We know from what is already proved, that there exists a subsequence (ψ_μ) of (φ_ν) such that $(\psi_\mu^{-1} \cdot p)$ converges for each $p \in M$. Let $p' = \lim (\psi_\mu^{-1} \cdot p)$. Then

$$\lim d(\psi_\mu \cdot p', p) = \lim d(p', \psi_\mu^{-1} \cdot p) = 0.$$

It follows that

$$p = \lim \psi_\mu p' = f \cdot p'.$$

Since $p \in M$ is arbitrary, we conclude that $fM = M$. From Theorem 11.1 Chapter I, we know that f is a diffeomorphism. Thus the lemma is proved.

Theorem 2.5.

(a) *Let M be a Riemannian manifold. The compact open topology of $I(M)$ turns $I(M)$ into a locally compact topological transformation group of M.*

(b) *Let $p \in M$ and let $\tilde{K}$ denote the subgroup of $I(M)$ which leaves p fixed. Then $\tilde{K}$ is compact.*

Proof. Let (f_n) be a sequence in $I(M)$ which converges to an element $f \in I(M)$. Then for each $p \in M$,

$$d(f_n^{-1} \cdot p, f^{-1} \cdot p) = d(p, f_n f^{-1} \cdot p) = d(f \cdot (f^{-1} \cdot p), f_n \cdot f^{-1} \cdot p),$$

which converges to 0 as $n \to \infty$. It follows from Lemma 2.4 that $\lim f_n^{-1} = f^{-1}$ in the compact open topology. Thus the inverse operation $f \to f^{-1}$ is continuous on $I(M)$. The continuity of the multiplication is proved similarly. Hence $I(M)$ is a topological group. Next, let $\lim f_n = f$ $(f_n \in I(M))$ and $\lim p_n = p$ $(p_n \in M)$. Let $\epsilon > 0$ be given and select an integer N such that $d(p_n, p) < \epsilon$ for $n \geqslant N$. The sequence $p_N, p_{N+1}, \ldots$ together with p form a compact set C. If $U = B_\epsilon(f \cdot p)$, then $f \in W(C, U)$. Let N_1 be an integer such that $f_n \in W(C, U)$ for $n \geqslant N_1$. Then $d(f_n \cdot p_n, f \cdot p) < \epsilon$ for $n \geqslant N_1 N$ so $I(M)$ is a topological transformation group of M. To show that $I(M)$ is locally compact and $\tilde{K}$ compact, let U be an open relatively compact neighborhood of p. Then $\tilde{K}$ is a closed subset of $W(\{p\}, U)$ and due to Theorem 2.2, $W(\{p\}, U)$ has compact closure. This finishes the proof.

§ 3. Riemannian Globally Symmetric Spaces

Let M be a Riemannian manifold with Riemannian structure Q; we recall that M is called an *analytic Riemannian manifold* if M and Q are both analytic. A mapping is called *involutive* if its square, but not the mapping itself, is the identity.

Definition. Let M be an analytic† Riemannian manifold; M is called *Riemannian globally symmetric* if each $p \in M$ is an isolated fixed point of an involutive isometry s_p of M.

Remark. It is obvious from the next lemma and Lemma 11.2, Chapter I, that there is only one such s_p.

Lemma 3.1. *Let M be Riemannian globally symmetric. For each $p \in M$ there exists a normal neighborhood N_p of $p \in M$ such that s_p is the geodesic symmetry on N_p.*

Let $A = (ds_p)_p$. Then $AM_p \subset M_p$ and $A^2 = I$. Writing

$$X = \tfrac{1}{2}(X - AX) + \tfrac{1}{2}(X + AX)$$

we see that $M_p = V^- + V^+$ (direct sum), where $V^{\pm} = \{X : AX = \pm X\}$. Suppose $X \neq 0$ belongs to V^+ and consider a geodesic γ tangent to X. Then s_p will leave γ pointwise fixed. This contradicts the assumption that p is an isolated fixed point. Thus $A = -I$ and the lemma follows.

Let M be a Riemannian globally symmetric space. Let γ be any geodesic in M. If $p \in \gamma$, $s_p \cdot \gamma$ gives an extension of γ, so that each maximal geodesic in M has infinite length. Thus M is complete and any two points $p, q \in M$ can be joined by a geodesic of length $d(p, q)$. If m is the midpoint of this geodesic then s_m interchanges p and q. In particular, the group $I(M)$ acts transitively on M. Owing to the theorems of §2, $I(M)$ has a countable base in the compact open topology and is a transitive, locally compact topological transformation group of M. Let $\tilde{K}$ be the subgroup of $I(M)$ which leaves some point p_0 of M fixed. Then $\tilde{K}$ is compact and due to Theorem 3.2 of Chapter II, $I(M)/\tilde{K}$ is homeomorphic to M under the mapping $g\tilde{K} \to g \cdot p_0$, $g \in I(M)$.

Lemma 3.2. *Let M be a Riemannian globally symmetric space. Then*

† This analyticity assumption is convenient; no loss of generality results from it as Prop. 5.5 shows.

$I(M)$ has an analytic structure compatible with the compact open topology in which it is a Lie transformation group of M.

Remark. The properties of $I(M)$ stated characterize its analytic structure uniquely. In fact, a topological group has at most one analytic structure compatible with its topology with respect to which it is a Lie group (Theorem 2.6, Chapter II).

Proof. Using the notation above, we consider the mapping $\sigma : k \rightarrow (dk)_{p_0}$ of $\tilde{K}$ into the orthogonal group $O(M_{p_0})$. Let $X_1, ..., X_m$ be an orthonormal basis of M_{p_0}, $\{x_1, ..., x_m\}$ the normal coordinate system with respect to this basis, valid on a convex normal ball $B_r(p_0)$. The expression of the mapping k in coordinates $(x_1, ..., x_m)$ is the same as the expression of $(dk)_{p_0}$ in terms of the Cartesian coordinates on M_{p_0}. Thus σ is continuous. Owing to Lemma 11.2, Chapter I, σ is one-to-one, hence a homeomorphism. The linear isotropy group $K^* = \sigma(\tilde{K})$ is a compact subgroup of $O(M_{p_0})$ and has therefore a unique differentiable structure compatible with the topology induced by $O(M_{p_0})$ in which it is a Lie subgroup of $O(M_{p_0})$. If we carry this differentiable structure over by σ^{-1}, $\tilde{K}$ becomes a compact Lie group.

Let π be the natural mapping $g \rightarrow g \cdot p_0$ of $I(M)$ onto M. We shall now construct a subset B of $I(M)$ (a certain local cross section), which π maps homeomorphically onto $B_r(p_0)$. Let $t \rightarrow p_t$ be a geodesic in $B_r(p_0)$ starting at p_0. For simplicity we put $s_{p_t} = s_t$. The mapping $T_t = s_{t/2}s_0$ is an isometry of M and sends p_0 into p_t. Owing to relation (6), §1, it is clear that $(dT_t)_{p_0}$ is the parallel translation from p_0 to p_t along the geodesic. Consider now the mapping $\psi : B_r(p_0) \rightarrow I(M)$ given by $\psi(p_t) = T_t$. The mapping ψ is of course one-to-one. In order to prove that ψ is continuous it suffices to prove that if a sequence $(q_n) \subset M$ converges to $q \in M$, then the corresponding symmetries s_{q_n} converge to s_q in $I(M)$. If p is sufficiently close to q then it is obvious that $(s_{q_n} \cdot p)$ converges to $s_q \cdot p$. Since an isometry is determined by its action on any open set it suffices, due to Lemma 2.4, to prove that the sequence $(s_{q_n} \cdot p)$ is convergent for each $p \in M$. Let S denote the set of points $p \in M$ for which $(s_{q_n} \cdot p)$ is convergent. The set S is open (Lemma 10.1, Chapter I) and not empty. It is also closed (Lemma 2.3), so $S = M$ and the continuity of ψ is established.

Let $B = \psi(B_r(p_0))$. The mapping π is one-to-one on B and $\pi \circ \psi = I$. Hence π is a homeomorphism of B onto $B_r(p_0)$. The set $B\tilde{K} = \{bk : b \in B, k \in \tilde{K}\}$ is the inverse image $\pi^{-1}(B_r(p_0))$, and is therefore an open subset of $I(M)$. Let $g \in B\tilde{K}$. Then $g = bk$ $(b \in B, k \in \tilde{K})$. It follows that $b = \psi(\pi(g))$ so the mapping $(b, k) \rightarrow bk$ is a homeomorphism of $B \times \tilde{K}$ onto $B\tilde{K}$. Hence if U is an open subset of $\tilde{K}$, the set BU

is open in $B\tilde{K}$, hence open in $I(M)$. In particular, let U be an open neighborhood of e in $\tilde{K}$ on which a system $\{y_1, ..., y_r\}$ of coordinates is valid. The mapping

$$\varphi_e : bu \to (x_1(\pi(b)), ..., x_m(\pi(b)), y_1(u), ..., y_r(u))$$

is a homeomorphism of BU onto an open subset of $\boldsymbol{R}^{m+r}$. For each $x \in I(M)$, the mapping $\varphi_x = \varphi_e \circ L_{x-1}$ is a homeomorphism of xBU onto an open subset of $\boldsymbol{R}^{m+r}$. In order that this should give an analytic structure on $I(M)$, it suffices to verify that $\varphi_x \circ \varphi_e^{-1}$ is analytic on $\varphi_e(BU \cap xBU)$. This, and the fact that $I(M)$ is a Lie group, will follow if we can prove the following statement:

If $b_1, b_2 \in B$, $u_1, u_2 \in U$ such that

$$b_1u_1b_2u_2 = bu,$$

where $b \in B$, $u \in U$, then the coordinates $y_\alpha(u)$, $x_i(\pi(b))$ are analytic functions of the coordinates $x_j(\pi(b_1))$, $x_k(\pi(b_2))$, $y_\beta(u_1), y_\gamma(u_2)$, $1 \leqslant i, j, k \leqslant m$, $1 \leqslant \alpha, \beta, \gamma \leqslant r$.

Let $k \in \tilde{K}$. Then the isometries ks_0k^{-1} and s_0 leave p_0 fixed and induce the same mapping of M_{p_0}. By Lemma 11.2, Chapter I, we have $ks_0k^{-1} = s_0$. In particular, if q is the midpoint of the geodesic from p_0 to $b_2 \cdot p_0$,

$$u_1b_2u_1^{-1} = u_1s_qs_0u_1^{-1} = u_1s_qu_1^{-1}s_0 = s_{q^*}s_0$$

if $q^* = u_1 \cdot q$. Hence $u_1b_2u_1^{-1}$ is an element b^* of B and the coordinates of b^* depend analytically on the coordinates of u_1 and b_2. We have now

$$b_1b^*u_1u_2 = bu$$

and

$$x_i(\pi(b_1b^*)) = x_i(\pi(b)), \qquad 1 \leqslant i \leqslant m.$$

Now, the point $b_1b^* \cdot p_0$ is determined from $b_1 \cdot p_0$ and $b^* \cdot p_0$ as follows: Let γ_1 and γ^* be the geodesics in $B_r(p_0)$ from p_0 to $b_1 \cdot p_0$ and from p_0 to $b^* \cdot p_0$. Let Y_1 and Y^* denote their unit tangent vectors at p_0. Let Y_3 be the parallel translate of Y^* along γ_1 and consider the geodesic γ_3 emanating from $b_1 \cdot p_0$ with tangent vector Y_3 and arc length $L(\gamma^*)$. The end point of γ_3 is $b_1b^* \cdot p_0$. This construction shows the analytic dependence of $x_i(\pi(b_1b^*))$ on $x_j(\pi(b_1))$ and $x_k(\pi(b^*))$, $1 \leqslant i, j, k \leqslant m$. Moreover, the coordinates of the tangent vector $d(u(u_1u_2)^{-1}) \cdot X_i = d(b^{-1}b_1b^*) \cdot X_i$ with respect to $X_1, ..., X_m$ depend analytically on the coordinates of b_1, b^*. Therefore, $\tilde{K}$ being a Lie group, it follows that the coordinates of u depend analytically on the coordinates of u_1, u_2, b_1, b^*. This shows that $I(M)$ is a Lie group. The argument above shows also that the mapping $(g, p) \to g \cdot p$ is an analytic mapping of $BU \times B_r(p_0)$ into M. Hence $I(M)$ is a Lie transformation group of M.

Theorem 3.3.

(i) *Let M be a Riemannian globally symmetric space and p_0 any point in M. If $G = I_0(M)$, and K is the subgroup of G which leaves p_0 fixed, then K is a compact subgroup of the connected group G and G/K is analytically diffeomorphic to M under the mapping $gK \to g \cdot p_0$, $g \in G$.*

(ii) *The mapping $\sigma : g \to s_{p_0} g s_{p_0}$ is an involutive automorphism of G such that K lies between the closed group K_σ of all fixed points of σ and the identity component of K_σ. The group K contains no normal subgroup of G other than $\{e\}$.*

(iii) *Let $\mathfrak{g}$ and $\mathfrak{k}$ denote the Lie algebras of G and K, respectively. Then $\mathfrak{k} = \{X \in \mathfrak{g} : (d\sigma)_e X = X\}$ and if $\mathfrak{p} = \{X \in \mathfrak{g} : (d\sigma)_e X = -X\}$ we have $\mathfrak{g} = \mathfrak{k} + \mathfrak{p}$ (direct sum). Let π denote the natural mapping $g \to g \cdot p_0$ of G onto M. Then $(d\pi)_e$ maps $\mathfrak{k}$ into $\{0\}$ and $\mathfrak{p}$ isomorphically onto M_{p_0}. If $X \in \mathfrak{p}$, then the geodesic emanating from p_0 with tangent vector $(d\pi)_e X$ is given by*

$$\gamma_{d\pi \cdot X}(t) = \exp tX \cdot p_0 \qquad (d\pi = (d\pi)_e).$$

Moreover, if $Y \in M_{p_0}$, then $(d \exp tX)_{p_0}(Y)$ is the parallel translate of Y along the geodesic.

Proof. The first part (i) follows from Prop. 4.3 in Chapter II. For (ii) and (iii) put $s_0 = s_{p_0}$. It is obvious that σ is an involutive automorphism of $I(M)$ and consequently maps the identity component G onto itself. If $k \in K$, the mappings k and $s_0 k s_0$ are isometries which induce the same mapping of M_{p_0}. From Lemma 11.2, Chapter I, we have $s_0 k s_0 = k$ for all $k \in K$. It follows that the automorphism $(d\sigma)_e$ of $\mathfrak{g}$ is identity on $\mathfrak{k}$. On the other hand, if $X \in \mathfrak{g}$ is left fixed by $(d\sigma)_e$, then $s_0 \exp tX s_0 = \exp tX$ for each $t \in \mathbf{R}$. This implies that $\exp tX \cdot p_0$ is a fixed point of s_0; since p_0 is an isolated fixed point of s_0 it follows that $\exp tX \cdot p_0 = p_0$ for all t so $X \in \mathfrak{k}$. Since $I(M)$ and G act effectively on G/K, K contains no normal subgroup $\neq \{e\}$ of G. This proves (ii).

The direct decomposition $\mathfrak{g} = \mathfrak{k} + \mathfrak{p}$ follows from the identity $X = \frac{1}{2}(X + d\sigma \cdot X) + \frac{1}{2}(X - d\sigma \cdot X)$. In the proof of Prop. 4.3, Chapter II, it is shown that $(d\pi)_e$ is a linear mapping of $\mathfrak{g}$ onto M_{p_0} with kernel $\mathfrak{k}$.

Finally, let $X \in \mathfrak{p}$; put $p_t = \gamma_{d\pi \cdot X}(t)$ and set as before $s_t = s_{p_t}$, $T_t = s_{t/2} s_0$. We have seen that $(dT_t)_{p_0}$ is the parallel translation along the geodesic from p_0 to p_t. Moreover, $s_\tau s_0 s_t = s_{\tau+t}$ because both sides of the equation are isometries which leave the point $p_{\tau+t}$ fixed and induce the same mapping in the tangent space $M_{p_{\tau+t}}$. It follows that $T_{2\tau+2t} = T_{2\tau}T_{2t}$ for all $\tau, t \in \boldsymbol{R}$. If t is sufficiently small, T_t lies in the local cross section B used in the proof of Lemma 3.2. Now π is an analytic diffeomorphism of B onto the normal neighborhood $B_r(p_0)$ and $\pi T_t = p_t$. Since the mapping $t \to p_t$ is analytic, it follows that $t \to T_t$ is a one-parameter subgroup. Hence $T_t = \exp tZ$ where $Z \in \mathfrak{g}$, $t \in \boldsymbol{R}$. Now $\sigma T_t = s_0 s_{t/2} = s_{-t/2}s_0 = T_{-t}$. Thus $d\sigma Z = -Z$ so $Z \in \mathfrak{p}$. But $\pi T_t = p_t$ so $d\pi Z = d\pi X$. Consequently, $X = Z$ and the theorem is proved.

Definition. Let G be a connected Lie group and H a closed subgroup. The pair (G, H) is called a *symmetric pair* if there exists an involutive analytic automorphism σ of G such that $(H_\sigma)_0 \subset H \subset H_\sigma$, where H_σ is the set of fixed points of σ and $(H_\sigma)_0$ is the identity component of H_σ.

If, in addition, the group† $\mathrm{Ad}_G(H)$ is compact, (G, H) is said to be a *Riemannian symmetric pair*.

As usual, we consider G as a Lie transformation group of the coset space G/H where each $g \in G$ gives rise to the diffeomorphism $\tau(g)$: $xH \to gxH$ of G/H onto itself.

Proposition 3.4. *Let (G, K) be a Riemannian symmetric pair. Let π denote the natural mapping of G onto G/K and put $o = \pi(e)$. Let σ be any analytic, involutive automorphism of G such that $(K_\sigma)_0 \subset K \subset K_\sigma$. In each G-invariant Riemannian structure Q on G/K (such Q exist) the manifold G/K is a Riemannian globally symmetric space. The geodesic symmetry s_o satisfies*

$$s_o \circ \pi = \pi \circ \sigma,$$

$$\tau(\sigma(g)) = s_o \tau(g)\, s_o, \qquad g \in G;$$

in particular, s_o is independent of the choice of Q.

Remark 1. We shall see later, that the Riemannian connection on G/K is independent of the choice of Q.

Proof. Let σ be an arbitrary analytic involutive automorphism of G such that $(K_\sigma)_0 \subset K \subset K_\sigma$. For simplicity we shall write $d\sigma$ and $d\pi$ instead of $(d\sigma)_e$ and $(d\pi)_e$. Let $\mathfrak{g}$ and $\mathfrak{k}$ denote the Lie algebras of G and K, respectively, and put $\mathfrak{p} = \{X \in \mathfrak{g} : d\sigma X = -X\}$. Then $\mathfrak{g} = \mathfrak{k} + \mathfrak{p}$ (direct sum). If $X \in \mathfrak{p}$ and $k \in K$, then $\sigma(\exp \mathrm{Ad}(k)\, tX) = k \exp(-tX)k^{-1}$

† Here $\mathrm{Ad}_G(H)$ means the Lie subgroup of $\mathrm{Ad}_G(G)$ which is the image of H under Ad_G. As in Prop. 5.4, Chapter II, we see that if $\mathrm{Ad}_G(H)$ is compact, then it is compact in the ordinary matrix topology.

so $d\sigma \operatorname{Ad}(k) X = -\operatorname{Ad}(k) X$. Thus $\mathfrak{p}$ is invariant under $\operatorname{Ad}_G(K)$. The mapping $d\pi$ maps $\mathfrak{g}$ onto T_o, the tangent space to G/K at o, and the kernel of $d\pi$ is $\mathfrak{k}$. The resulting isomorphism of $\mathfrak{p}$ onto T_o commutes with the action of K, that is,

$$d\pi \cdot \operatorname{Ad}(k) X = d\tau(k) \cdot d\pi(X), \qquad k \in K, X \in \mathfrak{p}. \tag{1}$$

In fact, this formula is an immediate consequence of the relation

$$\pi(\exp \operatorname{Ad}(k)\, tX) = \pi(k \exp tX\, k^{-1}) = \tau(k)\, \pi(\exp tX).$$

Since $\operatorname{Ad}_G(K)$ is a compact group in the relative topology of $\boldsymbol{GL}(\mathfrak{g})$, there exists a strictly positive definite quadratic form B on $\mathfrak{p}$ invariant under $\operatorname{Ad}_G(K)$. Then the form $Q_o = B \circ (d\pi)^{-1}$ on T_o is invariant under all the mappings $d\tau(k)$, $k \in K$. Let the corresponding symmetric bilinear form on $T_o \times T_o$ also be denoted by Q_o. For each $p \in G/K$ we define the bilinear form Q_p on $(G/K)_p \times (G/K)_p$ by

$$Q_p(d\tau(g)\, X_0, d\tau(g)\, Y_0) = Q_o(X_0, Y_0), \qquad X_0, Y_0 \in T_o,$$

where $g \in G$ is chosen such that $g \cdot o = p$. The invariance of B under $\operatorname{Ad}_G(K)$ guarantees that Q_p is well defined. Since each $\tau(g)$, $g \in G$, is an analytic diffeomorphism of G/K it follows that $p \to Q_p$ is an analytic Riemannian structure on G/K, invariant under the action of G. On the other hand, each G-invariant Riemannian structure on G/K arises in this fashion from an invariant quadratic form on $\mathfrak{p}$.

We now define a mapping s_o of G/K onto itself by the condition $s_o \circ \pi = \pi \circ \sigma$. Then s_o is an involutive diffeomorphism of G/K onto itself and $(ds_o)_o = -I$. To see that s_o is an isometry, let $g \in G$, $p = \tau(g) \cdot o$, and $X, Y \in (G/K)_p$. Then the vectors $X_0 = d\tau(g^{-1})\, X$, $Y_0 = d\tau(g^{-1})\, Y$ belong to T_o. The formula $s_o \circ \pi = \pi \circ \sigma$ implies for each $x \in G$,

$$s_o \circ \tau(g)\,(xK) = \sigma(gx)\, K = \sigma(g)\, \sigma(x)\, K = (\tau(\sigma(g)) \circ s_o)\,(xK),$$

so $s_o \circ \tau(g) = \tau(\sigma(g)) \circ s_o$. Hence

$$\begin{aligned} Q(ds_o\, X, ds_o\, Y) &= Q(ds_o\, d\tau(g)\, X_0, ds_o d\tau(g)\, Y_0) \\ &= Q(ds_o\, X_0, ds_o\, Y_0) = Q(X_0, Y_0) = Q(X, Y). \end{aligned}$$

Thus s_o is an isometry and near o it must coincide with the geodesic symmetry. For an arbitrary point $p = \tau(g) \cdot o$ in G/K, the geodesic symmetry is given by

$$s_p = \tau(g) \circ s_o \circ \tau(g^{-1}).$$

This being an isometry, the space G/K is a Riemannian globally symmetric space. This finishes the proof.

The formula $s_o \circ \pi = \pi \circ \sigma$ shows that the geodesic symmetry on G/K is the same for all G-invariant metrics.

We shall now derive some further properties of the Riemannian symmetric pair (G, K). Let Z denote the center of G and let N denote the set of $n \in G$ for which $\tau(n)$ is the identity mapping of G/K. Then Z and N are closed normal subgroups of G and $N \subset K$. Due to Lemma 5.1, Chapter II, the group $K/K \cap Z$ and the linear group $\mathrm{Ad}_G(K)$ are analytically isomorphic. Hence $K/K \cap Z$ is compact and since $K \cap Z \subset N$, K/N is compact. Let $I(G/K)$ denote the group of all isometries of G/K (with the Riemannian structure Q), and let $\tilde{K}$ denote the subgroup of $I(G/K)$ which leaves o fixed. Then, by Lemma 3.2 and Theorem 2.5, $I(G/K)$ and $\tilde{K}$ are Lie groups, $\tilde{K}$ compact.

Consider now the (algebraic) isomorphism $\beta : gN \to \tau(g)$ of G/N into $I(G/K)$. If a sequence (g_nN) converges to gN in G/N, then (g_nxN) converges to gxN for each $x \in G$ and therefore g_nxK converges to gxK in G/K. In view of Lemma 2.4 this proves the continuity of β. The restriction of β to K/N is a homeomorphism.

Remark 2. The group $\beta(G/N)$ is a closed subgroup of $I(G/K)$.

In fact, let $K_1 = \beta(K/N)$ and $G_1 = \beta(G/N)$. Then K_1 is a compact topological subgroup of $\tilde{K}$ and if the analytic structure of G/N is carried over on G_1 by β, G_1 is a Lie transformation group of G/K. Let (g_n) be a sequence in G_1 which converges in $I(G/K)$ to an element $g \in I(G/K)$. The sequence $(g_n \cdot o)$ converges to the point $p = g \cdot o$ in G/K. Select $g^* \in G_1$ such that $g^* \cdot o = p$. There exists a local cross section in G_1 through g^*, that is a submanifold B^* of G_1 containing g^* such that the natural mapping $x \to x \cdot o$ of G_1 onto G/K is a diffeomorphism of B^* onto an open neighborhood of p in G/K. If n is sufficiently large, there exists an element $k_n \in K_1$ such that $g_nk_n \in B^*$. It is clear that the sequence (g_nk_n) in B^* converges to g^*, and since K_1 is compact we may assume that the sequence (k_n) is convergent in K_1. The imbeddings $B^* \to G_1$ and $K_1 \to G_1$ being continuous the sequences (g_nk_n) and (k_n) converge in G_1. It follows that (g_n) converges in G_1. Finally, since the imbedding $G_1 \subset I(G/K)$ is continuous, $g \in G_1$, so G_1 is closed.

Now it follows from Theorem 2.3, Chapter II, that G_1 has a unique analytic structure in which it is a topological Lie subgroup G_2 of $I(G/K)$. The identity mapping $G_1 \to G_2$ is continuous, hence a homeomorphism (Cor. 3.3, Chapter II), hence an analytic isomorphism (Theorem 2.6, Chapter II). This means that β is an analytic isomorphism of G/N onto a closed, topological Lie subgroup of $I(G/K)$.

Let $\tilde{\mathfrak{g}}$ and $\tilde{\mathfrak{k}}$ denote the Lie algebras of $I(G/K)$ and $\tilde{K}$, respectively.

From Theorem 3.3 we obtain a subspace $\tilde{\mathfrak{p}} \subset \tilde{\mathfrak{g}}$ such that $\tilde{\mathfrak{g}} = \tilde{\mathfrak{k}} + \tilde{\mathfrak{p}}$ (direct sum) and

$$\tilde{\pi}(\exp X) = \operatorname{Exp} d\tilde{\pi}(X), \qquad X \in \tilde{\mathfrak{p}}, \tag{2}$$

$\tilde{\pi}$ denoting the natural mapping $g \to g \cdot o$ of $I(G/K)$ onto G/K. Let $\tilde{\sigma}$ denote the involutive automorphism $g \to s_o g s_o$ of $I(G/K)$ and let π_1 denote the natural mapping of G onto G/N. Then $\beta \circ \pi_1 = \tau$ and by Prop. 3.4

$$\beta(\pi_1(\sigma(g))) = \tau(\sigma(g)) = \tilde{\sigma}(\tau(g)), \qquad (g \in G),$$

so

$$d\beta \circ d\pi_1 \circ d\sigma = d\tilde{\sigma} \circ d\tau.$$

Consequently, the mapping $d\beta \circ d\pi_1$ maps $\mathfrak{p}$ onto $\tilde{\mathfrak{p}}$. Since $\tilde{\pi} \circ \beta \circ \pi_1 = \pi$ we get from (2) for $X \in \mathfrak{p}$

$$\begin{aligned} \pi(\exp X) &= \tilde{\pi}(\beta(\pi_1(\exp X)) = \tilde{\pi}(\exp d\beta(d\pi_1(X)) \\ &= \operatorname{Exp}\,(d\tilde{\pi}(d\beta(d\pi_1(X)) \end{aligned}$$

so

$$\pi(\exp X) = \operatorname{Exp}\,(d\pi(X)) \qquad X \in \mathfrak{p}. \tag{3}$$

Thus the geodesics in G/K are still orbits of suitable one-parameter subgroups in G.

We shall now prove that under very general conditions, the automorphism σ is completely determined by its fixed points.

Proposition 3.5. *Let (G, K) be a Riemannian symmetric pair. Let $\mathfrak{k}$ denote the Lie algebra of K and let $\mathfrak{z}$ denote the Lie algebra of the center of G. Assume that $\mathfrak{k} \cap \mathfrak{z} = \{0\}$. Then there exists exactly one involutive, analytic automorphism σ of G such that $(K_\sigma)_0 \subset K \subset K_\sigma$.*

Proof. Let σ_1, σ_2 be two automorphisms with the described properties. Then the Lie algebra $\mathfrak{g}$ of G has direct decompositions $\mathfrak{g} = \mathfrak{k} + \mathfrak{p}_i$ where $\mathfrak{p}_i$ is the eigenspace for the eigenvalue -1 of the automorphism $d\sigma_i$ $(i = 1, 2)$. Since the Killing form B of $\mathfrak{g}$ is invariant under σ_i, it follows that $\mathfrak{k}$ is orthogonal to $\mathfrak{p}_i$ with respect to B $(i = 1, 2)$. Let $X_1 \in \mathfrak{p}_1$. Then there exists an element $X_2 \in \mathfrak{p}_2$ such that $X_1 = T + X_2$ where $T \in \mathfrak{k}$. It follows that T is orthogonal to $\mathfrak{k}$. Since $\mathfrak{k} \cap \mathfrak{z} = \{0\}$, B is strictly negative definite on $\mathfrak{k}$ (Prop. 6.8, Chapter II). Thus $T = 0$ and $\mathfrak{p}_1 = \mathfrak{p}_2$.

For further study of symmetric spaces it is important to express the symmetry conditions in terms of Lie algebras rather than in terms of the groups. In Theorem 3.3 we have seen that a Riemannian globally symmetric space gives rise to a pair $(\mathfrak{g}, s)$ where:

(i) $\mathfrak{g}$ is a Lie algebra over $\boldsymbol{R}$.

(ii) s is an involutive automorphism of $\mathfrak{g}$.

(iii) $\mathfrak{k}$, the set of fixed points of s, is a compactly imbedded subalgebra of $\mathfrak{g}$.

(iv) $\mathfrak{k} \cap \mathfrak{z} = \{0\}$ if $\mathfrak{z}$ denotes the center of $\mathfrak{g}$.

Definition. A pair $(\mathfrak{g}, s)$ with the properties (i), (ii), (iii) is called an *orthogonal symmetric Lie algebra*. It is said to be *effective* if, in addition, (iv) holds. A *pair* (G, K), where G is a connected Lie group with Lie algebra $\mathfrak{g}$, and K is a Lie subgroup of G with Lie algebra $\mathfrak{k}$, is said to be *associated* with the orthogonal symmetric Lie algebra $(\mathfrak{g}, s)$.

Remark 1. If $\mathfrak{k}$ is a compactly imbedded subalgebra of a Lie algebra $\mathfrak{g}$ with center $\mathfrak{z}$ and $\mathfrak{k} \cap \mathfrak{z} = \{0\}$, then by the proof of Prop. 3.5 there exists at most one involutive automorphism of $\mathfrak{g}$ whose fixed point set is $\mathfrak{k}$.

Remark 2. It is important to distinguish between a symmetric pair as defined earlier (before Prop. 3.4) and a pair associated with an orthogonal symmetric Lie algebra.

Proposition 3.6. *Let* $(\mathfrak{g}, s)$ *be an orthogonal symmetric Lie algebra,* $\mathfrak{k}$ *the set of fixed points of* s. *Let* (G, K) *and* $(\tilde{G}, \tilde{K})$ *be two pairs associated with* $(\mathfrak{g}, s)$. *Suppose* K *and* $\tilde{K}$ *are connected and* $\tilde{G}$ *simply connected. Then* $\tilde{K}$ *is closed and* $(\tilde{G}, \tilde{K})$ *is a Riemannian symmetric pair. If* K *is closed in* G *(this is the case if the center of* $\mathfrak{g}$ *is* $\{0\}$*), then* G/K *is Riemannian locally symmetric for each* G*-invariant metric (such exist) and* $\tilde{G}/\tilde{K}$ *is the universal covering manifold of* G/K.

Proof. Since $\tilde{G}$ is simply connected, there exists an analytic homomorphism $\sigma : \tilde{G} \to \tilde{G}$ for which $(d\sigma)_e = s$. Since s is an involutive automorphism the same is true of σ. The group $\tilde{K}$ is the identity component of the group of fixed points of σ. In particular $\tilde{K}$ is closed in $\tilde{G}$. The space $\tilde{G}/\tilde{K}$ is simply connected. In fact, let $\gamma(t)$, $0 \leqslant t \leqslant 1$, be a continuous closed curve in $\tilde{G}/\tilde{K}$. Without loss of generality we can assume that $\gamma(0) = \gamma(1) = \tilde{\pi}(e)$ ($\tilde{\pi}$ being the natural mapping of $\tilde{G}$ onto $\tilde{G}/\tilde{K}$). Using local cross sections in $\tilde{G}$ (Lemma 4.1, Chapter II), it is easy to find a continuous curve $\tilde{\gamma}(t)$, $0 \leqslant t \leqslant 1$, in $\tilde{G}$ such that $\tilde{\pi}(\tilde{\gamma}(t)) = \gamma(t)$ for $0 \leqslant t \leqslant 1$. Then $\tilde{\gamma}(0)$ and $\tilde{\gamma}(1)$ belong to $\tilde{K}$ and can be joined by a continuous curve, $\tilde{K}$ being connected. The closed curve in $\tilde{G}$, so obtained, is homotopic to a point in $\tilde{G}$. It follows that the projection $\gamma(t)$ is also homotopic to a point in $\tilde{G}/\tilde{K}$.

The groups $\mathrm{Ad}_G(K)$ and $\mathrm{Ad}_{\tilde{G}}(\tilde{K})$ coincide because they are both analytic subgroups of $\mathrm{Int}(\mathfrak{g})$ and have the same Lie algebra. They are compact (and thus carry the relative topology of $GL(\mathfrak{g})$) since $\mathfrak{k}$ is compactly imbedded in $\mathfrak{g}$. The space $\mathfrak{p} = \{X \in \mathfrak{g} : s \cdot X = -X\}$ is invariant under $\mathrm{Ad}_G(K)$ and has a strictly positive definte quadratic form B invariant under $\mathrm{Ad}_G(K)$. As before, this form gives rise to a $\tilde{G}$-invariant Riemannian structure on $\tilde{G}/\tilde{K}$, and, if K is closed, to a G-invariant Riemannian structure on G/K. Let φ be the homomorphism of $\tilde{G}$ onto G such that $(d\varphi)_e$ is the identity mapping of $\mathfrak{g}$. Let $K^\natural$ denote the inverse image $\varphi^{-1}(K)$. Then $\tilde{K}$ is the identity component of $K^\natural$ and $\tilde{G}/\tilde{K}$ is a covering space (see Chevalley [2], p. 58) of $\tilde{G}/K^\natural$. If ψ denotes the mapping $g\tilde{K} \to \varphi(g)K$ of $\tilde{G}/\tilde{K}$ onto G/K, then the pair $(\tilde{G}/\tilde{K}, \psi)$ is the simply connected covering manifold of G/K (Lemma 13.4, Chapter I). Moreover, ψ is a local isometry. Since $\tilde{G}/\tilde{K}$ is globally symmetric, G/K is locally symmetric.

Let π denote the natural mapping of G onto G/K. Since one-parameter subgroups in $\tilde{G}$ and G correspond under φ and since geodesic in $\tilde{G}/\tilde{K}$ and G/K correspond under ψ, we obtain from (3),

$$\pi(\exp X) = \mathrm{Exp}\, d\pi X \qquad \text{for } X \in \mathfrak{p}, \tag{4}$$

exp and Exp referring to G and G/K, respectively. Relation (1) also holds here and allows us to identify $\mathfrak{p}$ and $(G/K)_{\pi(e)}$ whenever this is convenient.

Finally, let K^* denote the complete inverse image $\mathrm{Ad}_G^{-1}(\mathrm{Ad}_G(K))$. Since $\mathrm{Ad}_G(K)$ is closed in $\mathrm{Int}(\mathfrak{g})$, K^* is closed in G. If $\mathfrak{k}^*$ denotes the Lie algebra of K^* we have $\mathrm{ad}_{\mathfrak{g}}(\mathfrak{k}^*) = \mathrm{ad}_{\mathfrak{g}}(\mathfrak{k})$. If $\mathfrak{g}$ has center $\{0\}$, it follows that $\mathfrak{k}^* = \mathfrak{k}$; consequently K is the identity component of K^*, hence closed.

§ 4. The Exponential Mapping and the Curvature

The notation in this section will be as follows: Let $(\mathfrak{g}, s)$ be an orthogonal symmetric Lie algebra, $\mathfrak{k}$ the set of fixed points of s, and $\mathfrak{p}$ the subspace $\{X \in \mathfrak{g} : sX = -X\}$. For $X \in \mathfrak{p}$, let T_X denote the restriction of $(\mathrm{ad}\, X)^2$ to $\mathfrak{p}$. Then $T_X\mathfrak{p} \subset \mathfrak{p}$. Suppose the pair (G, K) is associated with $(\mathfrak{g}, s)$ and suppose that K is connected and closed in G. Let π be the natural mapping of G onto G/K and put $o = \pi(e)$. For $g \in G$, let $\tau(g)$ denote the mapping $xK \to gxK$ of G/K onto itself. The subspace $\mathfrak{p}$ will be identified with the tangent space $(G/K)_o$ by means of the mapping $d\pi$. Let Q be any G-invariant Riemannian structure on G/K. Then G/K is complete and locally symmetric.

We shall now describe the geometric concepts Exponential mapping and curvature for G/K in group theoretic terms.

Theorem 4.1. *The Exponential mapping of* $\mathfrak{p}$ *into* G/K *is independent of the choice of* Q. *Its differential is given by*

$$d\,\mathrm{Exp}_X = d\tau(\exp X)_o \circ \sum_{n=0}^{\infty} \frac{(T_X)^n}{(2n+1)!}, \qquad X \in \mathfrak{p}.$$

Here $\mathfrak{p}$ *is considered as a manifold in the usual way and whose tangent space at each point is identified with* $\mathfrak{p}$ *itself.*

Proof. Let $X, Y \in \mathfrak{p}$. From (4), §3, we have $\pi(\exp X) = \mathrm{Exp}\, X$ so from Theorem 1.7, Chapter II, we obtain

$$d\,\mathrm{Exp}_X(Y) = d\pi \circ d\exp_X(Y) = d\pi \circ dL_{\exp X} \circ \frac{1 - e^{-\mathrm{ad}\, X}}{\mathrm{ad}\, X}(Y)$$

$$= d\tau(\exp X) \circ d\pi \circ \sum_{m=0}^{\infty} \frac{(-\,\mathrm{ad}\, X)^m}{(m+1)!}(Y).$$

From the relations $[\mathfrak{k}, \mathfrak{p}] \subset \mathfrak{p}$, $[\mathfrak{p}, \mathfrak{p}] \subset \mathfrak{k}$ it follows that $d\pi\,(\mathrm{ad}\, X)^m(Y)$ is equal to $T_X^n(Y)$ if $m = 2n$ and 0 if m is odd. This proves the theorem.

Theorem 4.2. *Let* R *denote the curvature tensor of the space* G/K *corresponding to the Riemannian structure* Q. *Then, at the point* $o \in G/K$,

$$R_o(X, Y)Z = -\,[[X, Y], Z] \qquad \text{for } X, Y, Z \in \mathfrak{p}.$$

Proof. First we evaluate the sectional curvature directly and then use results from §12, Chapter I, to find the curvature tensor.

Assuming of course that $\dim \mathfrak{p} > 1$, let S be a two-dimensional subspace of $\mathfrak{p}$ and let $X_1, X_2, ..., X_m$ be an orthonormal basis of $\mathfrak{p}$ such that X_1 and X_2 belong to S. Each $X \in S$ can be written $X = x_1X_1 + x_2X_2$ where $x_1, x_2 \in \boldsymbol{R}$ and the Laplacian Δ on S is given by

$$\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2}.$$

We also put

$$A_X = \sum_{n=0}^{\infty} \frac{T_X^n}{(2n+1)!}, \qquad v_1 = A_X(X_1), \qquad v_2 = A_X(X_2).$$

If a and b are two vectors in a metric vector space, we denote by $a \vee b$ the parallelogram spanned by a and b and by $| a \vee b |$ the area.

Let N_0 be a normal neighborhood of 0 in the tangent space $\mathfrak{p}$. The submanifold $M_S = \operatorname{Exp}(S \cap N_0)$ of $M = G/K$ has a Riemannian structure induced by Q; a curve in M_S has the same arc length whether it is considered as a curve in M_S or M. If $p \in M_S$, the unique geodesic in $\operatorname{Exp} N_0$ from o to p is the shortest curve in M_S joining o and p. It follows that the Exponential mappings at o, for M_S and M, respectively, coincide on $S \cap N_0$. Thus, if $X \in S \cap N_0$, the vectors $d\tau(\exp X) \cdot v_1$ and $d\tau(\exp X) \cdot v_2$ are tangent vectors to M_S at $\operatorname{Exp} X$; the ratio of the surface elements in M_S and $S \cap N_0$ is therefore given by

$$f(X) = \frac{| d\tau(\exp X) v_1 \vee d\tau(\exp X) v_2 |}{| X_1 \vee X_2 |} = | v_1 \vee v_2 |,$$

since $\tau(\exp X)$ is an isometry of G/K. According to Lemma 12.1, Chapter I, the sectional curvature of M along the plane section S is given by

$$K(S) = -\frac{3}{2}\left[\left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2}\right) f\right](0, 0).$$

Let (A_{ij}) be the matrix expressing A_X in terms of the basis $X_1, \ldots, X_m$ of $\mathfrak{p}$, $A_X X_j = \sum_i A_{ij} X_i$. Then

$$\begin{aligned} f(X) &= | (A_{11}X_1 + \ldots + A_{m1}X_m) \vee (A_{12}X_1 + \ldots + A_{m2}X_m) | \\ &= \left\{ (A_{11}A_{22} - A_{12}A_{21})^2 + \sum_{1<i<j\leq m} (A_{i1}A_{j2} - A_{j1}A_{j2})^2 \right\}^{1/2}, \end{aligned}$$

since $| A_{i1}A_{j2} - A_{j1}A_{i2} |$ is the area of the projection of $v_1 \vee v_2$ on the (X_i, X_j)-plane. In computing $[\Delta f](0)$ from the expression for $f(X)$, we only have to consider terms of second order in x_1 or x_2. If $i \neq j$, A_{ij} only contains terms of second order and higher. Hence

$$[\Delta f](0) = [\Delta A_{11}A_{22}](0).$$

Now the matrix elements T_{ij} of $T_X = (\operatorname{ad}(x_1X_1 + x_2X_2))^2$ are of second order in x_1 and x_2. It follows that (writing Q for Q_o)

$$\begin{aligned} [\Delta f](0) &= \frac{1}{3!}[\Delta(T_{11} + T_{22})](0) = \frac{1}{3!}[\Delta\{Q(T_X X_1, X_1) + Q(T_X X_2, X_2)\}](0) \\ &= \frac{1}{3}\{Q(T_{X_1}X_1, X_1) + Q(T_{X_2}X_1, X_1) + Q(T_{X_1}X_2, X_2) + Q(T_{X_2}X_2, X_2)\}. \end{aligned}$$

In this expression the first and the last terms vanish. The two other terms are equal because of the invariance

$$Q_o(\operatorname{Ad}(k) X, \operatorname{Ad}(k) Y) = Q_o(X, Y), \qquad X, Y \in \mathfrak{p},$$

which implies

$$Q_o([Z, X], Y) + Q_o(X, [Z, Y]) = 0, \qquad X, Y \in \mathfrak{p},\ Z \in \mathfrak{k}. \tag{1}$$

We have thus proved

$$K(S) = -Q_o(T_{X_1}X_2, X_2) = Q_o(\operatorname{ad}([X_1, X_2]) \cdot X_1, X_2).$$

To prove the formula for R we consider the quadrilinear form

$$B(X, Y, Z, T) = Q_o((R(X, Y) + \operatorname{ad}[X, Y]) \cdot Z, T) \qquad (X, Y, Z, T \in \mathfrak{p}).$$

In view of Theorem 12.2, Chapter I, we know then that

$$B(X_1, X_2, X_1, X_2) = 0 \tag{2}$$

if X_1, X_2 are orthonormal vectors in $\mathfrak{p}$. We also have

$$B(X, Y, Z, T) = -B(Y, X, Z, T), \tag{3}$$

$$B(X, Y, Z, T) = -B(X, Y, T, Z), \tag{4}$$

$$B(X, Y, Z, T) + B(Y, Z, X, T) + B(Z, X, Y, T) = 0. \tag{5}$$

Here we have used the Jacobi identity and Lemma 12.5, Chapter I. Now if X, Y are arbitrary vectors in $\mathfrak{p}$, there exist orthonormal vectors X_1,X_2 in $\mathfrak{p}$ such that

$$X = x_1X_1 + x_2X_2, \qquad Y = y_1X_1 + y_2X_2, \qquad x_1, x_2, y_1, y_2 \in \mathbf{R}.$$

Then, using (3) and (4) we get

$$B(X, Y, X, Y) = (x_1y_2 - x_2y_1)^2\, B(X_1, X_2, X_1, X_2)$$

so

$$B(X, Y, X, Y) = 0 \qquad \text{for } X, Y \in \mathfrak{p}.$$

From Lemma 12.4, Chapter I we can conclude that $B \equiv 0$ and this proves Theorem 4.2.

The result shows that the curvature tensor is independent of the G-invariant Riemannian structure Q. In view of Lemma 1.2 we get the following:

Corollary 4.3. *The Riemannian connection on G/K is the same for all G-invariant Riemannian structures Q on G/K.*

§ 5. Locally and Globally Symmetric Spaces

Let M be a Riemannian manifold, p a point in M. In general it is impossible to find any neighborhood N of p which can be extended to a complete Riemannian manifold $\tilde{M}$. However, if M is locally symmetric then this turns out to be possible and $\tilde{M}$ can be taken globally symmetric. We shall also establish another relation between locally and globally symmetric spaces, namely that the universal covering manifold of a complete Riemannian locally symmetric space is globally symmetric.

Theorem 5.1. *Let M be a Riemannian locally symmetric space and p a point in M. There exists a Riemannian globally symmetric space $\tilde{M}$, an open neighborhood N_p of p in M and an isometry φ of N_p onto an open neighborhood of $\varphi(p)$ in $\tilde{M}$.*

The proof below is broken up into a few lemmas. Let the Riemannian structure of M be denoted by Q and let R denote the curvature tensor. Let $\mathfrak{p}$ denote the tangent space M_p. If A is an endomorphism of $\mathfrak{p}$, then A can be uniquely extended to the mixed tensor algebra $\mathfrak{D}(p)$ over $\mathfrak{p}$ as a derivation, preserving type of tensors and commuting with contractions. Denoting this extension again by A we have, if $X \in M_p$, $\omega \in M_p^*$,

$$A(\omega \otimes X) = A\omega \otimes X + \omega \otimes AX.$$

Applying contractions and noting that A annihilates scalars, we get

$$(A\omega)(X) = -\omega(AX).$$

It follows easily that

$$(A \cdot Q_p)(X, Y) = -Q_p(AX, Y) - Q_p(X, AY), \tag{1}$$

$$(A \cdot R_p)(X, Y) = [A, R_p(X, Y)] - R_p(AX, Y) - R_p(X, AY) \tag{2}$$

for $X, Y \in \mathfrak{p}$. Here the bracket $[E, F]$ of two endomorphisms denotes the endomorphism $EF - FE$.

Lemma 5.2. *Let $\mathfrak{k}$ denote the set of all endomorphisms of $\mathfrak{p}$, which, when extended to the mixed tensor algebra $\mathfrak{D}(p)$ as above, annihilate Q_p and R_p. Then $\mathfrak{k}$ is a Lie algebra with the bracket $[A, B] = AB - BA$; further, $R_p(X, Y) \in \mathfrak{k}$ for any $X, Y \in \mathfrak{p}$.*

Proof. By (1) and (2) above, $A \in \mathfrak{k}$ if and only if

$$Q_p(AX, Y) + Q_p(X, AY) = 0, \tag{3}$$

$$[A, R_p(X, Y)] = R_p(AX, Y) + R_p(X, AY) \tag{4}$$

for all $X, Y \in \mathfrak{p}$. We express (3) by saying that A is skew symmetric with respect to Q_p. Now, suppose $A, B \in \mathfrak{k}$; then

$$Q_p((AB - BA)X, Y) + Q_p(X, (AB - BA)Y)$$
$$= -Q_p(BX, AY) + Q_p(AX, BY) - Q_p(AX, BY) + Q_p(BX, AY) = 0,$$

so

$$Q_p([A, B] \cdot X, Y) + Q_p(X, [A, B] \cdot Y) = 0.$$

Similarly, by (4) and the Jacobi identity

$$\begin{aligned}[][A, B], R_p(X, Y)] &= -[[B, R_p(X, Y)], A] - [[R_p(X, Y), A], B] \\ &= [A, R_p(BX, Y) + R_p(X, BY)] - [B, R_p(AX, Y) + R_p(X, AY)] \\ &= R_p(ABX, Y) + R_p(BX, AY) + R_p(AX, BY) + R_p(X, ABY) \\ &\quad - R_p(BAX, Y) - R_p(AX, BY) - R_p(BX, AY) - R_p(X, BAY),\end{aligned}$$

so

$$[[A, B], R_p(X, Y)] = R_p([A, B] \cdot X, Y) + R_p(X, [A, B] \cdot Y).$$

Now, if $X, Y \in \mathfrak{p}$, the endomorphism $R_p(X, Y)$ of $\mathfrak{p}$ is skew symmetric with respect to Q_p (Lemma 12.5, Chapter I). Let X^* and Y^* be any vector fields on M such that $X_p^* = X$, $Y_p^* = Y$. In general, if D_1, D_2 are derivations of an algebra the same is true of $D_1D_2 - D_2D_1$. Thus the endomorphism

$$R(X^*, Y^*) = \nabla_{X^*}\nabla_{Y^*} - \nabla_{Y^*}\nabla_{X^*} - \nabla_{[X^*, Y^*]}$$

is a derivation of the mixed tensor algebra $\mathfrak{D}(M)$, preserving type of tensors and commuting with contractions. In addition, $R(fX^*, gY^*) \cdot hZ = fgh\, R(X^*, Y^*) \cdot Z$ for $f, g, h \in C^\infty(M)$ and $Z \in \mathfrak{D}^1(M)$. This implies that

$$(R(X^*, Y^*)\, T)_p$$

for a tensor field T only depends on the values $R_p(X, Y)$ and T_p. We can therefore put

$$R_p(X, Y) \cdot T_p = (R\,(X^*, Y^*)\, T)_p.$$

The mapping $T_p \to R_p(X, Y) \cdot T_p$ is the unique extension of the endomorphism $R_p(X, Y)$ of $\mathfrak{p}$ to a derivation of the mixed tensor algebra $\mathfrak{D}(p)$ commuting with contractions. Since M is locally symmetric, we have

$$R_p(X, Y) \cdot R_p = ((\nabla_{X^*}\nabla_{Y^*} - \nabla_{Y^*}\nabla_{X^*} - \nabla_{[X^*, Y^*]})\, R)_p = 0,$$

and this shows that $R_p(X, Y) \in \mathfrak{k}$, proving the lemma.

Consider now the direct sum $\mathfrak{g} = \mathfrak{k} + \mathfrak{p}$; we introduce a bracket operation in $\mathfrak{g}$ as follows:

For $X_1, X_2 \in \mathfrak{p}$, $\quad [X_1, X_2] = -R(X_1, X_2)$.

For $X \in \mathfrak{p}, T \in \mathfrak{k}$, $\quad [T, X] = -[X, T] = T \cdot X \quad$ (T operating on X).

For $T_1, T_2 \in \mathfrak{k}$, $\quad [T_1, T_2] = T_1T_2 - T_2T_1$.

The definition of the bracket $[X_1, X_2]$ is of course motivated by Theorem 4.2.

Lemma 5.3. *The bracket operation above turns* $\mathfrak{g}$ *into a Lie algebra.*

Proof. Since the bracket operation is skew symmetric, only the Jacobi identity

$$[Z_1, [Z_2, Z_3]] + [Z_2, [Z_3, Z_1]] + [Z_3, [Z_1, Z_2]] = 0 \tag{5}$$

has to be verified. If all Z_i belong to $\mathfrak{k}$, (5) is just the Jacobi identity for $\mathfrak{k}$. If $Z_1, Z_2 \in \mathfrak{k}$, $Z_3 \in \mathfrak{p}$, (5) is immediate from the definition of the bracket. If $Z_1, Z_2 \in \mathfrak{p}$, $Z_3 \in \mathfrak{k}$, then (5) reduces to (4). Finally, if all Z_i belong to $\mathfrak{p}$, the Jacobi identity is the Bianchi identity (Lemma 12.5, Chapter I).

Now we have the relations

$$[\mathfrak{k}, \mathfrak{k}] \subset \mathfrak{k}, \qquad [\mathfrak{k}, \mathfrak{p}] \subset \mathfrak{p}, \qquad [\mathfrak{p}, \mathfrak{p}] \subset \mathfrak{k},$$

which show that the mapping $s : T + X \to T - X$ ($T \in \mathfrak{k}, X \in \mathfrak{p}$) is an involutive automorphism of $\mathfrak{g}$. The set of fixed points of s coincides with $\mathfrak{k}$.

Lemma 5.4. *The pair* $(\mathfrak{g}, s)$ *is an effective orthogonal symmetric Lie algebra.*

Proof. Suppose $\mathfrak{n}$ is an ideal of $\mathfrak{g}$ contained in $\mathfrak{k}$. Then if $T \in \mathfrak{n}, X \in \mathfrak{p}$ we have

$$[T, X] \in \mathfrak{p} \cap \mathfrak{n} = \{0\}$$

so $T \cdot X = 0$; hence $T = 0$ and $\mathfrak{n} = \{0\}$. In particular, if $\mathfrak{z}$ denotes the center of $\mathfrak{g}$, then $\mathfrak{k} \cap \mathfrak{z} = \{0\}$.

The adjoint group $\mathrm{Int}(\mathfrak{g})$ has Lie algebra $\mathrm{ad}_{\mathfrak{g}}(\mathfrak{g})$. Let K denote the analytic subgroup of $\mathrm{Int}(\mathfrak{g})$ whose Lie algebra is $\mathrm{ad}_{\mathfrak{g}}(\mathfrak{k})$. Each member of K leaves $\mathfrak{p}$ and $\mathfrak{k}$ invariant, so K is a Lie subgroup of the product group $GL(\mathfrak{p}) \times GL(\mathfrak{k})$. The mappings of $GL(\mathfrak{p}) \times GL(\mathfrak{k})$ onto $GL(\mathfrak{p})$ and of $GL(\mathfrak{p}) \times GL(\mathfrak{k})$ onto $GL(\mathfrak{k})$, obtained by restriction to $\mathfrak{p}$ and $\mathfrak{k}$, respectively, are analytic homomorphisms. The images $K_{\mathfrak{p}}$ and $K_{\mathfrak{k}}$

of K under these mappings are analytic subgroups of $\boldsymbol{GL}(\mathfrak{p})$ and $\boldsymbol{GL}(\mathfrak{k})$, respectively. Their Lie algebras are obtained by restricting the endomorphisms in $\mathrm{ad}_{\mathfrak{g}}(\mathfrak{k})$ to $\mathfrak{p}$ and $\mathfrak{k}$, respectively. We see then that the Lie algebra of $K_{\mathfrak{p}}$ is exactly $\mathfrak{k}$ and the Lie algebra of $K_{\mathfrak{k}}$ is $\mathrm{ad}_{\mathfrak{k}}(\mathfrak{k})$. Thus $K_{\mathfrak{k}} = \mathrm{Int}(\mathfrak{k})$ (as Lie groups).

Now, each automorphism A of $\mathfrak{p}$ can be extended uniquely to a type preserving automorphism $\tilde{A}$ of the mixed tensor algebra $\mathfrak{D}(p)$ over $\mathfrak{p}$ such that $\tilde{A}$ coincides with $({}^tA)^{-1}$ on the dual space $\mathfrak{p}^\wedge$. Those automorphisms A of $\mathfrak{p}$ for which $\tilde{A}$ leaves invariant Q_p and R_p, form a compact Lie subgroup of $\boldsymbol{GL}(\mathfrak{p})$ with Lie algebra $\mathfrak{k}$. The identity component of this group must therefore coincide with $K_{\mathfrak{p}}$; thus the group $K_{\mathfrak{p}}$ is compact and so is its homomorphic image $K_{\mathfrak{k}}$.

Let (k_n) be a sequence in K. There exists a subsequence (k_ν) of (k_n) such that the corresponding sequences of restrictions to $\mathfrak{p}$ and $\mathfrak{k}$ are convergent; it follows that (k_ν) is convergent (in the relative topology of $\boldsymbol{GL}(\mathfrak{g})$) to an element $k \in K$. In particular, K is a closed subset of $\boldsymbol{GL}(\mathfrak{g})$; owing to Theorem 2.10, Chapter II, the Lie group K carries the relative topology of $\boldsymbol{GL}(\mathfrak{g})$. Thus K is a compact Lie group. This shows that $\mathfrak{k}$ is compactly imbedded in $\mathfrak{g}$ and the lemma is proved.

The center $\mathfrak{z}$ of $\mathfrak{g}$ is invariant under s; hence $\mathfrak{z} = (\mathfrak{k} \cap \mathfrak{z}) + (\mathfrak{p} \cap \mathfrak{z})$. Now, $\mathfrak{k} \cap \mathfrak{z} = \{0\}$ so $\mathfrak{z} \subset \mathfrak{p}$. Let $\mathfrak{p}'$ denote the orthogonal complement of $\mathfrak{z}$ in $\mathfrak{p}$ (with respect to Q_p). Then $[\mathfrak{k}, \mathfrak{p}'] \subset \mathfrak{p}'$ so $\mathfrak{k} + \mathfrak{p}'$ is an ideal of $\mathfrak{g}$, isomorphic to $\mathrm{ad}_{\mathfrak{g}}(\mathfrak{g})$. Let $c = \dim \mathfrak{z}$. Then the product group

$$G = \mathrm{Int}(\mathfrak{g}) \times \boldsymbol{R}^c$$

has Lie algebra $\mathrm{ad}_{\mathfrak{g}}(\mathfrak{g}) \times \mathfrak{z}$; the Lie group K is a compact Lie subgroup of G with Lie algebra $\mathrm{ad}_{\mathfrak{g}}(\mathfrak{k})$. The automorphism s of $\mathfrak{g}$ induces an automorphism σ of $\mathrm{Int}(\mathfrak{g})$ such that

$$\sigma \cdot e^{\mathrm{ad}\, X} = e^{\mathrm{ad}(s.X)}, \qquad X \in \mathfrak{g}$$

(Chapter II, §5). We extend σ to an automorphism of G, also denoted σ, by putting $\sigma \cdot a = a^{-1}$ for $a \in \boldsymbol{R}^c$. Then K is the identity component of the set of fixed points of σ. Thus (G, K) is a Riemannian symmetric pair.

If $X \in \mathfrak{g}$, let $X_{\mathfrak{z}}$ denote the component of X in $\mathfrak{z}$ according to the direct decomposition $\mathfrak{g} = (\mathfrak{k} + \mathfrak{p}') + \mathfrak{z}$. Then the mapping $X \to (\mathrm{ad}\, X, X_{\mathfrak{z}})$ is an isomorphism of $\mathfrak{g}$ onto $\mathrm{ad}_{\mathfrak{g}}(\mathfrak{g}) \times \mathfrak{z}$, carrying $\mathfrak{k}$ onto $\mathrm{ad}_{\mathfrak{g}}(\mathfrak{k})$ and such that the automorphisms s and $(d\sigma)_e$ correspond. Thus we can regard $\mathrm{Ad}_G(K)$ as a group of automorphisms of $\mathfrak{g}$; this group leaves invariant $\mathfrak{p}$ and the quadratic form Q_p on $\mathfrak{p}$. Identifying $\mathfrak{p}$ with $(G/K)_p$ we see that there exists a unique G-invariant Riemannian structure $\tilde{Q}$ on G/K

such that $\tilde{Q}_p = Q_p$. With this Riemannian structure, G/K is globally symmetric (Prop. 3.4) and this is the desired space $\tilde{M}$. In fact, the curvature tensor $\tilde{R}$ of G/K is given by Theorem 4.2,

$$\tilde{R}_p(X, Y) \cdot Z = -[[X, Y], Z], \qquad X, Y, Z \in \mathfrak{p},$$

and by the definition of the bracket in $\mathfrak{g}$,

$$-[[X, Y], Z] = -[-R_p(X, Y), Z] = R_p(X, Y) \cdot Z.$$

Hence

$$\tilde{R}_p = R_p, \qquad \tilde{Q}_p = Q_p,$$

and the theorem now follows from Lemma 1.2 and Lemma 1.4.

Remark. In the proof above, the group G was constructed by means of the adjoint group Int ($\mathfrak{g}$). A minor shortcut could be taken by making use of the theorem, that for every Lie algebra $\mathfrak{a}$ over $\boldsymbol{R}$ there exists a Lie group A whose Lie algebra is isomorphic to $\mathfrak{a}$. However, this theorem is neither proved nor used in this book.

Proposition 5.5. *A Riemannian locally symmetric space M is an analytic Riemannian manifold.*

Proof. In view of Theorem 5.1, there exists a covering $\{B_\alpha\}_{\alpha \in A}$ of M with open balls $B_\alpha = B_{\rho(\alpha)}(p_\alpha)$ such that for each α, $B_{3\rho(\alpha)}(p_\alpha)$ is a normal neighborhood of p_α, isometric to an open set in a Riemannian globally symmetric space. We have to show that for any $\alpha, \beta \in A$, the normal coordinates at p_α and p_β are analytically related on $B_\alpha \cap B_\beta$. We may assume for the radii, that $\rho(\alpha) \geqslant \rho(\beta)$. Then if $B_\alpha \cap B_\beta \neq \emptyset$, the ball $B_{3\rho(\alpha)}(p_\alpha)$ contains B_β. Since the Riemannian structure on M is analytic on $B_{3\rho(\alpha)}(p_\alpha)$, the normal coordinates at p_α and p_β are analytically related on B_β, in particular on $B_\alpha \cap B_\beta$.

As a consequence we note that the analyticity assumption in the definition of a Riemannian globally symmetric space can be dropped.

We shall now establish another connection between locally and globally symmetric Riemannian spaces.

Theorem 5.6. *Let M be a complete, simply connected Riemannian locally symmetric space. Then M is Riemannian globally symmetric.*

Proof. Let $p \in M$ and let $B_\rho(p)$ be a spherical normal neighborhood such that the geodesic symmetry s_p is an isometry of $B_\rho(p)$ onto itself. We define a mapping Φ of M into M as follows. Let $q \in M$ and let $\gamma(t)$, $0 \leqslant t \leqslant 1$, be a continuous curve joining p and q. Let φ_t be a continuation of s_p along γ as defined in §11, Chapter I, and put $\Phi(q) =$

$\varphi_1(q)$. Since M is simply connected it follows from Prop. 11.4 that $\Phi(q)$ does not depend on the choice of γ. For the same reason Φ coincides with φ_1 in a neighborhood of q. Hence Φ is a differentiable mapping of M into M such that for each $q \in M$, $d\Phi_q$ is an isometry. Since Φ reverses the direction of each geodesic starting at p, it is clear that $\Phi(M) = M$ and $\Phi \circ \Phi$ is the identity mapping. Being involutive, Φ must be one-to-one and the theorem is proved.

Corollary 5.7. *Let M be a complete Riemannian locally symmetric space. Let g denote the Riemannian structure on M and let (M^*, π) be the universal covering manifold of M. Then M^*, with the Riemannian structure π^*g, is a Riemannian globally symmetric space.*

In fact, M^* satisfies the hypothesis of Theorem 5.6.

§ 6. Compact Lie Groups

A compact connected Lie group G can always be regarded as a Riemannian globally symmetric space. The mapping $\sigma : (g_1, g_2) \rightarrow (g_2, g_1)$ is an involutive automorphism of the product group $G \times G$. The fixed points of σ constitute the diagonal G^* of $G \times G$; the pair $(G \times G, G^*)$ is a Riemannian symmetric pair and the coset space $G \times G/G^*$ is diffeomorphic to the group G under the mapping

$$(g_1, g_2)\, G^* \rightarrow g_1 g_2^{-1}.$$

A Riemannian structure on $G \times G/G^*$ is $G \times G$-invariant if and only if the corresponding Riemannian structure on G is invariant under right and left translations. Thus by Prop. 3.4, G is a Riemannian globally symmetric space in each bi-invariant Riemannian structure. The natural mapping of $G \times G$ onto $G \times G/G^*$ now becomes the mapping π : $G \times G \rightarrow G$ given by $\pi(g_1, g_2) = g_1 g_2^{-1}$. Recalling that the geodesic symmetry s_e is given by $s_e \circ \pi = \pi \circ \sigma$, we see that $s_e(g) = g^{-1}$ for $g \in G$.

Let $\mathfrak{g}$ denote the Lie algebra of G. Then the product algebra $\mathfrak{g} \times \mathfrak{g}$ is the Lie algebra of $G \times G$ and the identity

$$(X, Y) = (\tfrac{1}{2}(X + Y), \tfrac{1}{2}(X + Y)) + (\tfrac{1}{2}(X - Y), -\tfrac{1}{2}(X - Y))$$

gives the decomposition of $\mathfrak{g} \times \mathfrak{g}$ into the two eigenspaces of $d\sigma$. Since $\pi(g_1, g_2) = g_1 g_2^{-1}$, it follows that

$$d\pi(X, Y) = X - Y, \qquad X, Y \in \mathfrak{g}.$$

We now denote,

exp* : the exponential mapping of $\mathfrak{g} \times \mathfrak{g}$ into $G \times G$;

exp : the exponential mapping of $\mathfrak{g}$ into G;

Exp : the Exponential mapping of $\mathfrak{g}$ into G (G being considered as a Riemannian globally symmetric space).

Formula (3) in §3 then implies

$$\pi(\exp^* (X, -X)) = \operatorname{Exp} (d\pi(X, -X)) \qquad (X \in \mathfrak{g}).$$

Hence $\exp X \cdot (\exp (-X))^{-1} = \operatorname{Exp} 2X$, so we have

$$\exp X = \operatorname{Exp} X, \qquad X \in \mathfrak{g}.$$

The geodesics in G through e are therefore just the one-parameter subgroups. This fact could also have been verified by using (2) §9, Chapter I, as was done in the special case considered in Theorem 6.9 in Chapter II.

The orthogonal symmetric Lie algebra associated with $(G \times G, G^*)$ is $(\mathfrak{g} \times \mathfrak{g}, \tau)$ where τ is the automorphism $(X, Y) \to (Y, X)$ of $\mathfrak{g} \times \mathfrak{g}$.

§ 7. Totally Geodesic Submanifolds. Lie Triple Systems

In contrast to general Riemannian manifolds, globally symmetric spaces contain plenty of totally geodesic submanifolds. We shall now describe these in Lie algebra terms.

Proposition 7.1. *Let M be a Riemannian manifold and S a totally geodesic submanifold of M. If M is locally symmetric, the same holds for S.*

The proof is straightforward because each geodesic symmetry on S is obtained from a geodesic symmetry of M by restriction.

Definition. Let $\mathfrak{g}$ be a Lie algebra over R and let $\mathfrak{m}$ be a subspace of $\mathfrak{g}$; $\mathfrak{m}$ is called a *Lie triple system* if $X, Y, Z \in \mathfrak{m}$ implies $[X, [Y, Z]] \in \mathfrak{m}$.

Theorem 7.2. *Let M be a Riemannian globally symmetric space and let the notation be as in Theorem* 3.3. *Identifying as usual the tangent space M_{p_0} with the subspace $\mathfrak{p}$ of the Lie algebra of $I(M)$, let $\mathfrak{s}$ be a Lie triple system contained in $\mathfrak{p}$. Put $S = \operatorname{Exp} \mathfrak{s}$. Then S has a natural differentiable structure in which it is a totally geodesic submanifold of M satisfying $S_{p_0} = \mathfrak{s}$.*

On the other hand, if S is a totally geodesic submanifold of M and $p_0 \in S$, then the subspace $\mathfrak{s} = S_{p_0}$ of $\mathfrak{p}$ is a Lie triple system.

Proof. Suppose first that S is a totally geodesic submanifold and let X, Y be two vectors in the tangent space $\mathfrak{s} = S_{p_0}$. For each $t \in \boldsymbol{R}$, the vector $A = d\,\mathrm{Exp}_{tY}(X)$ is a tangent vector to S at $\mathrm{Exp}\,(tY)$. As we have seen earlier, the vector $d\tau(\exp\,(-\,tY)) \cdot A$ is M-parallel to A along the curve $\exp tY$ $(t \in \boldsymbol{R})$. Using Theorem 14.5, Chapter I, we conclude that $d\tau(\exp\,(-\,tY)) \cdot A \in \mathfrak{s}$. In view of Theorem 4.1, this means that

$$\sum_0^\infty \frac{(T_{tY})^n}{(2n+1)!}(X) \in \mathfrak{s}$$

for all $t \in \boldsymbol{R}$. This implies that $T_Y(X) \in \mathfrak{s}$. Now,

$$T_{Y+Z} = T_Y + T_Z + \mathrm{ad}\, Y\, \mathrm{ad}\, Z + \mathrm{ad}\, Z\, \mathrm{ad}\, Y.$$

Combining this with the Jacobi identity, we obtain

$$2[Y, [Z, X]] + [X, [Y, Z]] \in \mathfrak{s} \qquad \text{for } X, Y, Z \in \mathfrak{s}. \tag{1}$$

Interchange of X, Y gives the equation

$$4[X, [Z, Y]] + 2[Y, [X, Z]] \in \mathfrak{s},$$

which, added to (1), shows that $[X, [Y, Z]] \in \mathfrak{s}$.

On the other hand, suppose $\mathfrak{s}$ is a Lie triple system. Still using the notation of Theorem 3.3, we have $[\mathfrak{s}, \mathfrak{s}] \subset [\mathfrak{p}, \mathfrak{p}] \subset \mathfrak{k}$. Moreover, the subspace $[\mathfrak{s}, \mathfrak{s}]$ is a subalgebra of $\mathfrak{k}$; this follows from the identity

$$[[X, Y], [U, V]] + [U, [V, [X, Y]]] + [V, [[X, Y], U]] = 0$$

combined with the fact that $\mathfrak{s}$ is a Lie triple system. It follows immediately that the subspace $\mathfrak{g}' = \mathfrak{s} + [\mathfrak{s}, \mathfrak{s}]$ is a subalgebra of $\mathfrak{g}$. Let G' denote the analytic subgroup of G with Lie algebra $\mathfrak{g}'$, let M' denote the orbit $G' \cdot p_0$ and let K' denote the subgroup of G' leaving the point p_0 fixed. Since the identity mapping of G' into G is continuous, K' is a closed subgroup of G'. Since M' is in one-to-one correspondence $g \cdot p_0 \to gK'$ with G'/K' we can carry the topology and differentiable structure of G'/K' over on M'; by Prop. 4.4, Chapter II, M' is then a submanifold of M. Furthermore, $(M')_{p_0} = \mathfrak{s}$. The M-geodesics through p_0 have the form $\exp tX \cdot p_0$ $(t \in \boldsymbol{R})$ where X is a general vector in $\mathfrak{p}$. This geodesic is tangent to M' at p_0 if and only if $X \in \mathfrak{s}$; it follows that the submanifold M' of M is geodesic at p_0. Since G' is a group of isometries of M and

M', and acts transitively on M', it follows that M' is geodesic at each of its points, hence totally geodesic. Obviously $M' = \text{Exp}\ \mathfrak{s}$, and the theorem is proved.

Remark. Since the automorphism $g \to s_{p_0} g s_{p_0}$ of G leaves G' invariant, the pair (G', K') is a symmetric pair and the manifold $M' = \text{Exp}\ \mathfrak{s}$ is a Riemannian globally symmetric space.

EXERCISES

1. Let M be a Riemannian globally symmetric space, ω a differential form on M invariant under all isometries of M. Show that ω is closed, i.e., $d\omega = 0$. (Hint: Use the formula $d\omega = \Sigma_{i=1}^{m} \omega_i \wedge \nabla_{X_i}(\omega)$ from Exercise C.3 following Chapter I.)

2. Formulate and prove a generalization of Lemma 1.2 to arbitrary manifolds with affine connections.

3. Show that any two complete, simply connected Riemannian manifolds M_1, M_2 of the same constant sectional curvature are isometric. (Hint: Lemma 1.2 gives an isometry between two normal neighborhoods. The method of the proof of Theorem 5.6 gives a local isometry of M_1 onto M_2. Now use Lemma 13.4, Chapter I.)

4. A compact semisimple Lie group G has a bi-invariant Riemannian structure Q such that Q_e is the negative of the Killing form of the Lie algebra $\mathfrak{g}$ of G. If G is considered as a symmetric space $G \times G/G^*$ as in §6, it acquires a bi-invariant Riemannian structure Q^* from the Killing form of $\mathfrak{g} \times \mathfrak{g}$. Show that $Q = 2Q^*$.

5. Let M be a connected locally compact metric space. Let $I(M)$ denote the group of distance-preserving mappings of M onto itself, topologized by the compact open topology. Let H be a closed subgroup of $I(M)$. Then, for each $p \in M$, the orbit $H \cdot p$ is closed in M.

NOTES

§1. The material here is due to É. Cartan ([6] and [22], Chapter XI) for the Riemannian case. Concerning the affine case see Whitehead [1] for the local theory, Nomizu [2], Fedenko [1], Rozenfeld [2] and Berger [2] for the global theory.

The results of §2 (which actually apply to all (separable) connected, locally compact metric spaces M) are due to van Dantzig and van der Waerden [1], see also Arens [1].

§3. Lemma 3.2 was proved by É. Cartan [6], p. 230, by the use of differential equations. It was extended by Myers and Steenrod [1] to all Riemannian manifolds. In [16] É. Cartan points out the fact that the geodesics in M are orbits of one-parameter subgroups from $I(M)$. This property is further examined in Nomizu [2].

§4. The formula in Theorem 4.2 for the curvature tensor of a symmetric space is due to É. Cartan [6] and extended by Nomizu [2] to all reductive homogeneous spaces. The proof in the text is from Helgason [2] where the formula for the differential of Exp (Theorem 4.1) is also given.

§5. The relation between locally and globally symmetric spaces is not altogether clear from É. Cartan's work although his extensive paper [10] gives a global classification. Theorem 5.1 is a special case of a theorem of Nomizu [2] on reductive homogeneous spaces. The idea of the proof was already used by É. Cartan [6], p. 225, for the similar problem of constructing a locally symmetric space whose curvature tensor satisfies certain necessary conditions involving the holonomy group (see the introduction to this chapter). Theorem 5.6 is due to Borel and Lichnerowicz [1]; they outlined a proof based on results of Ehresmann [1] applied to the local group of local isometries of a Riemannian locally symmetric space. Ambrose has in [1] stated and proved an extension of Theorem 5.6 to arbitrary Riemannian manifolds.

§7. The connection between totally geodesic submanifolds and Lie triple systems is pointed out in É. Cartan [7], p. 133; see also Mostow [3].

CHAPTER V

DECOMPOSITION OF SYMMETRIC SPACES

In the previous chapter we have seen that a Riemannian globally symmetric space M gives rise to a pair $(\mathfrak{l},s)$ where $\mathfrak{l}$ is the Lie algebra of the group of isometries of $I(M)$ and s is an involutive automorphism of $\mathfrak{l}$ having a compactly imbedded subalgebra for the set of fixed points. This chapter is devoted to the study of such pairs. It is shown in §1 that they fall into three different categories, the compact type, the noncompact type, and the Euclidean type. An arbitrary pair $(\mathfrak{l},s)$ can be decomposed into three parts each of which is from the types above. Since the Euclidean type is uninteresting we are left with two types of pairs $(\mathfrak{l},s)$, the compact type and the noncompact type, both of which have $\mathfrak{l}$ semisimple. The symmetric spaces corresponding to these have positive sectional curvature and negative sectional curvature, respectively. There is a remarkable duality (§2) between the two types which for example provides two viewpoints of the classification problem and incidentally explains the formal analogy between spherical trigonometry and hyperbolic trigonometry.

The rank of a symmetric space M is an important invariant; it is defined as the maximum dimension of any flat totally geodesic subspace A of M. It is shown in §6 that each geodesic in M can be moved into A by an isometry of M. This means that for any two points in M one can speak of their complex distance; this is an l-tuple $(r_1, ..., r_l)$ of real numbers (l = rank of M) and has the property that two point-pairs in M are congruent under an isometry of M if and only if their complex distance is the same. Just as ordinary Euclidean distance is > 0, the l-tuple $(r_1, ..., r_l)$ is restricted to the fundamental domain of a certain discontinuous group. (For the noncompact type this is the Weyl group W (Chapter VII, §2); for the compact type the group is larger (Chapter VII, §7).)

§ 1. Orthogonal Symmetric Lie Algebras

We recall that an orthogonal symmetric Lie algebra is a pair $(\mathfrak{l}, s)$ where

(i) $\mathfrak{l}$ is a Lie algebra over $\boldsymbol{R}$.

(ii) s is an involutive† automorphism of $\mathfrak{l}$.

(iii) $\mathfrak{u}$, the set of fixed points of s, is a compactly imbedded subalgebra of $\mathfrak{l}$.

† That is, $s \neq I$ and $s^2 = I$.

If, in addition, $\mathfrak{u} \cap \mathfrak{z} = \{0\}$, where $\mathfrak{z}$ denotes the center of $\mathfrak{l}$, then $(\mathfrak{l}, s)$ is called *effective*. Two orthogonal symmetric Lie algebras $(\mathfrak{l}_1, s_1)$ and $(\mathfrak{l}_2, s_2)$ are called *isomorphic* if there exists an isomorphism φ of $\mathfrak{l}_1$ onto $\mathfrak{l}_2$ such that $\varphi \circ s_1 = s_2 \circ \varphi$.

Examples.

(a) Let $\mathfrak{l}$ be a compact semisimple Lie algebra and s any involutive automorphism of $\mathfrak{l}$. Then $(\mathfrak{l}, s)$ is an effective orthogonal symmetric Lie algebra.

(b) Let $\mathfrak{l}$ be a noncompact semisimple Lie algebra and let $\mathfrak{l} = \mathfrak{u} + \mathfrak{e}$ be any Cartan decomposition of $\mathfrak{l}$ (where $\mathfrak{u}$ is the subalgebra). Let s denote the automorphism of $\mathfrak{l}$ given by $s(T + X) = T - X$ $(T \in \mathfrak{u}, X \in \mathfrak{e})$. Then $(\mathfrak{l}, s)$ is an effective orthogonal symmetric Lie algebra (Prop. 7.4, Chapter III).

(c) Let $\mathfrak{e}$ be a finite-dimensional vector space over $\boldsymbol{R}$ and let $\mathfrak{u}$ be the Lie algebra of a compact Lie subgroup of $GL(\mathfrak{e})$. Let $\mathfrak{l}$ denote the direct sum $\mathfrak{l} = \mathfrak{u} + \mathfrak{e}$; $\mathfrak{l}$ can be turned into a Lie algebra by defining

$$
\begin{aligned}
&[X_1, X_2] = 0 && && \text{if } X_1, X_2 \in \mathfrak{e},\\
&[T, X] = -[X, T] = T \cdot X && (T \text{ acting on } X) && \text{if } T \in \mathfrak{u}, X \in \mathfrak{e},\\
&[T_1, T_2] = T_1T_2 - T_2T_1 && && \text{if } T_1, T_2 \in \mathfrak{u}.
\end{aligned}
$$

Then $\mathfrak{l}$ is a Lie algebra containing $\mathfrak{u}$ as a subalgebra. Assuming $\mathfrak{e} \neq \{0\}$, the mapping $s : T + X \to T - X$, $(T \in \mathfrak{u}, X \in \mathfrak{e})$, is an involutive automorphism of $\mathfrak{l}$. The pair $(\mathfrak{l}, s)$ is an effective orthogonal symmetric Lie algebra. The proof of this statement is the same as that of Lemma 5.4, Chapter IV.

Definition. Let $(\mathfrak{l}, s)$ be an effective orthogonal symmetric Lie algebra. Let $\mathfrak{l} = \mathfrak{u} + \mathfrak{e}$ be the decomposition of $\mathfrak{l}$ into the eigenspaces of s for the eigenvalue $+1$ and -1, respectively.

(a) If $\mathfrak{l}$ is compact and semisimple, $(\mathfrak{l}, s)$ is said to be of the *compact type*.

(b) If $\mathfrak{l}$ is noncompact and semisimple and $\mathfrak{l} = \mathfrak{u} + \mathfrak{e}$ is a Cartan decomposition of $\mathfrak{l}$, then $(\mathfrak{l}, s)$ is said to be of the *noncompact type*.

(c) If $\mathfrak{e}$ is an abelian ideal in $\mathfrak{l}$, then $(\mathfrak{l}, s)$ is said to be of the *Euclidean type*.

Definition. Let $(\mathfrak{l}, s)$ be an orthogonal symmetric Lie algebra and suppose the pair (L, U) is associated† with $(\mathfrak{l}, s)$. The pair (L, U) is

† See definition preceding Prop. 3.6, Chapter IV.

said to be of the compact type, noncompact type, or Euclidean type according to the type of $(\mathfrak{l}, s)$.

The next theorem shows that every effective orthogonal symmetric Lie algebra can be decomposed into three others, which are of the compact type, noncompact type, and Euclidean type, respectively.

Theorem 1.1. *Let* $(\mathfrak{l}, s)$ *be an effective, orthogonal symmetric Lie algebra. Then there exist ideals* $\mathfrak{l}_0$, $\mathfrak{l}_-$, *and* $\mathfrak{l}_+$ *in* $\mathfrak{l}$ *with the following properties*:

1. $\mathfrak{l} = \mathfrak{l}_0 + \mathfrak{l}_- + \mathfrak{l}_+$ *(direct sum)*.

2. $\mathfrak{l}_0$, $\mathfrak{l}_-$ *and* $\mathfrak{l}_+$ *are invariant under* s *and orthogonal with respect to the Killing form of* $\mathfrak{l}$.

3. *Let* s_0, s_-, *and* s_+ *denote the restrictions of* s *to* $\mathfrak{l}_0$, $\mathfrak{l}_-$, *and* $\mathfrak{l}_+$, *respectively. The pairs* $(\mathfrak{l}_0, s_0)$, $(\mathfrak{l}_-, s_-)$, *and* $(\mathfrak{l}_+, s_+)$ *are effective orthogonal symmetric Lie algebras of the Euclidean type, compact type, and noncompact type, respectively.*

The proof of this theorem will be broken up into a sequence of lemmas. Let $\mathfrak{u}$ and $\mathfrak{e}$ denote the eigenspaces of s for the eigenvalues $+1$ and -1, respectively. Then we have

$$\mathfrak{l} = \mathfrak{u} + \mathfrak{e} \text{ (direct sum)}, \quad [\mathfrak{u}, \mathfrak{u}] \subset \mathfrak{u}, \quad [\mathfrak{u}, \mathfrak{e}] \subset \mathfrak{e}, \quad [\mathfrak{e}, \mathfrak{e}] \subset \mathfrak{u}. \tag{1}$$

Let B denote the Killing form of $\mathfrak{l}$. Since B is invariant under each automorphism of $\mathfrak{l}$, in particular under s, it follows that the subspaces $\mathfrak{u}$ and $\mathfrak{e}$ are orthogonal with respect to B.

Lemma 1.2. *The Killing form* B *is strictly negative definite on* $\mathfrak{u}$.

This lemma is a special case of Prop. 6.8, Chapter II.

Now let U denote the analytic subgroup of the adjoint group $\mathrm{Int}(\mathfrak{l})$ with Lie algebra $\mathrm{ad}_{\mathfrak{l}}(\mathfrak{u})$. As a result of our assumptions, U is a compact Lie subgroup of $\boldsymbol{GL}(\mathfrak{l})$; thus it carries the relative topology of $\boldsymbol{GL}(\mathfrak{l})$. Since U is connected, relations (1) imply

$$u \cdot \mathfrak{u} \subset \mathfrak{u}, \qquad u \cdot \mathfrak{e} \subset \mathfrak{e} \qquad \text{for } u \in U.$$

Being a compact linear group, U leaves invariant a strictly positive definite, symmetric bilinear form Q on $\mathfrak{e} \times \mathfrak{e}$. There exists a basis $X_1, \ldots, X_n$ of $\mathfrak{e}$ and real numbers $\beta_1, \ldots, \beta_n$ such that

$$Q(X, X) = x_1^2 + \ldots + x_n^2,$$

$$B(X, X) = \beta_1 x_1^2 + \ldots + \beta_n x_n^2,$$

if $X = \sum_{i=1}^{n} x_i X_i$. Let

$$\mathfrak{e}_0 = \sum_{\beta_i = 0} RX_i, \qquad \mathfrak{e}_- = \sum_{\beta_i < 0} RX_i, \qquad \mathfrak{e}_+ = \sum_{\beta_i > 0} RX_i.$$

Then $\mathfrak{e}$ is the direct sum of the subspaces $\mathfrak{e}_0$, $\mathfrak{e}_-$, $\mathfrak{e}_+$; moreover, these subspaces are orthogonal with respect to Q and B and each one is invariant under s. If we define the endomorphism b of $\mathfrak{e}$ by $bX_i = \beta_i X_i$ $(1 \leqslant i \leqslant n)$, then

$$Q(bX, Y) = B(X, Y)$$

for $X, Y \in \mathfrak{e}$. Since B and Q are invariant under U, the endomorphism b commutes with the restriction of each $u \in U$ to $\mathfrak{e}$. It follows that the spaces $\mathfrak{e}_0$, $\mathfrak{e}_-$, and $\mathfrak{e}_+$ are invariant under U and under $\mathrm{ad}_{\mathfrak{l}}(\mathfrak{u})$.

Lemma 1.3. *The subspaces $\mathfrak{e}_0$, $\mathfrak{e}_-$, and $\mathfrak{e}_+$ satisfy the following relations:*

(i) $\mathfrak{e}_0 = \{X \in \mathfrak{l} : B(X, Y) = 0$ *for all* $Y \in \mathfrak{l}\}$.

(ii) $[\mathfrak{e}_0, \mathfrak{e}] = \{0\}$ *and* $\mathfrak{e}_0$ *is an abelian ideal in* $\mathfrak{l}$.

(iii) $[\mathfrak{e}_-, \mathfrak{e}_+] = \{0\}$.

Proof. Let $\mathfrak{n}$ denote the set on the right-hand side in (i). Then $\mathfrak{n}$ is invariant under s, so $\mathfrak{n} = \mathfrak{n} \cap \mathfrak{u} + \mathfrak{n} \cap \mathfrak{e}$ (direct sum). Now $\mathfrak{n} \cap \mathfrak{u} = \{0\}$ due to Lemma 1.2 so $\mathfrak{n} \subset \mathfrak{e}$. But $\mathfrak{n} \cap \mathfrak{e}_- = \mathfrak{n} \cap \mathfrak{e}_+ = \{0\}$ so $\mathfrak{n} \subset \mathfrak{e}_0$. On the other hand, if $X \in \mathfrak{e}_0$, then $B(X, Y) = 0$ for all $Y \in \mathfrak{e}_0$, hence for all $Y \in \mathfrak{l}$. This proves (i). As a result of (i), $\mathfrak{e}_0$ is an ideal in $\mathfrak{l}$, but $[\mathfrak{e}_0, \mathfrak{e}] \subset \mathfrak{u}$ by (1). This proves (ii). In order to prove (iii) we observe that $[\mathfrak{e}_-, \mathfrak{e}_+] \subset \mathfrak{u}$, so, owing to Lemma 1.2, it suffices to prove

$$B(\mathfrak{u}, [\mathfrak{e}_-, \mathfrak{e}_+]) = 0.$$

But if $T \in \mathfrak{u}$, $X_\pm \in \mathfrak{e}_\pm$, then

$$B(T, [X_-, X_+]) = B([T, X_-], X_+) = 0$$

and the lemma is proved.

We define now $\mathfrak{u}_+ = [\mathfrak{e}_+, \mathfrak{e}_+]$, $\mathfrak{u}_- = [\mathfrak{e}_-, \mathfrak{e}_-]$ and let $\mathfrak{u}_0$ denote the orthogonal complement (with respect to B) of the subspace of $\mathfrak{u}$ spanned by $\mathfrak{u}_+$ and $\mathfrak{u}_-$.

Lemma 1.4. *The subspaces $\mathfrak{u}_0$, $\mathfrak{u}_+$, and $\mathfrak{u}_-$ are ideals in $\mathfrak{u}$, orthogonal with respect to B, and $\mathfrak{u} = \mathfrak{u}_0 + \mathfrak{u}_- + \mathfrak{u}_+$ (direct sum).*

Proof. Since $[\mathfrak{u}, \mathfrak{e}_+] \subset \mathfrak{e}_+$, we have by the Jacobi identity $[\mathfrak{u}_+, \mathfrak{u}] = [[\mathfrak{e}_+, \mathfrak{e}_+], \mathfrak{u}] \subset \mathfrak{u}_+$. Similarly $[\mathfrak{u}_-, \mathfrak{u}] \subset \mathfrak{u}_-$. Now, let $X_\pm \in \mathfrak{e}_\pm$, $Y_\pm \in \mathfrak{e}_\pm$. Then

$$B([X_+, Y_+], [X_-, Y_-]) = B(X_+, [Y_+, [X_-, Y_-]]) = 0$$

due to the Jacobi identity and Lemma 1.3 (iii). Hence $\mathfrak{u}_+$ and $\mathfrak{u}_-$ are orthogonal and the sum $\mathfrak{u}_- + \mathfrak{u}_+$ is an ideal in $\mathfrak{u}$. The orthogonal complement $\mathfrak{u}_0$ is also an ideal and the lemma now follows from Lemma 1.2.

Lemma 1.5. *The following commutation relations hold*:

(i) $[\mathfrak{u}_0, \mathfrak{e}_-] = [\mathfrak{u}_0, \mathfrak{e}_+] = \{0\}$.

(ii) $[\mathfrak{u}_-, \mathfrak{e}_0] = [\mathfrak{u}_-, \mathfrak{e}_+] = \{0\}$.

(iii) $[\mathfrak{u}_+, \mathfrak{e}_0] = [\mathfrak{u}_+, \mathfrak{e}_-] = \{0\}$.

Proof. (i) Let $T \in \mathfrak{u}_0$, X, $Y \in \mathfrak{e}_+$. Then

$$B([T, X], Y) = B(T, [X, Y]) = 0.$$

Thus $[\mathfrak{u}_0, \mathfrak{e}_+]$ is orthogonal to $\mathfrak{e}_+$. Since B is strictly positive definite on $\mathfrak{e}_+$ and since $[\mathfrak{u}_0, \mathfrak{e}_+] \subset \mathfrak{e}_+$, it follows that $[\mathfrak{u}_0, \mathfrak{e}_+] = \{0\}$. Similarly $[\mathfrak{u}_0, \mathfrak{e}_-] = \{0\}$. For (ii) we have $[\mathfrak{u}_-, \mathfrak{e}_0] = [[\mathfrak{e}_-, \mathfrak{e}_-], \mathfrak{e}_0] = \{0\}$ by Lemma 1.3 (ii). Moreover, $[\mathfrak{u}_-, \mathfrak{e}_+] = [[\mathfrak{e}_-, \mathfrak{e}_-], \mathfrak{e}_+] = \{0\}$ by Lemma 1.3 (iii). The last part (iii) follows in the same way.

In order to prove Theorem 1.1, we have to distinguish between two cases: $\mathfrak{e}_0 = \{0\}$ and $\mathfrak{e}_0 \neq \{0\}$. Suppose first $\mathfrak{e}_0 \neq \{0\}$. Then we put

$$\mathfrak{l}_0 = \mathfrak{u}_0 + \mathfrak{e}_0, \qquad \mathfrak{l}_- = \mathfrak{u}_- + \mathfrak{e}_-, \qquad \mathfrak{l}_+ = \mathfrak{u}_+ + \mathfrak{e}_+.$$

Then $\mathfrak{l}$ is the direct sum of the subspaces $\mathfrak{l}_0$, $\mathfrak{l}_-$, $\mathfrak{l}_+$; these subspaces are invariant under s and orthogonal with respect to B. Using the lemmas above, we have

$$\begin{aligned}[\mathfrak{l}_0, \mathfrak{l}] &= [\mathfrak{u}_0, \mathfrak{l}] + [\mathfrak{e}_0, \mathfrak{l}] \\ &= [\mathfrak{u}_0, \mathfrak{u}] + [\mathfrak{u}_0, \mathfrak{e}_0] + [\mathfrak{u}_0, \mathfrak{e}_+] + [\mathfrak{u}_0, \mathfrak{e}_-] + [\mathfrak{e}_0, \mathfrak{l}] \\ &\qquad \subset \mathfrak{u}_0 + \mathfrak{e}_0 + \{0\} + \{0\} + \mathfrak{e}_0\end{aligned}$$

so $[\mathfrak{l}_0, \mathfrak{l}] \subset \mathfrak{l}_0$. Secondly

$$\begin{aligned}[\mathfrak{l}_+, \mathfrak{l}] &= [\mathfrak{u}_+, \mathfrak{l}] + [\mathfrak{e}_+, \mathfrak{l}] \\ &= [\mathfrak{u}_+, \mathfrak{u}] + [\mathfrak{u}_+, \mathfrak{e}_0] + [\mathfrak{u}_+, \mathfrak{e}_-] + [\mathfrak{u}_+, \mathfrak{e}_+] + [\mathfrak{e}_+, \mathfrak{u}] + [\mathfrak{e}_+, \mathfrak{e}_0] + [\mathfrak{e}_+, \mathfrak{e}_-] + [\mathfrak{e}_+, \mathfrak{e}_+] \\ &\qquad \subset \mathfrak{u}_+ + \{0\} + \{0\} + \mathfrak{e}_+ + \mathfrak{e}_+ + \{0\} + \{0\} + \mathfrak{u}_+\end{aligned}$$

so $[\mathfrak{l}_+, \mathfrak{l}] \subset \mathfrak{l}_+$. Similarly $[\mathfrak{l}_-, \mathfrak{l}] \subset \mathfrak{l}_-$ so the subspaces $\mathfrak{l}_0$, $\mathfrak{l}_-$, $\mathfrak{l}_+$ are ideals in $\mathfrak{l}$. This being so, their Killing forms are obtained from B by restriction. Since B is strictly negative on $\mathfrak{l}_-$, it follows (Prop. 6.6, Chapter II) that $\mathfrak{l}_-$ is a semisimple compact Lie algebra. Since B is strictly negative definite on $\mathfrak{u}_+$, and strictly positive definite on $\mathfrak{e}_+$, $\mathfrak{l}_+$ is semisimple and it follows (Prop. 7.4, Chapter III) that the decomposition $\mathfrak{l}_+ = \mathfrak{u}_+ + \mathfrak{e}_+$ is a Cartan decomposition of $\mathfrak{l}_+$. Finally we consider $\mathfrak{l}_0$. Since the center $\mathfrak{z}_0$ of $\mathfrak{l}_0$ coincides with the center $\mathfrak{z}$ of $\mathfrak{l}$ we have $\mathfrak{u}_0 \cap \mathfrak{z}_0 \subset \mathfrak{u} \cap \mathfrak{z} = \{0\}$.

In order to show that $\mathfrak{u}_0$ is compactly imbedded in $\mathfrak{l}_0$, we make use of the following lemma which was communicated to the author by J. Hano.

Lemma 1.6. *Let G_0 be a Lie group and $\mathfrak{g}_0$ its Lie algebra. Suppose $\mathfrak{g}_0$ is the direct sum of two ideals $\mathfrak{g}_1$ and $\mathfrak{g}_2$. Let $\mathfrak{k}_i$ be a subalgebra of $\mathfrak{g}_i$ $(i = 1, 2)$, and put $\mathfrak{k}_0 = \mathfrak{k}_1 + \mathfrak{k}_2$. Then $\mathfrak{k}_0$ is a compactly imbedded subalgebra of $\mathfrak{g}_0$, if and only if $\mathfrak{k}_1$ and $\mathfrak{k}_2$ are compactly imbedded in $\mathfrak{g}_1$ and $\mathfrak{g}_2$, respectively.*

Proof. Without loss of generality, we can assume that G_0 is simply connected and is the direct product $G_0 = G_1 \times G_2$ where G_i is a simply connected Lie group with Lie algebra $\mathfrak{g}_i$ $(i = 1, 2)$. Let K_i denote the analytic subgroup of G_i with Lie algebra $\mathfrak{k}_i$ $(i = 0, 1, 2)$. Then $K_0 = K_1 \times K_2$. If Z_i denotes the center of G_i $(i = 0, 1, 2)$, then $Z_0 = Z_1 \times Z_2$. The mapping $(k_1(K_1 \cap Z_1), k_2(K_2 \cap Z_2)) \to k_1k_2(K_0 \cap Z_0)$ is a topological isomorphism of the product group

$$(K_1/K_1 \cap Z_1) \times (K_2/K_2 \cap Z_2) \quad \text{onto} \quad K_0/K_0 \cap Z_0.$$

In view of Lemma 5.1, Chapter II, the Lie group $K_i/K_i \cap Z_i$ is analytically isomorphic to the Lie subgroup $\mathrm{Ad}_{G_i}(K_i)$ of $\mathrm{Int}(\mathfrak{g}_i)$ $(i = 0, 1, 2)$. The lemma now follows immediately.

Returning now to Theorem 1.1, we first note that there exists a Lie group L whose Lie algebra is isomorphic to $\mathfrak{l}$. In fact, if $c = \dim \mathfrak{z}$, then the product group $L = \mathrm{Int}(\mathfrak{l}) \times \boldsymbol{R}^c$ has for Lie algebra the product $\mathrm{ad}_{\mathfrak{l}}(\mathfrak{l}) \times \mathfrak{z}$. To see that this Lie algebra is isomorphic to $\mathfrak{l}$, we observe that $\mathfrak{z} \subset \mathfrak{e}$ and denote by $\mathfrak{e}'$ the orthogonal complement (with respect to Q) of $\mathfrak{z}$ in $\mathfrak{e}$. Then $[\mathfrak{u}, \mathfrak{e}'] \subset \mathfrak{e}'$ and $[\mathfrak{e}', \mathfrak{e}'] \subset \mathfrak{u}$ so the subspace $\mathfrak{u} + \mathfrak{e}'$ is an ideal in $\mathfrak{l}$, isomorphic to $\mathrm{ad}_{\mathfrak{l}}(\mathfrak{l})$.

From Lemma 1.6 it now follows that $\mathfrak{u}_0$ is compactly imbedded in $\mathfrak{l}_0$. Moreover, $\mathfrak{e}_0$ is an abelian ideal in $\mathfrak{l}_0$ so $(\mathfrak{l}_0, s_0)$ is an orthogonal symmetric Lie algebra of the Euclidean type.

It remains to consider the case $\mathfrak{e}_0 = \{0\}$. In this case $\mathfrak{u}_0$ is an ideal in $\mathfrak{l}$. Hence its Killing form is strictly negative definite so $\mathfrak{u}_0$ is compact and semisimple. We put

$$\mathfrak{l}_0 = \{0\}, \qquad \mathfrak{l}_- = \mathfrak{u}_0 + \mathfrak{u}_- + \mathfrak{e}_-, \qquad \mathfrak{l}_+ = \mathfrak{u}_+ + \mathfrak{e}_+ \qquad \text{if } \mathfrak{e}_- \neq \{0\};$$
$$\mathfrak{l}_0 = \{0\}, \qquad \mathfrak{l}_- = \{0\}, \qquad \mathfrak{l}_+ = \mathfrak{u}_0 + \mathfrak{u}_+ + \mathfrak{e}_+ \qquad \text{if } \mathfrak{e}_- = \{0\}.$$

In each case, Theorem 1.1 follows easily.

Corollary 1.7. *Suppose $X \in \mathfrak{e}$ commutes elementwise with $\mathfrak{u}$. Then $X \in \mathfrak{e}_0$.*

In fact, we can write $X = X_0 + X_- + X_+$ where $X_0 \in \mathfrak{e}_0$, $X_- \in \mathfrak{e}_-$, and $X_+ \in \mathfrak{e}_+$. Then the hypothesis implies that $\mathrm{ad}_{\mathfrak{l}}(X_+)$ and $\mathrm{ad}_{\mathfrak{l}}(X_-)$

map $\mathfrak{u}$ into $\{0\}$. Thus $(\mathrm{ad}_{\mathfrak{l}}(X_+))^2 = (\mathrm{ad}_{\mathfrak{l}}(X_-))^2 = 0$ and therefore $B(X_+, X_+) = B(X_-, X_-) = 0$. Hence $X_+ = X_- = 0$ and $X \in \mathfrak{e}_0$.

In the sequel, the Euclidean type will mostly be left out because the associated symmetric spaces are just spaces covered by Euclidean spaces.

§ 2. The Duality

There is a remarkable and important duality between the compact type and the noncompact type. Let $(\mathfrak{l}, s)$ be an orthogonal symmetric Lie algebra and put $\mathfrak{l} = \mathfrak{u} + \mathfrak{e}$ as in (1), §1. Let $\mathfrak{l}^*$ denote the subset $\mathfrak{u} + i\mathfrak{e}$ of the complexification $\mathfrak{l}^C$ of $\mathfrak{l}$. With the bracket operation inherited from $\mathfrak{l}^C$, $\mathfrak{l}^*$ is a Lie algebra over $\boldsymbol{R}$. The mapping $s^* : T + iX \to T - iX$ ($T \in \mathfrak{u}$, $X \in \mathfrak{e}$) is an involutive automorphism of $\mathfrak{l}^*$. As will be verified presently, $(\mathfrak{l}^*, s^*)$ is an orthogonal symmetric Lie algebra, called the *dual* of $(\mathfrak{l}, s)$. Then $(\mathfrak{l}, s)$ is the dual of $(\mathfrak{l}^*, s^*)$.

Proposition 2.1. *Let $(\mathfrak{l}, s)$ be an orthogonal symmetric Lie algebra. Then:*

(i) *The pair $(\mathfrak{l}^*, s^*)$ is an orthogonal symmetric Lie algebra.*

(ii) *If $(\mathfrak{l}, s)$ is of the compact type, then $(\mathfrak{l}^*, s^*)$ is of the noncompact type and conversely.*

(iii) *If $(\mathfrak{l}_1, s_1)$ is isomorphic to $(\mathfrak{l}_2, s_2)$, then $(\mathfrak{l}_1^*, s_{1^*})$ is isomorphic to $(\mathfrak{l}_2^*, s_{2^*})$.*

Proof. (i) Let $(\mathfrak{l}^C)^R$ denote the Lie algebra $\mathfrak{l}^C$ when considered as a Lie algebra over $\boldsymbol{R}$. Then the Lie algebra $(\mathfrak{l}^C)^R$ has a complex structure J given by the multiplication by i on $\mathfrak{l}^C$. Each endomorphism A of $\mathfrak{l}$ or $\mathfrak{l}^*$ extends uniquely to a linear transformation of $(\mathfrak{l}^C)^R$ commuting with J. In this way the Lie groups $\boldsymbol{GL}(\mathfrak{l})$ and $\boldsymbol{GL}(\mathfrak{l}^*)$ become closed Lie subgroups of $\boldsymbol{GL}((\mathfrak{l}^C)^R)$. Consequently, the adjoint groups Int $(\mathfrak{l})$ and Int $(\mathfrak{l}^*)$ are Lie subgroups of $\boldsymbol{GL}((\mathfrak{l}^C)^R)$. Let U denote the analytic subgroup of Int $(\mathfrak{l})$ with Lie algebra $\mathrm{ad}_{\mathfrak{l}}(\mathfrak{u})$. Now U is compact, so by Cor. 2.9, Chapter II, $\mathfrak{u}$ is compactly imbedded in $\mathfrak{l}^*$. For (ii) one just has to observe that $\mathfrak{l}$ and $\mathfrak{l}^*$ are real forms of $\mathfrak{l}^C$ so their Killing forms are obtained from the Killing form of $\mathfrak{l}^C$ by restriction. For (iii) suppose φ is an isomorphism of $\mathfrak{l}_1$ onto $\mathfrak{l}_2$ such that $\varphi \circ s_1 = s_2 \circ \varphi$. Then φ extends uniquely to an isomorphism $\tilde{\varphi}$ of $\mathfrak{l}_1^C$ onto $\mathfrak{l}_2^C$. The restriction of $\tilde{\varphi}$ to $\mathfrak{l}_1^*$ then sets up the required isomorphism between $(\mathfrak{l}_1^*, s_1^*)$ and $(\mathfrak{l}_2^*, s_2^*)$.

Let $\mathfrak{g}$ be a semisimple Lie algebra over $\boldsymbol{C}$. If $\mathfrak{l}$ runs through all compact real forms of $\mathfrak{g}$ and s runs through all involutive automorphisms of $\mathfrak{l}$, then $\mathfrak{l}^*$ runs through all noncompact real forms of $\mathfrak{g}$.

Proposition 2.2. *Let $\mathfrak{l}$ be a compact semisimple Lie algebra. Let s_1 and s_2 be two involutive automorphisms of $\mathfrak{l}$ and let $\mathfrak{l}_1^*$ and $\mathfrak{l}_2^*$ denote the corresponding real forms of $\mathfrak{l}^C$. Then s_1 and s_2 are conjugate within the group Aut $(\mathfrak{l})$ if and only if $\mathfrak{l}_1^*$ and $\mathfrak{l}_2^*$ are conjugate under an automorphism of $\mathfrak{l}^C$.*

Proof. Suppose first that there exists a $\sigma \in \text{Aut}(\mathfrak{l})$ such that $s_2 = \sigma s_1 \sigma^{-1}$. Let $\mathfrak{l} = \mathfrak{u}_1 + \mathfrak{e}_1$, $\mathfrak{l} = \mathfrak{u}_2 + \mathfrak{e}_2$ be the direct decompositions of $\mathfrak{l}$ into eigenspaces of s_1 and s_2, respectively. Then $\sigma \mathfrak{u}_1 = \mathfrak{u}_2$ and $\sigma \mathfrak{e}_1 = \mathfrak{e}_2$. Let Σ denote the unique extension of σ to a (complex) automorphism of $\mathfrak{l}^C$. Since $\mathfrak{l}_1^* = \mathfrak{u}_1 + i\mathfrak{e}_1$, $\mathfrak{l}_2^* = \mathfrak{u}_2 + i\mathfrak{e}_2$, it is obvious that $\Sigma \cdot \mathfrak{l}_1^* = \mathfrak{l}_2^*$.

For the converse (and nontrivial) part of Prop. 2.2 we shall use Theorem 7.2, Chapter III, stating that two Cartan decompositions of a semisimple Lie algebra are necessarily conjugate under an inner automorphism. Suppose then that there exists an automorphism Σ of $\mathfrak{l}^C$ such that $\Sigma \cdot \mathfrak{l}_1^* = \mathfrak{l}_2^*$. Then the two Cartan decompositions $\mathfrak{l}_2^* = \mathfrak{u}_2 + i\mathfrak{e}_2$ and $\mathfrak{l}_2^* = \Sigma \cdot \mathfrak{u}_1 + i\Sigma \cdot \mathfrak{e}_1$ are conjugate under an inner automorphism γ of $\mathfrak{l}_2^*$. Let Γ denote the unique extension of γ to an automorphism of $\mathfrak{l}^C$. Then the automorphism $\Gamma \circ \Sigma$ leaves $\mathfrak{l}$ invariant and its restriction to $\mathfrak{l}$ sets up the desired conjugacy of s_2 and s_1.

As shown in Chapter IV, any compact Lie group can be given the structure of a Riemannian globally symmetric space. We shall now see that the subclass of orthogonal symmetric Lie algebras of compact type, so obtained, corresponds, under the duality, to the class of orthogonal symmetric Lie algebras $(\mathfrak{l}, s)$ of noncompact type, where $\mathfrak{l}$ has complex structure and s is a conjugation.

Proposition 2.3. *Let $\mathfrak{l}_0$ be a compact semisimple Lie algebra and let s denote the automorphism $(X, Y) \to (Y, X)$ of the product algebra $\mathfrak{l} = \mathfrak{l}_0 \times \mathfrak{l}_0$. Then $(\mathfrak{l}, s)$ is an orthogonal symmetric Lie algebra of the compact type. If $(\mathfrak{l}^*, s^*)$ denotes the dual of $(\mathfrak{l}, s)$, then $\mathfrak{l}^*$ is isomorphic (as a real Lie algebra) to a complex subalgebra $\mathfrak{a}$ of $\mathfrak{l}^C$ in such a way that s^* corresponds to the conjugation of $\mathfrak{a}$ with respect to a compact real form of $\mathfrak{a}$.*

Proof. Let $\mathfrak{u} = \{(X, X) : X \in \mathfrak{l}_0\}$ and $\mathfrak{e} = \{(X, -X) : X \in \mathfrak{l}_0\}$. Then the direct decomposition $\mathfrak{l} = \mathfrak{u} + \mathfrak{e}$ is the usual decomposition of $\mathfrak{l}$ into eigenspaces of s. We have $\mathfrak{l}^* = \mathfrak{u} + i\mathfrak{e}$ and $\mathfrak{l}^C = \mathfrak{l} + i\mathfrak{l}$. The algebra $\mathfrak{a} = \mathfrak{u} + i\mathfrak{u}$ is a (complex) subalgebra of $\mathfrak{l}^C$ and the mapping

$$\varphi : (X, X) + i(Y, -Y) \to (X, X) + i(Y, Y), \qquad X, Y \in \mathfrak{l}_0,$$

is an isomorphism of $\mathfrak{l}^*$ onto $\mathfrak{a}$ (considered as real Lie algebras). Moreover, if μ denotes the conjugation of $\mathfrak{a}$ with respect to $\mathfrak{u}$, then $\varphi \circ s^* = \mu \circ \varphi$. Since $\mathfrak{u}$ is a compact real form of $\mathfrak{a}$, the proposition is proved.

Theorem 2.4. *Let* $(\mathfrak{l}, s)$ *be an orthogonal symmetric Lie algebra of the compact type and* $(\mathfrak{l}^*, s^*)$ *its dual. Then the Lie algebra* $\mathfrak{l}^*$ *has a complex structure if and only if* $\mathfrak{l}$ *can be written as a direct sum* $\mathfrak{l} = \mathfrak{l}_1 + \mathfrak{l}_2$, *where* $\mathfrak{l}_1$ *and* $\mathfrak{l}_2$ *are ideals in* $\mathfrak{l}$, *which are interchanged by* s.

Proof. Suppose first, that $\mathfrak{l}^*$ has a complex structure, which will be denoted by J in order to avoid confusion with the complex structure of $\mathfrak{l}^C$. Then J satisfies the relation $[X, JY] = J[X, Y]$ for $X, Y \in \mathfrak{l}^*$. Since $\mathfrak{l}^*$ can be considered as a semisimple Lie algebra over $\boldsymbol{C}$ (by means of J), it has a compact real form $\mathfrak{k}$. We have then the direct decomposition

$$\mathfrak{l}^* = \mathfrak{k} + J\mathfrak{k},$$

which is a Cartan decomposition of $\mathfrak{l}^*$. On the other hand, we have the decompositions

$$\mathfrak{l} = \mathfrak{u} + \mathfrak{e}, \qquad \mathfrak{l}^* = \mathfrak{u} + i\mathfrak{e}$$

into eigenspaces of s and s^*, respectively. Since all Cartan decompositions of $\mathfrak{l}^*$ are conjugate under an inner automorphism of $\mathfrak{l}^*$ there exists an element $\sigma \in \operatorname{Int}(\mathfrak{l}^*)$ such that $\sigma \cdot \mathfrak{k} = \mathfrak{u}$, $\sigma \cdot (J\mathfrak{k}) = i\mathfrak{e}$. Consider now the following mappings:

$$\mathfrak{e} \xrightarrow{i} i\mathfrak{e} \xrightarrow{\sigma^{-1}} J\mathfrak{k} \xrightarrow{-J} \mathfrak{k} \xrightarrow{\sigma} \mathfrak{u}$$

and put $\gamma(X) = -\sigma J \sigma^{-1} iX$ for $X \in \mathfrak{e}$. Then γ is a one-to-one linear mapping of $\mathfrak{e}$ onto $\mathfrak{u}$ and has the following properties:

$$\begin{aligned}
&\text{(a)} \quad [\gamma(X), \gamma(Y)] = [X, Y];\\
&\text{(b)} \quad [X, \gamma(Y)] = [\gamma(X), Y];\\
&\text{(c)} \quad \gamma([\gamma(X), Y]) = [X, Y]
\end{aligned}$$

for $X, Y \in \mathfrak{e}$. The last property is verified as follows:

$$\begin{aligned}
\gamma[\gamma X, Y] &= \sigma J \sigma^{-1} i([\sigma J \sigma^{-1} iX, Y]) = \sigma J \sigma^{-1}([\sigma J \sigma^{-1} iX, iY])\\
&= \sigma J[J\sigma^{-1} iX, \sigma^{-1} iY] = -\sigma[\sigma^{-1} iX, \sigma^{-1} iY] = [X, Y].
\end{aligned}$$

Property (a) is proved in the same way and (b) follows from (c). We define now the subsets $\mathfrak{l}_1$ and $\mathfrak{l}_2$ of $\mathfrak{l}$ by

$$\mathfrak{l}_1 = \{X + \gamma(X) : X \in \mathfrak{e}\}, \qquad \mathfrak{l}_2 = \{X - \gamma(X) : X \in \mathfrak{e}\}.$$

Then the following statements hold:

(i) $\mathfrak{l}_1 \cap \mathfrak{l}_2 = \{0\}$.
(ii) s interchanges $\mathfrak{l}_1$ and $\mathfrak{l}_2$.
(iii) $\mathfrak{l}_1$ and $\mathfrak{l}_2$ are ideals in $\mathfrak{l}$.

The first statement is obvious because the relation $X + \gamma(X) = Y - \gamma(Y)$ implies $\gamma(X + Y) = Y - X \in \mathfrak{e} \cap \mathfrak{u} = \{0\}$. The second statement is obvious since $s(X + \gamma(X)) = - X + \gamma(X)$. Finally, properties (a), (b), and (c) above show that $\mathfrak{l}_1$ and $\mathfrak{l}_2$ are subalgebras of $\mathfrak{l}$ and $[\mathfrak{l}_1, \mathfrak{l}_2] = \{0\}$, proving (iii). This proves the first half of the theorem.

Remark. The mapping $X + \gamma(X) \to 2\gamma(X)$ is an isomorphism of $\mathfrak{l}_1$ onto $\mathfrak{u}$.

In order to prove the second half of the theorem, we consider a Lie algebra $\mathfrak{l}_0$ isomorphic to both $\mathfrak{l}_1$ and $\mathfrak{l}_2$. Let I_i denote an isomorphism of $\mathfrak{l}_i$ onto $\mathfrak{l}_0$, $(i = 1, 2)$. In the following let X denote an arbitrary element in $\mathfrak{l}_1$ and Y an arbitrary element in $\mathfrak{l}_2$. Let $\bar{\mathfrak{l}}$ denote the product algebra $\mathfrak{l}_0 \times \mathfrak{l}_0$ and let $\bar{s}$, s_0 denote the automorphisms of $\bar{\mathfrak{l}}$ given by

$$\bar{s} \cdot (I_1X, I_2Y) = (I_1sY, I_2sX),$$
$$s_0 \cdot (I_1X, I_2Y) = (I_2Y, I_1X).$$

Then we have the isomorphisms

$$(\mathfrak{l}, s) \xrightarrow{I_0} (\bar{\mathfrak{l}}, \bar{s}) \xrightarrow{\sigma} (\bar{\mathfrak{l}}, s_0),$$

where

$$I_0(X + Y) = (I_1X, I_2Y),$$
$$\sigma(I_1X, I_2Y) = (I_1X, I_1sY).$$

Let $(\bar{\mathfrak{l}}^*, \bar{s}^*)$ denote the dual of $(\bar{\mathfrak{l}}, \bar{s})$ and let $(\mathfrak{g}, s_0^*)$ denote the dual of $(\bar{\mathfrak{l}}, s_0)$. Then, as a result of Prop. 2.1, the orthogonal symmetric Lie algebras

$$(\mathfrak{l}^*, s^*), \qquad (\bar{\mathfrak{l}}^*, \bar{s}^*), \qquad (\mathfrak{g}, s_0^*)$$

are all isomorphic. But due to Prop. 2.3, $\mathfrak{g}$ has a complex structure. Using Prop. 2.2 it follows that $\bar{\mathfrak{l}}^*$, and therefore $\mathfrak{l}^*$, has a complex structure. This concludes the proof of the theorem.

Example 1. *The coset spaces $SO(p + q)/SO(p) \times SO(q)$ and $SO_0(p, q)/SO(p) \times SO(q)$.*

Let $SO(p, q)$ denote the group of real quadratic matrices of determinant 1, leaving invariant the quadratic form

$$- x_1^2 - \dots - x_p^2 + x_{p+1}^2 + \dots + x_{p+q}^2 \qquad (p + q > 2).$$

Let I_n denote the unit matrix of order n and put

$$I_{p,q} = \begin{pmatrix} -I_p & 0 \\ 0 & I_q \end{pmatrix}.$$

Denoting by ${}^tA$ the transpose of a matrix A, we see that a matrix g of determinant 1 belongs to $SO(p, q)$ if and only if ${}^tgI_{p,q}g = I_{p,q}$. Thus $SO(p, q)$ is a closed subgroup of $GL(p + q, R)$, hence a topological Lie subgroup. Its Lie algebra, denoted $\mathfrak{so}(p, q)$, is a subalgebra of $\mathfrak{gl}(p + q, R)$. According to formula (2) in Chapter II, §2 we have

$$X \in \mathfrak{so}(p,q) \text{ if and only if } e^{sX} \in SO(p,q) \text{ for all } s \in R.$$

Now $e^X \in SO(p, q)$ if and only if ${}^t(e^X) = I_{p,q}\, e^{-X}I_{p,q}$; since ${}^t(e^X) = e^{{}^tX}$, it follows that

$$X \in \mathfrak{so}(p, q) \text{ if and only if } {}^tXI_{p,q} + I_{p,q}X = 0.$$

Thus $\mathfrak{so}(p, q)$ is the set of matrices

$$X = \begin{pmatrix} X_1 & X_2 \\ {}^tX_2 & X_3 \end{pmatrix}$$

where X_1 and X_3 are skew symmetric matrices of order p and q, respectively, and X_2 is an arbitrary $p \times q$ matrix. In particular, the Lie algebra $\mathfrak{so}(n)$ of the group $SO(n)$ ($= SO(n, 0)$) consists of all $n \times n$ skew symmetric matrices.

Now let $\mathfrak{l} = \mathfrak{so}(p + q)$ and let s denote the restriction to $\mathfrak{l}$ of the automorphism $\sigma_{p,q} : X \to I_{p,q}XI_{p,q}$ of $\mathfrak{gl}(p + q, C)$. Then $(\mathfrak{l}, s)$ is an orthogonal symmetric Lie algebra of the compact type. If $\mathfrak{l} = \mathfrak{u} + \mathfrak{e}$ is the usual decomposition into eigenspaces of s, it is easily seen that

$$\mathfrak{u} = \left\{ \begin{pmatrix} X_1 & 0 \\ 0 & X_3 \end{pmatrix} \middle| \begin{matrix} X_1 : p \times p \text{ skew symmetric matrix} \\ X_3 : q \times q \text{ skew symmetric matrix} \end{matrix} \right\},$$

$$\mathfrak{e} = \left\{ \begin{pmatrix} 0 & X_2 \\ -{}^tX_2 & 0 \end{pmatrix} \middle| X_2 : p \times q \text{ arbitrary matrix} \right\}.$$

The pair $(SO(p + q), SO(p) \times SO(q))$ is therefore associated with $(\mathfrak{l}, s)$. Let $(\mathfrak{l}^*, s^*)$ denote the dual of $(\mathfrak{l}, s)$. Then $\mathfrak{l}^*$ is the subalgebra of $\mathfrak{gl}(p+q,C)$ given by $\mathfrak{l}^* = \mathfrak{u} + i\mathfrak{e}$ and s^* is again the restriction of $\sigma_{p,q}$ to $\mathfrak{l}^*$. Now it is easy to verify that the mapping

$$\begin{pmatrix} X_1 & iX_2 \\ -i\,{}^tX_2 & X_3 \end{pmatrix} \to \begin{pmatrix} X_1 & X_2 \\ {}^tX_2 & X_3 \end{pmatrix} = \begin{pmatrix} -iI_p & 0 \\ 0 & I_q \end{pmatrix} \begin{pmatrix} X_1 & iX_2 \\ -i\,{}^tX_2 & X_3 \end{pmatrix} \begin{pmatrix} iI_p & 0 \\ 0 & I_q \end{pmatrix}$$

is an isomorphism of $\mathfrak{l}^*$ onto $\mathfrak{so}(p, q)$. Under this isomorphism, the automorphism s^* corresponds again to the automorphism $X \to I_{p,q}XI_{p,q}$ of $\mathfrak{so}(p, q)$. Let $SO_0(p, q)$ denote the identity component of $SO(p, q)$. Then the pair $(SO_0(p, q), SO(p) \times SO(q))$ is associated with the orthogonal symmetric Lie algebra $(\mathfrak{so}(p, q), \sigma_{p,q})$, which is isomorphic to the dual of $(\mathfrak{so}(p + q), s)$.

Example II. *The case* $p = 1$, $q = 3$.

Let Q denote the algebra of quaternions, Q_+ the subspace of pure quaternions, and G the multiplicative group of quaternions of norm 1. To each pair x, y from G we associate the endomorphism

$$T_{x,y} : u \to xuy^{-1}, \qquad u \in Q$$

of Q onto itself. Since $T_{x,y}$ is norm preserving, it belongs to the group of rotations of Q; since G is connected it follows that all $T_{x,y}$ belong to a connected part of the group of rotations. Hence $T_{x,y} \in SO(4)$. Each endomorphism $T_{x,x}$ leaves the subspace Q_+ invariant; let τ_x denote the restriction of $T_{x,x}$ to Q_+. Then $\tau_x \in SO(3)$. The following statements hold.

(a) *The mapping* $x \to \tau_x$ *is an analytic homomorphism of* G *onto* $SO(3)$ *and the kernel consists of* e *and* $-e$, e *denoting the identity element in* Q.

(b) *The mapping* $(x, y) \to T_{x,y}$ *is an analytic homomorphism of the product* $G \times G$ *onto* $SO(4)$ *and the kernel consists of* (e, e) *and* $(-e, -e)$.

The verification of (a) and (b) will be left to the reader. Passing to the Lie algebras we obtain an isomorphism φ of $\mathfrak{so}(3) \times \mathfrak{so}(3)$ onto $\mathfrak{so}(4)$. Let s_0 denote the automorphism $(X, Y) \to (Y, X)$ of $\mathfrak{so}(3) \times \mathfrak{so}(3)$. An elementary quaternion computation shows that $\sigma_{1,3}(T_{x,y}) = T_{y,x} (x, y \in G)$. Hence we obtain

(c) *The orthogonal symmetric Lie algebras*

$$(\mathfrak{so}(3) \times \mathfrak{so}(3), s_0) \qquad \textit{and} \qquad (\mathfrak{so}(4), \sigma_{1,3})$$

are isomorphic under φ.

Now by Prop. 2.1 (iii), the duals of these pairs are isomorphic. In view of Theorem 2.4 we can conclude that the Lie algebra $\mathfrak{so}(1, 3)$ has a complex structure. Since this Lie algebra has dimension 6, it is isomorphic to a semisimple, three dimensional complex Lie algebra (considered as a real Lie algebra). Such an algebra must have a Cartan subalgebra of dimension 1 and two nonzero roots α, $-\alpha$. In view of Theorem 5.4, Chapter III, there is therefore at most one three-dimensional complex semisimple Lie algebra. On the other hand, the Lie algebra $\mathfrak{sl}(2, C)$ of all complex 2×2 matrices of trace 0 is such an algebra. We can therefore conclude that $\mathfrak{so}(1, 3)$ is isomorphic to $\mathfrak{sl}(2, C)^R$.

§ 3. Sectional Curvature of Symmetric Spaces

The three classes of symmetric spaces can be distinguished by means of their curvature as shown in the following theorem.

Theorem 3.1. *Let $(\mathfrak{l}, s)$ be an orthogonal symmetric Lie algebra and suppose that the pair (L, U) is associated with $(\mathfrak{l}, s)$. We assume that U is connected and closed.*† *Let Q be an arbitrary L-invariant Riemannian structure on L/U.*

(i) *If (L, U) is of the compact type, then L/U has sectional curvature everywhere $\geqslant 0$.*

(ii) *If (L, U) is of the noncompact type, then L/U has sectional curvature everywhere $\leqslant 0$.*

(iii) *If (L, U) is of the Euclidean type, then L/U has sectional curvature everywhere $= 0$.*

Proof. The tangent space to L/U at the point $o = \{U\}$ can, as usual, be identified with $\mathfrak{e}$, the eigenspace of s for the eigenvalue -1. Let S be a two-dimensional subspace of $\mathfrak{e}$, and let X, Y be an orthonormal basis of S. Then, according to Theorem 4.2, Chapter IV, the curvature of L/U along the section S is given by

$$K(S) = -Q_o(R(X, Y)X, Y) = +Q_o([[X, Y], X], Y).$$

Part (iii) is now obvious, so we can assume that $\mathfrak{l}$ is semisimple. As in §1, let b denote the endomorphism of $\mathfrak{e}$ given by

$$Q_o(bX, Y) = B(X, Y), \qquad X, Y \in \mathfrak{e},$$

B denoting the Killing form of $\mathfrak{l}$. Since $Q_o(bX, Y) = Q_o(X, bY)$, the eigenvalues $\beta_1, \ldots, \beta_n$ of b are real. Let $\mathfrak{e}_1, \ldots, \mathfrak{e}_n$ be the corresponding eigenspaces of b. Then, if $i \neq j$, the spaces $\mathfrak{e}_i$ and $\mathfrak{e}_j$ are orthogonal with respect to B as well as Q_o. We shall prove that $[\mathfrak{e}_i, \mathfrak{e}_j] = \{0\}$. Let $\mathfrak{u}$ denote the Lie algebra of U. Then b commutes with each member of $\mathrm{ad}_{\mathfrak{l}}(\mathfrak{u})$. Hence $[\mathfrak{u}, \mathfrak{e}_i] \subset \mathfrak{e}_i$ for each i. Now let $X_i \in \mathfrak{e}_i$, $X_j \in \mathfrak{e}_j$, $T \in \mathfrak{u}$. Then $[X_i, X_j] \in \mathfrak{u}$ and

$$B(T, [X_i, X_j]) = B([T, X_i], X_j) = 0.$$

Owing to Lemma 1.2, B is strictly negative definite on $\mathfrak{u}$; consequently $[\mathfrak{e}_i, \mathfrak{e}_j] = \{0\}$.

† For the noncompact type this hypothesis is always satisfied (see Chapter VI). For the compact type, U is always closed but not necessarily connected (see Chapter VII).

Now, going back to the formula for $K(S)$, let X_i, Y_i $(1 \leqslant i \leqslant n)$ be the components of X and Y, respectively, in the eigenspaces $\mathfrak{e}_i$. Then

$$[X, Y] = \sum_{i=1}^{n} [X_i, Y_i], \qquad [[X_i, Y_i], X] = [[X_i, Y_i], X_i]$$

so

$$K(S) = \sum_{i=1}^{n} Q_0([[X_i, Y_i], X_i], Y_i) = \sum_{i=1}^{n} \frac{1}{\beta_i} B([[X_i, Y_i], X_i], Y_i)$$

and

$$K(S) = \sum_{i=1}^{n} \frac{1}{\beta_i} B([X_i, Y_i], [X_i, Y_i]). \tag{1}$$

Since $\beta_i < 0$ or $\beta_i > 0$ $(1 \leqslant i \leqslant n)$, in the cases (i) and (ii), respectively, the theorem follows.

Remark. Suppose $(\mathfrak{l}, s)$ is one of the types (i), (ii), or (iii). Then the curvature $K(S)$ along the section S is 0 if and only if S is an abelian subspace of $\mathfrak{e}$.

Example. As a special case of Example I, §2, we consider the spaces $\boldsymbol{SO}(p+1)/\boldsymbol{SO}(p)$ and $\boldsymbol{SO}_0(p, 1)/\boldsymbol{SO}(p)$ which correspond to each other under the duality. Here the linear isotropy group at a point p acts transitively on the set of two-dimensional subspaces of the tangent space at p. Hence these spaces have constant sectional curvature. In particular, for $p = 2$, the spaces are the two-dimensional sphere and the two-dimensional non-Euclidean space of Lobatschevsky.

For the sphere $\boldsymbol{SO}(3)/\boldsymbol{SO}(2)$ the formulas

$$\frac{\sin a}{\sin A} = \frac{\sin b}{\sin B} = \frac{\sin c}{\sin C},$$

$$\cos a = \cos b \cos c + \sin b \sin c \cos A$$

hold for a geodesic triangle with angles A, B, C and sides of length a, b, c. For the two-dimensional Lobatschevsky space $\boldsymbol{SO}_0(2, 1)/\boldsymbol{SO}(2)$ the formulas are

$$\frac{\sinh a}{\sin A} = \frac{\sinh b}{\sin B} = \frac{\sinh c}{\sin C},$$

$$\cosh a = \cosh b \cosh c - \sinh b \sinh c \cos A.$$

Since $\sinh iz = i \sin z$ and $\cosh iz = \cos z$, the two sets of formulas correspond under the substitution $a \to ia$, $b \to ib$, $c \to ic$. The duality for symmetric spaces gives a general explanation of this formal analogy between spherical trigonometry and non-Euclidean trigonometry.

§ 4. Symmetric Spaces with Semisimple Groups of Isometries

Theorem 4.1. *Let (G, K) be a Riemannian symmetric pair. Suppose that G is semisimple and acts effectively on the coset space $M = G/K$. Let $\underset{\sim}{Q}$ be any G-invariant Riemannian structure on M and let R denote the corresponding curvature tensor. Then:*

(i) *$G = I_0(M)$ (as Lie groups).*

(ii) *The linear isotropy subgroup of G at $o = \{K\}$ is a Lie subgroup K^* of $\mathbf{GL}(M_o)$, isomorphic to K. Its Lie algebra $\mathfrak{k}^*$ consists of all endomorphisms of M_o which, when extended to the mixed tensor algebra over M_o as derivations commuting with contractions, annihilate Q_o and R_o.*

(iii) *$\mathfrak{k}^*$ is spanned by the set $\{R_o(X, Y) : X, Y \in M_o\}$.*

Proof. According to Prop. 3.5, Chapter IV, there exists a unique analytic involutive automorphism σ of G such that $(K_\sigma)_0 \subset K \subset K_\sigma$. Here K_σ denotes the set of fixed points of σ and $(K_\sigma)_0$ is the identity component of K_σ. Let s_o denote the geodesic symmetry of G/K with respect to o. Then, as proved in Chapter IV (Prop. 3.4),

$$\sigma(g) = s_o g s_o, \qquad g \in G. \tag{1}$$

Let $\mathfrak{g} = \mathfrak{k} + \mathfrak{p}$ be the direct decomposition of the Lie algebra $\mathfrak{g}$ of G into the eigenspaces of $(d\sigma)_e$ for the eigenvalues $+1$ and -1, respectively. Let Z denote the center of G. According to Lemma 5.1, Chapter II, the group $\mathrm{Ad}_G(K)$ is analytically isomorphic to $K/K \cap Z$ which equals K, G being effective. Thus K is compact and isomorphic to the linear isotropy group K^*. Let $G' = I_0(M)$ and let K' denote the (compact) subgroup of G' leaving the point o fixed. Owing to Remark 2 following Prop. 3.4, Chapter IV, the group G is a closed subgroup of G'. Hence $\mathfrak{g}$ is a subalgebra of the Lie algebra $\mathfrak{g}'$ of G'. Let $\tilde{\sigma}$ denote the automorphism $g \to s_0 g s_0$ of G' and let

$$\mathfrak{g}' = \mathfrak{k}' + \mathfrak{p}'$$

be the decomposition of $\mathfrak{g}'$ into the eigenspaces of $(d\tilde{\sigma})_e$, $\mathfrak{k}'$ being the Lie algebra of K'. Here $\mathfrak{k} \subset \mathfrak{k}'$, $\mathfrak{p} = \mathfrak{p}'$. We now apply Theorem 1.1 and the terminology introduced there to the pair $(\mathfrak{g}', (d\tilde{\sigma})_e)$ which is an effective orthogonal symmetric Lie algebra. The subspace $(\mathfrak{p}')_0$ is an abelian ideal in $\mathfrak{g}'$ and $\mathfrak{g}$ (Lemma 1.3). Since $\mathfrak{g}$ is semisimple, $(\mathfrak{p}')_0 = \{0\}$. Hence $(\mathfrak{k}')_0$ is an ideal of $\mathfrak{g}'$ contained in $\mathfrak{k}'$; thus $(\mathfrak{k}')_0 = \{0\}$ and $[\mathfrak{p}', \mathfrak{p}'] = \mathfrak{k}'$. Since $\mathfrak{p} = \mathfrak{p}'$ and $[\mathfrak{p}, \mathfrak{p}] \subset \mathfrak{k}$, it follows that $\mathfrak{k}' = \mathfrak{k} = [\mathfrak{p}, \mathfrak{p}]$ and $\mathfrak{g} = \mathfrak{g}'$, proving (i). Moreover, the relation $[\mathfrak{p}, \mathfrak{p}] = \mathfrak{k}$ is equivalent to (iii), in view of the formula for the curvature tensor (Theorem 4.2, Chapter IV).

Finally, in order to prove (ii), let $\tilde{\mathfrak{k}}$ denote the Lie algebra of all endomorphisms of M_o which, when extended to the mixed tensor algebra over M_o as derivations commuting with contractions, annihilate Q_o and R_o. Then $\mathfrak{k} \subset \tilde{\mathfrak{k}}$ and the space $\tilde{g} = \tilde{\mathfrak{k}} + \mathfrak{p}$ is a Lie algebra if the bracket $[T, X]$ for $T \in \tilde{\mathfrak{k}}$, $X \in \mathfrak{p}$ is defined as $T \cdot X$ (T operating on X). As in the proof of $[\mathfrak{p}, \mathfrak{p}] = \mathfrak{k}'$ we see that $[\mathfrak{p}, \mathfrak{p}] = \tilde{\mathfrak{k}}$ so $\mathfrak{k} = \tilde{\mathfrak{k}}$ and (ii) follows.

Definition. Let M be a Riemannian globally symmetric space; M is said to be of the *compact type* or the *noncompact type* according to the type of the Riemannian symmetric pair $(I_0(M), K)$, K being the isotropy subgroup of $I_0(M)$ at some point in M. If $(\mathfrak{g}, \theta)$ is the corresponding orthogonal involutive Lie algebra, M is said to be *associated with* $(\mathfrak{g}, \theta)$.

Proposition 4.2. *Let M be a simply connected Riemannian globally symmetric space. Then M is a product*

$$M = M_0 \times M_- \times M_+,$$

where M_0 is a Euclidean space, M_- and M_+ are Riemannian globally symmetric of the compact and noncompact type, respectively.

Proof. Let $G = I_0(M)$ and let K denote the isotropy subgroup at some point o in M. Then $M = G/K$. Let $(\tilde{G}, \varphi)$ denote the universal covering group of G and let $\tilde{K}$ denote the identity component of $\varphi^{-1}(K)$. Then if ψ denotes the mapping $g\tilde{K} \to \varphi(g) K$ of $\tilde{G}/\tilde{K}$ onto G/K, the pair $(\tilde{G}/\tilde{K}, \psi)$ is a covering manifold of G/K. Since M is simply connected, $M = \tilde{G}/\tilde{K}$.

Let s denote the involutive automorphism of $\mathfrak{g}$, the Lie algebra of G (and $\tilde{G}$), which corresponds to the automorphism $g \to s_o g s_o$ of G. Then $(\mathfrak{g}, s)$ is an effective orthogonal symmetric Lie algebra. We decompose $\mathfrak{g}$ according to Theorem 1.1 and let $\tilde{G} = G_0 \times G_- \times G_+$ be the corresponding decomposition of $\tilde{G}$. If $\tilde{K} = K_0 \times K_- \times K_+$ is the decomposition induced on $\tilde{K}$, the spaces $M_0 = G_0/K_0$, $M_- = G_-/K_-$ and $M_+ = G_+/K_+$ have the required properties.

§ 5. Notational Conventions

In order to avoid repeated explanation of notation we shall now make some notational conventions which will be in force in the subsequent chapters of this book.

The symbol $\mathfrak{g}_0$ shall denote an arbitrary semisimple Lie algebra over $\boldsymbol{R}$, and $\mathfrak{g}$ its complexification. Let $\mathfrak{g}_0 = \mathfrak{k}_0 + \mathfrak{p}_0$ be any Cartan decomposition of $\mathfrak{g}_0$ ($\mathfrak{k}_0$ the algebra), and let $\mathfrak{u}$ denote the compact real form $\mathfrak{k}_0 + i\mathfrak{p}_0$ of $\mathfrak{g}$. Let B denote the Killing form of $\mathfrak{g}$. Its restrictions to

$\mathfrak{g}_0 \times \mathfrak{g}_0$ and $\mathfrak{u} \times \mathfrak{u}$ are the Killing forms of $\mathfrak{g}_0$ and $\mathfrak{u}$, respectively. Let σ and τ denote the conjugations of $\mathfrak{g}$ with respect to $\mathfrak{g}_0$ and $\mathfrak{u}$, respectively, and put $\theta = \sigma\tau = \tau\sigma$. Then θ is an involutive automorphism of $\mathfrak{g}$. Subspaces of $\mathfrak{g}_0$ will usually be denoted by the subscript 0; the corresponding subspace of $\mathfrak{g}$ will then be denoted by the same letter but without the subscript. According to this convention, $\mathfrak{k}$ and $\mathfrak{p}$ denote the eigenspaces of the automorphism θ. If $\mathfrak{e}_0$ is a subspace of $\mathfrak{p}_0$, the subspace $i\mathfrak{e}_0$ of $i\mathfrak{p}_0$ will often be denoted by $\mathfrak{e}_*$.

The adjoint groups Int ($\mathfrak{g}_0$), Int ($\mathfrak{u}$) are groups of endomorphisms of $\mathfrak{g}_0$ and $\mathfrak{u}$, which, however, can be extended to endomorphisms of $\mathfrak{g}$. The Lie algebras of Int ($\mathfrak{g}_0$) and Int ($\mathfrak{u}$) will be identified with the subalgebras $\mathfrak{g}_0$ and $\mathfrak{u}$ of $\mathfrak{g}^R$. The analytic subgroup of Int ($\mathfrak{g}_0$) whose Lie algebra is $\mathfrak{k}_0$ will be denoted by K^*. Then K^* is compact and a Lie subgroup of Int ($\mathfrak{u}$). We shall see in the following chapter that $K^* = \text{Int}(\mathfrak{g}_0) \cap \text{Int}(\mathfrak{u})$. All the groups K^*, Int ($\mathfrak{g}_0$), Int ($\mathfrak{u}$) are closed, topological Lie subgroups of $GL(\mathfrak{g}^R)$.

Let (G, K_1) and (U, K_2) be Riemannian symmetric pairs associated with the orthogonal symmetric Lie algebras $(\mathfrak{g}_0, \theta)$ and $(\mathfrak{u}, \theta)$, respectively. In general, the space G/K_1, (and similarly U/K_2), will have many G-invariant Riemannian structures which are not proportional. However, the Riemannian connection is the same for all of these. Accordingly, we shall always (unless the contrary is specified) give G/K_1 the unique G-invariant Riemannian structure induced by the restriction of the Killing form B to $\mathfrak{p}_0 \times \mathfrak{p}_0$. Similarly, the space U/K_2 will be given the unique U-invariant Riemannian structure induced by the restriction of $-B$ to $i\mathfrak{p}_0 \times i\mathfrak{p}_0$.

§ 6. Rank of Symmetric Spaces

Definition. A Riemannian manifold is said to be *flat* if its curvature tensor vanishes identically.

Definition. Let M be a Riemannian globally symmetric space. The *rank* of M is the maximal dimension of a flat, totally geodesic submanifold of M.

Proposition 6.1. *Let M be a Riemannian globally symmetric space of the compact type or the noncompact type. Let o be any point in M and as usual identify the tangent space M_o with the subspace $\mathfrak{p}_0$ (or $i\mathfrak{p}_0$) of the Lie algebra of $I(M)$. Let $\mathfrak{s}_0$ be a Lie triple system contained in M_o. Then the totally geodesic submanifold $S = Exp\, \mathfrak{s}_0$ (with the differentiable structure from Theorem 7.2, Chapter IV) is flat if and only if $\mathfrak{s}_0$ is abelian.*

Proof. As we have seen in §7, Chapter IV, the manifold S is globally symmetric and can be written $S = G'/K'$ where G' is an analytic subgroup of $I(M)$, invariant under the automorphism $g \to s_o g s_o$ of $I(M)$. The curvature tensor of S is given by the formula in Theorem 4.2, Chapter IV. This formula was derived under the assumption of a connected isotropy group. However, passing from K' to its identity component only amounts to passing from S to a covering manifold of S which is locally isometric to S. The proposition now follows immediately from formula (1), §3.

Theorem 6.2. *Let M be a Riemannian globally symmetric space of the compact type or the noncompact type. Let l denote the rank of M and let A and A' denote two flat, totally geodesic submanifolds of M of dimension l.*

(i) *Let $q \in A$, $q' \in A'$. Then there exists an element $x \in I_0(M)$ such that $x \cdot A = A'$, $x \cdot q = q'$.*

(ii) *Let $X \in M_q$. Then there exists an element $k \in I_0(M)$ such that $k \cdot q = q$ and $dk(X) \in A_q$.*

(iii) *The manifolds A and A' are closed topological subspaces of M.*

In order to prove this theorem we begin with some general remarks about totally geodesic submanifolds. Let M be any manifold, S a connected submanifold. Let X and Y be vector fields on M such that $X_s, Y_s \in S_s$ for each $s \in S$. Then it follows easily from Prop. 3.2, Chapter I, that the families $s \to X_s$ $(s \in S)$ and $s \to Y_s$ $(s \in S)$ are vector fields on S. We denote these vector fields by $\bar{X}$ and $\bar{Y}$. It follows from Prop. 3.3, Chapter I, that $[X, Y]_s \in S_s$ for each $s \in S$ and $[X, Y]^- = [\bar{X}, \bar{Y}]$.

Now suppose the manifold M is connected and has a Riemannian structure g. Let $\bar{g}$ denote the induced Riemannian structure on S. Let ∇ and $\bar{\nabla}$ denote the corresponding Riemannian connections. Let X, Y, Z be any vector fields on M for which $X_s, Y_s, Z_s \in S_s$ for each $s \in S$. Then we conclude from (2), §9, Chapter I, that

$$g(X, \nabla_Z(Y))(s) = \bar{g}(\bar{X}, \bar{\nabla}_{\bar{Z}}(\bar{Y}))(s) \qquad (s \in S). \tag{1}$$

Suppose now that S is totally geodesic. Then $\nabla_Z(Y)_s \in S_s$ $(s \in S)$ by Theorem 14.5, Chapter I. Equation (1) therefore implies

$$\bar{\nabla}_{\bar{Z}}(\bar{Y}) = (\nabla_Z(Y))^- \tag{2}$$

Let R and $\bar{R}$ denote the curvature tensors of M and S, respectively. Then by (2)

$$\bar{R}(\bar{X}, \bar{Y}) \cdot \bar{Z} = (R(X, Y) \cdot Z)^- \tag{3}$$

and the sectional curvature along a two-dimensional subspace of S_s is the same for M and S.

After these preliminary remarks let us turn to the proof of Theorem 6.2. Let q be any point in the flat totally geodesic subspace A of M. As usual we identify the tangent space M_q with a subspace of the Lie algebra $\mathfrak{l}$ of $I(M)$. Let X and Y be two vectors in the tangent space A_q. By the preceding remarks the sectional curvature of M at q along the plane section spanned by X and Y is 0. Using (1), §3, it follows that $[X, Y] = 0$, the bracket being that of the Lie algebra $\mathfrak{l}$. Then by Prop. 6.1 A_q is a maximal abelian subspace of M_q. Let G' denote the analytic subgroup of $I(M)$ corresponding to the subalgebra A_q of $\mathfrak{l}$. Let K' denote the subgroup of G' leaving q fixed. The totally geodesic submanifold $\text{Exp}_q(A_q)$ from Prop. 6.1 is the orbit $G' \cdot q$ with differentiable structure derived from G'/K'. Consider the automorphism $\sigma : g \to s_q g s_q$ of $I(M)$, s_q denoting the symmetry of M with respect to q. Then $M_q = \{X \in \mathfrak{l} : d\sigma(X) = -X\}$ so $\sigma(g) = g^{-1}$ for $g \in G'$. This relation also holds for the closure of G' which therefore has an abelian Lie algebra contained in M_q. On the other hand this Lie algebra contains A_q since G' is a Lie subgroup of its closure. By the maximality of A_q, G' is closed and thus carries the relative topology of $I(M)$. The group K' is therefore a closed subgroup of the isotropy subgroup of $I(M)$ at q, hence compact. Using now Prop. 4.4, Chapter II, we deduce that the submanifold $\text{Exp}_q(A_q)$ is a closed topological subspace of M. The identity mapping of A into this submanifold is therefore continuous and, by Lemma 14.1, Chapter I, differentiable. We can therefore state that $\text{Exp}_q(A_q)$ and A coincide as submanifolds of M. This proves (iii) and reduces (i) and (ii) to the following lemma (see §5 for notation).

Lemma 6.3. *Let $\mathfrak{a}_*$ and $\mathfrak{a}'_*$ be two maximal abelian subspaces of $\mathfrak{p}_*$. Then:*

(i) *There exists an element $H \in \mathfrak{a}_*$ whose centralizer in $\mathfrak{p}_*$ is $\mathfrak{a}_*$.*
(ii) *There exists an element $k \in K^*$ such that $k \cdot \mathfrak{a}_* = \mathfrak{a}'_*$.*
(iii) $\mathfrak{p}_* = \bigcup_{k \in K^*} k \cdot \mathfrak{a}_*$.

Proof. Let $P_* = \exp \mathfrak{p}_*$. Then $\text{Int}(\mathfrak{u}) = P_* K^*$ since the geodesics $\text{Exp}\, tX$ $(t \in \boldsymbol{R})$ cover the manifold $\text{Int}(\mathfrak{u})/K^*$ as X varies through $\mathfrak{p}_*$. Let $\tilde{\theta}$ denote the involutive automorphism of $\text{Int}(\mathfrak{u})$ which corresponds to the restriction of θ to $\mathfrak{u}$. Then

$$\tilde{\theta}((\exp X)\, k) = (\exp(-X))\, k$$

so

$$P_* = \{g\tilde{\theta}(g)^{-1} : g \in \text{Int}(\mathfrak{u})\}.$$

Consequently, P_* is compact and closed in $\text{Int}(\mathfrak{u})$.

Let A_* denote the closure of $\exp \mathfrak{a}_*$ in $\mathrm{Int}(\mathfrak{u})$. Then A_* is a torus, contained in P_*. Since $\tilde{\theta}(a) = a^{-1}$ for each $a \in A_*$, the Lie algebra of A_* is contained in $\mathfrak{p}_*$; being abelian this Lie algebra must coincide with $\mathfrak{a}_*$. Select $H \in \mathfrak{a}_*$ such that the one-parameter subgroup $\exp tH$ $(t \in \boldsymbol{R})$ is dense in A_*. Then the centralizer of H in $\mathfrak{p}_*$ is $\mathfrak{a}_*$.

In order to prove (iii) let X be an arbitrary element in $\mathfrak{p}_*$. The function $k \to B(H, k \cdot X)$ is a continuous function on the compact group K_*, and takes a minimum for $k = k_0$, say. If $T \in \mathfrak{k}_0$ we have therefore

$$\left\{\frac{d}{dt} B(H, (\exp tT)\, k_0 \cdot X)\right\}_{t=0} = 0.$$

This can be written

$$B(H, [T, k_0 \cdot X]) = 0.$$

Consequently, $B([k_0 \cdot X, H], T) = 0$ for all $T \in \mathfrak{k}_0$. Since $[k_0 \cdot X, H] \in \mathfrak{k}_0$ it follows that $[k_0 \cdot X, H] = 0$ and by (i), $k_0 \cdot X \in \mathfrak{a}_*$. This proves (iii).

Finally, using (iii) on $\mathfrak{a}'_*$, there exists a $k \in K^*$ such that $H \in k \cdot \mathfrak{a}'_*$. Each element in $k \cdot \mathfrak{a}'_*$ commutes with H; since $\mathfrak{a}_*$ is the centralizer of H in $\mathfrak{p}_*$ it follows that $k \cdot \mathfrak{a}'_* \subset \mathfrak{a}_*$. This finishes the proof of the lemma.

We shall now give two other applications of Lemma 6.3.

Theorem 6.4. *Let G be a connected, compact Lie group. Let $\mathfrak{g}$ denote the Lie algebra of G and let $\mathfrak{t}$ and $\mathfrak{t}'$ denote two maximal abelian subalgebras of $\mathfrak{g}$. Then*

(i) *There exists an element $H \in \mathfrak{t}$ whose centralizer in $\mathfrak{g}$ is $\mathfrak{t}$.*

(ii) *There exists an element $g \in G$ such that $Ad(g)\,\mathfrak{t} = \mathfrak{t}'$.*

(iii) $\mathfrak{g} = \bigcup_{g \in G} Ad(g)\,\mathfrak{t}$.

Proof. The group G can be written $G \times G/G^*$ where G^* is the diagonal in $G \times G$. If G is semisimple, the pair $(G \times G, G^*)$ is a Riemannian symmetric pair of the compact type with the involutive automorphism $\sigma : (x, y) \to (y, x)$ of $G \times G$. In this case Theorem 6.4 is a special instance of Lemma 6.3. If however, G is not semisimple a slight extension of Lemma 6.3 is necessary. In the proof of Lemma 6.3 the semisimplicity of $\mathfrak{k}_0 + \mathfrak{p}_*$ was never used and the form B could be replaced by any strictly negative definite bilinear form Q on $\mathfrak{u} \times \mathfrak{u}$ satisfying the invariance condition

$$Q(X, [Y, Z]) = Q([X, Y], Z)$$

for all $X, Y, Z \in \mathfrak{u}$. In the case of the Riemannian symmetric pair $(G \times G, G^*)$ such a bilinear form Q exists due to the compactness of G. Hence Theorem 6.4 holds for any connected compact Lie group G.

Theorem 6.5. *Let* $\mathfrak{g}$ *be a semisimple Lie algebra over* C, $\mathfrak{h}_1$ *and* $\mathfrak{h}_2$ *two Cartan subalgebras of* $\mathfrak{g}$. *Then there exists an automorphism* σ *of* $\mathfrak{g}$ *such that* $\sigma\mathfrak{h}_1 = \mathfrak{h}_2$.

Proof. Each Cartan subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ determines a compact real form $\mathfrak{u}$ of $\mathfrak{g}$ such that $i\mathfrak{h}^* \subset \mathfrak{u}$ (Theorem 6.3, Chapter III). Let $\mathfrak{u}_1$ and $\mathfrak{u}_2$ denote two compact real forms of $\mathfrak{g}$ arising in this manner from $\mathfrak{h}_1$ and $\mathfrak{h}_2$, respectively. Then there exists an automorphism of $\mathfrak{g}$ carrying $\mathfrak{u}_1$ onto $\mathfrak{u}_2$. Hence we may assume $\mathfrak{u}_1 = \mathfrak{u}_2$ without loss of generality. The subspaces $i\mathfrak{h}_1^*$ and $i\mathfrak{h}_2^*$ are then maximal abelian subalgebras of $\mathfrak{u}_1$. By Theorem 6.4 these subalgebras are conjugate under an element $\sigma \in \mathrm{Int}(\mathfrak{u}_1)$. But σ extends uniquely to an automorphism of $\mathfrak{g}$ and $\mathfrak{h}_1$ and $\mathfrak{h}_2$ are conjugate under this automorphism.

EXERCISES

1. In §2 it was shown that $\mathfrak{so}(1, 3)$ is isomorphic to $\mathfrak{sl}(2, C)^R$. Exhibit this isomorphism explicitly, for example, by means of the stereographic projection of the sphere S^2 onto the complex plane.

2. Let G be a connected Lie group which contains a compact subgroup of dimension $\geqslant 1$ but has center $\{e\}$. Show that G has a subgroup H such that (G, H) is a symmetric pair.

3. State and prove a uniqueness property for the decomposition in Theorem 1.1.

4. Let M be a Riemannian globally symmetric space, $I(M)$ its group of isometries. Prove that (i) M is of the compact type if and only if $I(M)$ is semisimple and compact; (ii) M is of the noncompact type if and only if the Lie algebra of $I(M)$ is semisimple and has no compact ideal $\neq \{0\}$.

NOTES

The results of §1-§4 are, for the most part, due to É. Cartan [16].

The conjugacy of maximal tori (or more precisely Theorem 6.4 (iii)) was first proved by Weyl [1], Kap. IV, Satz 1. The simple proof given here is due to Hunt [2] and this proof applies equally well to Lemma 6.3 first proved by É. Cartan in [10]. The conjugacy statement for Cartan subalgebras (Theorem 6.5) has been extended and sharpened by Chevalley [6], Chapter VI, §4, Théorème 4.

CHAPTER VI

SYMMETRIC SPACES OF THE NONCOMPACT TYPE

Having in the last chapter dealt with analogies and common properties of the two types of symmetric spaces we shall now study the two types separately and start with the noncompact type.

In §1 it is shown that for a given noncompact simple Lie algebra $\mathfrak{g}_0$ over $\boldsymbol{R}$ there exists a unique Riemannian globally symmetric space M of the noncompact type such that $I(M)$ has Lie algebra $\mathfrak{g}_0$. This M is diffeomorphic to a Euclidean space. Section 2 contains a proof of É. Cartan's conjugacy theorem for maximal compact subgroups. The subsequent sections deal with topics connected with the Iwasawa decomposition $G = KAN$ of a semisimple connected Lie group G into an (essentially) maximal compact subgroup K, an abelian group A, and a nilpotent group N. The group N is studied separately in §4; particularly instructive is the proof of Theorem 4.7, due to Harish-Chandra.

§ 1. Decomposition of a Semisimple Lie Group

Let $\mathfrak{g}_0$ be a noncompact semisimple Lie algebra over $\boldsymbol{R}$ and let $\mathfrak{g}_0 = \mathfrak{k}_0 + \mathfrak{p}_0$ be a Cartan decomposition of $\mathfrak{g}_0$. The mapping $\theta : T + X \to T - X$ ($T \in \mathfrak{k}_0$, $X \in \mathfrak{p}_0$) is an involutive automorphism of $\mathfrak{g}_0$ and the pair $(\mathfrak{g}_0, \theta)$ is an orthogonal symmetric Lie algebra of the noncompact type. We recall that a pair (G, K) is said to be associated with $(\mathfrak{g}_0, \theta)$ if G is a connected Lie group with Lie algebra $\mathfrak{g}_0$ and K is a Lie subgroup of G with Lie algebra $\mathfrak{k}_0$. Such a pair is said to be of the noncompact type. This pair is said to be a Riemannian symmetric pair if K is closed, $\mathrm{Ad}_G(K)$ compact and there exists an analytic involutive automorphism $\tilde{\theta}$ of G such that $(K_{\tilde{\theta}})_0 \subset K \subset K_{\tilde{\theta}}$. Such a $\tilde{\theta}$ is necessarily unique and $d\tilde{\theta} = \theta$ (Prop. 3.5, Chapter IV). Finally, a Riemannian globally symmetric space M is said to be of the noncompact type if the pair $(I_0(M), H)$ is of the noncompact type, H being the isotropy subgroup of $I_0(M)$ at some point $o \in M$.

Theorem 1.1. *With the notation above, suppose* (G, K) *is any pair associated with* $(\mathfrak{g}_0, \theta)$. *Then*:

(i) *K is connected, closed, and contains the center Z of G. Moreover, K is compact if and only if Z is finite.*

(ii) *There exists an involutive, analytic automorphism* $\tilde{\theta}$ *of* G *whose fixed point set is* K *and whose differential at* e *is* θ; *the pair* (G, K) *is a Riemannian symmetric pair.*

(iii) *The mapping* $\varphi : (X, k) \to (\exp X)k$ *is a diffeomorphism of* $\mathfrak{p}_0 \times K$ *onto the group* G *and the mapping Exp is a diffeomorphism of* $\mathfrak{p}_0$ *onto the globally symmetric space* G/K.

Let Ad and ad denote the adjoint representations of G and $\mathfrak{g}_0$, respectively. Before starting on Theorem 1.1 we prove a simple lemma.

Lemma 1.2. *Let* (X_i) *be a basis of* $\mathfrak{p}_0$, *orthonormal with respect to* B *and let* (T_α) *be a basis of* $\mathfrak{k}_0$, *orthonormal with respect to* $-B$. *With respect to the basis* (X_i), (T_α) *of* $\mathfrak{g}_0$, *ad* X $(X \in \mathfrak{p}_0)$ *is expressed by a symmetric matrix, and ad* T $(T \in \mathfrak{k}_0)$ *is expressed by a skew symmetric matrix.*

Proof. We can write

$$\operatorname{ad} X(X_j) = \sum_\alpha a_{\alpha j} T_\alpha, \qquad \operatorname{ad} X(T_\alpha) = \sum_j a_{j\alpha} X_j$$

for suitable real numbers $a_{\alpha j}$, $a_{k\alpha}$. Here

$$a_{\alpha j} = -B([X, X_j], T_\alpha) = B(X_j, [X, T_\alpha]) = a_{j\alpha},$$

so ad X is expressed by a symmetric matrix. Similarily, we have

$$\operatorname{ad} T(X_j) = \sum_k b_{kj} X_k, \qquad \operatorname{ad} T(T_\alpha) = \sum_\beta c_{\beta\alpha} T_\beta,$$

where b_{kj} and $c_{\beta\alpha}$ are real numbers. Here

$$b_{kj} = B([T, X_j], X_k) = -B(X_j, [T, X_k]) = -b_{jk},$$
$$c_{\beta\alpha} = -B([T, T_\alpha], T_\beta) = B(T_\alpha, [T, T_\beta]) = -c_{\beta\alpha},$$

so ad T is expressed by a skew symmetric matrix.

Passing now to the proof of Theorem 1.1, let K_0 denote the identity component of K. Then owing to Prop. 3.6, Chapter IV, K_0 is closed in G, the coset space G/K_0 is Riemannian locally symmetric and $\pi(\exp X) = \operatorname{Exp} X$ for $X \in \mathfrak{p}_0$ if π denotes the natural mapping of G onto G/K_0. Thus G/K_0 is complete and consequently Exp maps $\mathfrak{p}_0$ onto G/K_0. It follows that φ maps $\mathfrak{p}_0 \times K_0$ onto G. To see that φ is one-to-one on $\mathfrak{p}_0 \times K$, suppose that $X_1, X_2 \in \mathfrak{p}_0$, $k_1, k_2 \in K$ such that

$$(\exp X_1)\, k_1 = (\exp X_2)\, k_2. \tag{1}$$

Applying Ad to this relation we obtain

$$e^{\operatorname{ad} X_1} \circ \operatorname{Ad}(k_1) = e^{\operatorname{ad} X_2} \circ \operatorname{Ad}(k_2). \tag{2}$$

The decomposition $m = po$ of a nonsingular matrix m, where p is symmetric and positive definite and o is orthogonal, is unique. Thus, it follows from (2) that

$$e^{\operatorname{ad} X_1} = e^{\operatorname{ad} X_2},$$

$$\operatorname{Ad}(k_1) = \operatorname{Ad}(k_2).$$

The exponential mapping is one-to-one on the set of symmetric matrices; hence, $\operatorname{ad} X_1 = \operatorname{ad} X_2$ and $X_1 = X_2$ since $\mathfrak{g}_0$ has center $\{0\}$. It follows from (1) that $k_1 = k_2$ so φ is one-to-one on $\mathfrak{p}_0 \times K$. Since we have proved that $\varphi(\mathfrak{p}_0 \times K_0) = \varphi(\mathfrak{p}_0 \times K)$, we conclude that $K_0 = K$. Now let as usual K^* denote the analytic subgroup of $\operatorname{Int}(\mathfrak{g}_0)$ with Lie algebra $\mathfrak{k}_0$. Then the pair $(G, \operatorname{Ad}^{-1}(K^*))$ is associated with $(\mathfrak{g}_0, \theta)$; from what is already proved, the group $\operatorname{Ad}^{-1}(K^*)$ is connected. Having Lie algebra equal to $\mathfrak{k}_0$, it must coincide with K. Hence $Z \subset K$. Since $K^* = K/Z$ is compact, and Z is discrete (§2, Chapter II), it follows that K is compact if and only if Z is finite.

The automorphism θ of $\mathfrak{g}_0$ induces an automorphism Θ of the universal covering group $\tilde{G}$ of G such that $d\Theta = \theta$. The fixed points of Θ form a subgroup $\tilde{K}$, which, by the above, must contain the center $\tilde{Z}$ of $\tilde{G}$. The kernel of the covering mapping of $\tilde{G}$ onto G is a discrete normal subgroup N of $\tilde{G}$ and must therefore belong to $\tilde{Z}$. The automorphism $\tilde{\theta}$ of $G = \tilde{G}/N$ induced by Θ then turns (G, K) into a Riemannian symmetric pair. Using (i) we see that the fixed point set of $\tilde{\theta}$ coincides with K.

Since the mapping φ is one-to-one, the mapping Exp is a one-to-one differentiable mapping of $\mathfrak{p}_0$ onto G/K. In order to prove that it is regular (and thus a diffeomorphism) it suffices, due to the formula for $d\operatorname{Exp}_X$ (Theorem 4.1, Chapter II), to prove that

$$\det\left(\sum_0^\infty \frac{(T_X)^n}{(2n+1)!}\right) \neq 0 \qquad \text{for } X \in \mathfrak{p}_0. \tag{3}$$

The relations

$$B(T_X Y, Z) = B(Y, T_X Z),$$

$$B(T_X Y, Y) = -B([X, Y], [X, Y]) \geqslant 0,$$

valid for $X, Y, Z \in \mathfrak{p}_0$, show that T_X is symmetric and positive definite with respect to B. The validity of (3) is therefore obvious.

It remains only to prove that φ is everywhere regular. A general tangent vector to $\mathfrak{p}_0 \times K$ at the point (X, k) has the form $(Y, dL_k \cdot T)$ where $Y \in \mathfrak{p}_0$ and $T \in \mathfrak{k}_0$, L_x denoting left translation by the group element x.

Since

$$\varphi(X + tY, k) = \exp(X + tY)\, k = k \exp(\operatorname{Ad}(k^{-1})(X + tY))$$

and

$$\varphi(X, k \exp tT) = (\exp X)\, k(\exp tT)$$

it follows from Theorem 1.7, Chapter II, that

$$d\varphi_{(X,k)}(Y, dL_k \cdot T) = dL_{(\exp X)k} \cdot \left(\frac{1 - e^{-\operatorname{ad} X'}}{\operatorname{ad} X'}(Y') + T\right), \tag{4}$$

where $X' = \operatorname{Ad}(k^{-1})\, X$, $Y' = \operatorname{Ad}(k^{-1})\, Y$. The $\mathfrak{p}_0$-component of the vector $(1 - e^{-\operatorname{ad} X'})(\operatorname{ad} X')^{-1}(Y')$ is

$$\sum_0^\infty \frac{(T_{X'})^n}{(2n+1)!}(Y'),$$

which, due to (3), vanishes only if $Y' = 0$. Consequently the right-hand side of (4) is $\neq 0$ unless $T = Y = 0$. This shows that φ is regular and completes the proof of Theorem 1.1.

Corollary 1.3. *Let M and M' be two Riemannian globally symmetric spaces of the noncompact type such that the groups $I(M)$ and $I(M')$ have the same Lie algebra $\mathfrak{g}_0$. As usual,* (Chapter V, §5), *suppose the Riemannian structures on M and M' arise from the Killing form of $\mathfrak{g}_0$. Then M and M' are isometric.*

In fact, the spaces M and M' arise from two Cartan decompositions of $\mathfrak{g}_0$. If these Cartan decompositions coincide, the corollary follows from Theorem 1.1. Now, any two Cartan decompositions of $\mathfrak{g}_0$ are conjugate under an inner automorphism σ of $\mathfrak{g}_0$. It is easy to set up an isometry between M and M' by means of σ.

If the Lie algebra $\mathfrak{g}_0$ is simple (or more generally if all its simple factors are noncompact), then by Theorem 4.1, Chapter V, the space $M = \operatorname{Int}(\mathfrak{g}_0)/K^*$ is the unique Riemannian globally symmetric space for which $I(M)$ has Lie algebra $\mathfrak{g}_0$. For compact semisimple Lie algebras the situation is quite different as we shall see in the next chapter.

Let $\mathfrak{s}_0$ be a Lie triple system contained in $\mathfrak{p}_0$. Then, according to Theorem 7.2, Chapter IV, and Theorem 1.1, $\operatorname{Exp} \mathfrak{s}_0$ is a closed totally geodesic submanifold of G/K. We can therefore apply Theorem 14.6, Chapter I, and get the following extension of Theorem 1.1.

Theorem 1.4. *In the notation of Theorem* 1.1, *let* $\mathfrak{s}_0$ *be a Lie triple system contained in* $\mathfrak{p}_0$ *and let* $\mathfrak{t}_0$ *denote the orthogonal complement of* $\mathfrak{s}_0$ *in* $\mathfrak{p}_0$. *Let* $S_0 = \exp \mathfrak{s}_0$ *and* $T_0 = \exp \mathfrak{t}_0$ *be given the relative topology of* G. *Then* G *decomposes topologically*

$$G = S_0 \cdot T_0 \cdot K.$$

Proof. In view of Theorem 14.6, Chapter I, the mapping $(X, Y) \to \tau(\exp X) \cdot \operatorname{Exp} Y$, $(X \in \mathfrak{s}_0,\ Y \in \mathfrak{t}_0)$, is a continuous one-to-one mapping of $\mathfrak{s}_0 \times \mathfrak{t}_0$ onto G/K. Each bounded set in M is the image of a bounded set in $\mathfrak{s}_0 \times \mathfrak{t}_0$ (Cor. 13.2 (i), Chapter I). Hence the mapping is a homeomorphism. If we state this fact in terms of G and make use of Theorem 1.1, the present theorem follows.

§ 2. Maximal Compact Subgroups and Their Conjugacy

The symbols $\mathfrak{g}_0$, $\mathfrak{k}_0$, $\mathfrak{p}_0$, G, and K have the same meaning as in §1.

It was proved in Chapter III that all Cartan decompositions of a semisimple Lie algebra over $\boldsymbol{R}$ are conjugate under an inner automorphism. Using differential geometric results, we shall now prove a stronger theorem, namely, that all maximal compactly imbedded subalgebras of a semisimple Lie algebra over $\boldsymbol{R}$ are conjugate under an inner automorphism. Consequently, each maximal compactly imbedded subalgebra $\mathfrak{u}$ of a semisimple Lie algebra $\mathfrak{l}$ is a part of a Cartan decomposition and its orthogonal complement $\mathfrak{e}$ with respect to the Killing form of $\mathfrak{l}$ satisfies $[\mathfrak{e}, \mathfrak{e}] \subset \mathfrak{u}$.

Theorem 2.1. *Let* (G, K) *be a Riemannian symmetric pair of the noncompact type. Let* K_1 *be any compact subgroup of* G. *Then there exists an element* $x \in G$ *such that* $x^{-1} K_1 x \subset K$.

Proof. The relation $x^{-1}K_1x \subset K$ means that xK is a fixed point under the action of the group K_1 on the coset space G/K. Since the space G/K is a simply connected Riemannian manifold of negative curvature, the existence of the fixed point is assured by Theorem 13.5, Chapter I.

Theorem 2.2.

(i) *Let* (G, K) *be a Riemannian symmetric pair of the noncompact type. Then* K *has a unique maximal compact subgroup* K' *and this group is maximal compact in* G.

(ii) *All maximal compact subgroups of a connected semisimple Lie group* G *are connected and conjugate under an inner automorphism of* G.

(iii) *Let K' be any maximal compact subgroup of a connected semisimple Lie group G. Then there exists a submanifold E of G, diffeomorphic to a Euclidean space such that the mapping $(e, k) \to ek$ is a diffeomorphism of $E \times K'$ onto G.*

Proof. The group $\mathrm{Ad}_G(K)$ is compact and has Lie algebra $\mathfrak{k}_0$. Thus $\mathfrak{k}_0$ is a compact Lie algebra. According to Prop. 6.6, Chapter II, $\mathfrak{k}_0$ can be written as a direct sum $\mathfrak{k}_0 = \mathfrak{k}_s + \mathfrak{k}_a$ where the ideals $\mathfrak{k}_s$ and $\mathfrak{k}_a$ are semisimple and abelian, respectively. Let K_s and K_a denote the corresponding analytic subgroups of K. The group K_a is a direct product $K_a = T \times V$ of analytic subgroups T, V of G, where T is a torus and V is analytically isomorphic to Euclidean space. Now we put

$$K' = K_sT = \{kt : k \in K_s, t \in T\}.$$

As a result of Theorem 6.9, Chapter II, the group K_s, and therefore the group K', is compact. The groups K' and V commute elementwise; the group $K' \cap V$ is a compact subgroup of the Euclidean group V, hence $K' \cap V = \{e\}$. It follows that $K = K' \times V$ (direct product) and K' is the unique maximal compact subgroup of K. Combining this with Theorem 2.1, we see that each compact subgroup of G is conjugate to a subgroup of K'. This proves (i) and (ii).

Consider now the set

$$E = \{(\exp X)\, v : X \in \mathfrak{p}_0, v \in V\}.$$

Since $\mathfrak{p}_0 \times V$ is a submanifold of $\mathfrak{p}_0 \times K$, it follows from Theorem 1.1, that E is a submanifold of G, diffeomorphic to Euclidean space. Finally, the mappings

$$((\exp X)\, v, k) \to (\exp X, vk) \to \exp X\, kv,$$

$$E \times K' \to (\exp \mathfrak{p}_0) \times K \to G$$

yield the desired diffeomorphism of $E \times K'$ onto G.

§ 3. The Iwasawa Decomposition

This decomposition results from combining the Cartan decomposition of a semisimple Lie algebra and the root space decomposition of its complexification.

Lemma 3.1. *Let $\mathfrak{g}$ be a complex semisimple Lie algebra, $\mathfrak{g}_k$ any compact real form of $\mathfrak{g}$, and let η denote the conjugation of $\mathfrak{g}$ with respect to $\mathfrak{g}_k$. Let*

$\mathfrak{h}$ be a Cartan subalgebra of $\mathfrak{g}$ which is invariant under η. Let Δ denote the set of nonzero roots of $\mathfrak{g}$ with respect to $\mathfrak{h}$ and let $\mathfrak{h}^* = \sum_{\alpha \in \Delta} \boldsymbol{R} H_\alpha$. Then:

(i) $\mathfrak{h}^* \subset i\mathfrak{g}_k$.

(ii) *There exists a vector $E_\alpha \in \mathfrak{g}^\alpha$ such that for all $\alpha \in \Delta$*

$$(E_\alpha - E_{-\alpha}), \qquad i(E_\alpha + E_{-\alpha}) \in \mathfrak{g}_k,$$
$$[E_\alpha, E_{-\alpha}] = (2/\alpha(H_\alpha))\, H_\alpha.$$

Proof. The Killing form B of $\mathfrak{g}$ is strictly negative definite on $\mathfrak{g}_k \times \mathfrak{g}_k$, and moreover

$$B(X, \eta X) < 0 \qquad \text{for } X \neq 0 \text{ in } \mathfrak{g}.$$

For each $\alpha \in \Delta$ we can define the complex linear function α^η on $\mathfrak{h}$ by

$$\alpha^\eta(H) = \overline{(\alpha(\eta \cdot H))}, \qquad H \in \mathfrak{h},$$

the bar denoting complex conjugation. Then if $Y \in \mathfrak{g}^\alpha$, $H \in \mathfrak{h}$, we find

$$[\eta H, \eta Y] = \eta[H, Y] = \overline{\alpha(h)}\, \eta Y = \alpha^\eta(\eta H)\, \eta Y,$$

which shows that $\alpha^\eta \in \Delta$. Also, if $H \in \mathfrak{h}$ we have

$$B(\eta H_\alpha, H) = \overline{B(H_\alpha, \eta H)} = \overline{\alpha(\eta H)} = \alpha^\eta(H) = B(H_{\alpha^\eta}, H)$$

so

$$\eta H_\alpha = H_{\alpha^\eta}.$$

This shows that $\mathfrak{h}^*$ is invariant under η. Since $H = \frac{1}{2}(H + \eta H) + \frac{1}{2}(H - \eta H)$ we have the direct decomposition $\mathfrak{h}^* = \mathfrak{h}^+ + \mathfrak{h}^-$ where $\eta(H) = \pm H$ for $H \in \mathfrak{h}^\pm$. Here $\mathfrak{h}^+ = \{0\}$; in fact if $H \neq 0$ in $\mathfrak{h}^+$ then $B(H, H) = B(H, \eta H) < 0$ which contradicts the fact that B is strictly positive definite on $\mathfrak{h}^* \times \mathfrak{h}^*$ (Theorem 4.4, Chapter III). This proves (i) and $\alpha^\eta = -\alpha$.

To prove (ii) we turn the dual space of $\mathfrak{h}^*$ into an ordered vector space. Since each $\alpha \in \Delta$ is real valued on $\mathfrak{h}^*$, Δ has now become an ordered set. Let Δ^+ denote the set of positive roots. Since $\alpha(H_\alpha) > 0$ we can to each $\alpha \in \Delta^+$ select $E_\alpha \in \mathfrak{g}^\alpha$ such that

$$B(E_\alpha, \eta E_\alpha) = -\frac{2}{\alpha(H_\alpha)}.$$

Since $\alpha^\eta = -\alpha$ it is easy to see that $\eta E_\alpha \in \mathfrak{g}^{-\alpha}$ for $\alpha \in \Delta^+$. We put $E_{-\alpha} = -\eta E_\alpha$ for $\alpha \in \Delta^+$. Then $\eta E_\alpha = -E_{-\alpha}$ for all $\alpha \in \Delta$ and consequently the elements $(E_\alpha - E_{-\alpha})$ and $i(E_\alpha + E_{-\alpha})$ belong to $\mathfrak{g}_k$ for each $\alpha \in \Delta$. Finally

$$B(H, [E_\alpha, E_{-\alpha}]) = B([H, E_\alpha], E_{-\alpha}) = \frac{2\alpha(H)}{\alpha(H_\alpha)}$$

and since $[E_\alpha, E_{-\alpha}]$ is a scalar multiple of H_α (Theorem 4.2, Chapter III), it follows that $[E_\alpha, E_{-\alpha}] = (2/\alpha(H_\alpha))\, H_\alpha$.

Suppose now $\mathfrak{g}_0$ is a semisimple Lie algebra over $\boldsymbol{R}$ and that $\mathfrak{g}_0 = \mathfrak{k}_0 + \mathfrak{p}_0$ is a Cartan decomposition of $\mathfrak{g}_0$ as usual. Let $\mathfrak{g}$ be the complexification of $\mathfrak{g}_0$, put $\mathfrak{u} = \mathfrak{k}_0 + i\mathfrak{p}_0$ and let σ and τ denote the conjugations of $\mathfrak{g}$ with respect to $\mathfrak{g}_0$ and $\mathfrak{u}$, respectively. The automorphism $\sigma\tau$ of $\mathfrak{g}$ will be denoted by θ. Let ad denote the adjoint representation of $\mathfrak{g}$.

Let $\mathfrak{h}_{\mathfrak{p}_0}$ denote any maximal abelian subspace of $\mathfrak{p}_0$. Let $\mathfrak{h}_0$ be any maximal abelian subalgebra of $\mathfrak{g}_0$ containing $\mathfrak{h}_{\mathfrak{p}_0}$. The existence of $\mathfrak{h}_0$ is obvious from Zorn's lemma. If $X \in \mathfrak{h}_0$ and $Y \in \mathfrak{h}_{\mathfrak{p}_0}$ we have

$$[X - \theta X, Y] = [X, Y] - \theta[X, \theta Y] = [X, Y] + \theta[X, Y] = 0 + 0.$$

Since $X - \theta X \in \mathfrak{p}_0$ it follows, in view of the maximality of $\mathfrak{h}_{\mathfrak{p}_0}$, that $X - \theta X \in \mathfrak{h}_{\mathfrak{p}_0}$. Thus $\theta\mathfrak{h}_0 \subset \mathfrak{h}_0$ so we have the direct decomposition $\mathfrak{h}_0 = \mathfrak{h}_0 \cap \mathfrak{k}_0 + \mathfrak{h}_0 \cap \mathfrak{p}_0$. Obviously $\mathfrak{h}_{\mathfrak{p}_0} = \mathfrak{h}_0 \cap \mathfrak{p}_0$. We put $\mathfrak{h}_{\mathfrak{k}_0} = \mathfrak{h}_0 \cap \mathfrak{k}_0$. Let $\mathfrak{h}$, $\mathfrak{h}_\mathfrak{p}$, $\mathfrak{h}_\mathfrak{k}$, $\mathfrak{k}$, and $\mathfrak{p}$ denote the subspaces of $\mathfrak{g}$ generated by $\mathfrak{h}_0$, $\mathfrak{h}_{\mathfrak{p}_0}$, $\mathfrak{h}_{\mathfrak{k}_0}$, $\mathfrak{k}_0$ and $\mathfrak{p}_0$, respectively.

Lemma 3.2. *The algebra $\mathfrak{h}$ is a Cartan subalgebra of $\mathfrak{g}$ and $\mathfrak{h}^* = \mathfrak{h}_{\mathfrak{p}_0} + i\mathfrak{h}_{\mathfrak{k}_0}$.*

It is obvious that $\mathfrak{h}$ is a maximal abelian subalgebra of $\mathfrak{g}$. Now, the Hermitian form $B_\tau(X, Y) = -B(X, \tau Y)$ on $\mathfrak{g} \times \mathfrak{g}$ is strictly positive definite and if $Z \in \mathfrak{u}$ we have

$$B_\tau([Z, X], Y) + B_\tau(X, [Z, Y]) = 0.$$

If ad Z leaves a subspace V of $\mathfrak{g}$ invariant then the orthogonal complement $V^\perp$ (with respect to B_τ) is also invariant and $\mathfrak{g} = V + V^\perp$ (direct sum). Hence ad Z is semisimple. Thus ad H is semisimple if $H \in \mathfrak{h}_\mathfrak{k} \cup \mathfrak{h}_\mathfrak{p}$. Since ad H_1 and ad H_2 commute if $H_1 \in \mathfrak{h}_\mathfrak{k}$, $H_2 \in \mathfrak{h}_\mathfrak{p}$, it follows that ad $(H_1 + H_2)$ is semisimple and $\mathfrak{h}$ is a Cartan subalgebra.

As a result of its definition, $\mathfrak{h}$ is invariant under σ and θ. Thus it is also invariant under τ. By Lemma 3.1 we have $\mathfrak{h}^* \subset \mathfrak{h} \cap (i\mathfrak{u})$. But $\theta\mathfrak{h} \subset \mathfrak{h}$ implies $\mathfrak{h} \cap (i\mathfrak{u}) = \mathfrak{h} \cap i\mathfrak{k}_0 + \mathfrak{h} \cap \mathfrak{p}_0 = i\mathfrak{h}_{\mathfrak{k}_0} + \mathfrak{h}_{\mathfrak{p}_0}$. Since $\dim \mathfrak{h}^* = \dim \mathfrak{h}_0$, the lemma follows.

Let V be a finite-dimensional vector space over $\boldsymbol{R}$, W a subspace of V. Let $V^\wedge$ and $W^\wedge$ denote their duals and suppose that $V^\wedge$ and $W^\wedge$

have been turned into ordered vector spaces. The orderings are said to be *compatible* (Harish-Chandra [10], p. 195) if $\lambda \in V^\wedge$ is positive whenever its restriction $\bar{\lambda}$ to W is positive. Compatible orderings can for example be constructed as follows: Let $X_1, \ldots, X_n$ be a basis of V such that $X_1, \ldots, X_m$ is a basis of W. Then the lexicographic orderings of $W^\wedge$ and $V^\wedge$ with respect to these bases are compatible.

Now select compatible orderings in the dual spaces of $\mathfrak{h}_{\mathfrak{p}_0}$ and $\mathfrak{h}^*$, respectively. Since each root $\alpha \in \Delta$ is real valued on $\mathfrak{h}^*$ we get in this way an ordering of Δ. Let Δ^+ denote the set of positive roots. Now for each $\alpha \in \Delta$ the linear functions α^τ, α^σ, and α^θ defined by

$$\alpha^\tau(H) = \overline{\alpha(\tau H)}, \quad \alpha^\sigma(H) = \overline{\alpha(\sigma H)}, \quad \alpha^\theta(H) = \alpha(\theta H) \qquad (H \in \mathfrak{h}),$$

are again members of Δ. The root α vanishes identically on $\mathfrak{h}_{\mathfrak{p}_0}$ if and only if $\alpha = \alpha^\theta$. We divide the positive roots in two classes as follows:

$$P_+ = \{\alpha : \alpha \in \Delta^+, \alpha \neq \alpha^\theta\},$$
$$P_- = \{\alpha : \alpha \in \Delta^+, \alpha = \alpha^\theta\}.$$

Lemma 3.3.

(i) *If* $\alpha \in P_+$, *then* $-\alpha^\theta \in P_+$, $\alpha^\sigma \in P_+$, $\alpha^\tau = -\alpha$.

(ii) *If* $\beta \in P_-$, *then* $\beta^\theta = \beta$, $\beta^\sigma = -\beta$, $\beta^\tau = -\beta$, *and* $\mathfrak{g}^\beta + \mathfrak{g}^{-\beta} \in \mathfrak{k}$.

Proof. The restrictions to $\mathfrak{h}_{\mathfrak{p}_0}$ of α and α^θ have sum 0. By the compatibility of the orderings, we have $\alpha^\theta < 0$. Since α^σ and α agree on $\mathfrak{h}_{\mathfrak{p}_0}$ we have $\alpha^\sigma \in P_+$. The relations $\alpha^\tau = -\alpha$, $\beta^\tau = -\beta$ were established during the proof of Lemma 3.1. The relation $\beta^\theta = \beta$ implies $H_\beta \in \mathfrak{k} \cap \mathfrak{h}^* = i\mathfrak{h}_{\mathfrak{k}_0}$ so $\beta^\sigma = -\beta$. Since $\theta\mathfrak{g}^\beta = \mathfrak{g}^\beta$, $\theta^2 = 1$ and $\dim(\mathfrak{g}^\beta) = 1$, it is clear that $\theta Z = -Z$ or $\theta Z = Z$ for each $Z \in \mathfrak{g}^\beta$. If $\theta Z = -Z$, then $Z \in \mathfrak{p}$. For $H \in \mathfrak{h}_\mathfrak{p}$ we have $[H, Z] = \beta(H) Z = 0$ and since $\mathfrak{h}_\mathfrak{p}$ is a maximal abelian subspace of $\mathfrak{p}$ it follows that $Z = 0$ and $\mathfrak{g}^\beta \in \mathfrak{k}$. Similarly $\mathfrak{g}^{-\beta} \in \mathfrak{k}$.

Theorem 3.4. *Let* $\mathfrak{n} = \sum_{\alpha \in P_+} \mathfrak{g}^\alpha$, $\mathfrak{n}_0 = \mathfrak{g}_0 \cap \mathfrak{n}$, $\mathfrak{s}_0 = \mathfrak{h}_{\mathfrak{p}_0} + \mathfrak{n}_0$. *Then* $\mathfrak{n}$ *and* $\mathfrak{n}_0$ *are nilpotent Lie algebras,* $\mathfrak{s}_0$ *is a solvable Lie algebra, and*

$$\mathfrak{g}_0 = \mathfrak{k}_0 + \mathfrak{h}_{\mathfrak{p}_0} + \mathfrak{n}_0 \quad (\textit{direct vector space sum}).$$

Proof. Let $\alpha, \beta \in P_+$. If $\alpha + \beta \in \Delta$, then $\alpha + \beta \in P_+$ and $\mathfrak{n}$ is a subalgebra of $\mathfrak{g}$ which obviously is nilpotent. Hence $\mathfrak{n}_0$ is a nilpotent subalgebra of $\mathfrak{g}_0$. From the relation $[\mathfrak{n}_0 + \mathfrak{h}_{\mathfrak{p}_0}, \mathfrak{n}_0 + \mathfrak{h}_{\mathfrak{p}_0}] \subset \mathfrak{n}_0$ we see that $\mathfrak{h}_{\mathfrak{p}_0} + \mathfrak{n}_0$ is a solvable Lie algebra. To see that the sum $\mathfrak{k}_0 + \mathfrak{h}_{\mathfrak{p}_0} + \mathfrak{n}_0$ is direct, suppose $T \in \mathfrak{k}_0$, $H \in \mathfrak{h}_{\mathfrak{p}_0}$ and $X \in \mathfrak{n}_0$ such that $T + H + X = 0$.

Applying θ we find that $T - H + \theta X = 0$ so $2H + X - \theta X = 0$. Now by Lemma 3.3

$$\theta X \in \sum_{\alpha \in P_+} \mathfrak{g}^{-\alpha}$$

and since the sum $\mathfrak{h} + \sum_{\alpha\in\Delta} \mathfrak{g}^{\alpha} = \mathfrak{g}$ is direct we conclude that $H = 0$ and $X - \theta X = 0$. But $\mathfrak{k}_0 \cap \mathfrak{n}_0 = \{0\}$ so $X = T = 0$.

Now let $X \in \mathfrak{g}_0$. Since $X = \frac{1}{2}(X + \sigma X)$ it follows that X can be written

$$X = H + \sum_{\alpha\in\Delta} (X_\alpha + \sigma X_\alpha),$$

where $H \in \mathfrak{h}_0$, $X_\alpha \in \mathfrak{g}^\alpha$ for each $\alpha \in \Delta$. If α or $-\alpha$ belongs to P_-, then $X_\alpha + \sigma X_\alpha \in \mathfrak{k}_0$ due to Lemma 3.3'. If $\alpha \in P_+$ then $X_\alpha + \sigma X_\alpha \in \mathfrak{n}_0$ by the same lemma. Finally, if $-\alpha \in P_+$ then $\tau(X_\alpha + \sigma X_\alpha) \in \mathfrak{g}^{-\alpha} + \mathfrak{g}^{\alpha^\theta} \subset \mathfrak{n}$ by Lemma 3.3. Consequently

$$X_\alpha + \sigma X_\alpha = \{(X_\alpha + \sigma X_\alpha) + \tau(X_\alpha + \sigma X_\alpha)\} - \tau(X_\alpha + \sigma X_\alpha) \in \mathfrak{u} \cap \mathfrak{g}_0 + \mathfrak{n} \cap \mathfrak{g}_0$$

so $X_\alpha + \sigma X_\alpha \in \mathfrak{k}_0 + \mathfrak{n}_0$. This proves the theorem.

Lemma 3.5. *There exists a basis (X_i) of $\mathfrak{g}$ such that the matrices representing $ad(\mathfrak{g})$ have the following properties:*

(i) *The matrices $ad\,\mathfrak{u}$ are skew Hermitian.*

(ii) The *matrices $ad\,\mathfrak{n}$ are lower triangular with zeros in the diagonal.*

(iii) *The matrices $ad\,\mathfrak{h}_{\mathfrak{p}_0}$ are diagonal matrices with a real diagonal.*

Proof. Let $\alpha_1 < \alpha_2 < \ldots$ be the roots in Δ^+ in increasing order. Let $H_1, \ldots, H_l$ be any basis of $\mathfrak{h}_{\mathfrak{p}_0}$, orthonormal with respect to B_τ. Select $E_{\alpha_i} \in \mathfrak{g}^{\alpha_i}$ such that $B_\tau(E_{\alpha_i}, E_{\alpha_i}) = 1$ $(i = 1, 2, \ldots)$. Since $\tau E_{\alpha_i} \in \mathfrak{g}^{-\alpha_i}$, the vectors $\ldots \tau E_{\alpha_2}, \tau E_{\alpha_1}, H_1, \ldots, H_l, E_{\alpha_1}, E_{\alpha_2}, \ldots$ form an orthonormal basis of $\mathfrak{g}$. This basis has the properties (i), (ii), and (iii).

Let $\mathfrak{m}_0$ denote the centralizer of $\mathfrak{h}_{\mathfrak{p}_0}$ in $\mathfrak{k}_0$. Let $\mathfrak{l}_0$ and $\mathfrak{q}_0$, respectively, denote the orthogonal complements of $\mathfrak{m}_0$ in $\mathfrak{k}_0$ and of $\mathfrak{h}_{\mathfrak{p}_0}$ in $\mathfrak{p}_0$. Here "orthogonal" refers to the positive definite form B_τ. Let $\mathfrak{m}$, $\mathfrak{l}$, and $\mathfrak{q}$ denote the subspaces of $\mathfrak{g}$ generated by $\mathfrak{m}_0$, $\mathfrak{l}_0$, and $\mathfrak{q}_0$, respectively.

Lemma 3.6. *The direct decompositions*

$$\mathfrak{g}_0 = \mathfrak{l}_0 + \mathfrak{m}_0 + \mathfrak{h}_{\mathfrak{p}_0} + \mathfrak{q}_0,$$

$$\mathfrak{g} = \mathfrak{l} + \mathfrak{m} + \mathfrak{h}_{\mathfrak{p}} + \mathfrak{q}$$

are orthogonal with respect to B_τ *and invariant under* θ. *Moreover, if* $X_\alpha \neq 0$ *is arbitrary in* $\mathfrak{g}^\alpha$ $(\alpha \in \Delta)$,

$$\mathfrak{m} = \mathfrak{h}_\mathfrak{k} + \sum_{\alpha \in P_-} (\mathfrak{g}^\alpha + \mathfrak{g}^{-\alpha}),$$

$$\mathfrak{l} = \sum_{\alpha \in P_+} C(X_\alpha + \theta X_\alpha),$$

$$\mathfrak{q} = \sum_{\alpha \in P_+} C(X_\alpha - \theta X_\alpha).$$

Proof. Each $X \in \mathfrak{g}$ can be written

$$X = H^* + \sum_{\alpha \in \Delta} c_\alpha X_\alpha \qquad (H^* \in \mathfrak{h},\ c_\alpha \in \boldsymbol{C}).$$

Hence

$$[H, X] = \sum_{\alpha \in \Delta} c_\alpha \alpha(H) X_\alpha \tag{1}$$

for each $H \in \mathfrak{h}$. Thus X commutes with $H \in \mathfrak{h}$ if and only if $c_\alpha \alpha(H) = 0$ for all $\alpha \in \Delta$. It follows that $\mathfrak{h} + \sum_{\alpha \in P_-} (\mathfrak{g}^\alpha + \mathfrak{g}^{-\alpha})$ is the centralizer of $\mathfrak{h}_\mathfrak{p}$ in $\mathfrak{g}$. Since $\mathfrak{m}$ is the centralizer of $\mathfrak{h}_\mathfrak{p}$ in $\mathfrak{k}$, the expression for $\mathfrak{m}$ follows.

To prove the formula for $\mathfrak{l}$ let $H \in \mathfrak{h}_\mathfrak{k}$, $\alpha \in P_+$, and $\beta \in P_-$ or $-\beta \in P_-$. Then, using Theorem 4.2, Chapter III,

$$-B_\tau(X_\alpha + \theta X_\alpha, H) = B(X_\alpha, \tau H) + B(\theta X_\alpha, \tau H) = 0,$$

$$-B_\tau(X_\alpha + \theta X_\alpha, X_\beta) = B(X_\alpha, \tau X_\beta) + B(\theta X_\alpha, \tau X_\beta) = 0,$$

since $\tau X_\beta \in \mathfrak{g}^{-\beta}$, $\theta X_\alpha \in \mathfrak{g}^{\alpha^\theta}$, and $\alpha^\theta + \beta \neq 0$. This proves that

$$\sum_{\alpha \in P_+} C(X_\alpha + \theta X_\alpha) \subset \mathfrak{l}$$

and the inclusion

$$\sum_{\alpha \in P_+} C(X_\alpha - \theta X_\alpha) \subset \mathfrak{q}$$

is proved in exactly the same way. Now let α be a fixed element in P_+ and let $c \in \boldsymbol{C}$ be determined by $\theta X_{-\alpha} = c X_{-\alpha^\theta}$. Then

$$X_\alpha = \tfrac{1}{2}(X_\alpha + \theta X_\alpha) + \tfrac{1}{2}(X_\alpha - \theta X_\alpha) \in \mathfrak{l} + \mathfrak{q};$$

$$\begin{aligned} X_{-\alpha} &= \tfrac{1}{2}(\theta X_{-\alpha} + \theta(\theta X_{-\alpha})) + \tfrac{1}{2}(\theta(\theta X_{-\alpha}) - \theta X_{-\alpha}) \\ &= \tfrac{1}{2}c(X_{-\alpha^\theta} + \theta X_{-\alpha^\theta}) - \tfrac{1}{2}c(X_{-\alpha^\theta} - \theta X_{-\alpha^\theta}) \in \mathfrak{l} + \mathfrak{q}. \end{aligned}$$

Consequently the element X above belongs to $\mathfrak{m} + \mathfrak{h}_\mathfrak{p} + \sum_{\alpha \in P_+} C(X_\alpha + \theta X_\alpha) + \sum_{\alpha \in P_+} C(X_\alpha - \theta X_\alpha)$ and the lemma is proved.

The following corollary is an immediate consequence of (1).

Corollary 3.7. *If* $H \in \mathfrak{h}_{\mathfrak{p}}$ *and* $\alpha(H) \neq 0$ *for all* $\alpha \in P_+$, *then the centralizer of* H *in* $\mathfrak{g}$ *is* $\mathfrak{h}_{\mathfrak{p}} + \mathfrak{m}$.

§ 4. Nilpotent Lie Groups

To begin with we establish certain facts concerning the exponential mapping of a nilpotent Lie group. We apply these to a more detailed study of the nilpotent Lie algebra which arises in the Iwasawa decomposition.

Let L be a Lie group with Lie algebra $\mathfrak{l}$. Let $T(\mathfrak{l})$ denote the tensor algebra over $\mathfrak{l}$ considered as a vector space. Let $X \to \tilde{X}$ denote the identity mapping of $\mathfrak{l}$ into $T(\mathfrak{l})$ (this makes it unnecessary to denote the multiplication in $T(\mathfrak{l})$ by a separate symbol). Similarly, if $M = (m_1 ... ,m_n)$ is a positive integral n-tuple we denote by $\tilde{X}(M)$ the coefficient to $t_1^{m_1} ... t_n^{m_n}$ in the product $(|M|!)^{-1} (t_1\tilde{X}_1 + ... + t_n\tilde{X}_n)^{|M|}$, where $X_1, ..., X_n$ is a basis of $\mathfrak{l}$ and $|M| = m_1 + ... + m_n$. An element of the form $\sum a_{e_1 ... e_n} \tilde{X}_1^{e_1} ... \tilde{X}_n^{e_n}$ will be called a *canonical polynomial.* As before, let J denote the two-sided ideal in $T(\mathfrak{l})$ generated by the set of all elements of the form $\tilde{X}\tilde{Y} - \tilde{Y}\tilde{X} - ([X, Y])^{\sim}$, $X, Y \in \mathfrak{l}$. The factor algebra $T(\mathfrak{l})/J$ is the universal enveloping algebra $U(\mathfrak{l})$ of $\mathfrak{l}$. Let $X(M)$ be the image of $\tilde{X}(M)$ under the canonical mapping of $T(\mathfrak{l})$ onto $U(\mathfrak{l})$. As proved earlier, the elements $X(M)$ form a basis of $U(\mathfrak{l})$. From Cor. 1.10, Chapter II, we obtain the following statement: To each $\tilde{X}(M)$ corresponds a unique canonical polynomial $\tilde{R}_M$ such that

$$\tilde{X}(M) \equiv \tilde{R}_M \qquad \text{mod } J.$$

Suppose now $\mathfrak{l}$ is nilpotent. In the central decending series $\mathscr{C}^0\mathfrak{l} \supset \mathscr{C}^1\mathfrak{l} \supset ...$ let $\mathscr{C}^{m-1}\mathfrak{l}$ denote the last nonzero term. The basis $X_1, ..., X_n$ of $\mathfrak{l}$ is said to be *compatible* with the central descending series if there exist integers $r_0 = 1 < r_1 < ... < r_m = n + 1$ such that $X_{r_i}, X_{r_i+1}, ..., X_{r_{i+1}-1}$ is a basis of a complementary subspace of $\mathscr{C}^{i+1}\mathfrak{l}$ in $\mathscr{C}^i\mathfrak{l}$ $(0 \leqslant i \leqslant m - 1)$. We put

$$w(\tilde{X}_i) = p$$

if X_i lies in $\mathscr{C}^{p-1}\mathfrak{l}$ but not in $\mathscr{C}^p\mathfrak{l}$. In particular, $w(\tilde{X}_i) = m$ if $X_i \in \mathscr{C}^{m-1}\mathfrak{l}$. We also put

$$w(c\tilde{X}_{i_1} ... \tilde{X}_{i_r}) = \sum_{k=1}^{r} w(\tilde{X}_{i_k}), \qquad d(c\tilde{X}_{i_1} ... \tilde{X}_{i_r}) = r,$$

c being any real number $\neq 0$. We shall call w the *weight* and d the *degree.* The terms in $\tilde{X}(M)$ all have the same weight, denoted $w(M)$.

Lemma 4.1. *Let L be a nilpotent Lie group with Lie algebra $\mathfrak{l}$. Let $X_1, \ldots, X_n$ be a basis of $\mathfrak{l}$ compatible with the central descending series*

$$\mathscr{C}^0\mathfrak{l} \supset \mathscr{C}^1\mathfrak{l} \supset \ldots \supset \mathscr{C}^m\mathfrak{l} = \{0\}, \qquad \mathscr{C}^{m-1}\mathfrak{l} \neq \{0\},$$

of $\mathfrak{l}$. Let the constants C^P_{MN} be determined by

$$X(M)\,X(N) = \sum_P C^P_{MN}\,X(P).$$

Let $[k]$ denote the n-tuple $(\delta_{k1}, \ldots, \delta_{kn})$. Then

$$C^{[k]}_{MN} = 0 \qquad \text{for each } k, \qquad 1 \leqslant k \leqslant n,$$

provided $|M| + |N| > m$.

Proof. Consider the structural constants $c^k{}_{ij}$ defined by

$$[X_i, X_j] = \sum_{k=1}^{n} c^k{}_{ij} X_k, \qquad 1 \leqslant i, j \leqslant n.$$

Then we have

$$w(\bar{X}_k) = w(\bar{X}_i) + w(\bar{X}_j) \qquad \text{if } c^k{}_{ij} \neq 0. \tag{1}$$

Now let M and N be any positive integral n-tuples. Then $\bar{R}_M$ is obtained from $\bar{X}(M)$ by finitely many replacements

$$\bar{X}_i\bar{X}_j \to \bar{X}_j\bar{X}_i + \sum_{k=1}^{n} c^k{}_{ij}\bar{X}_k + u, \qquad u \in J, \tag{2}$$

followed by reduction mod J. From (1) it follows that each term in $\bar{R}_M$ has weight $w(M)$. We have now

$$\bar{R}_M\bar{R}_N \equiv \bar{R}_{MN} \qquad \mod J,$$

where $\bar{R}_{MN}$ is a uniquely determined canonical polynomial. This polynomial is obtained from $\bar{R}_M\bar{R}_N$ by finitely many replacements (2) followed by reduction mod J. It follows that each term in $\bar{R}_{MN}$ has weight $w(M) + w(N)$. On the other hand, we have

$$\bar{R}_M\bar{R}_N \equiv \sum_P C^P_{MN}\bar{R}_P \qquad \mod J, \tag{3}$$

and consequently

$$\bar{R}_{MN} = \sum_P C^P_{MN}\bar{R}_P. \tag{4}$$

Now the various elements $\bar{X}_1^{e_1} \dots \bar{X}_n^{e_n} \in T(\mathfrak{l})$ are obviously linearly independent. As remarked earlier, each term in $\bar{R}_P$ has weight $w(P)$. From (4) we can therefore conclude that if $C_{MN}^P \neq 0$ then each term in $\bar{R}_P$ has weight $w(M) + w(N)$.

Now suppose $|M| + |N| > m$. Then $w(M) + w(N) > m$. If $C_{MN}^{[k]}$ were $\neq 0$, then $\bar{X}_k$ would have weight $w(M) + w(N) > m$. Since $\mathscr{C}^m\mathfrak{l} = \{0\}$ this is impossible and the lemma is proved.

Definition. Let V and W be two finite dimensional vector spaces over a field K. A *polynomial function* P on V is a function which can be put in the form $P = p(f_1, \dots, f_n)$ where each f_i is a linear function on V with values in K and p is a polynomial (with coefficients in K). A mapping $\varphi : V \to W$ is said to be a *polynomial mapping* if $P \circ \varphi$ is a polynomial function on V whenever P is a polynomial function on W.

Suppose we have chosen bases for V and W. Then the mapping $\varphi : V \to W$ is a polynomial mapping if and only if the coordinates of $\varphi(X) \in W$ are polynomials p_i in the coordinates of $X \in V$. The highest degree of the polynomials p_i is a number, independent of the choice of bases. We shall call this number the *degree* of the polynomial mapping φ.

Theorem 4.2. *Let L be a Lie group with Lie algebra $\mathfrak{l}$. A necessary and sufficient condition for $\mathfrak{l}$ to be nilpotent is the existence of a polynomial mapping $P : \mathfrak{l} \times \mathfrak{l} \to \mathfrak{l}$ such that*

$$\exp X \exp Y = \exp P(X, Y) \qquad \text{for } X, Y \in \mathfrak{l}. \tag{5}$$

In this case P has degree $\leqslant \dim \mathfrak{l}$.

Proof. Assume first that L is nilpotent and let $X_1, \dots, X_n$ be a basis of $\mathfrak{l}$ compatible with the central descending series. Combining Lemma 4.1 with Cor. 2.7, Chapter III, we see that $C_{MN}^{[k]} = 0$ for all $k = 1, \dots, n$ provided $|M| + |N| > \dim \mathfrak{l}$. Denoting canonical coordinates with subscripts we have relation (14) in Chapter II, §1:

$$(xy)_k = \sum_{M,N} C_{MN}^{[k]} x^M y^N,$$

where $x^M = x_1^{m_1} \dots x_n^{m_n}$, etc. We can therefore state: There exist n polynomials $P_i(x_1, \dots, x_n, y_1, \dots, y_n)$, $1 \leqslant i \leqslant n$, of degree $\leqslant n$ and a number $a > 0$ such that

$$\exp(x_1X_1 + \dots + x_nX_n) \exp(y_1X_1 + \dots + y_nX_n) = \exp(P_1X_1 + \dots + P_nX_n) \tag{6}$$

for $|x_i| < a$, $|y_i| < a$, $1 \leqslant i \leqslant n$. However, making use of the following lemma we conclude that (6) holds for all x_i and all y_i. Hence the condition of the theorem is necessary.

Lemma 4.3. *Let M and N be analytic manifolds, M connected. Let φ and ψ be two analytic mappings of M into N. Suppose $\varphi(p) = \psi(p)$ for all p in an open subset of M. Then $\varphi(p) = \psi(p)$ for all $p \in M$.*

Proof. Let $q \in M$. Let us say that φ and ψ have the same partial derivatives at q if (1) $\varphi(q) = \psi(q)$. (2) If $\{x_1, \ldots, x_m\}$ and $\{y_1, \ldots, y_n\}$ are coordinate systems valid near q and $\varphi(q)$, respectively, then the expressions of φ and ψ in these coordinates have the same partial derivatives (of all orders) at the point $(x_1(q), \ldots, x_m(q))$. Let M' be the subset of M consisting of all the points $q \in M$ such that φ and ψ have the same partial derivatives at q. Then M' is obviously closed in M. But M' is also open in M because the partial derivatives of an analytic function at a point determine the power series expansion of the function. Due to the connectedness of M we have $M' = M$ as desired.

To prove the second half of Theorem 4.2 suppose (5) holds for some polynomial mapping $P: \mathfrak{l} \times \mathfrak{l} \to \mathfrak{l}$. Let N_0 be a star-shaped open neighborhood of 0 in $\mathfrak{l}$ such that the mapping exp is a diffeomorphism of N_0 onto a neighborhood N_e of e in L. Let $\{x_1, \ldots, x_n\}$ be a system of canonical coordinates on N_e. Let $X, Y \in \mathfrak{l}$ and assume $X \in N_0$. Then

$$\exp X \exp tY = \exp (X + tZ_t) \qquad (t \in \boldsymbol{R}), \tag{7}$$

where $Z_t = Z_0 + tZ_1 + t^2Z_2 + \ldots$, each Z_i being of the form $Z_i = Q_i(X, Y)$ where Q_i is a polynomial mapping of $\mathfrak{l} \times \mathfrak{l}$ into $\mathfrak{l}$. Since

$$X + tZ_t - (X + tZ_0) = O(t^2)$$

it follows that

$$\left\{\frac{d}{dt} f(\exp (X + tZ_t))\right\}_{t=0} = \left\{\frac{d}{dt} f(\exp (X + tZ_0))\right\}_{t=0},$$

whenever f is one of the coordinate functions x_i. Using (7), we conclude

$$[dL_{\exp X}(Y)f](\exp X) = \left\{\frac{d}{dt} f(\exp (X + tZ_0)\right\}_{t=0}.$$

On the other hand, we know from Theorem 1.7, Chapter II, that

$$\left\{\frac{d}{dt} f(\exp (X + tZ_0))\right\}_{t=0} = \left[\left\{dL_{\exp X} \circ \frac{1 - e^{-\operatorname{ad} X}}{\operatorname{ad} X}(Z_0)\right\} f\right](\exp X).$$

Since $Z_0 = Q_0(X, Y)$ we get from these two equations

$$\frac{\operatorname{ad} X}{1 - e^{-\operatorname{ad} X}}(Y) = Q_0(X, Y)$$

if X is sufficiently small, Y arbitrary. Now if $|x| < 2\pi$ we have an absolutely convergent series $\Sigma_0^\infty a_n x^n$ satisfying

$$\left(\sum_0^\infty a_n x^n\right)\left(\sum_1^\infty \frac{(-1)^{n-1}}{n!} x^{n-1}\right) = 1 \tag{8}$$

and infinitely many a_n are $\neq 0$. Equation (8) simply amounts to an infinite set of equations for the coefficients a_n. If X is sufficiently small, the series $\Sigma_0^\infty a_n(\operatorname{ad} X)^n$ and $\Sigma_1^\infty (-1)^{n-1}/n!\,(\operatorname{ad} X)^{n-1}$ converge absolutely† and can be multiplied together, term by term. Thus relation (8) remains true if we replace x by $\operatorname{ad} X$. It follows that

$$\sum_0^\infty a_n(\operatorname{ad} X)^n (Y) = Q_0(X, Y).$$

Since infinitely many coefficients a_n are $\neq 0$, we conclude that $\operatorname{ad} X$ is nilpotent, hence $\mathfrak{l}$ is nilpotent.

Corollary 4.4. *Let L be a connected nilpotent Lie group with Lie algebra $\mathfrak{l}$. Then the exponential mapping is a regular mapping of $\mathfrak{l}$ onto L.*

If $X \in \mathfrak{l}$, then $\operatorname{ad} X$ is nilpotent so there exists a basis of $\mathfrak{l}$ such that the corresponding matrix expression for $\operatorname{ad} X$ has zeros on and below the diagonal. Consequently

$$\det\left(\frac{1 - e^{-\operatorname{ad} X}}{\operatorname{ad} X}\right) \neq 0.$$

In view of Theorem 1.7, Chapter II, this means that exp is regular everywhere on $\mathfrak{l}$. On the other hand, Theorem 4.2 shows that $\exp(\mathfrak{l})$ is a subgroup of L, which, due to the regularity of exp, is an open subgroup. An open subgroup is always closed and due to the connectedness of L we find $\exp(\mathfrak{l}) = L$.

Remark. This corollary can also be proved more directly and without using Theorem 4.2. If $\mathfrak{l} \neq \{0\}$ the center $\mathfrak{c}$ of $\mathfrak{l}$ is not zero and the factor algebra $\mathfrak{l}/\mathfrak{c}$ is again nilpotent and has dimension less than $\dim \mathfrak{l}$. Corollary 4.4 can now be proved by induction. The details are left to the reader.

† Let $|\ |$ be some norm on a finite-dimensional vector space V over $\boldsymbol{R}$. If A is an endomorphism of V, put $\|A\| = \sup(|Ax|/|x|)$. If A_n $(n = 0,1, \ldots)$ is an endomorphism of V, the series $\Sigma_0^\infty A_n$ is said to be absolutely convergent if $\Sigma_0^\infty \|A_n\|$ is convergent.

Let N be a nilpotent endomorphism of a finite-dimensional vector space V over $\boldsymbol{R}$. We put $\log(1 + N) = \Sigma_{n \geqslant 1}(-1)^{n-1} N^n/n$ (finite series). It is clear that $\log(1+N)$ and $e^N - 1$ are also nilpotent.

Lemma 4.5.

$$\log e^N = N, \qquad e^{\log(1+N)} = 1 + N.$$

Proof. Let x be a real number and let the coefficients $a_{m,n}$ be determined by

$$\left(\sum_{r=1}^{\infty} \frac{1}{r!} x^r\right)^n = \sum_{m=1}^{\infty} a_{m,n} x^m \qquad \text{for all } x. \tag{9}$$

If $|\exp x - 1| < 1$, then

$$x = \log((\exp x - 1) + 1) = \sum_{n=1}^{\infty} (-1)^{n-1}/n \left(\sum_{r=1}^{\infty} \frac{1}{r!} x^r\right)^n$$

$$= \sum_{n=1}^{\infty} (-1)^{n-1}/n \left(\sum_{m=1}^{\infty} a_{m,n} x^m\right).$$

Owing to Weierstrass' theorem on double series, the summations can be interchanged so

$$x = \sum_{m=1}^{\infty} \left(\sum_{n=1}^{\infty} (-1)^{n-1}/n \, a_{m,n}\right) x^m. \tag{10}$$

If A is any endomorphism of V, the series $\Sigma_1^{\infty} (1/r!) A^r$ is absolutely convergent. Relation (9) remains true if we replace x by A. Consequently,

$$\sum_{n=1}^{\infty} (-1)^{n-1}/n \left(\sum_{r=1}^{\infty} \frac{1}{r!} N^r\right)^n = \sum_{n=1}^{\infty} (-1)^{n-1}/n \sum_{m=1}^{\infty} a_{m,n} N^m.$$

Since N is nilpotent, the series are actually finite and the summations can of course be interchanged. Considering (10) it follows that

$$\log e^N = N.$$

The second relation can be proved in the same way.

We consider now an Iwasawa decomposition $\mathfrak{g}_0 = \mathfrak{k}_0 + \mathfrak{h}_{\mathfrak{p}_0} + \mathfrak{n}_0$ of an arbitrary semisimple Lie algebra $\mathfrak{g}_0$ over $\boldsymbol{R}$. Let G be any connected Lie group with Lie algebra $\mathfrak{g}_0$ and let K, $A_{\mathfrak{p}}$, and N be the analytic subgroups of G with Lie algebras $\mathfrak{k}_0$, $\mathfrak{h}_{\mathfrak{p}_0}$, and $\mathfrak{n}_0$, respectively. We shall now prove two results about N which will be used in Chapter X. W preserve the notation from §3. Moreover, we put

$$D(H) = \prod_{\alpha \in P_+} \sinh \frac{\alpha(H)}{2} \qquad \text{for } H \in \mathfrak{h}_0$$

and let $\mathfrak{h}_0'$ denote the set of points $H \in \mathfrak{h}_0$ for which $D(H) \neq 0$. Let Ad denote the adjoint representation of G and let ad denote the adjoint representation of $\mathfrak{g}_0$.

Lemma 4.6. *Let H be an element in $\mathfrak{h}_0'$. Then the mapping $\varphi : n \to Ad\,(n)\,H - H$ is an analytic diffeomorphism of N onto $\mathfrak{n}_0$.*

Proof. Let $X \in \mathfrak{n}_0$. Then $\mathrm{Ad}\,(\exp X)\,H - H = e^{\mathrm{ad}\,X}\,H - H \in \mathfrak{n}_0$ since $[\mathfrak{h}_0, \mathfrak{n}_0] \subset \mathfrak{n}_0$. Since exp maps $\mathfrak{n}_0$ onto N (Cor. 4.4), it is clear that $\varphi(N) \subset \mathfrak{n}_0$. Next we prove that φ is one-to-one. Suppose $n_1, n_2 \in N$ such that $\mathrm{Ad}\,(n_1)\,H - H = \mathrm{Ad}\,(n_2)\,H - H$. There exists an element $X \in \mathfrak{n}_0$ such that $\exp X = n_2^{-1}n_1$. Then $e^{\mathrm{ad}\,X}\,H = H$ which by Lemma 4.5 implies $\mathrm{ad}\,X(H) = 0$ or equivalently $\mathrm{ad}\,H(X) = 0$. Since $\alpha(H) \neq 0$ for each $\alpha \in P_+$ it follows that the restriction of ad H to $\mathfrak{n}_0$ is nonsingular. Consequently, $X = 0$ and $n_2 = n_1$.

We next prove that $\varphi(N) = \mathfrak{n}_0$. If this were not the case let Z be an element in $\mathfrak{n}_0$ such that $Z \notin \varphi(N)$. Then $Z = \sum_{\alpha \in P_+} c_\alpha X_\alpha$ where the c_α are complex numbers, not all zero. Let β be the lowest root in P_+ for which $c_\beta \neq 0$. We can assume Z chosen in $\mathfrak{n}_0 - \varphi(N)$ in such a manner that β is as high as possible. Since ad H is nonsingular on $\mathfrak{n}_0$ there exists an element $Z_1 \in \mathfrak{n}_0$ such that $[H, Z_1] = Z$. If $n_1 = \exp Z_1$ we have

$$\mathrm{Ad}\,(n_1)\,(H + Z) - H \equiv [Z_1, H] + Z \qquad \mod \sum_{\alpha > \beta} \mathfrak{g}^\alpha$$

so

$$\mathrm{Ad}\,(n_1)\,(H + Z) - H = Z' \qquad \text{where } Z' \in \mathfrak{n}_0 \cap \sum_{\alpha > \beta} \mathfrak{g}^\alpha.$$

Owing to the choice of Z there exists an element $n' \in N$ such that $\mathrm{Ad}\,(n')\,H - H = Z'$. This implies $\mathrm{Ad}\,(n)\,H - H = Z$ if $n = n_1^{-1}n'$. This contradiction shows that $\varphi(N) = \mathfrak{n}_0$.

Finally, to show that φ is regular we compute its differential $d\varphi_n$ at an arbitrary point $n \in N$. Each tangent vector in N_n has the form dL_nX where $X \in \mathfrak{n}_0$ and L_n denotes as usual the left translation by n. Since

$$\begin{aligned}\varphi(n \exp tX) &= \mathrm{Ad}\,(n)\,e^{\mathrm{ad}\,tX}\,H - H \\ &= \mathrm{Ad}\,(n)\,H - H + \mathrm{Ad}\,(n)\,(t[X, H] + \frac{t^2}{2}[X, [X, H]] + \ldots)\end{aligned}$$

it follows that

$$d\varphi_n(dL_n X) = -\operatorname{Ad}(n)\operatorname{ad} H(X), \qquad X \in \mathfrak{n}_0. \tag{11}$$

Since $\operatorname{Ad}(n)$ and $\operatorname{ad} H$ are both nonsingular on $\mathfrak{n}_0$, the regularity of φ follows and the lemma is proved.

Theorem 4.7. *Let $H \in \mathfrak{h}_0'$. Then the mapping*

$$\psi : X \to \exp(-H)\exp(H + X)$$

is an analytic diffeomorphism of $\mathfrak{n}_0$ onto N.

Proof. Let $X \in \mathfrak{n}_0$. Then by Lemma 4.6 $H + X = \operatorname{Ad}(n) H$ for some $n \in N$ and therefore $\psi(X) = \exp(-H)\exp(H + X) = \exp(-H)n \exp H\, n^{-1} \in N$. Next we prove that ψ is one-to-one. If this were not so, select $X_1, X_2 \in \mathfrak{n}_0$ such that $X_1 \neq X_2$ and $\psi(X_1) = \psi(X_2)$. By Lemma 4.6 there exist elements $n_1, n_2 \in N$ such that $\operatorname{Ad}(n_i) H - H = X_i$ $(i = 1, 2)$. Put $n = n_2^{-1}n_1$ and let $X \in \mathfrak{n}_0$ be determined by $\operatorname{Ad}(n) H - H = X$. Then the equation $\psi(X_1) = \psi(X_2)$ implies $\exp(\operatorname{Ad}(n_1)H) = \exp(\operatorname{Ad}(n_2)H)$ so $\exp(H + X) = \exp H$. Now $X = \Sigma_{\alpha \in P_+} a_\alpha X_\alpha$ where the a_α are complex numbers, not all zero. Let β be the lowest root in P_+ such that $a_\beta \neq 0$ and put $\mathfrak{g}(\beta) = \Sigma_{\alpha > \beta}\, \mathfrak{g}^\alpha$. Then, if $H' \in \mathfrak{h}_0'$

$$\begin{aligned}
&\operatorname{Ad}(\exp(H + X))\, H' = e^{\operatorname{ad}(H+X)} H' \\
&\equiv H' + [X, H'] + \frac{1}{2!}[H, [X, H']] + \dots && \operatorname{mod} \mathfrak{g}(\beta) \\
&\equiv H' + \frac{e^{\operatorname{ad} H} - 1}{\operatorname{ad} H}([X, H']) && \operatorname{mod} \mathfrak{g}(\beta) \\
&\equiv H' + a_\beta \beta(H')\frac{1 - e^{\beta(H)}}{\beta(H)} X_\beta && \operatorname{mod} \mathfrak{g}(\beta)
\end{aligned}$$

so $\operatorname{Ad}(\exp(H + X))\, H' \neq H'$, contradicting $\exp(H + X) = \exp H$.

Next we find the differential $d\psi_X$ at an arbitrary point $X \in \mathfrak{n}_0$. Since ψ is the composite of three mappings, namely $\psi = L_{\exp(-H)} \circ \exp \circ L_H$, we have

$$d\psi_X = dL_{\exp(-H)} \circ d\exp_{(H+X)} \circ dL_H.$$

Using the formula for $d\exp_Z$ (Theorem 1.7, Chapter II), we obtain

$$d\psi_X(Y) = dL_{\psi(X)} \circ \frac{1 - e^{-\operatorname{ad}(H+X)}}{\operatorname{ad}(H + X)}(Y), \qquad Y \in \mathfrak{n}_0. \tag{12}$$

The restriction of $\operatorname{ad}(H + X)$ to $\mathfrak{n}_0$ is an endomorphism of $\mathfrak{n}_0$. We extend this endomorphism to an endomorphism D of $\mathfrak{n}$. If D is expressed by means of the basis $X_\alpha(\alpha \in P_+)$ of $\mathfrak{n}$, we get a triangular matrix with elements $\alpha(H)$ $(\alpha \in P_+)$ in the diagonal. It follows that

$$\det\left(\frac{1 - e^{-D}}{D}\right) = \prod_{\alpha \in P_+} \frac{1 - e^{-\alpha(H)}}{\alpha(H)} \neq 0$$

since $H \in \mathfrak{h}_0'$. This proves the regularity of ψ.

Now, let $V = \psi(\mathfrak{n}_0)$ and let U be the set of $X \in \mathfrak{n}_0$ such that $\exp X \in V$. From the regularity of ψ it follows that V is open in N; hence U is open in $\mathfrak{n}_0$. To prove the lemma it remains to show that $U = \mathfrak{n}_0$. The mapping $\zeta : X \to \psi^{-1}(\exp X)$ $(X \in U)$ is an analytic mapping of U onto $\mathfrak{n}_0$. Since $\psi \circ \zeta = \exp$ we get the differential $d\zeta_X$ from (12)

$$d\zeta_X(Y) = \frac{\operatorname{ad}(H + \zeta(X))}{1 - \exp(-\operatorname{ad}(H + \zeta(X)))} \circ \frac{1 - \exp(-\operatorname{ad} X)}{\operatorname{ad} X}(Y) \quad (13)$$

if $X \in U$, $Y \in \mathfrak{n}_0$. Now fix $X \in U$ and let T denote the set of $t \in \boldsymbol{R}$ such that $tX \in U$. Let T_0 denote the component of T containing 0. Consider now the curve $Z(t)$ in $\mathfrak{n}_0$ given by

$$Z(t) = \zeta(tX) \qquad (t \in T_0).$$

Then $dZ(t)/dt = d\zeta_{tX}(X)$ and using (13) we get the differential equation

$$\frac{dZ(t)}{dt} = \frac{\operatorname{ad}(H + Z(t))}{1 - \exp(-\operatorname{ad}(H + Z(t)))} X, \qquad t \in T_0.$$

The restriction of $1 - \operatorname{Ad}((\exp H \exp tX)^{-1}))$ to $\mathfrak{n}_0$ is (for each $t \in \boldsymbol{R}$) an endomorphism $D(t)$ of $\mathfrak{n}_0$. If $t \in T_0$, then $D(t)$ coincides with $1 - \exp(-\operatorname{ad}(H + Z(t)))$. If $D(t)$ $(t \in \boldsymbol{R})$ is extended to $\mathfrak{n}$ and expressed in matrix form by means of the basis X_α $(\alpha \in P_+)$ then we obtain a triangular matrix with diagonal elements $1 - e^{-\alpha(H)}$, $(\alpha \in P_+)$. Hence $\det D(t) = \prod_{\alpha \in P_+} (1 - e^{-\alpha(H)}) \neq 0$ so $D(t)$ has an inverse. Let $|\ |$ be any norm on $\mathfrak{g}_0$ and let $\| D(t)^{-1} \| = \sup (| D(t)^{-1}(Z) | / | Z |)$ as Z varies through $\mathfrak{n}_0$. Now, obviously

$$\frac{d}{dt} | Z(t) | \leqslant \left| \frac{dZ(t)}{dt} \right|,$$

$$| \operatorname{ad}(H + Z(t)) X | = | \operatorname{ad} X(H) + \operatorname{ad} X(Z(t)) | \leqslant p(1 + | Z(t) |)$$

for all $t \in T_0$, p being a positive constant. It follows that

$$\frac{d}{dt} | Z(t) | \leqslant p \| D(t)^{-1} \| (1 + | Z(t) |), \qquad t \in T_0.$$

The set T_0 is an open interval. In order to prove that $U = \mathfrak{n}_0$ it suffices, X being arbitrary, to prove that $T_0 = \boldsymbol{R}$. If this were not the case, let t_0 be an end point of the interval T_0. Then $t_0 \notin T_0$ and without loss of generality we can assume that $t_0 > 0$. There exists a constant M such that $p \, \| D(t)^{-1} \| \leqslant M$ for $0 \leqslant t < t_0$. The resulting inequality

$$\frac{d}{dt} | Z(t) | \leqslant M(1 + | Z(t) |), \qquad 0 \leqslant t < t_0,$$

yields by repeated integration

$$| Z(t) | \leqslant e^{Mt} - 1 + M \frac{(tM)^n}{n!} \max_{0 \leqslant s \leqslant t} | Z(s) |, \qquad 0 \leqslant t < t_0.$$

Letting $n \rightarrow \infty$ we obtain

$$| Z(t) | \leqslant e^{Mt}, \qquad 0 \leqslant t < t_0.$$

Let (t_k) be a sequence converging to t_0 from below. Then the sequence $Z(t_k)$ is bounded and, passing to a subsequence of (t_k) if necessary, we can assume $Z(t_k)$ converges to an element $Z \in \mathfrak{n}_0$. Then

$$\exp t_0 X = \lim_{k \to \infty} \psi(Z(t_k)) = \psi(Z) \in V$$

so $t_0 \in T_0$; this contradiction proves the theorem.

Corollary 4.8. *Suppose $H \in \mathfrak{h}_0'$ and $h = \exp H$. Then the mapping $\xi : n \rightarrow h^{-1}nhn^{-1}$ is an analytic diffeomorphism of N onto itself.*

This is an immediate consequence of the last two lemmas. In fact $\xi = \psi \circ \varphi$.

Remark. The Lie algebra $\mathfrak{g}_0$ plays little role in the three last results. These can be generalized accordingly.

§ 5. Global Decompositions

Theorem 5.1. *Let $\mathfrak{g}_0 = \mathfrak{k}_0 + \mathfrak{h}_{\mathfrak{p}_0} + \mathfrak{n}_0$ be an Iwasawa decomposition of a semisimple Lie algebra $\mathfrak{g}_0$ over $\boldsymbol{R}$. Let G be any connected Lie group with Lie algebra $\mathfrak{g}_0$ and let K, $A_{\mathfrak{p}}$, N be the analytic subgroups of G with Lie algebras $\mathfrak{k}_0$, $\mathfrak{h}_{\mathfrak{p}_0}$, and $\mathfrak{n}_0$, respectively. Then the mapping*

$$\Phi : (k, a, n) \rightarrow kan \qquad (k \in K,\ a \in A_{\mathfrak{p}},\ n \in N),$$

is an analytic diffeomorphism of the product manifold $K \times A_{\mathfrak{p}} \times N$ onto G.

The groups $A_{\mathfrak{p}}$ and N are simply connected.

We begin by proving a general lemma (Harish-Chandra [4], p. 213), which will also be useful later.

Lemma 5.2. *Let U be a Lie group with Lie algebra $\mathfrak{u}$. Suppose $\mathfrak{u}$ is a direct sum $\mathfrak{u} = \mathfrak{m} + \mathfrak{h}$ where $\mathfrak{m}$ and $\mathfrak{h}$ are subalgebras of $\mathfrak{u}$ (not necessarily ideals). Let M and H be the analytic subgroups of U with Lie algebras $\mathfrak{m}$ and $\mathfrak{h}$, respectively. Then the mapping $\alpha : (m, h) \to mh$ $(m \in M, h \in H)$ of $M \times H$ into U is everywhere regular.*

Proof. As usual we denote by L_x the left translation by a group element x. The tangent vector to the curve $x \exp tX$ at x is $dL_x(X)$. We identify H and M, respectively, with the subgroups (e, H) and (M, e) of the product group $M \times H$. Also, the tangent space $(M \times H)_{(m,h)}$ is identified with the direct sum $M_m + H_h$ $(m \in M, h \in H)$.

Let $Y \in \mathfrak{m}$, $Z \in \mathfrak{h}$. Then

$$\alpha(m \exp tY, h) = mh \exp (t \operatorname{Ad}(h^{-1}) Y), \qquad t \in \boldsymbol{R},$$

$$\alpha(m, h \exp tZ) = mh \exp tZ.$$

It follows that

$$d\alpha_{(m,h)}(dL_m Y, dL_h Z) = dL_{mh}(\operatorname{Ad}(h^{-1}) Y + Z). \tag{1}$$

Now suppose $\operatorname{Ad}(h^{-1}) Y + Z = 0$; then $Y + \operatorname{Ad}(h) Z = 0$ and since $\operatorname{Ad}(h) Z \in \mathfrak{h}$ we have $Y = Z = 0$. This proves the lemma.

Let G_0 be the adjoint group of $\mathfrak{g}_0$. As usual we identify $\operatorname{ad}(\mathfrak{g}_0)$ and $\mathfrak{g}_0$. Let K_0, A_0, N_0, and S_0 denote the analytic subgroups of G_0 with Lie algebras $\mathfrak{k}_0$, $\mathfrak{h}_{\mathfrak{p}_0}$, $\mathfrak{n}_0$, and $\mathfrak{s}_0 = \mathfrak{h}_{\mathfrak{p}_0} + \mathfrak{n}_0$, respectively. We shall begin by proving Theorem 5.1 for the group G_0. The elements of G_0 are endomorphisms of $\mathfrak{g}_0$ which we extend to the complex algebra $\mathfrak{g}$. In terms of the basis (X_i) of $\mathfrak{g}$ from Lemma 3.5, the elements of K_0 are represented by unitary matrices, the group A_0 consists of positive diagonal matrices and the elements of N_0 are represented by lower triangular matrices with diagonal elements equal to 1. Now if a triangular matrix with positive diagonal elements is unitary, it must be the unit matrix. It follows that the mapping

$$(k, a, n) \to kan, \qquad k \in K_0,\ a \in A_0,\ n \in N_0,$$

of $K_0 \times A_0 \times N_0$ into G_0 is one-to-one. The group A_0 is obviously a simply connected closed subgroup of G_0. From Cor. 4.4 and Lemma 4.5 it follows immediately that the exponential mapping for matrices is an analytic diffeomorphism of the Lie algebra of all lower triangular matrices with zeros in the diagonal onto the Lie group of all lower triangular matrices with all diagonal elements equal to 1. This being so, it follows that N_0 is a closed, simply connected subgroup of G_0. The set $A_0N_0 = \{an : a \in A_0, n \in N_0\}$ is a subgroup of G_0. Since a represents the diagonal in the matrix an, and since A_0 is closed in G_0, it is obvious that the group A_0N_0 is a closed subgroup of G_0, hence an analytic subgroup of G_0. Now, by Lemma 5.2, the mapping $(a, n) \to an$ is a diffeomorphism of $A_0 \times N_0$ onto A_0N_0; hence $\dim A_0N_0 = \dim S_0$. But obviously $A_0N_0 \subset S_0$; hence $A_0N_0 = S_0$.

Consider now the Riemannian globally symmetric space $M = G_0/K_0$. Let R be the subgroup consisting of those elements in G_0 which leave every point of M fixed. Then we know (Remark 2, Prop. 3.4, Chapter IV) that the factor group G_0/R is a closed subgroup of the group $I(M)$ of all isometries of M. The natural mapping of G_0 onto G_0/R maps S_0 onto a subgroup S_* of G_0/R. Since S_0 is closed in G_0 and R compact, it follows that S_* is closed in G_0/R, hence a closed subgroup of $I(M)$. Let p denote the point in M given by the coset $\{K_0\}$. It is clear that $\dim S_* = \dim M$ so the orbit $S_* \cdot p$ is an open subset of M due to Lemma 4.1, Chapter II. But being the orbit of a closed subgroup of $I(M)$, $S_* \cdot p$ is a closed subset of M (Theorem 2.2, Chapter IV).† Hence $S_* \cdot p = M$. In terms of G_0, this result means that each $g \in G_0$ can be written $g = sk$, where $s \in S_0$, $k \in K_0$. Taking inverses it follows that each $g_1 \in G_0$ can be written $g_1 = k_1s_1$, $(k_1 \in K_0, s_1 \in S_0)$. The mapping $(k, a, n) \to kan$ of $K_0 \times A_0 \times N_0$ into G_0 is therefore one-to-one, onto and regular, hence a diffeomorphism.

To prove Theorem 5.1 in full generality, let π denote the natural mapping of G onto G_0. The kernel of π is the center Z of G. Since Z is discrete, (G, π) is a covering group of G_0. The identity component of the groups $\pi^{-1}(A_0)$ and $\pi^{-1}(N_0)$ coincides with $A_\mathfrak{p}$ and N, respectively. The groups A_0 and N_0 are simply connected and have covering groups $(A_\mathfrak{p}, \pi)$ and (N, π). Hence $A_\mathfrak{p} \cap Z = N \cap Z = \{e\}$ and $A_\mathfrak{p}$ and N are simply connected. If we put $\tilde{K} = \pi^{-1}(K_0)$ we have evidently $G = \tilde{K}A_\mathfrak{p}N$ and each g can be written uniquely $g = kan$ $(k \in \tilde{K}, a \in A_\mathfrak{p}, n \in N)$. Here a and n depend continuously on g, because π is a homeomorphism of $A_\mathfrak{p}$ onto A_0 and of N onto N_0. Hence k depends continuously on g; hence $A_\mathfrak{p}N$ and $KA_\mathfrak{p}N$ are closed in G. The regularity of Φ follows by applying Lemma 5.2 twice: first on the subgroups $A_\mathfrak{p}$ and N of $A_\mathfrak{p}N$, next on the subgroups K and $A_\mathfrak{p}N$ of G. Thus $KA_\mathfrak{p}N$ is open and closed in G so $G = KA_\mathfrak{p}N$; this finishes the proof of Theorem 5.1.

Proposition 5.3. *In the notation of Theorem* 5.1, *let* $S = A_\mathfrak{p}N$, $P = exp\, \mathfrak{p}_0$. *Then* S *is a closed solvable subgroup of* G, P *is a closed*

† Or by Prop. 4.4, Chapter II.

submanifold of G. Let $\tilde{\theta}$ denote the automorphism of G for which $d\tilde{\theta} = \theta$. Then the mapping

$$\psi : s \to \tilde{\theta}(s)\, s^{-1}, \qquad s \in S,$$

is a diffeomorphism of S onto P.

Proof. Only the last statement has not been proved already. Each $g \in G$ can be written $g = pk$, $p \in P$, $k \in K$. Then $\tilde{\theta}(g) = p^{-1}k$ so $\tilde{\theta}(g)g^{-1} = p^{-2} \in P$. In particular, $\psi(S) \subset P$. The mapping ψ is one-to-one. In fact, if $\tilde{\theta}(s_1)\, s_1^{-1} = \tilde{\theta}(s_2)s_2^{-1}$, $(s_1, s_2 \in S)$, then $\tilde{\theta}(s_2^{-1}s_1) = s_2^{-1}s_1$; hence $s_2^{-1}s_1 \in K \cap S = \{e\}$ so $s_1 = s_2$. Furthermore, $\psi(S) = P$; in fact, given $p \in P$ there exists a unique $X \in \mathfrak{p}_0$ such that $p = \exp X$. By Theorem 5.1 there exist unique elements $k \in K$ and $s \in S$ such that $\exp \frac{1}{2} X = ks^{-1}$. Then $p = \tilde{\theta}(s)\, s^{-1}$ as desired. This mapping $\psi^{-1} : p \to s$ is differentiable because it is composed of the mappings

$$p \xrightarrow{\exp^{-1}} X \to \frac{X}{2} \to \exp \frac{X}{2} \to s.$$

§ 6. The Complex Case

It will be convenient later to have the Iwasawa decomposition (Theorems 3.4 and 5.1) restated for the case when the semisimple Lie algebra in question has a complex structure.

Suppose $\mathfrak{g}_0$ is a semisimple Lie algebra over $\boldsymbol{R}$ with complex structure J. This simply means (Chapter III, §6) that there exists an endomorphism J of $\mathfrak{g}_0$ such that

$$J^2 = -I,$$
$$(\mathrm{ad}_{\mathfrak{g}_0} X)\, J = J\, \mathrm{ad}_{\mathfrak{g}_0} X, \qquad X \in \mathfrak{g}_0.$$

As shown in §6, Chapter III, the set $\mathfrak{g}_0$ can be regarded as a Lie algebra $\tilde{\mathfrak{g}}_0$ over $\boldsymbol{C}$. The Lie algebra $\mathfrak{g}_0$ is obtained from $\tilde{\mathfrak{g}}_0$ by restricting the scalars to $\boldsymbol{R}$, in other words $(\tilde{\mathfrak{g}}_0)^R = \mathfrak{g}_0$. Let $\mathfrak{c}$ be any compact real form of the semi-simple Lie algebra $\tilde{\mathfrak{g}}_0$. Then

$$\mathfrak{g}_0 = \mathfrak{c} + J\mathfrak{c}$$

is a Cartan decomposition of $\mathfrak{g}_0$ (Cor. 7.5, Chapter III). We can therefore carry through the construction in §3 for $\mathfrak{k}_0 = \mathfrak{c}$ and $\mathfrak{p}_0 = J\mathfrak{c}$. The maximal abelian subspace $\mathfrak{h}_{\mathfrak{p}_0}$ has the form $J\mathfrak{a}_0$ where $\mathfrak{a}_0$ is a maximal abelian subalgebra of $\mathfrak{c}$. Then $\mathfrak{h}_0 = \mathfrak{a}_0 + J\mathfrak{a}_0$ is a maximal abelian subalgebra of $\mathfrak{g}_0$. Since $J\mathfrak{h}_0 \subset \mathfrak{h}_0$, the Lie algebra $\mathfrak{h}_0$ has a complex structure inherited from $\mathfrak{g}_0$. The corresponding Lie algebra $\tilde{\mathfrak{h}}_0$ over C is a Cartan subalgebra of $\tilde{\mathfrak{g}}_0$. Let $\tilde{\Delta}$ denote the set of nonzero roots of $\tilde{\mathfrak{g}}_0$ with respect to $\tilde{\mathfrak{h}}_0$. In accordance with previous terminology we put

$$(\tilde{\mathfrak{h}}_0)^* = \sum_{\alpha \in \tilde{\Delta}} \boldsymbol{R} H_\alpha.$$

Owing to Lemma 3.1 we have $(\tilde{\mathfrak{h}}_0)^* = J\mathfrak{a}_0$. Suppose now that the dual of the (real) vector space $J\mathfrak{a}_0$ has been turned into an ordered vector space. Let $(\tilde{\Delta})^+$ denote the set of positive roots in $\tilde{\Delta}$ according to this ordering. We put

$$\tilde{\mathfrak{n}}_+ = \sum_{\alpha \in (\tilde{\Delta})^+} \tilde{\mathfrak{g}}_0^\alpha \qquad \text{and} \qquad \mathfrak{n}_+ = (\tilde{\mathfrak{n}}_+)^R.$$

Let $\mathfrak{g} = \mathfrak{g}_0 + i\mathfrak{g}_0$ be the complexification of the real Lie algebra $\mathfrak{g}_0$ and extend J to a (C-linear) endomorphism of $\mathfrak{g}$. The algebra $\mathfrak{h} = \mathfrak{h}_0 + i\mathfrak{h}_0$ is a Cartan subalgebra of $\mathfrak{g}$ (Lemma 3.2). Let Δ denote the set of nonzero roots of $\mathfrak{g}$ with respect to $\mathfrak{h}$ and as before we put

$$\mathfrak{h}^* = \sum_{\alpha \in \Delta} \boldsymbol{R} H_\alpha.$$

Then according to Lemma 3.2 we have $\mathfrak{h}^* = J\mathfrak{a}_0 + i\mathfrak{a}_0$. It is obviously possible to turn the dual space of $\mathfrak{h}^*$ into an ordered vector space such that the orderings in the duals of $(\tilde{\mathfrak{h}}_0)^*$ and $\mathfrak{h}^*$ are compatible. Let Δ^+ denote the set of positive roots in Δ with respect to this ordering. In §3 we have divided the set Δ^+ into two subsets P_- and P_+, P_- containing exactly those roots in Δ^+ which vanish identically on $\mathfrak{h}_{\mathfrak{p}_0}$.

Lemma 6.1. *In the present case, $\Delta^+ = P_+$ so P_- is empty.*

Proof. The equation $J[X, Y] = [JX, Y]$ holds for all $X, Y \in \mathfrak{g}_0$, hence for all $X, Y \in \mathfrak{g}$, J being C-linear. Now let α be a nonzero root of $\mathfrak{g}$ with respect to $\mathfrak{h}$ and select a nonzero vector $X_\alpha \in \mathfrak{g}^\alpha$. Let $H \in \mathfrak{h}$. The equation $[H, X_\alpha] = \alpha(H)\, X_\alpha$ implies $[H, JX_\alpha] = \alpha(H)\, JX_\alpha$. Hence $JX_\alpha = cX_\alpha$ where c is a complex number. On the other hand, $J\mathfrak{h} \subset \mathfrak{h}$ and $\alpha(JH)\, X_\alpha = [JH, X_\alpha] = [H, JX_\alpha] = c\alpha(H)\, X_\alpha$. Hence $\alpha(JH) = c\alpha(H)$ for all $H \in \mathfrak{h}$. From this equation it is obvious that α cannot vanish identically on the space $J\mathfrak{a}_0$ (which plays the role of $\mathfrak{h}_{\mathfrak{p}_0}$). This proves the lemma.

Lemma 6.2. *Let $\mathfrak{n} = \sum_{\alpha \in \Delta^+} \mathfrak{g}^\alpha$. Then $\mathfrak{n} \cap \mathfrak{g}_0 = \mathfrak{n}_+$.*

Proof. Let $\gamma \in (\tilde{\Delta})^+$ and $Z_\gamma \neq 0$ in $\tilde{\mathfrak{g}}_0^\gamma$. Then $[H, Z_\gamma] = \gamma(H)\, Z_\gamma$ for all $H \in \mathfrak{h}_0$. If we extend γ to a C-linear function γ^* on $\mathfrak{h}$ then $[H, Z_\gamma] =$

$\gamma^*(H) Z_\gamma$ for all $H \in \mathfrak{h}$. Thus $\gamma^* \in \Delta$ and due to the compatibility of the orderings we have $\gamma^* \in \Delta^+$. Hence $Z_\gamma \in \mathfrak{n} \cap \mathfrak{g}_0$ and $\mathfrak{n}_+ \subset \mathfrak{n} \cap \mathfrak{g}_0$. On the other hand, the number of elements in Δ^+ is twice the number of elements in $(\tilde{\Delta})^+$. It follows that

$$\dim_R \mathfrak{n}_+ = \dim_C \mathfrak{n} = \dim_R \mathfrak{n} \cap \mathfrak{g}_0.$$

This proves the lemma. Theorems 3.4 and 5.1 can now be restated (in simplified notation).

Theorem 6.3. *Let $\mathfrak{g}$ be a semisimple Lie algebra over C, $\mathfrak{g}^R$ the Lie algebra $\mathfrak{g}$ considered as a Lie algebra over R. Let J be the complex structure on $\mathfrak{g}^R$ which corresponds to multiplication by i on $\mathfrak{g}$. Let $\mathfrak{u}$ be any compact real form of $\mathfrak{g}$ and let $\mathfrak{a}$ be any maximal abelian subalgebra of $\mathfrak{u}$. Then the algebra $\mathfrak{h} = \mathfrak{a} + i\mathfrak{a}$ is a Cartan subalgebra of $\mathfrak{g}$. Let Δ be the set of roots of $\mathfrak{g}$ with respect to $\mathfrak{h}$ and let Δ^+ be the set of positive roots with respect to some ordering of Δ. If $\mathfrak{n}_+$ denotes the space $\sum_{\alpha \in \Delta^+} \mathfrak{g}^\alpha$ considered as a real subspace of $\mathfrak{g}^R$ the following direct decomposition is valid*

$$\mathfrak{g}^R = \mathfrak{u} + J\mathfrak{a} + \mathfrak{n}_+.$$

Let G_c be any connected Lie group with Lie algebra $\mathfrak{g}^R$ and let U, A^*, and N_+ denote the analytic subgroups of G_c with Lie algebras $\mathfrak{u}$, $J\mathfrak{a}$, and $\mathfrak{n}_+$, respectively. Then the mapping

$$(u, a, n) \to uan, \qquad u \in U, a \in A^*, n \in N_+,$$

is an analytic diffeomorphism of $U \times A^* \times N_+$ onto G_c. The groups A^* and N_+ are simply connected.

EXERCISES

1. Let G be a connected semisimple Lie group whose Lie algebra has a complex structure. Show that G has finite center.

2. (See Chapter V, §2 for the notation.) The group $SO(p) \times SO(2)$ is a maximal compact subgroup of $SO_0(p, 2)$. Deduce from Theorem 1.1 that the universal covering group of $SO_0(p, 2)$ has infinite center ($p \geqslant 1$).

3. Let V be a finite-dimensional vector space over R and let G be a semisimple analytic subgroup of $GL(V)$. Then (Harish-Chandra [2]):

(i) G has finite center.

(ii) Every continuous homomorphism of G onto itself is an analytic isomorphism.

(iii) The semisimplicity assumption of G can neither be dropped in (i) nor (ii).

4. Let A be a set of isometries of a complete simply connected Riemannian manifold M of negative curvature. Then

(i) If not empty, the set of fixed points under A forms a connected totally geodesic submanifold of M (for a generalization, see Kobayashi [3]).

(ii) Suppose M is a Riemannian globally symmetric space of the non-compact type, $M = G/K$ $(G = I_0(M))$, and that A is a closed connected subgroup of K with Lie algebra $\mathfrak{a}$. Show that the fixed point set of A is $\operatorname{Exp} \mathfrak{b}$ where

$$\mathfrak{b} = \{X \in \mathfrak{p}_0 : \operatorname{Ad}_G(\exp(-X))\, \mathfrak{a} \subset \mathfrak{k}_0\}.$$

(iii) Show that the set $\mathfrak{b}$ in (ii) can be written

$$\mathfrak{b} = \{X \in \mathfrak{p}_0 : [\boldsymbol{R}X, \mathfrak{a}] = \{0\}\}.$$

5. Let M be a Riemannian globally symmetric space of the non-compact type, o any point in M. Let σ denote the automorphism $g \to s_0 g s_0$ of $I_0(M)$. Let $I_0(M)$ be given the pseudo-Riemannian structure induced by the Killing form. Then the mapping

$$p \to g\sigma(g^{-1}) \qquad (g \cdot o = p),$$

is a diffeomorphism of M onto a closed totally geodesic submanifold S of $I_0(M)$. Under this mapping the action of an element $x \in I_0(M)$ on M corresponds to the diffeomorphism $s \to xs\sigma(x^{-1})$ of S.

NOTES

§1-§2. Most of the results here are due to É. Cartan [10]; see also Mostow [1, 3]. Theorem 1.4 was first proved by Mostow in [3]; his proof differs somewhat from that of the text and as several results of the same paper it is based on the imbedding φ: $gK \to \operatorname{Ad}_G(g\sigma(g^{-1}))$ of G/K into the space P of positive definite symmetric matrices. The space P is itself Riemannian globally symmetric; in fact $P = \boldsymbol{GL}(n, \boldsymbol{R})/\boldsymbol{O}(n)$. For a further study and application of this imbedding see Satake [3]. É. Cartan's conjugacy theorem (Theorem 2.1) has been extended by Iwasawa [1] to all connected locally compact groups.

§3-§5. Theorem 3.4 and its global analog Theorem 5.1 are due to Iwasawa [1]; see also Harish-Chandra [4], p. 223. In these theorems the Cartan subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ is chosen such that $\mathfrak{h} \cap \mathfrak{p}$ has maximal dimension. The theorems have been extended by Harish-Chandra [10], p. 212, to other Cartan subalgebras; see also Satake [3], §3. Concerning Theorem 4.2 see Chen [1]. Lemma 4.6 and Theorem 4.7 are due to Harish-Chandra [5].

CHAPTER VII

SYMMETRIC SPACES OF THE COMPACT TYPE

In contrast to the situation for the noncompact type there may be several Riemannian globally symmetric spaces U/K associated with the same orthogonal symmetric Lie algebra $(\mathfrak{u}, \theta)$ of the compact type. These spaces are finite in number and have the same universal covering space. They can all be described by means of the center of the universal covering group of U (Theorem 8.1).

The entire chapter centers around the maximal abelian subspace $\mathfrak{h}_{\mathfrak{p}_*}$ together with the root system $\Delta_{\mathfrak{p}}$. These define the diagram $D(U,K)$ in which certain information about the space U/K is contained, for example, the location of its conjugate points. Elementary properties of the Weyl group $W(U,K)$ and the diagram are developed in §2-§3 but the classification of the spaces U/K requires description of the fundamental group of U in terms of $\mathfrak{u}$. Here the singular set in U interferes and a rigorous proof of the fact that the singular set has no influence on the fundamental group requires some tools from dimension theory, collected in §9. The theory in §2-§3 can be applied to U itself (considered as a symmetric space) and gives a method for reading some global properties of U from the diagram $D(U)$ together with the unit lattice $\mathfrak{t}_e$. The fundamental Theorem 7.2 which underlies the description of all the spaces U/K now results from comparison of the diagrams $D(U)$ and $D(U,K)$.

§ 1. The Contrast between the Compact Type and the Noncompact Type

Let $\mathfrak{u}$ be a compact semisimple Lie algebra and θ an involutive automorphism of $\mathfrak{u}$. Then θ extends uniquely to a (complex) involutive automorphism of $\mathfrak{g}$, the complexification of $\mathfrak{u}$. We denote this extension also by θ. Let $(\mathfrak{g}_0, s)$ denote the orthogonal symmetric Lie algebra which is dual to $(\mathfrak{u}, \theta)$. Then $\mathfrak{g}_0$ is a real form of $\mathfrak{g}$ and s is just the restriction of θ to $\mathfrak{g}_0$. As usual we adopt the notational conventions in §5, Chapter V. We have then the direct decompositions

$$\mathfrak{u} = \mathfrak{k}_0 + \mathfrak{p}_*,$$

$$\mathfrak{g}_0 = \mathfrak{k}_0 + \mathfrak{p}_0$$

into eigenspaces for θ. We recall that a pair (U, K) is said to be associated with $(\mathfrak{u}, \theta)$ if U is a connected Lie group with Lie algebra $\mathfrak{u}$ and K is a

Lie subgroup of U with Lie algebra $\mathfrak{k}_0$. Such a pair is said to be of the compact type. This pair is said to be a Riemannian symmetric pair if K is closed, $\mathrm{Ad}_U(K)$ compact and if there exists an analytic involutive automorphism $\tilde{\theta}$ of U such that $(K_{\tilde{\theta}})_0 \subset K \subset K_{\tilde{\theta}}$. Such a $\tilde{\theta}$ is necessarily unique and $d\tilde{\theta} = \theta$. Finally, a Riemannian globally symmetric space M is said to be of the compact type if the pair $(I_0(M), H)$ is of the compact type, H being the isotropy subgroup of $I_0(M)$ at some point $o \in M$.

Proposition 1.1. *Let $(\mathfrak{u}, \theta)$ be an orthogonal symmetric Lie algebra of the compact type. Let (U, K) be an arbitrary pair associated with $(\mathfrak{u}, \theta)$. Then K is compact and $Ad_U(K)\, \mathfrak{p}_* \subset \mathfrak{p}_*$. The restriction of $-B$ to $\mathfrak{p}_* \times \mathfrak{p}_*$ gives rise to a U-invariant Riemannian structure on U/K, which turns U/K into a Riemannian locally symmetric space.*

Proof. Let K_0 denote the identity component of K and $N(K_0)$ the normalizer of K_0 in U, that is, the set of $u \in U$ such that $uK_0u^{-1} \subset K_0$. The group K_0 is a closed subgroup of U (Prop. 3.6, Chapter IV); hence $N(K_0)$ is closed in U. The group U is compact (Theorem 6.9, Chapter II), so $N(K_0)$ is compact. The Lie algebra of $N(K_0)$ is the normalizer $\mathfrak{n}(\mathfrak{k}_0)$ of $\mathfrak{k}_0$ in $\mathfrak{u}$, that is, the set of elements $X \in \mathfrak{u}$ such that $[RX, \mathfrak{k}_0] \subset \mathfrak{k}_0$. If $X \in \mathfrak{n}(\mathfrak{k}_0) \cap \mathfrak{p}_*$, then $[RX, \mathfrak{k}_0] \subset \mathfrak{k}_0 \cap \mathfrak{p}_* = \{0\}$. Using Corollary 1.7, Chapter V, it follows that $X = 0$ so $\mathfrak{n}(\mathfrak{k}_0) = \mathfrak{k}_0$. The group $N(K_0)$ has finitely many components and the same is true of K since $K \subset N(K_0)$. Hence K is compact. The group $\mathrm{Ad}_U(K)$ leaves $\mathfrak{k}_0$, and therefore its orthogonal complement, $\mathfrak{p}_*$, invariant. The local symmetry of U/K is clear from Prop. 3.6, Chapter IV (even if K is not connected).

The question is now: When is U/K Riemannian globally symmetric? The answer will be given in the present chapter (Theorem 8.1). It is more complicated than the answer to the analogous question for the noncompact type, where G/K is always globally symmetric (Theorem 1.1, Chapter VI). To begin with we establish a negative result which indicates what kind of complications the compact type presents.

Proposition 1.2. *Let $(\mathfrak{u}, \theta)$ be an orthogonal symmetric Lie algebra of the compact type. Let (U, K) be an arbitrary pair associated with $(\mathfrak{u}, \theta)$. Then*

(i) *The center of U does not in general belong to K.*

(ii) *Even if U/K is Riemannian globally symmetric, K is not necessarily connected.*

(iii) *Even if U/K is Riemannian globally symmetric, the automorphism θ does not necessarily correspond to an automorphism of U.*

Proof. An example for (i) is given by† $U = SU(n)$ $(n \geqslant 3)$, $K = SO(n)$ with the involutive automorphism $u \to \bar{u}$ (complex conjugation) of U. For (ii) we consider the two-dimensional real projective space P^2, that is, S^2 with all antipodal points identified. Then $P^2 = U/K$ where $U = SO(3)$ and K is the subgroup of U leaving a line l through 0 invariant. The group K is generated by the rotations around l and the reflection in a line through 0, perpendicular to l. Here U/K is Riemannian globally symmetric and K has two components. For (iii) let $\tilde{U} = SU(2) \times SU(2)$ and let $\tilde{\theta}$ denote the automorphism $(g_1, g_2) \to (g_2, g_1)$ of $\tilde{U}$. The subgroup $\tilde{K}$ of fixed points is isomorphic to $SU(2)$. The center of $SU(2)$ is a cyclic group of order 2; let z be the generator and let S denote the subgroup of the center $\tilde{Z}$ of $\tilde{U}$ consisting of the two elements (e, e), (e, z). Let $U = \tilde{U}/S$ and let $K = \pi(\tilde{K})$, π denoting the natural mapping of $\tilde{U}$ onto U. The pair (U, K) is associated with $(\mathfrak{u}, \theta)$ where $\mathfrak{u}$ is the Lie algebra of $\tilde{U}$ and θ is the automorphism of $\mathfrak{u}$ induced by $\tilde{\theta}$. The homomorphism Ad_U has a kernel consisting of two elements (since $\tilde{Z}$ has four elements). These elements are S and $(z, z)\,S$ both of which belong to K. Hence $U/K = \mathrm{Ad}_U(U)/\mathrm{Ad}_U(K)$ and this last space is globally symmetric since the automorphism θ induces an automorphism of $\mathrm{Int}\,(\mathfrak{u}) = \mathrm{Ad}_U\,(U)$. On the other hand, since $\tilde{\theta}(S) \not\subset S$, the following lemma shows that θ does not correspond to an automorphism of U.

Lemma 1.3.§ *Let L be a connected Lie group and let (L^*, π) denote the universal covering group of L. Let Z^* denote the kernel of π. Let σ be any analytic automorphism of L^* and $d\sigma$ the corresponding automorphism of $\mathfrak{l}$, the Lie algebra of L. Then $d\sigma$ corresponds to an analytic automorphism of L if and only if $\sigma(Z^*) \subset Z^*$.*

Proof. If $\sigma(Z^*) \subset Z^*$, then σ induces the desired automorphism of $L^*/Z^* = L$. On the other hand, suppose λ is an automorphism of L such that $d\lambda = d\sigma$. Consider the mappings $\varphi = \pi \circ \sigma$ and $\psi = \lambda \circ \pi$ of L^* onto L. Then φ and ψ are homomorphism and $d\varphi = d\psi$. Consequently $\varphi = \psi$. Since Z^* is the kernel of π it follows that $\sigma(Z^*) \subset Z^*$ and the lemma is proved.

§ 2. The Weyl Group

Let $(\mathfrak{u}, \theta)$ be an orthogonal symmetric Lie algebra of the compact type and let (U, K) be any pair associated with $(\mathfrak{u}, \theta)$. The notation

† For the notation $SU(n)$ see Chapter IX, §4.

§ See Berger [2], p. 162.

of §1 will be preserved but for simplicity we shall now write Ad instead of Ad_U. For each $X \in \mathfrak{p}_*$, let T_X denote the restriction of $(\mathrm{ad}\, X)^2$ to $\mathfrak{p}_*$.

Let $\mathfrak{h}_{\mathfrak{p}_*}$ denote an arbitrary maximal abelian subspace of $\mathfrak{p}_*$. Then the space $\mathfrak{h}_{\mathfrak{p}_0} = i\mathfrak{h}_{\mathfrak{p}_*}$ is a maximal abelian subspace of $\mathfrak{p}_0$. Let $\mathfrak{h}_0$ be any maximal abelian subalgebra of $\mathfrak{g}_0$ containing $\mathfrak{h}_{\mathfrak{p}_0}$ and let $\mathfrak{h}$ denote the subalgebra of $\mathfrak{g}$ generated by $\mathfrak{h}_0$. Then $\mathfrak{h}$ is a Cartan subalgebra of $\mathfrak{g}$ (Lemma 3.2, Chapter VI), and as in §3, Chapter VI, we can form the subspaces $\mathfrak{h}_\mathfrak{p}$, $\mathfrak{h}_{\mathfrak{k}_0}$, $\mathfrak{h}_\mathfrak{k}$, and $\mathfrak{h}^*$. Let Δ denote the set of nonzero roots of $\mathfrak{g}$ with respect to $\mathfrak{h}$. Let $\Delta_\mathfrak{p}$ denote the set of roots in Δ which do not vanish identically on $\mathfrak{h}_\mathfrak{p}$. As in Chapter VI, §3, let Δ^+ denote the subset of Δ formed by the positive roots with respect to an ordering of Δ given by any compatible orderings in the dual spaces of $\mathfrak{h}_{\mathfrak{p}_0}$ and $\mathfrak{h}^*$, respectively. Let $P_+ = \Delta^+ \cap \Delta_\mathfrak{p}$. Finally, $\mathfrak{m}_0$ shall denote the centralizer of $\mathfrak{h}_{\mathfrak{p}_0}$ (or $\mathfrak{h}_{\mathfrak{p}_*}$) in $\mathfrak{k}_0$.

Let M and M', respectively, denote the centralizer and normalizer of $\mathfrak{h}_{\mathfrak{p}_*}$ in K. In other words,

$$M = \{k \in K : \mathrm{Ad}\,(k)\, H = H \text{ for each } H \in \mathfrak{h}_{\mathfrak{p}_*}\},$$

$$M' = \{k \in K : \mathrm{Ad}\,(k)\, \mathfrak{h}_{\mathfrak{p}_*} \subset \mathfrak{h}_{\mathfrak{p}_*}\}.$$

It is clear that M is a normal subgroup of M'.

Proposition 2.1. *The groups M and M' are compact and have the same Lie algebra, namely* $\mathfrak{m}_0$.

Proof. The groups M and M' are closed subgroups of K, hence compact. It is obvious that M has Lie algebra $\mathfrak{m}_0$. On the other hand, let Y belong to $\mathfrak{L}(M')$, the Lie algebra of M'. Then $[Y, H] \in \mathfrak{h}_{\mathfrak{p}_*}$ for each $H \in \mathfrak{h}_{\mathfrak{p}_*}$ so

$$B\,(\mathrm{ad}\, H(Y), \mathrm{ad}\, H(Y)) = -\, B((\mathrm{ad}\, H)^2\, Y, Y) = 0.$$

Hence $[H, Y] = 0$ for each $H \in \mathfrak{h}_{\mathfrak{p}_*}$ so $Y \in \mathfrak{m}_0$.

Definition. The factor group M'/M is called the *Weyl group* of the pair (U, K). It is denoted by $W(U, K)$ (or simply W).

Remark. It is clear from Prop. 2.1 that $W(U, K)$ is a finite group. The mapping $k \to \mathrm{Ad}\,(k)$ $(k \in M')$ induces an isomorphism of $W(U, K)$ into $GL(\mathfrak{h}_{\mathfrak{p}_*})$. We can therefore regard $W(U, K)$ as a group of (complex) linear transformations of $\mathfrak{h}_\mathfrak{p}$. It will be shown later that for a fixed $\mathfrak{h}_{\mathfrak{p}_*}$, $W(U, K)$ only depends on $(\mathfrak{u}, \theta)$. On the other hand, it is clear from Lemma 6.3, Chapter V, that different choices of $\mathfrak{h}_{\mathfrak{p}_*}$ lead to isomorphic Weyl groups.

Proposition 2.2. *Let $\mathfrak{a}$ be a subset of $\mathfrak{h}_{\mathfrak{p}_*}$ and suppose k is an element of K such that $Ad(k)\,\mathfrak{a} \subset \mathfrak{h}_{\mathfrak{p}_*}$. Then there exists an element $s \in W(U, K)$ such that $s \cdot A = Ad(k)\,A$ for each $A \in \mathfrak{a}$.*

Proof. The centralizer $Z_{\mathfrak{a}}$ of $\mathfrak{a}$ in U is a closed subgroup of U. Its Lie algebra is $\mathfrak{z}_{\mathfrak{a}}$, the centralizer of $\mathfrak{a}$ in $\mathfrak{u}$. Since $\mathfrak{z}_{\mathfrak{a}}$ is invariant under θ we have the direct decomposition

$$\mathfrak{z}_{\mathfrak{a}} = \mathfrak{z}_{\mathfrak{a}} \cap \mathfrak{k}_0 + \mathfrak{z}_{\mathfrak{a}} \cap \mathfrak{p}_*.$$

The spaces $\mathfrak{h}_{\mathfrak{p}_*}$ and $\mathrm{Ad}\,(k^{-1})\,\mathfrak{h}_{\mathfrak{p}_*}$ are maximal abelian subspaces of $\mathfrak{z}_{\mathfrak{a}} \cap \mathfrak{p}_*$. In view of Lemma 6.3, Chapter V, there exists an element $H \in \mathfrak{h}_{\mathfrak{p}_*}$ whose centralizer in $\mathfrak{p}_*$ coincides with $\mathfrak{h}_{\mathfrak{p}_*}$. Let X be any fixed element in $\mathfrak{z}_{\mathfrak{a}} \cap \mathfrak{p}_*$. The function $z \to B(H, \mathrm{Ad}\,(z)\,X)$, $(z \in Z_{\mathfrak{a}} \cap K)$, is real and attains its minimum, the group $Z_{\mathfrak{a}} \cap K$ being compact. If z_0 is a minimum point, we have

$$\left\{\frac{d}{dt} B(H, \mathrm{Ad}\,(\exp tT)\,\mathrm{Ad}\,(z_0)\,X)\right\}_{t=0} = 0$$

for each $T \in \mathfrak{z}_{\mathfrak{a}} \cap \mathfrak{k}_0$. It follows that

$$B(H, [T, \mathrm{Ad}\,(z_0)\,X]) = -B([H, \mathrm{Ad}\,(z_0)\,X], T) = 0 \tag{1}$$

for each $T \in \mathfrak{z}_{\mathfrak{a}} \cap \mathfrak{k}_0$. Since $[H, \mathrm{Ad}\,(z_0)\,X] \in \mathfrak{z}_{\mathfrak{a}} \cap \mathfrak{k}_0$ we conclude from (1) that $[H, \mathrm{Ad}\,(z_0)\,X] = 0$, so, due to the choice of H, $\mathrm{Ad}\,(z_0)\,X \in \mathfrak{h}_{\mathfrak{p}_*}$. In particular, let $X = H'$ where H' is an element in $\mathrm{Ad}\,(k^{-1})\,\mathfrak{h}_{\mathfrak{p}_*}$ whose centralizer in $\mathfrak{p}_*$ is $\mathrm{Ad}\,(k^{-1})\,\mathfrak{h}_{\mathfrak{p}_*}$. Then from the above follows that $H' \in \mathrm{Ad}\,(z_0^{-1})\,\mathfrak{h}_{\mathfrak{p}_*}$ so

$$\mathrm{Ad}\,(z_0^{-1})\,\mathfrak{h}_{\mathfrak{p}_*} = \mathrm{Ad}\,(k^{-1})\,\mathfrak{h}_{\mathfrak{p}_*}.$$

Consequently $kz_0^{-1} \in M'$. Since $z_0 \in Z_{\mathfrak{a}} \cap K$, the restriction of $\mathrm{Ad}\,(kz_0^{-1})$ to $\mathfrak{h}_{\mathfrak{p}_*}$ is the desired element $s \in W(U, K)$.

The Killing form B is nondegenerate on $\mathfrak{h}_{\mathfrak{p}} \times \mathfrak{h}_{\mathfrak{p}}$. For each $\alpha \in \Delta_{\mathfrak{p}}$ there exists a unique vector† $H_\alpha \in \mathfrak{h}_{\mathfrak{p}}$ such that $B(H, H_\alpha) = \alpha(H)$ for all $H \in \mathfrak{h}_{\mathfrak{p}}$. Since α is real on $\mathfrak{h}_{\mathfrak{p}_0}$ it follows that $H_\alpha \in \mathfrak{h}_{\mathfrak{p}_0}$.

Lemma 2.3. *For each $\alpha \in \Delta_{\mathfrak{p}}$ there exist nonzero vectors $Y_\alpha \in \mathfrak{p}_0$, $Z_\alpha \in \mathfrak{k}_0$ such that*

$$[H, Y_\alpha] = \alpha(H)\,Z_\alpha,$$

$$[H, Z_\alpha] = \alpha(H)\,Y_\alpha$$

for all $H \in \mathfrak{h}_{\mathfrak{p}}$.

† This vector is of course in general different from the vector introduced in Theorem 4.2, Chapter III, under the same name.

Proof. There exists a vector $X_\alpha \neq 0$ in $\mathfrak{g}$ such that $[H, X_\alpha] = \alpha(H) X_\alpha$ for all $H \in \mathfrak{h}$. Writing $X_\alpha = X_1 + iX_2$ where $X_1, X_2 \in \mathfrak{g}_0$, and noting that α is real on $\mathfrak{h}_{\mathfrak{p}_0}$, we obtain $[H, X_i] = \alpha(H) X_i$ for $H \in \mathfrak{h}_{\mathfrak{p}}$, $i = 1, 2$. At least one of the vectors X_i is nonzero and can be decomposed $Y_\alpha + Z_\alpha$ where $Y_\alpha \in \mathfrak{p}_0$, $Z_\alpha \in \mathfrak{k}_0$. Equating the components in $\mathfrak{p}_0$ and $\mathfrak{k}_0$, respectively, we get

$$[H, Y_\alpha] = \alpha(H) Z_\alpha$$
$$[H, Z_\alpha] = \alpha(H) Y_\alpha$$

for $H \in \mathfrak{h}_{\mathfrak{p}}$. At least one of the vectors Y_α, Z_α is $\neq 0$; it follows that both are $\neq 0$, since α does not vanish identically on $\mathfrak{h}_{\mathfrak{p}}$.

The form $-B$ induces a positive definite quadratic form on $\mathfrak{h}_{\mathfrak{p}_*}$. For each $\alpha \in \Delta_{\mathfrak{p}}$, let s_α denote the reflection of $\mathfrak{h}_{\mathfrak{p}_*}$ in the hyperplane $\alpha(H) = 0$ of $\mathfrak{h}_{\mathfrak{p}_*}$. Then s_α extends uniquely to a complex linear transformation of $\mathfrak{h}_{\mathfrak{p}}$ and

$$s_\alpha(H) = H - 2\frac{\alpha(H)}{\alpha(H_\alpha)} H_\alpha, \qquad H \in \mathfrak{h}_{\mathfrak{p}}.$$

Lemma 2.4. *Let $\alpha \in \Delta_{\mathfrak{p}}$. Then $s_\alpha \in W(U, K)$.*

Proof. Consider the vectors Y_α, Z_α from Lemma 2.3. We can assume, changing $Y_\alpha + Z_\alpha$ by a real factor if necessary, that $B(Z_\alpha, Z_\alpha) = -1$. Then, if $H \in \mathfrak{h}_{\mathfrak{p}}$,

$$B(H, [Z_\alpha, Y_\alpha]) = -B([H, Y_\alpha], Z_\alpha) = \alpha(H),$$
$$[H, [Z_\alpha, Y_\alpha]] = -[Z_\alpha, [Y_\alpha, H]] - [Y_\alpha, [H, Z_\alpha]] = 0.$$

The last relation implies that $[Z_\alpha, Y_\alpha] \in \mathfrak{h}_{\mathfrak{p}}$ and the first then shows that $[Z_\alpha, Y_\alpha] = H_\alpha$. It follows that

$$\operatorname{Ad}(\exp tZ_\alpha) H_\alpha = \sum_0^\infty \frac{1}{(2n)!} (t \operatorname{ad} Z_\alpha)^{2n} (H_\alpha) + \sum_0^\infty \frac{1}{(2n+1)!} (t \operatorname{ad} Z_\alpha)^{2n+1} (H_\alpha)$$
$$= \sum_0^\infty \frac{1}{(2n)!} t^{2n} (-\alpha(H_\alpha))^n H_\alpha + \sum_0^\infty \frac{1}{(2n+1)!} t^{2n+1} (-\alpha(H_\alpha))^{n+1} Y_\alpha.$$

Since $\alpha(H_\alpha) = B(H_\alpha, H_\alpha) > 0$, there exists a number $t_0 \in \boldsymbol{R}$ such that $t_0 \sqrt{\alpha(H_\alpha)} = \pi$. Putting $k_0 = \exp t_0 Z_\alpha$ we obtain

$$\operatorname{Ad}(k_0) H_\alpha = -H_\alpha.$$

Moreover, since $[Z_\alpha, H] = 0$ if $\alpha(H) = 0$, the hyperplane $\alpha(H) = 0$ in $\mathfrak{h}_{\mathfrak{p}_*}$ is left pointwise fixed by $\operatorname{Ad}(k_0)$. Hence s_α is the restriction of $\operatorname{Ad}(k_0)$ to $\mathfrak{h}_{\mathfrak{p}_*}$. This proves the lemma.

Now we need some facts about toral subgroups (i.e., compact, abelian, connected subgroups) of compact Lie groups.

Theorem 2.5. *Let S be a compact connected Lie group and T a toral subgroup of S. Suppose a is an element in S which commutes with each member of T. Then there exists a torus $T' \subset S$ such that $T \subset T'$ and $a \in T'$.*

For the proof we need a lemma concerning monothetic groups. A topological group S is called *monothetic* if there exists an element $x \in S$ such that the sequence $e, x, x^2, \ldots,$ is dense in S. In this case, the element x is called a generator. Any torus is monothetic due to the classical theorem of Kronecker.

Lemma 2.6. *Let A be a compact abelian Lie group such that A/A_0 is cyclic, A_0 denoting the identity component of A. Then A is monothetic.*

Proof. The group A_0 is a torus, hence monothetic. Let a_0 be a generator for A_0 and let N denote the number of elements in A/A_0. Select a generator B for A/A_0 and an element b in the coset B. Then $b^N \in A_0$ and there exists an element $c \in A_0$ such that $b^N c^N = a_0$. Then bc is a generator of A.

In order to prove Theorem 2.5, let A denote the closed subgroup of S generated by T and a. The identity component A_0 of A is a torus containing T and the group $\bigcup_{n \in Z} A_0 a^n$ equals A. Since A is compact, some positive power of a lies in A_0. If N is the smallest such power of a, the group A/A_0 is cyclic of order N. By Lemma 2.6, A is monothetic. If b is a generator of A, let $\exp tX$ ($t \in \mathbf{R}$) be a one-parameter subgroup γ of S passing through b. The closure of γ in S is a torus containing a and T.

Corollary 2.7. *A maximal torus in a compact, connected Lie group is a maximal abelian subgroup.*

Corollary 2.8. *Let T be a toral subgroup of a compact connected Lie group S. The centralizer of T in S is connected.*

In fact, it is the union of the tori containing T.

Each root $\alpha \in \Delta_{\mathfrak{p}}$ defines a hyperplane $\alpha(H) = 0$ in the vector space $\mathfrak{h}_{\mathfrak{p}_*}$. These hyperplanes divide the space $\mathfrak{h}_{\mathfrak{p}_*}$ into finitely many connected components, called the *Weyl chambers*. These are open, convex subsets of $\mathfrak{h}_{\mathfrak{p}_*}$.

Lemma 2.9. *Let $H \in \mathfrak{h}_{\mathfrak{p}_*}$. Then the eigenvalues of T_H are:* (1) 0 *with multiplicity dim* $\mathfrak{h}_{\mathfrak{p}_*}$; (2) *the numbers* $\alpha(H)^2$ *as* α *runs through* P_+ (*with the right multiplicity*).

Proof. As proved in Chapter VI (Lemma 3.6), the space $\mathfrak{p}$ has a direct decomposition

$$\mathfrak{p} = \mathfrak{h}_{\mathfrak{p}} + \mathfrak{q} \qquad \text{where } \mathfrak{q} = \sum_{\alpha \in P_+} C(X_\alpha - \theta X_\alpha).$$

Now $[H, X_\alpha] = \alpha(H)\, X_\alpha$, $[H, \theta X_\alpha] = -\alpha(H)\, \theta X_\alpha$, so

$$(\text{ad}\, H)^2 (X_\alpha - \theta X_\alpha) = \alpha(H)^2 (X_\alpha - \theta X_\alpha).$$

Let V be the eigenspace of $(\text{ad}\, H)^2$ in $\mathfrak{p}$ for the eigenvalue $\alpha(H)^2$. Since $\alpha(H)^2$ is real, we have $V = V \cap \mathfrak{p}_0 + V \cap \mathfrak{p}_*$ and the complex dimension of V equals the real dimension of $V \cap \mathfrak{p}_*$. The lemma now follows immediately. Lemma 2.3 gives another proof.

The maximal abelian subspace $\mathfrak{h}_{\mathfrak{p}}$ of $\mathfrak{p}$ can in general be extended to a Cartan subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ in many different ways. However, the restrictions of the roots to $\mathfrak{h}_{\mathfrak{p}}$ do not depend on the choice of $\mathfrak{h}$, as Lemma 2.9 shows. More precisely, we have

Corollary 2.10. *Let λ be a real linear function on $\mathfrak{h}_{\mathfrak{p}_0}$. Then λ is the restriction of a root* (*of* $\mathfrak{g}$ *with respect to* $\mathfrak{h}$) *if and only if there exists a vector $X \neq 0$ in $\mathfrak{p}_0$ such that*

$$(\text{ad}\, H)^2\, X = \lambda(H)^2\, X \qquad \text{for } H \in \mathfrak{h}_{\mathfrak{p}_0}.$$

As an immediate consequence, we obtain

Corollary 2.11. *Let $k \in M'$ and for each real linear function λ on $\mathfrak{h}_{\mathfrak{p}_0}$, put $\lambda^k(H) = \lambda(Ad\,(k^{-1})\, H)$, $H \in \mathfrak{h}_{\mathfrak{p}_0}$. Then λ is the restriction of a root* (*of* $\mathfrak{g}$ *with respect to* $\mathfrak{h}$) *if and only if λ^k is the restriction of a root.*

Theorem 2.12. *Each $s \in W(U, K)$ permutes the Weyl chambers. The Weyl group is simply transitive on the set of Weyl chambers in $\mathfrak{h}_{\mathfrak{p}_*}$.*

Proof. Select $k \in M'$ such that s coincides with the restriction of Ad (k) to $\mathfrak{h}_{\mathfrak{p}_*}$. It follows from Cor. 2.11 that if some root in $\Delta_{\mathfrak{p}}$ vanishes at a point $H \in \mathfrak{h}_{\mathfrak{p}_*}$ then some root in $\Delta_{\mathfrak{p}}$ vanishes at Ad (k) H. Consequently, s permutes the Weyl chambers. Next we show that $W(U, K)$ is transitive. Let W' denote the subgroup of $W(U, K)$ generated by all s_α, $\alpha \in \Delta_{\mathfrak{p}}$. Let C_1 and C_2 be two arbitrary Weyl chambers and select $H_1 \in C_1$, $H_2 \in C_2$. If the segment $\overrightarrow{H_1 H_2}$ intersects a hyperplane $\alpha(H) = 0$, then it is clear that

$$| H_1 - H_2 | > | H_1 - s_\alpha \cdot H_2 |,$$

the norm in $\mathfrak{h}_{\mathfrak{p}_*}$ being denoted by $|\ |$. As s varies over the finite group W', the distance $|H_1 - s \cdot H_2|$ reaches a minimum, say for $s = s_0$. Then the segment from H_1 to $s_0 \cdot H_2$ intersects no hyperplane $\alpha(H) = 0$ and H_1 and $s_0\ H_2$ lie in the same Weyl chamber. Hence $C_1 = s_0 C_2$ so the group W', and therefore $W(U, K)$ is transitive.

Suppose now that an element $s \in W(U, K)$ maps a chamber C into itself. Select any $H_0 \in C$ and let $H = N^{-1}(H_0 + sH_0 + \ldots + s^{N-1}H_0)$, where N is the order of s. Then $sH = H$ and, since C is convex, $H \in C$. In view of Cor. 3.7, Chapter VI, the centralizer $\mathfrak{z}_H$ of H in $\mathfrak{u}$ coincides with the centralizer of $\mathfrak{h}_{\mathfrak{p}_*}$ in $\mathfrak{u}$. Moreover, the centralizer Z_γ in U of .he one-parameter subgroup $\gamma = \{\exp tH : t \in \boldsymbol{R}\}$ has Lie algebra $\mathfrak{z}_H$. The closure of γ in U is a torus, so by Cor. 2.8, Z_γ is connected. Select $k \in K$ such that s coincides with the restriction of $\operatorname{Ad}(k)$ to $\mathfrak{h}_{\mathfrak{p}_*}$. Then $\operatorname{Ad}(k)\, tH = tH$ for all $t \in \boldsymbol{R}$ so $k \in Z_\gamma$. Since Z_γ is generated by $\exp(\mathfrak{z}_H)$, it follows that the restriction of $\operatorname{Ad}(k)$ to $\mathfrak{h}_{\mathfrak{p}_*}$, that is, s, is the identity. This proves that $W(U, K)$ is simply transitive.

Corollary 2.13. *The Weyl group is generated by the reflections s_α, $\alpha \in \Delta_{\mathfrak{p}}$. Thus, for a fixed $\mathfrak{h}_{\mathfrak{p}_*}$, $W(U, K)$ depends only on* $(\mathfrak{u}, \theta)$.

In view of this corollary we shall often refer to $W(U, K)$ as the Weyl group of $(\mathfrak{u}, \theta)$ and denote it by $W(\mathfrak{u}, \theta)$.

Lemma 2.14. *Let $\mathfrak{a}$ be a subspace of $\mathfrak{h}_{\mathfrak{p}_*}$ and let $P_{\mathfrak{a}}$ denote the set of roots in P_+ which vanish identically on $\mathfrak{a}$. Let $\tilde{\mathfrak{a}}$ denote the subset of $\mathfrak{h}_{\mathfrak{p}_*}$ consisting of the points where all the roots in $P_{\mathfrak{a}}$ vanish. Then the centralizers $Z_{\mathfrak{a}}$ and $Z_{\tilde{\mathfrak{a}}}$ of $\mathfrak{a}$ and $\tilde{\mathfrak{a}}$ in U coincide, i.e.,*

$$Z_{\mathfrak{a}} = Z_{\tilde{\mathfrak{a}}}.$$

Proof. These centralizers are connected, due to Cor. 2.8. Therefore we only have to prove that their Lie algebras are the same. Each element $X \in \mathfrak{g}$ can be written

$$X = H_0 + \sum_{\alpha \in \Delta} a_\alpha X_\alpha \qquad (a_\alpha \in \boldsymbol{C}),$$

where $X_\alpha \in \mathfrak{g}^\alpha$, $H_0 \in \mathfrak{h}$. Then X commutes with each element in $\mathfrak{a}$ if and only if $a_\alpha \alpha(H) = 0$ for each $H \in \mathfrak{a}$ and each $\alpha \in \Delta$. Since α vanishes on $\mathfrak{a}$ if and only if it vanishes on $\tilde{\mathfrak{a}}$, it follows that $\mathfrak{a}$ and $\tilde{\mathfrak{a}}$ have the same centralizers in $\mathfrak{g}$ and also in $\mathfrak{u}$. This proves the lemma.

Theorem 2.15. *Let $\mathfrak{a}$ be a subspace of $\mathfrak{h}_{\mathfrak{p}_*}$ and let $W_{\mathfrak{a}}$ denote the group of elements in $W(\mathfrak{u}, \theta)$ which leave $\mathfrak{a}$ pointwise fixed. Then $W_{\mathfrak{a}}$ is generated by those reflections s_α $(\alpha \in \Delta_{\mathfrak{p}})$ for which α vanishes identically on $\mathfrak{a}$.*

Proof. Each element in $W_{\mathfrak{a}}$ leaves the set $\tilde{\mathfrak{a}}$ from Lemma 2.14 pointwise fixed. We can therefore assume, without restriction of generality, that $\mathfrak{a} = \tilde{\mathfrak{a}}$. Let $\mathfrak{z}_{\mathfrak{a}}$ and $Z_{\mathfrak{a}}$ denote the centralizers of $\mathfrak{a}$ in $\mathfrak{u}$ and U, respectively. Then $\mathfrak{z}_{\mathfrak{a}}$ is a compact Lie algebra invariant under θ so $\mathfrak{z}_{\mathfrak{a}} = \mathfrak{k}_0 \cap \mathfrak{z}_{\mathfrak{a}} + \mathfrak{p}_* \cap \mathfrak{z}_{\mathfrak{a}}$. Let $\mathfrak{c}$ denote the center of $\mathfrak{z}_{\mathfrak{a}}$. Then $\mathfrak{c} \cap \mathfrak{p}_* = \mathfrak{a}$. Let $\mathfrak{k}_1$ and $\mathfrak{p}_1$ denote the orthogonal complements of $\mathfrak{k}_0 \cap \mathfrak{c}$ in $\mathfrak{k}_0 \cap \mathfrak{z}_{\mathfrak{a}}$ and of $\mathfrak{a}$ in $\mathfrak{p}_* \cap \mathfrak{z}_{\mathfrak{a}}$. We put $\mathfrak{u}_1 = \mathfrak{k}_1 + \mathfrak{p}_1$ and let $\mathfrak{a}_1$ denote the orthogonal complement of $\mathfrak{a}$ in $\mathfrak{h}_{\mathfrak{p}_*}$. Since

$$\mathfrak{p}_1 = \mathfrak{a}_1 + \mathfrak{p}_* \cap \sum_{\alpha \in P_{\mathfrak{a}}} C(X_\alpha - \theta X_\alpha),$$

$\mathfrak{a}_1$ is a maximal abelian subspace of $\mathfrak{p}_1$. Also, the pair $(\mathfrak{u}_1, \theta)$ is an orthogonal symmetric Lie algebra of the compact type. Its Weyl group is generated by the reflections $s_{\bar{\alpha}}$ of $\mathfrak{a}_1$, if $\bar{\alpha}$ denotes the restriction of $\alpha \in P_{\mathfrak{a}}$ to $\mathfrak{a}_1$. If $s_{\bar{\alpha}}$ is extended to $\mathfrak{h}_{\mathfrak{p}_*}$ by defining it to be the identity mapping on $\mathfrak{a}$, then we obtain a member of $W_{\mathfrak{a}}$. It remains to be proved that the elements in $W_{\mathfrak{a}}$ thus obtained generate the whole of $W_{\mathfrak{a}}$.

Let $\tilde{Z}$ denote the universal covering group of $Z_{\mathfrak{a}}$. If $k \in Z_{\mathfrak{a}} \cap M'$ and $\tilde{k}$ is an element in $\tilde{Z}$ over k, then $\mathrm{Ad}\,(k)$ and $\mathrm{Ad}_{\tilde{Z}}\,(\tilde{k})$ agree on $\mathfrak{a}_1$. The group $\tilde{Z}$ decomposes according to the direct decomposition $\mathfrak{z}_{\mathfrak{a}} = \mathfrak{c} + \mathfrak{u}_1$

$$\tilde{Z} = C \times U_1,$$

where the groups C and U_1 have Lie algebras $\mathfrak{c}$ and $\mathfrak{u}_1$, respectively. If $\tilde{k}$ decomposes accordingly, $\tilde{k} = (c, k_1)$, then $\mathrm{Ad}_{\tilde{Z}}\,(\tilde{k})$ and $\mathrm{Ad}_{U_1}\,(k_1)$ agree on $\mathfrak{a}_1$. Since $k \in K \cap Z_{\mathfrak{a}}$, k_1 lies in a Lie subgroup of U_1 with Lie algebra $\mathfrak{k}_1$. Thus $\mathrm{Ad}_{U_1}\,(k_1)$ restricted to $\mathfrak{a}_1$ coincides with an element of the Weyl group of $(\mathfrak{u}_1, \theta)$. The same is therefore true of $\mathrm{Ad}\,(k)$ if $k \in Z_{\mathfrak{a}} \cap M'$. This proves the theorem.

§3. Conjugate Points. Singular Points. The Diagram

Again, let $(\mathfrak{u}, \theta)$ be an orthogonal symmetric Lie algebra of the compact type and let (U, K) be any pair associated with $(\mathfrak{u}, \theta)$. The notation of the preceding section will be preserved.

The manifold U/K is a Riemannian locally symmetric space whose tangent space at o (the point K in U/K) is identified with $\mathfrak{p}_*$. Let $X \in \mathfrak{p}_*$. The formula for $d\,\mathrm{Exp}_X$ (Theorem 4.1, Chapter IV) is clearly valid here, even if K is not necessarily connected. By this formula, X is conjugate to o if and only if

$$\det\left(\sum_0^\infty \frac{1}{(2n+1)!}\,(T_X)^n\right) = 0. \tag{1}$$

According to Lemma 6.3, Chapter V, each $X \in \mathfrak{p}_*$ can be expressed $X = \mathrm{Ad}\,(k)\,H$ where $k \in K$, $H \in \mathfrak{h}_{\mathfrak{p}_*}$. Since $T_{\mathrm{Ad}(k)H} = \mathrm{Ad}\,(k) \circ T_H \circ \mathrm{Ad}(k^{-1})$, we obtain from Lemma 2.9,

$$\det\Big(\sum_0^\infty \frac{1}{(2n+1)!}(T_X)^n\Big) = \prod_{\alpha \in P_+} \frac{\sin \alpha(iH)}{\alpha(iH)}. \tag{2}$$

From this formula follows:

Proposition 3.1. *The point $X = Ad\,(k)\,H$ is conjugate to o if and only if $\alpha(H) \in \pi i(Z - 0)$ for some $\alpha \in \Delta_{\mathfrak{p}}$.*

Consider now the coset space K/M, where M as before denotes the centralizer of $\mathfrak{h}_{\mathfrak{p}_*}$ in K. The mapping

$$\Phi : (kM, H) \to \mathrm{Exp}\,\mathrm{Ad}\,(k)\,H, \qquad k \in K,\ H \in \mathfrak{h}_{\mathfrak{p}_*},$$

is a differentiable mapping of $K/M \times \mathfrak{h}_{\mathfrak{p}_*}$ onto U/K. The mapping Φ can be decomposed $\Phi = \mathrm{Exp} \circ \beta$ where β is the mapping of $K/M \times \mathfrak{h}_{\mathfrak{p}_*}$ onto $\mathfrak{p}_*$ given by

$$\beta : (kM, H) \to \mathrm{Ad}\,(k)\,H, \qquad k \in K,\ H \in \mathfrak{h}_{\mathfrak{p}_*}.$$

As usual $\tau(x)$ ($x \in K$) denotes the mapping $kM \to xkM$ of K/M onto itself. As in Chapter VI, let $\mathfrak{l}_0$ denote the orthogonal complement of $\mathfrak{m}_0$ in $\mathfrak{k}_0$. According to Lemma 3.6, Chapter VI, the subspace $\mathfrak{l}$ of $\mathfrak{g}$ generated by $\mathfrak{l}_0$ is given by

$$\mathfrak{l} = \sum_{\alpha \in P_+} C(X_\alpha + \theta X_\alpha). \tag{3}$$

The natural mapping of K onto K/M induces an isomorphism of $\mathfrak{l}_0$ onto the tangent space to K/M at $\{M\}$. We shall therefore denote this tangent space also by $\mathfrak{l}_0$. As usual (Remark in §2, No. 1, Chapter I), a finite-dimensional vector space will be identified with its tangent space at each point.

Let (k_0M, H_0) be an arbitrary point in $K/M \times \mathfrak{h}_{\mathfrak{p}_*}$. If L runs through $\mathfrak{l}_0$ and H runs through $\mathfrak{h}_{\mathfrak{p}_*}$, then $(d\tau(k_0) \cdot L, H)$ runs through the tangent space to $K/M \times \mathfrak{h}_{\mathfrak{p}_*}$ at (k_0M, H_0). Since

$$\beta(k_0\,(\exp tL)\,M, H_0) = \mathrm{Ad}\,(k_0)\,\mathrm{Ad}\,(\exp tL)\,H_0,$$

$$\beta(k_0, H_0 + tH) = \mathrm{Ad}\,(k_0)\,(H_0 + tH),$$

we find

$$d\beta_{(k_0M, H_0)}(d\tau(k_0)\,L, H) = \mathrm{Ad}\,(k_0)\,([L, H_0] + H). \tag{4}$$

Since

$$B([L, H_0], H) = B(L, [H_0, H]) = 0,$$

it is clear that the right-hand side of (4) vanishes if and only if $[L, H_0] = H = 0$. Using (3) we may write

$$L = \sum_{\alpha \in P_+} l_\alpha(X_\alpha + \theta X_\alpha), \qquad l_\alpha \in \boldsymbol{C},$$

so

$$[H_0, L] = \sum_{\alpha \in P_+} l_\alpha \alpha(H_0)\,(X_\alpha - \theta X_\alpha).$$

Consequently, the mapping β is regular at (k_0M, H_0) if and only if $\alpha(H_0) \neq 0$ for all $\alpha \in \Delta_{\mathfrak{p}}$. Combining this result with Prop. 3.1, we can state:

Proposition 3.2. *The mapping Φ is regular at the point (k_0M, H_0) if and only if $\alpha(iH_0)/\pi$ is not an integer for any $\alpha \in \Delta_{\mathfrak{p}}$.*

Definition. The set

$$\{H \in \mathfrak{h}_{\mathfrak{p}_*} : \alpha(H) \in \pi i \boldsymbol{Z} \qquad \text{for some } \alpha \in \Delta_{\mathfrak{p}}\}$$

is called the *diagram* of the pair (U, K). It will be denoted by $D(U, K)$ or $D(\mathfrak{u}, \theta)$.

The diagram is therefore the union of finitely many families of equispaced hyperplanes. The complement $\mathfrak{h}_{\mathfrak{p}_*} - D(U, K)$ will be denoted by $(\mathfrak{h}_{\mathfrak{p}_*})_r$. It is obvious from Prop. 3.2 that $D(U, K)$ is invariant under each $s \in W(U, K)$.

Definition. The set $S_{U/K} = \Phi(K/M \times D(U, K))$ is called the *singular set* in U/K. The complement $U/K - S_{U/K}$ will be denoted $(U/K)_r$.

The topological dimension (see §9 in this chapter) of a subset S of a separable metric space will be denoted by dim S. This notation is permissible since the dimension of a separable C^∞ manifold coincides with its topological dimension (see Hurewicz-Wallman [1], Chapter IV).

For each $\alpha \in \Delta_{\mathfrak{p}}$ we put

$$\mathfrak{h}_\alpha = \{H \in \mathfrak{h}_{\mathfrak{p}_*} : \alpha(H) \in \pi i \boldsymbol{Z}\},$$

$$M_\alpha = \{k \in K : \operatorname{Exp} \operatorname{Ad}(k)\, H = \operatorname{Exp} H \text{ for all } H \in \mathfrak{h}_\alpha\}.$$

It is obvious that M_α is a closed subgroup of U containing M. We shall now prove the following stronger statement:

$$\dim M_\alpha > \dim M. \tag{5}$$

In fact, consider the vector $Z_\alpha \in \mathfrak{k}_0$ from Lemma 2.3. This lemma implies that

$$\mathrm{Ad}\,(\exp H)\, Z_\alpha = \cos \alpha(iH)\, Z_\alpha - i \sin \alpha(iH)\, Y_\alpha \qquad \text{for } H \in \mathfrak{h}_{\mathfrak{p}_*}.$$

This implies that

$$\exp \mathrm{Ad}\,(\exp tZ_\alpha)\, H = \exp H, \qquad \text{if } \cos \alpha(iH) = +1,$$

$$\exp \mathrm{Ad}\,(\exp tZ_\alpha)\, H = \exp H \exp(-2tZ_\alpha), \qquad \text{if } \cos \alpha(iH) = -1$$

for all $t \in \boldsymbol{R}$. In any case, we have

$$\mathrm{Exp}\, \mathrm{Ad}\,(\exp tZ_\alpha)\, H = \mathrm{Exp}\, H$$

for all $t \in \boldsymbol{R}$, if $H \in \mathfrak{h}_\alpha$. This means that $\exp tZ_\alpha \in M_\alpha$ for all t so Z_α belongs to the Lie algebra of M_α. Since $Z_\alpha \notin \mathfrak{m}_0$, relation (5) follows.

Theorem 3.3. *Let (U, K) be any pair associated with $(\mathfrak{u}, \theta)$. Then the singular set is closed and*

$$\dim S_{U/K} \leqslant \dim U/K - 2.$$

The same statement holds for the set of points in U/K which are conjugate to o.

Proof. *Let* $\alpha \in \Delta_{\mathfrak{p}}$. Consider the mapping Φ_α of $K/M_\alpha \times \mathfrak{h}_\alpha$ into U/K given by

$$\Phi_\alpha(kM_\alpha, H) = \mathrm{Exp}\,\mathrm{Ad}\,(k)\, H.$$

Then

$$\Phi_\alpha(K/M_\alpha \times \mathfrak{h}_\alpha) = \Phi(K/M \times \mathfrak{h}_\alpha). \tag{6}$$

Using Lemma 3.6, Chapter VI and relation (5) above we find

$$\dim K/M = \dim \mathfrak{p}_* - \dim \mathfrak{h}_{\mathfrak{p}_*},$$

$$\begin{aligned}\dim (K/M_\alpha \times \mathfrak{h}_\alpha) &\leqslant \dim \mathfrak{p}_* - \dim \mathfrak{h}_{\mathfrak{p}_*} - 1 + (\dim \mathfrak{h}_{\mathfrak{p}_*} - 1) \\ &= \dim \mathfrak{p}_* - 2.\end{aligned}$$

Using (6) and Lemma 9.3 in the Appendix we conclude

$$\dim \Phi(K/M \times \mathfrak{h}_\alpha) \leqslant \dim U/K - 2.$$

Let $(\mathfrak{h}_{\mathfrak{p}_*})_e$ denote the set of $H \in \mathfrak{h}_{\mathfrak{p}_*}$ for which $\exp H = e$. As noted in the proof of Lemma 6.3, Chapter V, the subset $\exp \mathfrak{h}_{\mathfrak{p}_*}$ of U is compact; it follows that the factor space $\mathfrak{h}_{\mathfrak{p}_*}/(\mathfrak{h}_{\mathfrak{p}_*})_e$ is compact. Using this fact, it is easily seen that $\Phi(K/M \times \mathfrak{h}_\alpha)$ is closed in U/K. From the sum theorem in dimension theory (Appendix, Theorem 9.2), it follows that

$$\dim \Phi(K/M \times D(U, K)) \leqslant \dim U/K - 2.$$

Remark. For the example $S^2 = \mathbf{SO}(3)/\mathbf{SO}(2)$ the set of points conjugate to o consists of two points, namely, the antipodal point to o and o. Thus the inequality in Theorem 3.3 is the best possible. On the other hand, the equality sign does not in general hold in Theorem 3.3. This is seen from the example $S^3 = \mathbf{SO}(4)/\mathbf{SO}(3)$, the group of unit quaternions (Theorem 4.7).

§ 4. Applications to Compact Groups

As pointed out in §6, Chapter IV, a compact, connected Lie group is a Riemannian globally symmetric space when provided with a bi-invariant Riemannian structure.

Let U be a compact, connected semisimple Lie group. Let $\mathfrak{u}$ denote the Lie algebra of U. Let U^* denote the subgroup $\{(u, u) : u \in U\}$ of the product group $U \times U$. Then $(U \times U, U^*)$ is a Riemannian symmetric pair of the compact type associated with the orthogonal symmetric Lie algebra $(\mathfrak{u} \times \mathfrak{u}, d\sigma)$ where $d\sigma$ is the differential of the automorphism $\sigma : (u_1, u_2) \rightarrow (u_2, u_1)$ of $U \times U$. The coset space $U \times U/U^*$ is diffeomorphic to U under the mapping

$$(u_1, u_2)\, U^* \rightarrow u_1 u_2^{-1} \qquad (u_1, u_2 \in U).$$

Under this correspondence, the natural mapping of $U \times U$ onto $U \times U/U^*$ corresponds to the mapping

$$\pi : (u_1, u_2) \rightarrow u_1 u_2^{-1}$$

of $U \times U$ onto U, whose differential is

$$d\pi : (X, Y) \rightarrow X - Y, \qquad X, Y \in \mathfrak{u}.$$

The Lie algebra $\mathfrak{u} \times \mathfrak{u}$ decomposes into the eigenspaces of $d\sigma$ for the eigenvalues $+ 1$ and $- 1$:

$$\mathfrak{u} \times \mathfrak{u} = \mathfrak{u}^* + \mathfrak{v}^*,$$

where $\mathfrak{u}^*$ equals $\{(X, X) : X \in \mathfrak{u}\}$, the Lie algebra of U^*, and $\mathfrak{v}^* = \{(X, - X) : X \in \mathfrak{u}\}$.

Let T be a maximal torus in U and let $\mathfrak{t}_0$ denote the Lie algebra of T. Then the space

$$\mathfrak{t}^* = \{(H, -H) : H \in \mathfrak{t}_0\}$$

is a maximal abelian subspace of $\mathfrak{v}^*$. Let $\mathfrak{g}$ be the complexification of $\mathfrak{u}$ and let $\mathfrak{t}$ denote the subalgebra of $\mathfrak{g}$ generated by $\mathfrak{t}_0$. We shall now investigate the Weyl group of the symmetric pair $(U \times U, U^*)$ defined by means of the maximal abelian subspace $\mathfrak{t}^*$ of $\mathfrak{v}^*$. The spaces $\mathfrak{u}^*$, $\mathfrak{v}^*$, $\mathfrak{t}^*$, and $\mathfrak{t} \times \mathfrak{t}$, respectively, play the role of the spaces $\mathfrak{k}_0$, $\mathfrak{p}_*$, $\mathfrak{h}_{\mathfrak{p}_*}$, and $\mathfrak{h}$ from §2. In particular, $\mathfrak{t} \times \mathfrak{t}$ is a Cartan subalgebra of $\mathfrak{g} \times \mathfrak{g}$, the complexification of $\mathfrak{u} \times \mathfrak{u}$.

Let Δ^* denote the set of nonzero roots of $\mathfrak{g}$ with respect to $\mathfrak{t}$. Let $\alpha \in \Delta^*$. Then the linear functions α' and α'' on $\mathfrak{t} \times \mathfrak{t}$ given by

$$\alpha'(H_1, H_2) = \alpha(H_1),$$
$$\alpha''(H_1, H_2) = \alpha(H_2), \qquad H_1, H_2 \in \mathfrak{t},$$

are roots of $\mathfrak{g} \times \mathfrak{g}$ with respect to $\mathfrak{t} \times \mathfrak{t}$. In fact, if the vector $X_\alpha \in \mathfrak{g}$ satisfies $[H, X_\alpha] = \alpha(H) X_\alpha$ for all $H \in \mathfrak{t}$, then

$$[(H_1, H_2), (X_\alpha, 0)] = \alpha(H_1)(X_\alpha, 0), \quad [(H_1, H_2), (0, X_\alpha)] = \alpha(H_2)(0, X_\alpha)$$

for all $H_1, H_2 \in \mathfrak{t}$. By counting, it is clear that each nonzero root of $\mathfrak{g} \times \mathfrak{g}$ with respect to $\mathfrak{t} \times \mathfrak{t}$ arises in this manner from a member of Δ^*. The roots α' and α'' cannot vanish identically on $\mathfrak{t}^*$, moreover, their values on $\mathfrak{t}^*$ are purely imaginary.

Let $X \in \mathfrak{u}$. As in §2 we consider now the operator $T_{(X,-X)}$, which is the restriction of $(\mathrm{ad}\,(X, -X))^2$ to $\mathfrak{v}^*$. From Lemma 2.9 we know that if $H \in \mathfrak{t}_0$, the operator $T_{(H,-H)}$ has eigenvalues 0 (with multiplicity $\dim \mathfrak{t}_0$) and the numbers $\alpha(H)^2$ as α runs through Δ^*.

The manifold U has a bi-invariant Riemannian structure (for example, the one induced by the negative of the Killing form of $\mathfrak{u}$). The corresponding affine connection is always the same and $\mathrm{Exp}\, X = \exp X$ for $X \in \mathfrak{u}$. According to Theorem 6.4, Chapter V, each $X \in \mathfrak{u}$ can be written $X = \mathrm{Ad}\,(u) H$ where $u \in U$ and $H \in \mathfrak{t}_0$.

Proposition 4.1. *The point $X = Ad\,(u) H$ is conjugate to e if and only if $\alpha(H) \in 2\pi i(Z - 0)$ for some $\alpha \in \Delta^*$.*

Proof. From the formula for $d \exp_X$ (Theorem 1.7, Chapter II), it follows that X is conjugate to e if and only if

$$\det_{\mathfrak{u}} \left(\frac{1 - e^{-\mathrm{ad}\, X}}{\mathrm{ad}\, X} \right) = 0.$$

Since $\operatorname{ad} X = \operatorname{Ad}(u) \circ \operatorname{ad} H \circ \operatorname{Ad}(u^{-1})$, we have

$$\det_{\mathfrak{u}}\left(\frac{1-e^{-\operatorname{ad} X}}{\operatorname{ad} X}\right) = \det_{\mathfrak{u}}\left(\frac{1-e^{-\operatorname{ad} H}}{\operatorname{ad} H}\right).$$

The endomorphism $(1 - e^{-adH})/\operatorname{ad} H$ of $\mathfrak{g}$ has obviously determinant

$$\prod_{\alpha \in \Delta^*} \frac{1 - e^{-\alpha(H)}}{\alpha(H)}. \tag{1}$$

The restriction to $\mathfrak{u}$ has the same determinant so the proposition follows immediately.

Remark. Proposition 4.1 above is of course a special case of Prop. 3.1. The reason for the appearance of the factor 2 in Prop. 4.1 is that $d\pi(X/2, -X/2) = X$ so X is conjugate to e if and only if

$$\det\left(\sum_0^\infty \frac{1}{(2n+1)!}(T_{(X/2,-X/2)})^n\right) = 0.$$

This determinant, however, equals

$$\prod_{\alpha \in \Delta^*} \frac{\sin \frac{1}{2}\alpha(iH)}{\frac{1}{2}\alpha(iH)} \tag{2}$$

and Prop. 4.1 follows again. Expressions (1) and (2) both give the determinant of $d\operatorname{Exp}_X$ (evaluated by orthonormal basis). It follows that these expressions are equal (as is easily seen anyway).

Consider now the diagram $D(U \times U, U^*) \subset \mathfrak{t}^*$. Let $D(U)$ denote the image of $D(U \times U, U^*)$ under $d\pi$. Since $d\pi(H, -H) = 2H$, $(H \in \mathfrak{t}_0)$ and $\alpha'(H, -H) = -\alpha''(H, -H) = \alpha(H)$, $(\alpha \in \Delta^*)$, it follows that

$$D(U) = \{H \in \mathfrak{t}_0 : \alpha(H) \in 2\pi i Z \text{ for some } \alpha \in \Delta^*\}.$$

The set $D(U)$ is a union of a finite number of families of equispaced hyperplanes of $\mathfrak{t}_0$. It will be called the *diagram of* U. Under the mapping $d\pi : \mathfrak{t}^* \to \mathfrak{t}_0$ the Weyl group $W(U \times U, U^*)$ corresponds to a group $W(U)$ of endomorphisms of $\mathfrak{t}_0$. Since $W(U \times U, U^*)$ leaves $D(U \times U, U^*)$ invariant, it is clear that $W(U)$ leaves $D(U)$ invariant. Moreover, $W(U)$ is generated by the reflexions in the hyperplanes of $D(U)$ which pass through 0. On the other hand, the centralizer of $\mathfrak{t}^*$ in U^* is $T^* = \{(t, t) : t \in T\}$ and the normalizer of $\mathfrak{t}^*$ in U^* is $\{(n, n) : n \in N_T\}$ where N_T denotes the normalizer of T in U. It follows that the group N_T/T, considered as a group of linear transformations of $\mathfrak{t}_0$, coincides with $W(U)$. The *Weyl chambers* in $\mathfrak{t}_0$ are the components of the open subset of $\mathfrak{t}_0$ where all $\alpha \in \Delta^*$ are $\neq 0$. From Theorem 2.12 follows immediately that $W(U)$ is simply transitive on the set of Weyl chambers. Since $W(U)$ and $D(U)$ depend only on $\mathfrak{u}$, they will sometimes be denoted by $W(\mathfrak{u})$ and $D(\mathfrak{u})$, respectively.

In §3 we considered a mapping Φ of $K/M \times \mathfrak{h}_{\mathfrak{p}_*}$ onto U/K. In the present situation this is a mapping of $U^*/T^* \times \mathfrak{t}^*$ onto $U \times U/U^*$. We compare Φ with the mapping $\Psi : (uT, H) \to \exp \operatorname{Ad}(u) H$ which maps $(U/T \times \mathfrak{t}_0)$ onto U and consider the diagram

$$\begin{array}{ccc} U^*/T^* \times \mathfrak{t}^* & \xrightarrow{\Phi} & U \times U/U^* \\ f \downarrow & & \downarrow g \\ U/T \times \mathfrak{t}_0 & \xrightarrow{\Psi} & U \end{array} \tag{3}$$

where the mappings f and g are given by

$$f : ((u, u)\, T^*, (H, -H)) \to (uT, 2H),$$

$$g : (u_1, u_2)\, U^* \to u_1 u_2^{-1}.$$

The diagram is then commutative. Since f and g are diffeomorphisms, we conclude from Prop. 3.2:

Proposition 4.2. Let $u_0 \in U$, $H_0 \in \mathfrak{t}_0$. *The mapping* $\Psi : (uT, H) \to$ *exp Ad* (u) H *which maps* $U/T \times \mathfrak{t}_0$ *onto* U *is regular at* (u_0T, H_0) *if and only if* $\alpha(iH_0)/2\pi$ *is not an integer for any* $\alpha \in \Delta^*$.

In analogy with the notations in §3 we make now the following definition.

Definition. The set $S = \Psi(U/T \times D(U))$ will be called the *singular set* in U and its elements will be called the *singular elements*. The complement $U - S$ will be denoted by U_r, and its elements will be called the *regular elements*. Finally $\mathfrak{t}_r$ shall denote the complement $\mathfrak{t}_0 - D(U)$. It is obvious from Prop. 4.4 below that $\Psi(U/T \times \mathfrak{t}_r) = U_r$.

Lemma 4.3. *Let* $H \in \mathfrak{t}_0$ *and put* $t = \exp H$. *Let* Z_t *denote the centralizer of* t *in* U. *Then*

$$\dim Z_t = \dim T + n,$$

where n *denotes the number of roots* $\alpha \in \Delta^*$ *for which* $\alpha(H) \in 2\pi i \boldsymbol{Z}$.

Proof. The Cartan subalgebra $\mathfrak{t}$ of $\mathfrak{g}$ is invariant under the conjugation of $\mathfrak{g}$ with respect to $\mathfrak{u}$. Using Lemma 3.1, Chapter VI, we can for each $\alpha \in \Delta^*$ select a vector $X_\alpha \neq 0$ in $\mathfrak{g}$ such that $[H, X_\alpha] = \alpha(H) X_\alpha$ for all $H \in \mathfrak{t}$ and such that the vectors $E_\alpha = X_\alpha - X_{-\alpha}$, $F_\alpha = i(X_\alpha + X_{-\alpha})$ belong to $\mathfrak{u}$. Let Δ^* be ordered in some way and let $(\Delta^*)_+$ be the set of positive roots with respect to this ordering. Then E_α, F_α, $(\alpha \in (\Delta^*)_+)$, is a basis of $\mathfrak{u}$ (mod $\mathfrak{t}_0$) and

$$[H, E_\alpha] = - i\alpha(H) F_\alpha, \qquad (4)$$
$$[H, F_\alpha] = i\alpha(H) E_\alpha$$

for all $H \in \mathfrak{t}_0$. It follows that

$$\begin{aligned} \mathrm{Ad}\,(t) E_\alpha &= \cos (i\alpha(H)) E_\alpha - \sin (i\alpha(H)) F_\alpha, \\ \mathrm{Ad}\,(t) F_\alpha &= \sin (i\alpha(H)) E_\alpha + \cos (i\alpha(H)) F_\alpha. \end{aligned} \qquad (5)$$

The Lie algebra $\mathfrak{L}(Z_t)$ of Z_t is given by

$$\mathfrak{L}(Z_t) = \{X \in \mathfrak{u} : (\exp sX)\, t = t \exp sX \text{ for all } s \in \boldsymbol{R}\}.$$

Since $t \exp sX t^{-1} = \exp s \,\mathrm{Ad}\,(t) X$, we find

$$\mathfrak{L}(Z_t) = \{X \in \mathfrak{u} : \mathrm{Ad}\,(t) X = X\}.$$

Consequently, dim Z_t equals the dimension of the eigenspace of Ad (t) for the eigenvalue 1. From (5) we see that this eigenspace has dimension dim $\mathfrak{t}_0$ + twice the number of $\alpha \in (\Delta^*)_+$ for which $\alpha(H) \in 2\pi i\boldsymbol{Z}$.

Proposition 4.4. *Let $x \in U$ and let Z_x denote the centralizer of x in U. Then x is singular (regular) if and only if dim $Z_x >$ dim T (dim $Z_x =$ dim T).*

The element x can be written $x = utu^{-1}$ where $u \in U$, $t \in T$. Then dim $Z_x =$ dim Z_t and the proposition follows from Lemma 4.3.

For each $\alpha \in \Delta^*$, put

$$\mathfrak{t}_\alpha = \{H \in \mathfrak{t}_0 : \alpha(H) \in 2\pi i\boldsymbol{Z}\},$$
$$T_\alpha = \{u \in U : \exp \mathrm{Ad}\,(u) H = \exp H \text{ for each } H \in \mathfrak{t}_\alpha\}.$$

Then T_α is a closed subgroup of U containing T. Let $(\mathfrak{t}_\alpha)_0$ denote the hyperplane $\alpha(H) = 0$ in $\mathfrak{t}_0$.

Lemma 4.5. *The group T_α is connected and dim $T_\alpha =$ dim $T + 2$.*

Proof. It follows from (5) that if $H \in \mathfrak{t}_\alpha$,

$$\mathrm{Ad}\,(\exp H)\, sE_\alpha = sE_\alpha, \qquad \mathrm{Ad}\,(\exp H)\, sF_\alpha = sF_\alpha \qquad (6)$$

for all $s \in \boldsymbol{R}$. Now, (6) implies that

$$\exp \operatorname{Ad}(\exp sE_\alpha) H = \exp H, \qquad \exp \operatorname{Ad}(\exp \ F_\alpha) H = \exp H$$

for all $H \in \mathfrak{t}_\alpha$ and all $s \in \boldsymbol{R}$. It follows that E_α and F_α lie in the Lie algebra of T_α so $\dim T_\alpha \geqslant \dim T + 2$. On the other hand, $T_\alpha \subset (T_\alpha)_0$ if $(T_\alpha)_0$ denotes the centralizer of $(\mathfrak{t}_\alpha)_0$ in U. Then $(T_\alpha)_0$ is the centralizer in U of the closure of $\exp (\mathfrak{t}_\alpha)_0$. By Cor. 2.8, $(T_\alpha)_0$ is connected. Furthermore, $(t_\alpha)_0$ contains an element H such that $\beta(H) \notin 2\pi i \boldsymbol{Z}$ for all $\beta \in \Delta^*$ different from $\pm\alpha$. By Lemma 4.3, $\dim (T_\alpha)_0 \leqslant \dim T + 2$ and Lemma 4.5 follows.

Let $(\mathfrak{t}_\alpha)_r$ denote the subset of $\mathfrak{t}_\alpha$ given by

$$(\mathfrak{t}_\alpha)_r = \{H \in \mathfrak{t}_\alpha : \beta(H) \notin 2\pi i \boldsymbol{Z} \text{ for } \beta \in \Delta^* - \{\alpha \cup -\alpha\}$$

and consider the mapping

$$\Psi_\alpha : (uT_\alpha, H) \to \exp \operatorname{Ad}(u) H$$

of $U/T_\alpha \times \mathfrak{t}_\alpha$ into U.

Lemma 4.6. *The mapping Ψ_α is regular on the subset $U/T_\alpha \times (\mathfrak{t}_\alpha)_r$ of $U/T_\alpha \times \mathfrak{t}_\alpha$.*

Proof. Consider the subspace $\mathfrak{u}_\alpha$ of $\mathfrak{u}$ spanned by all the vectors E_β, F_β as β varies through the positive roots in $\Delta^* - \{\alpha \cup -\alpha\}$. The natural mapping of U onto U/T_α has a differential which identifies $\mathfrak{u}_\alpha$ with the tangent space to U/T_α at $\{T_\alpha\}$. If $u \in U$, then as usual, $\tau(u)$ shall denote the mapping $xT_\alpha \to uxT_\alpha$ of U/T_α onto itself. The tangent space to the product $U/T_\alpha \times \mathfrak{t}_\alpha$ can be identified with the subspace $\mathfrak{u}_\alpha + (\mathfrak{t}_\alpha)_0$ of $\mathfrak{u}$, the subspaces $\mathfrak{u}_\alpha$ and $(\mathfrak{t}_\alpha)_0$ of $\mathfrak{u}$ being orthogonal. Now let $u_0 \in U$, $H_0 \in (\mathfrak{t}_\alpha)_r$, $X \in \mathfrak{u}_\alpha$. Then

$$\begin{aligned}\Psi_\alpha(u_0(\exp tX)\, T_\alpha, H_0) &= \exp (\operatorname{Ad}(u_0 \exp tX)\, H_0)\\ &= u_0 \exp (H_0 + t[X, H_0] + O(t^2))\, u_0^{-1}.\end{aligned}$$

Using Theorem 1.7, Chapter II, we obtain

$$(d\Psi_\alpha)_{(u_0T_\alpha, H_0)}\,(d\tau(u_0)\, X, 0) = dL_{u_0 \exp H_0 u_0^{-1}} \circ \operatorname{Ad}(u_0) \frac{1 - e^{-\operatorname{ad} H_0}}{\operatorname{ad} H_0}([X, H_0]). \qquad (7)$$

Similarly, if $H \in (\mathfrak{t}_\alpha)_0$,

$$\Psi_\alpha(u_0T_\alpha, H_0 + tH) = \exp (\operatorname{Ad}(u_0)\,(H_0 + tH))$$

so

$$(d\Psi_\alpha)_{(u_0T_\alpha, H_0)}\,(0, H) = dL_{u_0 \exp H_0 u_0^{-1}} \circ \operatorname{Ad}(u_0)\, H. \qquad (8)$$

Combining (7) and (8) we get

$$(d\Psi_\alpha)_{(u_0T_\alpha, H_0)}(d\tau(u_0)X, H) = dL_{u_0 \exp H_0 u_0^{-1}} \circ \mathrm{Ad}(u_0)\{(\mathrm{Ad}(\exp - H_0) - 1)X + H\}.$$

Since $(\mathrm{Ad}(\exp(-H_0)) - 1)X \in \mathfrak{u}_\alpha$, the right-hand side of this formula can vanish only if $H = 0$ and $\mathrm{Ad}(\exp H_0)X = X$. In view of (5), this would either require $X = 0$ or $\beta(H_0) \in 2\pi i \boldsymbol{Z}$ for some $\beta \in \Delta^* - \{\alpha \cup -\alpha\}$. But this last possibility is excluded by the assumption that $H_0 \in (\mathfrak{t}_\alpha)_r$. This proves the lemma.

Theorem 4.7. *Let S denote the singular set in U and let conj (U) denote the set of points in U which are conjugate to e. Then*

$$\dim S = \dim \mathrm{conj}\,(U) = \dim U - 3.$$

Proof. Let $\alpha \in \Delta^*$. It is obvious that

$$\Psi_\alpha(U/T_\alpha \times \mathfrak{t}_\alpha) = \Psi(U/T \times \mathfrak{t}_\alpha).$$

Using Lemmas 4.5 and 4.6 we find

$$\begin{aligned}\dim \Psi_\alpha(U/T_\alpha \times (\mathfrak{t}_\alpha)_r) &\geqslant \dim U/T_\alpha + \dim \mathfrak{t}_\alpha \\ &= \dim U - (\dim T + 2) + \dim T - 1 = \dim U - 3.\end{aligned}$$

Using Lemma 9.3 and Prop. 9.1 in the Appendix we have

$$\dim \Psi_\alpha(U/T_\alpha \times (\mathfrak{t}_\alpha)_r) \leqslant \dim \Psi_\alpha(U/T_\alpha \times \mathfrak{t}_\alpha) \leqslant \dim U - 3$$

so

$$\dim \Psi(U/T \times \mathfrak{t}_\alpha) = \dim U - 3.$$

Since

$$S = \bigcup_{\alpha \in \Delta^*} \Psi(U/T \times \mathfrak{t}_\alpha)$$

and since $\Psi(U/T \times \mathfrak{t}_\alpha)$ is closed in U, it follows from the sum theorem (Appendix, Theorem 9.2) that $\dim S = \dim U - 3$. The statement about conj (U) is proved in the same manner.

§ 5. Control over the Singular Set

We recall that an element $x \in U$ is regular or singular according to whether $\dim Z_x$ is equal or larger than $\dim T$. In view of Lemma 4.3 we can make the following definition.

Definition. An element $x \in U$ is called *singular of order $n/2$* if $\dim Z_x = \dim T + n$ and $n > 0$.

Let k be an integer > 0, and let S_k denote the set of singular elements in U of order k. Then $S = \bigcup_{k>0} S_k$.

Consider now a fixed element $x_0 \in S_k$ and select $u \in U$, $H_0 \in \mathfrak{t}_0$ such that $x_0 = \exp \operatorname{Ad}(u) H_0$. Then there exist exactly k positive roots in Δ^*, say $\alpha_1, \ldots, \alpha_k$, for which $\alpha_j(H_0) \in 2\pi i \boldsymbol{Z}$ $(1 \leqslant j \leqslant k)$. Consider the group

$$T_{\alpha_1 \ldots \alpha_k} = \{u \in U : \exp \operatorname{Ad}(u) H = \exp H \text{ for each } H \in \mathfrak{t}_{\alpha_1} \cap \ldots \cap \mathfrak{t}_{\alpha_k}\}.$$

Lemma 5.1.

$$\dim T_{\alpha_1 \ldots \alpha_k} = \dim T + 2k.$$

Proof. We imitate the proof of Lemma 4.5. Just as there it can be verified that the vectors E_{α_j}, F_{α_j} $(1 \leqslant j \leqslant k)$ belong to the Lie algebra of $T_{\alpha_1 \ldots \alpha_k}$. Consequently,

$$\dim T_{\alpha_1 \ldots \alpha_k} \geqslant \dim T + 2k.$$

On the other hand,

$$\dim Z_{\exp H_0} = \dim T + 2k$$

according to Lemma 4.3. Since $T_{\alpha_1 \ldots \alpha_k} \subset Z_{\exp H_0}$, the lemma is proved.

Consider now the subset $(\mathfrak{t}_{\alpha_1 \ldots \alpha_k})_r$ of $\mathfrak{t}_{\alpha_1} \cap \ldots \cap \mathfrak{t}_{\alpha_k}$ consisting of all points H such that $\beta(H) \notin 2\pi i \boldsymbol{Z}$ unless β is among the roots $\pm \alpha_j$ $(1 \leqslant j \leqslant k)$. Then $(\mathfrak{t}_{\alpha_1 \ldots \alpha_k})_r$ is an open subset of $\mathfrak{t}_{\alpha_1} \cap \ldots \cap \mathfrak{t}_{\alpha_k}$ containing H_0. Consider the mapping

$$\Psi_{\alpha_1 \ldots \alpha_k} : (uT_{\alpha_1 \ldots \alpha_k}, H) \rightarrow \exp \operatorname{Ad}(u) H$$

of $(U/T_{\alpha_1 \ldots \alpha_k}) \times (\mathfrak{t}_{\alpha_1} \cap \ldots \cap \mathfrak{t}_{\alpha_k})$ into U.

Lemma 5.2. *The mapping $\Psi_{\alpha_1 \ldots \alpha_k}$ is regular on the subset*

$$(U/T_{\alpha_1 \ldots \alpha_k}) \times (\mathfrak{t}_{\alpha_1 \ldots \alpha_k})_r.$$

The proof is an immediate extension of that of Lemma 4.6 and can be omitted.

Definition. Let N be a subset of a manifold M; N is called a *quasi-submanifold* of M if there exists a connected manifold N^* and a regular differentiable mapping $f : N^* \rightarrow M$ such that $f(N^*) = N$.

If $\mathfrak{s}$ is a connected component of $(\mathfrak{t}_{\alpha_1 \ldots \alpha_k})_r$, then the image

$$\Psi(U/T \times \mathfrak{s}) = \Psi_{\alpha_1 \ldots \alpha_k}((U/T_{\alpha_1 \ldots \alpha_k}) \times \mathfrak{s})$$

is a quasisubmanifold of U due to Lemma 5.2.

Now let

$$\mathfrak{t}_e = \{H \in \mathfrak{t}_0 : \exp H = e\}.$$

The set $\mathfrak{t}_e$ is called the *unit lattice* in $\mathfrak{t}_0$. Then clearly $\alpha(H) \in 2\pi i Z$ for all $H \in \mathfrak{t}_e$ and all $\alpha \in \Delta^*$. Therefore, if we consider $\mathfrak{t}_e$ as a group of translations of $\mathfrak{t}_0$, it leaves the diagram $D(U)$ invariant. Moreover, each transformation from $\mathfrak{t}_e$ maps $(\mathfrak{t}_{\alpha_1 \ldots \alpha_k})_r$ onto itself and therefore permutes the various components $\mathfrak{s}$ of $(\mathfrak{t}_{\alpha_1 \ldots \alpha_k})_r$.

Lemma 5.3. *There are only finitely many components of* $(\mathfrak{t}_{\alpha_1 \ldots \alpha_k})_r$ *which are incongruent modulo* $\mathfrak{t}_e$.

Proof. The factor space $\mathfrak{t}_0/\mathfrak{t}_e$ can be identified with T so the group $\mathfrak{t}_e$ has a bounded fundamental domain in $\mathfrak{t}_0$. Now on each component $\mathfrak{s}$ of $(\mathfrak{t}_{\alpha_1 \ldots \alpha_k})_r$ the roots $\alpha_1, \ldots, \alpha_k$ are constants and the other positive roots vary through an interval of $2\pi i$. It follows that the components $\mathfrak{s}$ are uniformly bounded. Consequently, there exists a closed ball $\mathfrak{b}$ in $\mathfrak{t}_0$ such that for each component $\mathfrak{s}$ of $(\mathfrak{t}_{\alpha_1 \ldots \alpha_k})_r$ there exists a vector $H \in \mathfrak{t}_e$ such that the translate of $\mathfrak{s}$ by $-H$, that is, $\mathfrak{s} - H$, lies in $\mathfrak{b}$. Since each root $\alpha \in \Delta^*$ is bounded on $\mathfrak{b}$ there can only be finitely many such sets $\mathfrak{s} - H$ in $\mathfrak{b}$. This proves the lemma.

The roots $\alpha_1, \ldots, \alpha_k$ were obtained by means of an arbitrary point $H_0 \in \mathfrak{t}_0$ which lies on exactly k hyperplanes in the diagram. As H_0 varies through all such points we get finitely many systems $(\alpha_1, \ldots, \alpha_k)$ of k positive roots. Let S_k^i $(i = 1, 2, \ldots)$ denote the images $\Psi(U/T \times \mathfrak{s})$ as $\mathfrak{s}$ varies through the components of $(\mathfrak{t}_{\alpha_1 \ldots \alpha_k})_r$ for the various systems $(\alpha_1, \ldots, \alpha_k)$.

Lemma 5.4. *The set of singular points of order k is a finite union*

$$S_k = \bigcup_i S_k^i,$$

where each S_k^i is a quasisubmanifold of U. If S_k^i is given the relative topology of U, its boundary is contained in $\bigcup_{p>k} S_p$.

Proof. The finiteness statement is a consequence of Lemma 5.3 since the components of $(\mathfrak{t}_{\alpha_1 \ldots \alpha_k})_r$ which are congruent mod $\mathfrak{t}_e$ give rise to the same set S_k^i. The last statement follows from the fact that the boundary points of $(\mathfrak{t}_{\alpha_1 \ldots \alpha_k})_r$ lie on more than k hyperplanes of the diagram.

Proposition 5.5. *Let $\gamma(t)$ and $\gamma'(t)$ $(0 \leqslant t \leqslant 1)$ be two continuous curves in U_r. Then γ is homotopic to γ' in U_r if and only if they are homotopic in U.*

Proof. We may assume that γ and γ' are homotopic in U. Let p and q, respectively, denote the beginning and end of γ. Let f denote the mapping of the unit square $\square$ $(0 \leqslant s \leqslant 1,\ 0 \leqslant t \leqslant 1)$ into U which sets up the assumed homotopy $\gamma \sim \gamma'$. In other words, $f(0, t) = \gamma(t)$, $f(1, t) = \gamma'(t)$ for $0 \leqslant t \leqslant 1$ and $f(s, 0) = p$, $f(s, 1) = q$ for $0 \leqslant s \leqslant 1$. We have to deform $f(\square)$ into U_r in such a way that the points p and q remain fixed. The deformations considered below are always understood to take place with p and q kept fixed.

Let $2m$ denote the number of elements in Δ^*. Then S_m is the center of U and S_p is empty if $p > m$. Since S_m is finite and $\dim U \geqslant 3$, we can, using Prop. 9.6 (Appendix), deform $f(\square)$ such that the resulting deformation, say $f_1(\square)$, satisfies

$$f_1(\square) \cap S_m = \emptyset.$$

Here f_1 denotes a new homotopy between γ and γ'. It follows that there exists a compact neighborhood N_1 of $f_1(\square)$ such that

$$N_1 \cap S_m = \emptyset.$$

Suppose now that we have found deformations $f_1(\square), \ldots, f_{m-k}(\square)$ of $f(\square)$ and compact neighborhoods $N_1, \ldots, N_{m-k}$ of $f_1(\square), \ldots, f_{m-k}(\square)$, respectively, such that

$$\begin{aligned} &N_1 \supset \ldots \supset N_{m-k}, \\ &N_1 \cap S_m = \ldots = N_{m-k} \cap S_{k+1} = \emptyset. \end{aligned} \tag{1}$$

We shall then show that there exists a deformation $f_{m-k+1}(\square)$ of $f_{m-k}(\square)$ and a compact neighborhood N_{m-k+1} of $f_{m-k+1}(\square)$ such that (1) holds with k replaced by $k - 1$. Then the validity of (1) for $k = m - 1$ implies its validity for $k = 0$, proving the proposition.

Starting now from (1), we consider the set $N_{m-k} \cap S_k^i$ where S_k^i is one of the sets from Lemma 5.4. The last statement of that lemma, together with (1), implies that $N_{m-k} \cap S_k^i$ is compact. For a suitable system $(\alpha_1, \ldots, \alpha_k)$ of positive roots, the set $N_{m-k} \cap S_k^i$ is the image, under $\Psi_{\alpha_1 \ldots \alpha_k}$, of a compact subset of $U/T_{\alpha_1 \ldots \alpha_k} \times \mathfrak{s}$. This compact set can be covered by finitely many sets $U_j \times \mathfrak{b}_j$ $(j \in J)$, where the U_j are open subsets of $U/T_{\alpha_1 \ldots \alpha_k}$ and the $\mathfrak{b}_j$ are open balls whose closure is contained in $\mathfrak{s}$ such that for each $j \in J$, $\Psi_{\alpha_1 \ldots \alpha_k}$ maps some neighborhood of $\bar{U}_j \times \bar{\mathfrak{b}}_j$ in a one-to-one manner into $S_k{}^i$. We put now

$$\Psi_{\alpha_1 \ldots \alpha_k}(U_j \times \mathfrak{b}_j) = B_j \qquad (j \in J),$$

and turn B_j into a manifold diffeomorphic to $U_j \times \mathfrak{b}_j$. Then:

(a) Each B_j is a submanifold and a topological subspace of U.

(b) $B_j \subset S_k{}^i$ for each $j \in J$.

(c) $N_{m-k} \cap (\bigcup_{j \in J} B_j) = N_{m-k} \cap S_k{}^i$.

Consider now a fixed B_j. Using Props. 9.4 and 9.6 in the Appendix we can deform $f_{m-k}(\square)$ in the interior of N_{m-k} such that the resulting deformation, say $g_{m-k}(\square)$, is disjoint from the closure of B_j. Hence we can surround $g_{m-k}(\square)$ with a compact neighborhood $V_{m-k} \subset N_{m-k}$ disjoint from B_j. We treat the (finitely many) B_j $(j \in J)$ successively in the same manner. The resulting deformation $G_{m-k}(\square)$ is then enclosed in a compact neighborhood W_{m-k} which is contained in N_{m-k} and is disjoint from $\bigcup_{j \in J} B_j$. It follows from (*c*) that $W_{m-k} \cap S_k{}^i = \emptyset$. The preceding process can be applied to the sets $S_k{}^1, S_k{}^2, \ldots$ successively. Since these are only finite in number, the result is a deformation $f_{m-k+1}(\square)$ enclosed in a compact neighborhood $N_{m-k+1} \subset N_{m-k}$ such that

$$N_{m-k+1} \cap S_k = \emptyset.$$

This concludes the proof.

§ 6. The Fundamental Group and the Center

Consider now the open set $\mathfrak{t}_r = \mathfrak{t}_0 - D(U)$. Let P_0 denote a component of $\mathfrak{t}_r$ whose closure $\bar{P}_0$ contains the origin. The polyhedron P_0 is an intersection of half-spaces in $\mathfrak{t}_0$, hence P_0 is an open, convex set. Since $\alpha(H) \in 2\pi i Z$ for all $\alpha \in \Delta^*$ and all $H \in \mathfrak{t}_e$, it is clear that each point in $\mathfrak{t}_e \cap \bar{P}_0$ is a vertex of P_0.

Theorem 6.1. *The number of points in* $\mathfrak{t}_e \cap \bar{P}_0$ *equals the order of the fundamental group* $\pi_1(U)$ *of* U.

The proof will require a few lemmas.

Lemma 6.2. *The coset space* U/T *is simply connected.*

Proof. Let $(\tilde{U}, \gamma)$ denote the simply connected covering group of U and put $\tilde{T} = \gamma^{-1}(T)$. Then $\tilde{U}$ is compact and the Lie algebra of $\tilde{T}$ is a maximal abelian subalgebra of the Lie algebra of $\tilde{U}$. Since the centralizer of a torus is connected, $\tilde{T}$ is a maximal torus in $\tilde{U}$. The space $\tilde{U}/\tilde{T}$ is simply connected and homeomorphic to U/T under the mapping $u\tilde{T} \to \gamma(u)\, T$, $(u \in \tilde{U})$.

Lemma 6.3. *Let ψ denote the restriction of Ψ to $U/T \times P_0$. Then $(U/T \times P_0, \psi)$ is the universal covering space of U_r.*

Proof. The connectedness of U_r is clear from Cor. 9.5 in the Appendix. Since P_0 is convex, it is simply connected. Hence $U/T \times P_0$ is simply connected. Let P_1 be an arbitrary component of $\mathfrak{t}_r$. If the images $\psi(U/T \times P_0)$ and $\psi(U/T \times P_1)$ are not disjoint there exist elements $H_0 \in P_0$, $H_1 \in P_1$, $x \in U$ such that $x \exp H_0 x^{-1} = \exp H_1$. It follows that the automorphism $u \to xux^{-1}$ $(u \in U)$ maps the centralizer $Z_{\exp H_0}$ onto the centralizer $Z_{\exp H_1}$. Owing to Lemma 4.3, T is the identity component of these centralizers. Consequently $xTx^{-1} = T$ and there exists an element $s \in W(U)$ which coincides with the restriction $\operatorname{Ad}(x)$ to $\mathfrak{t}_0$. Hence $\exp sH_0 = \exp H_1$, $sH_0 \in \mathfrak{t}_0$ and there exists a vector $A \in \mathfrak{t}_e$ such that $H_1 = sH_0 + A$. Since the groups $W(U)$ and $\mathfrak{t}_e$ leave the diagram invariant it follows that the transformation $H \to sH + A$ of $\mathfrak{t}_0$ maps P_0 onto P_1. Consequently $\psi(U/T \times P_0) = \psi(U/T \times P_1)$. The connectedness of U_r now implies that ψ maps $U/T \times P_0$ onto U_r.

For each $x \in U_r$, the inverse image $\psi^{-1}(x)$ is finite. In fact suppose the contrary; then there exists a convergent sequence $(Q_n) \subset U/T \times \bar{P}_0$ such that all Q_n are different and $\psi(Q_n) = x$ for all n. Since ψ is the restriction of the mapping $\Psi : U/T \times \mathfrak{t}_0 \to U$, the limit

$$(u_0T, H_0) = \lim Q_n$$

satisfies $u_0 \exp H_0 u_0^{-1} = x$ and $H_0 \in \bar{P}_0$. From the regularity of x follows that $H_0 \in P_0$. Now Prop. 4.2 implies that ψ is a local homeomorphism, so there exists a neighborhood N of (u_0T, H_0) in $U/T \times P_0$ which is mapped homeomorphically under ψ. Since N contains infinitely many Q_n we have a contradiction. Now, $\psi^{-1}(x)$ being finite and ψ being a local homeomorphism, it follows that $(U/T \times P_0, \psi)$ is a covering space of U_r.

Lemma 6.4. *The number of points in $\mathfrak{t}_e \cap \bar{P}_0$ equals the number of elements in $\pi_1(U_r)$, the fundamental group of U_r.*

Proof. Consider a neighborhood

$$V = \{X \in \mathfrak{u} : -B(X, X) < \rho^2\}$$

of 0 in $\mathfrak{u}$.

We can select $\rho > 0$ so small that:

(a) $|\alpha(H)| < 2\pi$ *for* $H \in V \cap \mathfrak{t}_0$ *and* $\alpha \in \Delta^*$.
(b) $\exp$ *is one-to-one on* $2V$.
(c) $\exp(V \cap \mathfrak{t}_0) = (\exp V) \cap T$.

It is trivial to satisfy (a) and (b). For (c) one just has to make use of the fact that T is a topological subgroup of U (Lemma 2.5, Chapter II).

Fix an element $x \in U_r \cap (\exp V)$ and consider the inverse image

$$\psi^{-1}(x) = \{(u_1T, H_1), ..., (u_rT, H_r)\}.$$

For each i, $1 \leqslant i \leqslant r$, we have

$$\exp H_i \in u_i^{-1} (\exp V)\, u_i \subset \exp V.$$

From (c) and (b) it follows that there exists a *unique* vector $A_i \in \mathfrak{t}_e$ such that $H_i - A_i \in V$.

On the other hand, since

$$u_i \exp H_i\, u_i^{-1} = u_j \exp H_j\, u_j^{-1}, \qquad 1 \leqslant i, j \leqslant r,$$

it follows as in the last lemma, that $uTu^{-1} = T$ if $u = u_i^{-1}u_j$. Let s_{ij} denote the restriction of $\mathrm{Ad}(u)$ to $\mathfrak{t}_0$. Then there exists a vector $A_{ij} \in \mathfrak{t}_e$ such that $H_i = s_{ij}H_j + A_{ij}$. Since s_{ij} leaves $\mathfrak{t}_e$ invariant, there exists a vector $A^* \in \mathfrak{t}_e$ such that $s_{ij}A^* = A_i - A_{ij}$. Then $H_i - A_i = s_{ij}(H_j - A^*)$ so $H_j - A^* \in V$. By the uniqueness above, $A^* = A_j$ so

$$H_i - A_i = s_{ij}(H_j - A_j) \qquad (1 \leqslant i, j \leqslant r). \tag{1}$$

We shall now draw some consequences of this relation.

(d) *The points* $H_1, ..., H_r$ *are all different.*

In fact, if for example $H_1 = H_2$, then $A_1 = A_2$. But $W(U)$ is simply transitive on the set of Weyl chambers so (1) implies that $s_{12} = I$. Hence $u_1T = u_2T$, which is a contradiction.

(e) *The points* $A_1, ..., A_r$ *are all different.*

Suppose to the contrary that for example $A_1 = A_2$. The segment l joining H_1 and H_2 lies in P_0. It follows that the translated segment $l - A_1$ lies in $\mathfrak{t}_r$; in particular $H_1 - A_1$ and $H_2 - A_1$ lie in the same Weyl chamber. On the other hand, (1) implies $H_1 - A_1 = s_{12}(H_2 - A_1)$ so again by the simple transitivity $s_{12} = I$ which is a contradiction.

(f) *The points* $A_1, ..., A_r$ *belong to* $\bar{P}_0$.

In fact, if $A_i \notin \bar{P}_0$, then the interior of the segment from H_i to A_i intersects the boundary of P_0. Therefore, there exists a root $\alpha \in \Delta^*$ such that $|\alpha(H_i - A_i)| > 2\pi$. In view of (a), this contradicts $H_i - A_i \in V$.

(g) $\mathfrak{t}_e \cap \bar{P}_0$ *consists of precisely the points* $A_1, ..., A_r$.

Suppose to the contrary, that there were a point $A \in \mathfrak{t}_e \cap \bar{P}_0$ which does not occur among $A_1, \ldots, A_r$. Let C denote the Weyl chamber in $\mathfrak{t}_0$ with the property that P_0 is contained in the translated set $C + A$. There exists a unique element $s \in W(U)$ such that $s(H_1 - A_1) \in C$. Then $A + s(H_1 - A_1) \in C + A$. Moreover, due to (a), the open segment from 0 to $s(H_1 - A_1)$ lies in $\mathfrak{t}_r$; the same is true of the translated segment from A to $A + s(H_1 - A_1)$. Consequently the point

$$H_A = A + s(H_1 - A_1)$$

lies in P_0. Let u be an element in U such that $\mathrm{Ad}(u)$ and s^{-1} coincide on $\mathfrak{t}_0$; then

$$\begin{aligned}\psi(u_1uT, H_A) &= u_1u \exp(A + s(H_1 - A_1))\, u^{-1}u_1^{-1} \\ &= u_1 \exp(H_1 - A_1)\, u_1^{-1} = \psi(u_1T, H_1) = x.\end{aligned}$$

On the other hand, $H_A \neq H_i$ for $i = 1, \ldots, r$ because the relation $H_A = H_i$ implies $| H_i - A | = | H_1 - A_1 |$ which equals $| H_i - A_i |$ due to (1). But $| H_i - A | = | H_i - A_i |$ implies $A - A_i \in 2V$ which contradicts (b).

This finishes the proof of Lemma 6.4.

Now let u_0 be a point in U_r and γ a curve in U beginning and ending at u_0. Due to Prop. 9.4 in the Appendix, γ is homotopic to a curve $\gamma' \subset U_r$. Hence it follows from Prop. 5.5 that $\pi_1(U)$ and $\pi_1(U_r)$ are isomorphic. This concludes the proof of Theorem 6.1.

Let (M^*, π) be a simply connected covering space of a topological space M. A homeomorphism φ of M^* such that $\pi \circ \varphi = \pi$ is called a *covering transformation* of M^*. These homeomorphisms form a group which is isomorphic with the fundamental group $\pi_1(M)$ of M.

What are the covering transformations corresponding to the covering space $(U/T \times P_0, \psi)$ of U_r? In terms of the notation of Lemma 6.4, the transformation

$$(vT, H) \to (vu_1^{-1}u_iT, A_i + s_{i1}(H - A_1))$$

is a covering transformation of $U/T \times P_0$. In fact, due to (1) we have

$$A_i + s_{i1}(H - A_1) \in P_0 \qquad \text{for } H \in P_0$$

and

$$\psi(vu_1^{-1}u_iT, A_i + s_{i1}(H - A_1)) = \psi(vT, H)$$

for $H \in P_0$, $v \in U$. These covering transformations, $\pi_1(U)$ in number, are all different because the images $H_i = A_i + s_{i1}(H_1 - A_1)$ are all different as shown above. The group $\pi_1(U)$ is isomorphic with the group of transformations

$$\varphi_i : H \to A_i + s_{i1}(H - A_1)$$

of $\mathfrak{t}_0$, each of which maps P_0 into itself. The orbit of A_1 under this group is $\mathfrak{t}_e \cap \bar{P}_0$.

Lemma 6.5. *Let Z denote the center of U and let $\mathfrak{t}(\mathfrak{u})$ denote the set of points $H \in \mathfrak{t}_0$ for which exp $H \in Z$. Then*

$$\mathfrak{t}(\mathfrak{u}) = \{H \in \mathfrak{t}_0 : \alpha(H) \in 2\pi i \boldsymbol{Z} \text{ for each } \alpha \in \Delta^*\}.$$

Proof. It is obvious that $\mathfrak{t}(\mathfrak{u})$ is the set of $H \in \mathfrak{t}_0$ for which $\text{Ad}(\exp H) = I$. The lemma now follows from relations (5), §4. The notation $\mathfrak{t}(\mathfrak{u})$ is to indicate that the set in question is the same for all groups that have Lie algebra $\mathfrak{u}$.

Corollary 6.6. *Let $\tilde{U}$ denote the universal covering group of U and let $\tilde{Z}$ denote the center of $\tilde{U}$. The number of points in $\mathfrak{t}(\mathfrak{u}) \cap \bar{P}_0$ (in geometric terms: the number of vertices of P_0) equals the order of $\tilde{Z}$.*

Proof. This corollary results from using Theorem 6.1 on the group $\text{Ad}(\tilde{U}) = \tilde{U}/\tilde{Z}$. For this group, the unit lattice $\mathfrak{t}_e$ coincides with $\mathfrak{t}(\mathfrak{u})$ and the fundamental group $\pi_1(\text{Ad}(\tilde{U}))$ is isomorphic to $\tilde{Z}$.

Example. Let $U = \boldsymbol{SU}(n)$, the group of unitary $n \times n$ matrices of determinant one. The Lie algebra $\mathfrak{u} = \mathfrak{su}(n)$ consists of all $n \times n$ skew Hermitian matrices of trace 0, and the complexification $\mathfrak{g}$ is the Lie algebra $\mathfrak{sl}(n, \boldsymbol{C})$ of all $n \times n$ matrices of trace 0. The subset $\mathfrak{t}_0$ of $\mathfrak{u}$ consisting of all $n \times n$ purely imaginary diagonal matrices of trace 0 is a maximal abelian subalgebra of $\mathfrak{u}$. The subspace $\mathfrak{t}$ of $\mathfrak{g}$ generated by $\mathfrak{t}_0$ is a Cartan subalgebra of $\mathfrak{g}$ and consists of all diagonal matrices of trace 0. Let E_{ij} denote the matrix $(\delta_{ai}\delta_{bj})_{1 \leq a \leq n, 1 \leq b \leq n}$, and for each $H \in \mathfrak{t}$ let $e_i(H)$ denote the ith diagonal element in H. Then

$$[H, E_{ij}] = (e_i(H) - e_j(H))\, E_{ij}, \qquad H \in \mathfrak{t}, \tag{2}$$

so the linear function $\alpha_{ij} : H \to e_i(H) - e_j(H)$ is a root of $\mathfrak{g}$ with respect to $\mathfrak{t}$. In this manner we obtain $n(n-1)$ nonzero roots; on the other hand, the set Δ^* of all nonzero roots (of $\mathfrak{g}$ with respect to $\mathfrak{t}$) contains $\dim \mathfrak{g} - \dim \mathfrak{t} = n^2 - 1 - (n-1) = n^2 - n$ elements. Consequently

$$\Delta^* = \{\alpha_{ij} : 1 \leqslant i \neq j \leqslant n\}.$$

If $H, H' \in \mathfrak{t}$, then (2) implies that

$$B(H, H') = \operatorname{Tr}(\operatorname{ad} H \operatorname{ad} H') = \sum_{1 \le i,j \le n} (e_i(H) - e_j(H))(e_i(H') - e_j(H'))$$

$$= \sum_{i,j} e_i(HH') + e_j(HH') - \sum_{i,j} e_j(H)\, e_i(H') + e_i(H)\, e_j(H')$$

so

$$B(H, H') = 2n \operatorname{Tr}(HH'). \tag{3}$$

Now given a matrix $X \in \mathfrak{su}(n)$, there exists an element $u \in \boldsymbol{SU}(n)$ such that $uXu^{-1} \in \mathfrak{t}_0$. This well-known fact about matrices is a special case of Theorem 6.4, Chapter V. Since the mapping $X \to uXu^{-1}$ is an automorphism of $\mathfrak{su}(n)$ it follows that

$$B(X, Y) = 2n \operatorname{Tr}(XY)$$

for all $X, Y \in \mathfrak{su}(n)$, hence for all $X, Y \in \mathfrak{sl}(n, \boldsymbol{C})$. Let H_{ij} denote the vector in $\mathfrak{t}$ determined by

$$B(H_{ij}, H) = \alpha_{ij}(H), \qquad H \in \mathfrak{t}.$$

Then

$$H_{ij} = \frac{1}{2n}(E_{ii} - E_{jj}).$$

The case $n = 3$. Here $\dim \mathfrak{t}_0 = 2$ and the roots in Δ^* are given by $\pm\alpha_{12}$, $\pm\alpha_{13}$, $\pm\alpha_{23}$. The angle θ between the vectors H_{ij} and H_{kl} is given by

$$\cos\theta = \frac{B(H_{ij}, H_{kl})}{(B(H_{ij}, H_{ij})\, B(H_{kl}, H_{kl}))^{1/2}}.$$

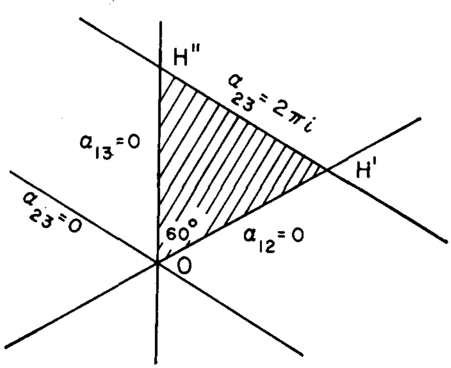

FIG. 2

The lines $\alpha_{12} = 0$, $\alpha_{13} = 0$, $\alpha_{23} = 0$ in $\mathfrak{t}_0$ are therefore situated as in Fig. 2. For the polyhedron P_0 we can select the triangle formed by the lines $\alpha_{12} = 0$, $\alpha_{13} = 0$ and $\alpha_{23} = 2\pi i$. The vectices of this triangle are the origin and the points

$$H' = i\frac{2\pi}{3}(E_{11} + E_{22} - 2E_{33})$$

$$H'' = i\frac{2\pi}{3}(- E_{11} + 2E_{22} - E_{33}).$$

If follows from Cor. 6.6 that the center $\tilde{Z}$ of $\tilde{U}$ has order 3 in this case. Moreover, since $\exp H' \neq e$ and $\exp H'' \neq e$, we see that $\mathfrak{t}_e \cap \bar{P}_0$ consists of the origin alone. In view of Theorem 6.1, the group $\boldsymbol{SU}(3)$ is simply connected.

Corollary 6.6 shows, that the order of $\tilde{Z}$ (in É. Cartan's terminology "indice de connexion") can be determined from the root pattern of $\mathfrak{u}$ as is done above in a very simple case. In Cartan's paper "La géométrie des groupes simples," *Annali di Mat.* **4** (1927), this method is used to determine the order of $\tilde{Z}$ corresponding to all simple compact Lie algebras. The result is restated in Chapter IX.

A simple complement to Theorem 6.1 and Corollary 6.6 shows that the group $\pi_1(U)$ itself and not just its order, can be read off from the diagram and $\mathfrak{t}_e$.

Theorem 6.7. *Let $\tilde{\mathfrak{t}}_e$ denote the unit lattice for the group $\tilde{U}$. Considering $\tilde{\mathfrak{t}}_e$, $\mathfrak{t}_e$ and $\mathfrak{t}(\mathfrak{u})$ as groups of translations of $\mathfrak{t}_0$, the following isomorphisms hold:*

$$\pi_1(U) \approx \mathfrak{t}_e/\tilde{\mathfrak{t}}_e, \qquad \tilde{Z} \approx \mathfrak{t}(\mathfrak{u})/\tilde{\mathfrak{t}}_e.$$

Proof. Let $\tilde{T}$ denote the analytic subgroup of $\tilde{U}$ with Lie algebra $\mathfrak{t}_0$. Then the mapping $\exp : \mathfrak{u} \to \tilde{U}$ induces a homomorphism α of $\mathfrak{t}_0$ onto $\tilde{T}$. Then $\alpha(\mathfrak{t}(\mathfrak{u})) = \tilde{Z}$, $\tilde{\mathfrak{t}}_e = \alpha^{-1}(e)$ and $\alpha(\mathfrak{t}_e) = Z_1$ where Z_1 is the subgroup of $\tilde{Z}$ such that $U = \tilde{U}/Z_1$. Since $\pi_1(U) \approx Z_1$ the theorem follows immediately.

Corollary 6.8. *The diagram $D(\mathfrak{u})$ and the unit lattice $\mathfrak{t}_e$ determine the group U up to isomorphism.*

In fact, the Lie algebra $\mathfrak{u}$ is determined by $\mathfrak{t}_0$ and $D(\mathfrak{u})$ up to isomorphism (Theorem 5.4, Cor. 7.3, Chapter III). The group $\tilde{U}$ is determined by $\mathfrak{u}$ up to isomorphism and $U = \tilde{U}/\alpha(\mathfrak{t}_e)$ in the notation above.

§ 7. Application to the Symmetric Space U/K

We return now to the situation in §1-§3 where $(\mathfrak{u}, \theta)$ is an arbitrary orthogonal symmetric Lie algebra of the compact type. The notations of §1-§3 will be preserved. In particular exp maps $\mathfrak{u}$ onto U and Exp maps $\mathfrak{p}_*$ onto U/K.

The *unit lattice in* $\mathfrak{h}_{\mathfrak{p}_*}$ is defined as the set

$$(\mathfrak{h}_{\mathfrak{p}_*})_e = \{H \in \mathfrak{h}_{\mathfrak{p}_*} : \exp H = e\}.$$

If H belongs to this set then $\operatorname{Ad}(\exp H)$ is the identity mapping of $\mathfrak{u}$ so $\alpha(H) \in 2\pi i \boldsymbol{Z}$ for each root $\alpha \in \Delta$. Hence if we consider $(\mathfrak{h}_{\mathfrak{p}_*})_e$ as a group of translations of $\mathfrak{h}_{\mathfrak{p}_*}$, each element in this group leaves the diagram $D(U, K)$ invariant and permutes the components of $(\mathfrak{h}_{\mathfrak{p}_*})_r = \mathfrak{h}_{\mathfrak{p}_*} - D(U, K)$. The mapping

$$\Phi : (kM, H) \to \operatorname{Exp} \operatorname{Ad}(k) H$$

maps $K/M \times \mathfrak{h}_{\mathfrak{p}_*}$ onto U/K. We recall that the singular set in U/K is defined as the image

$$S_{U/K} = \Phi(U/K \times D(U, K))$$

and the regular set is defined as the complement

$$(U/K)_r = U/K - S_{U/K}.$$

Lemma 7.1. *Suppose the pair* (U, K) *associated with* $(\mathfrak{u}, \theta)$ *is a symmetric pair. Then the mapping* Φ *maps the subset* $K/M \times (\mathfrak{h}_{\mathfrak{p}_*})_r$ *regularly onto* $(U/K)_r$.

Proof. Since the regularity is already established (Prop. 3.2), we only have to prove that if $H_s \in D(U, K)$ and $H_r \in (\mathfrak{h}_{\mathfrak{p}_*})_r$, then the relation

$$\operatorname{Exp} \operatorname{Ad}(k_s) H_s = \operatorname{Exp} \operatorname{Ad}(k_r) H_r \tag{1}$$

is impossible for $k_r, k_s \in K$. Relation (1) implies for $k = k_r^{-1} k_s$ that

$$k \exp H_s k_1 = \exp H_r \tag{2}$$

for a suitable $k_1 \in K$. Applying the automorphism of U which corresponds to θ we obtain

$$k \exp 2H_s k^{-1} = \exp 2H_r.$$

The centralizers of the elements $\exp 2H_r$ and $\exp 2H_s$ in U are therefore isomorphic. This contradicts Lemma 4.3 because $\alpha(2H_s) \in 2\pi i\boldsymbol{Z}$ whenever $\alpha(2H_r) \in 2\pi i\boldsymbol{Z}$ but not conversely.

The lemma shows that Φ is a local homeomorphism of $K/M \times (\mathfrak{h}_{\mathfrak{p}_*})_r$ onto $(U/K)_r$. It follows easily that if $p(t)$, $0 \leqslant t \leqslant 1$, is a continuous curve in $(U/K)_r$ and if the point $q_0 \in K/M \times (\mathfrak{h}_{\mathfrak{p}_*})_r$ satisfies $\Phi(q_0) = p(0)$ then there exists a unique continuous curve $q(t)$, $0 \leqslant t \leqslant 1$, in $K/M \times (\mathfrak{h}_{\mathfrak{p}_*})_r$ such that $\Phi(q(t)) = p(t)$ for all t and $q(0) = q_0$. In analogy with the terminology of covering spaces, $q(t)$ is called a *lift* of $p(t)$.

Theorem 7.2. *Let σ be an analytic involutive automorphism of a compact simply connected Lie group. Then the set of fixed points of σ is connected.*

Proof. A compact simply connected Lie group is necessarily semisimple (Prop. 6.6, Chapter II). In terms of the notation of Lemma 7.1, it suffices therefore to prove that U/K is simply connected if U is simply connected as we now assume.

Let Q_0 denote a component of $(\mathfrak{h}_{\mathfrak{p}_*})_r$ whose closure contains the origin. Let (k^*M, H^*) be a fixed point in $K/M \times Q_0$ and put $p^* = \Phi(k^*M, H^*)$. Let $\gamma_0(t)$ $(0 \leqslant t \leqslant 1)$ be a continuous curve in U/K which begins and ends at the point $o = \{K\}$. Due to Prop. 9.4 in the Appendix, γ_0 is homotopic to a path $\gamma(t)$ $(0 \leqslant t \leqslant 1)$ which lies in $(U/K)_r$ except for the point $o = \gamma(0) = \gamma(1)$. Since $(U/K)_r$ is connected (Cor. 9.5), we may assume that $\gamma(\frac{1}{2}) = p^*$. Now let $0 < \epsilon < \frac{1}{2}$ and consider the path γ_ϵ given by $\gamma_\epsilon(t) = \gamma(t)$, $(\epsilon \leqslant t \leqslant 1 - \epsilon)$. Let $\Gamma_\epsilon(t)$ be the lift of γ_ϵ to $K/M \times (\mathfrak{h}_{\mathfrak{p}_*})_r$ such that $\Gamma_\epsilon(\frac{1}{2}) = (k^*M, H^*)$. Put $\Gamma_\epsilon(t) = (k_tM, H_t)$ for $\epsilon \leqslant t \leqslant 1 - \epsilon$. Then, since Γ_ϵ is connected, $H_t \in Q_0$ for $\epsilon \leqslant t \leqslant 1 - \epsilon$. Let $0 < \epsilon_1 < \epsilon_2 < \frac{1}{2}$. By the uniqueness of the lift we have

$$\Gamma_{\epsilon_1}(t) = \Gamma_{\epsilon_2}(t) \qquad \text{for } \epsilon_2 \leqslant t \leqslant 1 - \epsilon_2.$$

Consequently we can define a continuous curve

$$\Gamma(t) = (k_tM, H_t), \qquad 0 < t < 1,$$

in $K/M \times Q_0$ such that $\Phi(\Gamma(t)) = \gamma(t)$. The set Q_0 is bounded since all roots are bounded on it. Let H_0 be any limit point of $\{H_t\}$ as $t \to 0$. Then there exists a sequence $t_n \to 0$ such that $k_{t_n} \to k_0 \in K$ and $H_{t_n} \to H_0$. It follows that

$$\gamma(t_n) = \operatorname{Exp} \operatorname{Ad}(k_{t_n}) H_{t_n} \to \operatorname{Exp}(\operatorname{Ad}(k_0)) H_0.$$

On the other hand, $\gamma(t_n) \to o$ so

$$\operatorname{Exp} \operatorname{Ad}(k_0) H_0 = o, \qquad \text{whence } \exp H_0 \in K.$$

Consequently $\exp 2H_0 = e$, i.e., $2H_0 \in (\mathfrak{h}_{\mathfrak{p}_*})_e$.

We shall now apply Theorem 6.1 for $\mathfrak{t}_0 = \mathfrak{h}_{\mathfrak{k}_0} + \mathfrak{h}_{\mathfrak{p}_*}$. Since H_0 lies in the closure of Q_0 each root $\alpha \in \Delta^*$ has its values in the closed interval $[0, 2\pi i]$ on the segment joining 0 and $2H_0$. It follows that this segment is contained in a closed polyhedron $\bar{P}_0$, if P_0 is a suitably chosen component of $\mathfrak{t}_r$. Since $\alpha(2H_0) \in 2\pi i \boldsymbol{Z}$ for all $\alpha \in \Delta^*$, $2H_0$ is a vertex of $\bar{P}_0$. The group U being simply connected, Theorem 6.1 implies that $H_0 = 0$. Consequently $\lim_{t \to 0} H_t = 0$. In the same way we find that the limit $H_1 = \lim_{t \to 1} H_t$ exists and $H_1 = 0$. The curve H_t $(0 \leqslant t \leqslant 1)$ is therefore a closed curve in $\bar{Q}_0$. The mapping

$$\alpha(s, t) = sH_t, \qquad 0 \leqslant s \leqslant 1, \qquad 0 \leqslant t \leqslant 1,$$

is a homotopy of this curve and 0. The mapping β given by

$$\beta(s, t) = \operatorname{Exp}(\operatorname{Ad}(k_t)\, sH_t), \qquad 0 \leqslant s \leqslant 1, \qquad 0 < t < 1,$$
$$\beta(s, 0) = \beta(s, 1) = o, \qquad 0 \leqslant s \leqslant 1,$$

is then a homotopy of γ and o. This proves the theorem.

Remark. Theorem 7.2 does not hold if the assumption of simple connectedness is dropped. As an example let

$$s_0 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and let σ denote the involutive automorphism $u \to s_0 u s_0$ of $\boldsymbol{SO}(3)$. The set K of fixed points consists of two components.

§ 8. Classification of Locally Isometric Spaces

Let $(\mathfrak{u}, \theta)$ be an orthogonal symmetric Lie algebra of the compact type. We recall that a Riemannian globally symmetric space M is said to be associated with $(\mathfrak{u}, \theta)$ if the Riemannian symmetric pair $(I_0(M), K)$ is associated with $(\mathfrak{u}, \theta)$, K being the isotropy subgroup of $I_0(M)$ at some point in M. We shall now return to the problem stated in §1, namely, to find all M associated with the given $(\mathfrak{u}, \theta)$. For this purpose we shall use Theorem 7.2.

Theorem 8.1. *Let* $(\mathfrak{u}, \theta)$ *be an orthogonal symmetric Lie algebra of the compact type and suppose that* $\mathfrak{k}_0$*, the fixed point set of* θ*, contains no ideal* $\neq \{0\}$ *in* $\mathfrak{u}$*. Let* $\tilde{U}$ *denote the simply connected Lie group with Lie algebra* $\mathfrak{u}$*, let* $\tilde{\theta}$ *denote the automorphism of* $\tilde{U}$ *such that* $d\tilde{\theta} = \theta$ *and let* $\tilde{K}$ *denote the set of fixed points of* $\tilde{\theta}$*. Let* $\tilde{Z}$ *denote the center of* $\tilde{U}$*.*

Let S *be any subgroup of* $\tilde{Z}$ *and put*

$$K_S = \{u \in \tilde{U} : u^{-1}\tilde{\theta}(u) \in S\}.$$

The Riemannian globally symmetric spaces M *associated with* $(\mathfrak{u}, \theta)$ *are exactly the spaces* $M = U/K$ *with any* U*-invariant metric, where*

$$U = \tilde{U}/S, \qquad K = K^*/K^* \cap S. \tag{1}$$

Here S *varies through all subgroups of* $\tilde{Z}$ *and* K^* *varies through all subgroups of* $\tilde{U}$ *such that* $\tilde{K} \subset K^* \subset K_S$*.*

Proof. Consider first a pair (U, K) given by (1).

$$\begin{array}{ccc}
\tilde{U} & & K^* \\
\downarrow{\scriptstyle\varphi} & & \downarrow{\scriptstyle\varphi} \\
U = \tilde{U}/S & & K = K^*/K^* \cap S \\
\downarrow{\scriptstyle\psi} & & \downarrow{\scriptstyle\psi} \\
U' = U/N & & K' = K/N
\end{array}$$

The groups K^* and K_S are clearly compact. Let φ denote the natural projection of $\tilde{U}$ onto U. Then $\varphi(K^*) = K$. The group $N = \varphi(\tilde{K} \cap \tilde{Z})$ is a normal subgroup of U contained in K. Let $U' = U/N$, $K' = K/N$. We shall now construct an involutive automorphism θ' of U' such that $d\theta' = \theta'$ and θ leaves K' pointwise fixed. Since $U/K = U'/K'$, this will prove that U/K is globally symmetric. Let ψ denote the natural projection of U onto U/N.

Let $u' \in U'$; choose $\tilde{u} \in \tilde{U}$ such that $\psi\varphi\tilde{u} = u'$ and put $\theta' u' = \psi\varphi\tilde{\theta}\tilde{u}$. To show that this is a valid definition, suppose $\tilde{u}_1, \tilde{u}_2 \in \tilde{U}$ such that $\psi\varphi\tilde{u}_1 = \psi\varphi\tilde{u}_2 = u'$. Then $\varphi\tilde{u}_1 = \varphi(\tilde{u}_2)\varphi(k)$ where $k \in \tilde{K} \cap \tilde{Z}$ so $\tilde{u}_1 = \tilde{u}_2 ks$ where $s \in S$. It follows that $\psi\varphi\tilde{\theta}\tilde{u}_1 = (\psi\varphi\tilde{\theta}\tilde{u}_2)(\psi\varphi\tilde{\theta}s)$ so we just have to prove

$$\varphi\tilde{\theta}s \in \varphi(\tilde{K} \cap \tilde{Z}).$$

However, $\varphi(s) = e$ and $s\tilde{\theta}(s) \in \tilde{K} \cap \tilde{Z}$ so $\varphi\tilde{\theta}s = \varphi(s\tilde{\theta}s) \in \varphi(\tilde{K} \cap \tilde{Z})$. This shows that θ' is a well defined mapping of U' into itself. Moreover, θ' is an involutive automorphism and $d\theta' = d\tilde{\theta} = \theta$. Finally, let $k' \in K'$.

Then there exists an element $k^* \in K^*$ such that $k' = \psi\varphi k^*$. Then since $K^* \subset K_S$ we have for a suitable element $s \in S$,

$$\theta'(k') = \psi\varphi\tilde{\theta}(k^*) = \psi\varphi(k^*s) = \psi\varphi k^* = k'.$$

This proves that the manifold U/K is globally symmetric. Furthermore, the set of elements in U which induce the identity mapping of U/K is a closed subgroup D of K whose Lie algebra is 0, being an ideal of $\mathfrak{u}$ contained in $\mathfrak{k}_0$. It follows that U/D is a semisimple subgroup of the isometry group $I(U/K)$, so by Theorem 4.1, Chapter V, $U/D = I_0(U/K)$. In particular, $I(U/K)$ has Lie algebra $\mathfrak{u}$, so the space U/K is associated with $(\mathfrak{u}, \theta)$.

On the other hand, suppose M is a Riemannian globally symmetric space associated with $(\mathfrak{u}, \theta)$. This means that there exists a point $o \in M$ such that the automorphism $\sigma : u \to s_o u s_o$ of $I_0(M)$ has differential $d\sigma = \theta$. If $U = I_0(M)$ and K denotes the isotropy subgroup of $I_0(M)$ at o, then the pair (U, K) is a Riemannian symmetric pair associated with $(\mathfrak{u}, \theta)$. There exists a subgroup S of $\tilde{Z}$ such that $U = \tilde{U}/S$. Let φ denote the natural mapping $\tilde{U} \to U$ and put $K^* = \varphi^{-1}(K)$. Then $K = K^*/K^* \cap S$. Since $\sigma \circ \varphi = \varphi \circ \tilde{\theta}$, we find for $k \in K^*$,

$$\varphi(\tilde{\theta}k) = \sigma\varphi(k) = \varphi(k).$$

Consequently, $k^{-1}\tilde{\theta}(k) \in S$ so $K^* \subset K_S$. On the other hand, $\tilde{K} \subset K^*$ since $\tilde{K}$ is connected. This finishes the proof.

Taking $K^* = \tilde{K}$ in (1) yields the following:

Corollary 8.2. *Let* $(\mathfrak{u}, \theta)$ *be as in Theorem 8.1 and let* (U, K) *be any pair associated with* $(\mathfrak{u}, \theta)$, *for which* K *is connected. Then* U/K *is globally symmetric.*

Corollary 8.3. *If* $\tilde{Z}$ *consists of the identity element alone,*† *then* $\tilde{U}/\tilde{K}$ *is the only Riemannian globally symmetric space* M *associated with* $(\mathfrak{u}, \theta)$.

§ 9. Appendix. Results from Dimension Theory

In this section we collect some results from dimension theory which have been used earlier in this chapter. The dimension concept is here the topological dimension of Brouwer, Menger, and Urysohn, defined for all separable metric spaces. This definition assigns to the empty set dimension -1 and by induction the dimension of an arbitrary

† This is the case for the exceptional algebras $\mathfrak{e}_8$, $\mathfrak{f}_4$, and $\mathfrak{g}_2$.

separable metric space M is defined as the smallest integer n for which each point $p \in M$ has arbitrarily small neighborhoods with boundaries of dimension less than n. Whenever possible, we refer to the book of Hurewicz and Wallman [1] for proofs of the results below. *All spaces considered are assumed to be separable metric spaces.* An n-dimensional manifold has topological dimension n (Hurewicz and Wallman [1], p. 46).

Proposition 9.1. *If M is a subspace of N then*

$$\dim M \leqslant \dim N.$$

For the proof, see Hurewicz and Wallman [1], p. 26.

Theorem 9.2. *Suppose a space M is a countable union $M = \bigcup_n M_n$ of closed subspaces M_n. Then*

$$\dim M \leqslant \sup_n M_n.$$

For the proof, see Hurewicz and Wallman [1], p. 30.

Lemma 9.3.† *Let M and N be differentiable manifolds and f a differentiable mapping of M into N. Then*

$$\dim f(Q) \leqslant \dim Q$$

for each subset $Q \subset M$.

Proof. Let m and n denote the dimensions of M and N, respectively. Let $p \in M$ and (B, φ) a local chart around p. The set B is called an open ball if φ can be chosen such that $\varphi(B)$ is an open ball in $\boldsymbol{R}^m$ with center $\varphi(p)$. Since f is continuous there exists a countable family $B_1, B_2, \ldots$ of open balls in M such that $M = \bigcup_i B_i$ and for each i, $f(B_i)$ is contained in an open ball B_i' in N. Now $Q = \bigcup_i (Q \cap B_i)$ and $f(Q) = \bigcup_i (f(Q) \cap B_i')$ so due to Theorem 9.2 we may assume that $M = \boldsymbol{R}^m$ and $N = \boldsymbol{R}^n$ and that Q is a bounded subset in $\boldsymbol{R}^m$. If $|\;|$ denotes the norm in $\boldsymbol{R}^m$ (and $\boldsymbol{R}^n$) we have

$$|f(x) - f(y)| < c\,|x - y| \tag{1}$$

for all x, y in some cube containing Q, c being a constant. Let $q = \dim Q$. Then the $(q+1)$-dimensional Hausdorf measure of Q is 0 (Hurewicz and Wallman [1], p. 105). From (1) follows that the $(q+1)$-dimensional Hausdorf measure of $f(Q)$ is 0 and therefore $\dim f(Q) \leqslant q$.

† Harish-Chandra [6], p. 615.

Proposition 9.4.† *Let M be a connected manifold and let S be a closed subset of M such that dim $S \leqslant$ dim $M - 2$. Let $\gamma(t)$, $0 \leqslant t \leqslant 1$, be a continuous curve in M. Then γ is homotopic to a continuous curve $\gamma'(t), 0 \leqslant t \leqslant 1$, such that $\gamma'(t) \in M - S$ for $0 < t < 1$.*

Proof. Let B be an open ball in M and let a, b be two points in $B - S$. We shall first show that a and b can be joined by means of a continuous curve, not intersecting S. Let B' be an open ball, "concentric" to B such that $a, b \in B'$ and $\bar{B}' \subset B$. The "central projection" with respect to b gives a differentable mapping of B onto the boundary of B'. The image of $B \cap S$ under this mapping contains no open subset of the boundary of B' (Lemma 9.3). Therefore, if N_a is a neighborhood of a in B such that $N_a \cap S = \emptyset$, there exists a point $a' \in N_a$ such that the segment from b to a' is disjoint from S. Combining this with the segment aa' we obtain the desired path between a and b.

Suppose now that the end points $\gamma(0)$ and $\gamma(1)$ do not belong to S. Then there exists a positive number $\epsilon = 1/(2n)$ ($n =$ integer) such that each segment $\gamma(t)$, $| t - t_0 | \leqslant \epsilon$ is contained in an open ball $B(t_0)$ ($| t_0 | \leqslant 1$). Let $0 < j < n$. Then the point $\gamma(2j\epsilon)$ belongs to the intersection $B((2j - 1)\,\epsilon) \cap B((2j + 1)\,\epsilon)$. We replace $\gamma(2j\epsilon)$ by another point $\gamma'(2j\epsilon)$ in this intersection, not belonging to S. Finally we put $\gamma'(0) = \gamma(0)$ and $\gamma'(1) = \gamma(1)$. From the first part of the proof follows that the points $y'(2j\epsilon)$ and $\gamma'((2j + 2)\,\epsilon)$ can be connected by a path in $B((2j + 1)\,\epsilon)$, not intersecting S. The desired path γ' is obtained by combining these small paths. Then γ is homotopic to γ', since each ball is simply connected.

Finally suppose that at least one of the points $\gamma(0)$, $\gamma(1)$ belongs to S. Suppose for example that $\gamma(0) \in S$, $\gamma(1) \notin S$. Then there exists a sequence $x_1, x_2, \ldots$ in $M - S$, converging to $\gamma(0)$. Select N so large that all x_n ($n \geqslant N$) and all $\gamma(t)$ ($0 \leqslant t \leqslant 1/N$) belong to a ball B around $\gamma(0)$. Combining the part of γ from $\gamma(1/N)$ to $\gamma(1)$ with a curve from x_N to $\gamma(1/N)$ we obtain a curve δ from x_N to $\gamma(1)$ whose end points do not lie in S. In view of the result just proved there exists a curve δ' in $M - S$ homotopic to δ. If we combine δ' with a sequence of suitable paths in $B - S$ joining x_n and x_{n+1} ($n \geqslant N$), we obtain the desired path γ'.

Corollary 9.5. *Under the assumptions of Prop. 9.4, the set $M - S$ is connected.*

According to Hurewicz and Wallman [1], p. 48, this corollary holds even if S is not closed.

† Compare Pontrjagin [1], p. 263, Teil 2.

Proposition 9.6. *Let M be a connected manifold and S a connected submanifold of M. We assume that $\dim S \leqslant \dim M - 3$ and that S is a topological subspace of M. Let $\gamma(t)$ and $\gamma'(t)$ $(0 \leqslant t \leqslant 1)$ be two continuous curves in the complement $M - S$. Then if γ and γ' are homotopic in M they are also homotopic in $M - S$.*

Proof. The homotopy $\gamma \sim \gamma'$ can be broken up into a sequence of homotopies

$$\gamma = \Gamma_0 \sim \Gamma_1 \sim \ldots \sim \Gamma_{n-1} \sim \Gamma_n = \gamma',$$

where, for each i, the curves $\Gamma_{i-1}(t)$ and $\Gamma_i(t)$ coincide except on a subinterval I_i of $0 \leqslant t \leqslant 1$ for which $\Gamma_{i-1}(I_i)$ and $\Gamma_i(I_i)$ lie in the same open ball (compare Seifert and Threlfall [1], §44). This means, roughly speaking, that every deformation is a finite sequence of small deformations. We can therefore assume that γ lies in an open ball V and that γ' reduces to a point. Since S is a topological subspace of M we may also assume (Prop. 3.2, Chapter I), that V and the coordinates $\{x_1, \ldots, x_m\}$ on V are such that $S \cap V$ is the submanifold of V given by $x_1 = x_2 = x_3 = 0$. Thus it can be assumed that $M = \boldsymbol{R}^m$ and that S is the subspace given by $x_1 = 0$, $x_2 = 0$, $x_3 = 0$. If $| \ |$ denotes the norm in M let $C(S)$ denote the set of points in M whose distance from S equals $|\gamma(0)|$. Then $C(S)$ is homeomorphic to $\boldsymbol{S}^2 \times \boldsymbol{R}^{m-3}$, in particular, $C(S)$ is simply connected. Now $M - S$ can be mapped onto $C(S)$ by "central projection" φ from S. This mapping is defined as follows: if $p \in M - S$ let $s(p)$ denote the unique point in S at shortest distance from p. The ray from $s(p)$ through p intersects $C(S)$ at a point which we call $\varphi(p)$. The image of γ under φ is homotopic in $M - S$ to γ and since $C(S)$ is simply connected, $\varphi \cdot \gamma$ is homotopic in $C(S) \subset M - S$ to the point $\gamma(0)$. This finishes the proof.

EXERCISES

1. Find the centralizer $\mathfrak{m}_0$ and the Weyl group $W(U, K)$ for the Riemannian symmetric pair (U, K) where $U = \boldsymbol{SU}(n)$, $K = \boldsymbol{SO}(n)$.

2*. The number of walls of a Weyl chamber in $\mathfrak{h}_{\mathfrak{p}_*}$ equals $\dim(\mathfrak{h}_{\mathfrak{p}_*})$ (É. Cartan [11], Weyl [2]).

3. Let U be a compact semisimple connected Lie group. Let the notation be as in §4-§5.

(i) Deduce from Lemma 4.5 that

$$\exp(s_\alpha H_0 - H_0) = e \qquad \text{if} \quad \alpha \in \Delta^*,\ H_0 \in \mathfrak{t}_\alpha,$$

and s_α denotes the reflection in the plane $\alpha(H) = 0$.

(ii) Let n be an integer and suppose H_0 lies on the plane $\alpha(H) = 2\pi i n$. Then the reflection σ_α in this plane can be written

$$\sigma_\alpha(H) = s_\alpha(H) - s_\alpha(H_0) + H_0.$$

(iii) Consider the unit lattice $\mathfrak{t}_e$ as a group of translations of $\mathfrak{t}_0$ and let Γ denote the group generated by the Weyl group $W(U)$ and $\mathfrak{t}_e$. By (i) and (ii) $\sigma_\alpha \in \Gamma$.

(iv) The group generated by the reflections in the planes of the diagram acts transitively on the set of components of $\mathfrak{t}_r$. (Proceed as in the first part of Theorem 2.12.)

4. The notation being the same as in Exercise 3, suppose now that U is simply connected.

(i) Deduce from Theorem 2.12 and Theorem 6.1 that Γ is simply transitive on the set of components of $\mathfrak{t}_r$.

(ii) The group Γ coincides with the group generated by the reflections in the planes of the diagram.

(iii) Each component of $\mathfrak{t}_r$ is a fundamental domain of Γ acting on $\mathfrak{t}_r$.

5. Let σ be an involutive automorphism of a compact connected Lie group U. Let H denote the set of fixed points. Let U be given any two-sided invariant Riemannian structure. The mapping

$$uH \rightarrow u\sigma(u^{-1})$$

is a diffeomorphism of U/H onto a closed totally geodesic submanifold of U. This submanifold is Riemannian globally symmetric with respect to the induced Riemannian structure, (É. Cartan [16]).

6. Let Σ denote the set of restrictions to $\mathfrak{h}_{\mathfrak{p}_0}$ of the roots in $\Delta_{\mathfrak{p}}$. Let Σ_1 and Σ_2 denote the set of positive elements in Σ given by means of two different orderings of the dual of $\mathfrak{h}_{\mathfrak{p}_0}$. Show that there exists a unique element s in the Weyl group $W(\mathfrak{u}, \theta)$ such that $s \cdot \Sigma_1 = \Sigma_2$. (Hint: Ordering amounts to the selection of a Weyl chamber.)

7. Each ordering of the dual of $\mathfrak{h}_{\mathfrak{p}_0}$ gives rise to an Iwasawa decomposition $\mathfrak{g}_0 = \mathfrak{k}_0 + \mathfrak{h}_{\mathfrak{p}_0} + \mathfrak{n}_0$. Using Exercise 6 deduce a conjugacy theorem for such decompositions.

8.* Let Σ be as in Exercise 6. Suppose $\mu, \lambda \in \Sigma$ such that $\mu = c\lambda$ ($c \in \mathbf{C}$). Then $c = \pm \frac{1}{2}, \pm 1, \pm 2$ (É. Cartan [10], p. 430, Harish-Chandra [10] I, p. 197).

NOTES

§2. The group $W(U)$ (see §4) was introduced by Weyl in [1], Kap. III, §4, where it was used for determining the characters of irreducible representations of U. The more general group $W(U,K)$ was defined in É. Cartan [10] and determined in the cases when U is a classical compact group; Cor. 2.13 was also verified for these cases (see also Weyl [2]). Theorem 2.15 is used without proof by Harish-Chandra [8] (where it is attributed to Chevalley) and by Kostant [1]. Theorem 2.5 is due to Hopf [1].

§3. The results of §3 are due to É. Cartan [10]; see also Harish-Chandra [6], VI, §12.

§4-§7. The dimension of the singular set S ($\dim S = \dim U - 3$) is determined in Weyl [1], Kap. IV, §1. This equality was extensively used by H. Weyl and É. Cartan. Weyl used it to prove the conjugacy theorem (Theorem 6.4 (iii), Chapter V) and the compactness of the universal covering group (Theorem 6.9, Chapter II). For these applications it would be sufficient to know that S is closed and has dimension $\leqslant \dim U - 2$ because only Prop. 9.4 is needed (see Pontrjagin [1], §64). Cartan, on the other hand, used the relation $\dim S = \dim U - 3$ to prove the more delicate Theorem 6.1, ([9], pp. 217-218), and Theorem 7.2, ([10], p. 430), which rely on the equality $\pi_1(U) = \pi_1(U - S)$. Here Cartan used Prop. 9.6, but did not enter into the difficulties which stem from the fact that S is not a manifold. That these difficulties are present can be seen from the fact (mentioned to the author by G. W. Whitehead) that Prop. 9.6 does not hold for a suitable 0-dimensional subset (Antoine's necklace) of $\boldsymbol{R}^3$. The reasoning actually gives $\pi_2(U) = 0$, (É. Cartan [20], cf. Borel [1]). It is also known that if U is simple then $\pi_3(U) = \boldsymbol{Z}$ (Bott [1]), and $\pi_4(U)$ can be read off from the diagram $D(U)$ and the unit lattice $\mathfrak{t}_e$ (Bott and Samelson [1]). The result is that $\pi_4(U)$ has two elements if the plane $\mu(H) = 2\pi i$ in $\mathfrak{t}_0$ contains a member of $\mathfrak{t}_e$, μ being the highest root with respect to a lexicographic ordering; otherwise $\pi_4(U) = 0$. In Bott and Samelson [1], Theorem 7.2 is reduced to Theorem 6.1 in a different way. Bott has also (unpublished) extended Theorem 7.2 to all automorphisms, involutive or not. A different proof (also based on Theorem 6.1) is given by Borel [8].

CHAPTER VIII

HERMITIAN SYMMETRIC SPACES

A Hermitian symmetric space is a Riemannian globally symmetric space which has a complex structure invariant under each geodesic symmetry. Examples are provided by all simply connected two-dimensional Riemannian globally symmetric spaces. We shall mostly be concerned with Hermitian symmetric spaces of the compact type and the noncompact type. These are always simply connected and have the characteristic property that their isotropy groups are not semisimple and therefore have nondiscrete centers. In §7 it is shown that the Hermitian symmetric spaces G_0/K_0 of the noncompact type are exactly the bounded symmetric domains in the space of several complex variables. Moreover, the space G_0/K_0 can always be imbedded to the compact dual U/K_0 as an open subset. The simplest instance of this imbedding is the unit disk $|z| < 1$ situated in the extended complex plane.

The three first sections deal with some basic facts concerning complex manifolds. The main notions treated are Hermitian and Kählerian structures, Ricci curvature, and the Bergman kernel function.

§ 1. Almost Complex Manifolds

Definition. Let M be a C^∞ manifold. An *almost complex structure* on M is a tensor field J of type (1, 1) such that $J(JX) = -X$ for each vector field X on M.

An almost complex structure on M thus amounts to a rule which in a differentiable fashion assigns to each $p \in M$ an endomorphism $J_p : M_p \to M_p$ such that $(J_p)^2 = -I$ for each $p \in M$. An *almost complex manifold* is a pair (M, J) where M is a C^∞ manifold and J is an almost complex structure on M.

For reasons given in Example II below it is important to consider the mapping $S : \mathfrak{D}^1(M) \times \mathfrak{D}^1(M) \to \mathfrak{D}^1(M)$ given by

$$S(X, Y) = [X, Y] + J[JX, Y] + J[X, JY] - [JX, JY] \qquad (1)$$

for $X, Y \in \mathfrak{D}^1(M)$. Using the relation

$$[fX, gY] = fg[X, Y] + f(Xg)\,Y - g(Yf)\,X$$

for $f, g \in C^\infty(M)$ it follows easily that $S(fX, gY) = fgS(X, Y)$. As customary, we identify S with the multilinear mapping $(\omega, X, Y) \to$

$\omega(S(X, Y))$ of $\mathfrak{D}_1 \times \mathfrak{D}^1 \times \mathfrak{D}^1$ into $C^\infty(M)$. Thus S is a tensor field of type $(1, 2)$, called the *torsion tensor* of the almost complex structure J. Obviously S is skew symmetric, that is, $S(X, Y) = -S(Y, X)$. If $S = 0$, the almost complex structure is said to be *integrable.*

Example I. Let $M = \boldsymbol{R}^2$, considered as a manifold with local coordinates the ordinary Cartesian coordinates (x, y). For each $p \in M$ the endomorphism of M_p given by

$$J_p : a\left(\frac{\partial}{\partial x}\right)_p + b\left(\frac{\partial}{\partial y}\right)_p \to -b\left(\frac{\partial}{\partial x}\right)_p + a\left(\frac{\partial}{\partial y}\right)_p$$

for $a, b \in \boldsymbol{R}$, has square $-I$. The tensor field $p \to J_p$, $p \in M$ is an almost complex structure on M. This almost complex structure is integrable, since $S(\partial/\partial x, \partial/\partial y) = 0$.

Example II. Let M be a complex manifold of dimension m as defined in Chapter I, §1. There exists a covering $M = \bigcup_{\alpha \in A} U_\alpha$ of M by open subsets U_α each of which is homeomorphic to an open subset of $\boldsymbol{C}^m$ under a mapping φ_α such that for each pair $\alpha, \beta \in A$, $\varphi_\beta \circ \varphi_\alpha^{-1}$ is a holomorphic mapping† of $\varphi_\alpha(U_\alpha \cap U_\beta)$ onto $\varphi_\beta(U_\alpha \cap U_\beta)$. As remarked in Chapter I, §1, we can always assume that the system $(U_\alpha, \varphi_\alpha)_{\alpha \in A}$ is maximal with this property. In that case, the system is said to be a *complex structure* on the underlying topological space M.

Let $p \in M$ and let α be an index in A such that $p \in U_\alpha$. If $q \in U_\alpha$ then $\varphi_\alpha(q) = (z_1(q), \ldots, z_m(q))$ where each $z_j(q)$ is a complex number $x_j(q) + iy_j(q)$. The mapping

$$\psi_\alpha : q \to (x_1(q), y_1(q), \ldots, x_m(q), y_m(q)), \qquad q \in U_\alpha,$$

is a homeomorphism of U_α onto an open subset of $\boldsymbol{R}^{2m}$. The collection of open charts $(U_\alpha, \psi_\alpha)_{\alpha \in A}$ on M turns M into an analytic manifold whose analytic structure is said to be *underlying* the complex structure above. Thus a complex structure has a definite underlying analytic structure. On the other hand, it can happen that two different complex structures have the same underlying analytic structure.

The tangent space M_p of the analytic manifold M has a basis given by the vectors

$$\left(\frac{\partial}{\partial x_1}\right)_p, \left(\frac{\partial}{\partial y_1}\right)_p, \ldots, \left(\frac{\partial}{\partial x_m}\right)_p, \left(\frac{\partial}{\partial y_m}\right)_p.$$

† If O and O' are open subsets in $\mathbf{C}^m$ and $\mathbf{C}^n$, respectively, then a mapping f: $O \to O'$ is called *holomorphic* if the coordinates of $f(z_1, \ldots, z_m)$ are holomorphic functions of $z_1, \ldots, z_m$.

The endomorphism $J^\alpha : M_p \to M_p$ given by

$$J^\alpha\left(\frac{\partial}{\partial x_i}\right)_p = \left(\frac{\partial}{\partial y_i}\right)_p, \qquad J^\alpha\left(\frac{\partial}{\partial y_i}\right)_p = -\left(\frac{\partial}{\partial x_i}\right)_p$$

for $1 \leqslant i \leqslant m$ satisfies $(J^\alpha)^2 = -I$. Suppose now β is another index in A such that $p \in U_\beta$. For $q \in U_\beta$ we denote the complex coordinates of $\varphi_\beta(q)$ by $(w_1(q), ..., w_m(q))$ and put $w_j(q) = u_j(q) + iv_j(q)$ $(1 \leqslant j \leqslant m)$. If we consider $(u_1, v_1, ..., u_m, v_m)$ as local coordinates on the analytic manifold M, the vectors

$$\left(\frac{\partial}{\partial u_1}\right)_p, \left(\frac{\partial}{\partial v_1}\right)_p, ..., \left(\frac{\partial}{\partial u_m}\right)_p, \left(\frac{\partial}{\partial v_m}\right)_p$$

form a basis of the tangent space M_p. The endomorphism $J^\beta : M_p \to M_p$ given by

$$J^\beta\left(\frac{\partial}{\partial u_i}\right)_p = \left(\frac{\partial}{\partial v_i}\right)_p, \qquad J^\beta\left(\frac{\partial}{\partial v_i}\right)_p = -\left(\frac{\partial}{\partial u_i}\right)_p$$

for $1 \leqslant i \leqslant m$ satisfies $(J^\beta)^2 = -I$.

Lemma 1.1. *The endomorphisms J^α and J^β are identical.*

Proof. Since the mapping $\varphi_\beta \circ \varphi_\alpha^{-1}$ is a holomorphic mapping of $\varphi_\alpha(U_\alpha \cap U_\beta)$ onto $\varphi_\beta(U_\alpha \cap U_\beta)$, each function $w_i(z_1, ..., z_m)$, $1 \leqslant i \leqslant m$, is a holomorphic function in a neighborhood of $\varphi_\alpha(p)$. This being so, the corresponding real functions $u_i(x_1, y_1, ..., x_m, y_m)$, $v_i(x_1, y_1, ..., x_m, y_m)$ satisfy the Cauchy-Riemann equations

$$\frac{\partial u_i}{\partial x_j} - \frac{\partial v_i}{\partial y_j} = 0, \qquad \frac{\partial u_i}{\partial y_j} + \frac{\partial v_i}{\partial x_i} = 0, \qquad 1 \leqslant i, j \leqslant m.$$

It follows that

$$\frac{\partial}{\partial x_j} = \sum_i \left(\frac{\partial u_i}{\partial x_j}\frac{\partial}{\partial u_i} + \frac{\partial v_i}{\partial x_j}\frac{\partial}{\partial v_i}\right) = \sum_i \left(\frac{\partial u_i}{\partial x_j}\frac{\partial}{\partial u_i} - \frac{\partial u_i}{\partial y_j}\frac{\partial}{\partial v_i}\right),$$

$$\frac{\partial}{\partial y_j} = \sum_i \left(\frac{\partial u_i}{\partial y_j}\frac{\partial}{\partial u_i} + \frac{\partial v_i}{\partial y_j}\frac{\partial}{\partial v_i}\right) = \sum_i \left(\frac{\partial u_i}{\partial y_j}\frac{\partial}{\partial u_i} + \frac{\partial u_i}{\partial x_j}\frac{\partial}{\partial v_i}\right),$$

and consequently

$$J^\beta\left(\frac{\partial}{\partial x_j}\right)_p = \left(\frac{\partial}{\partial y_j}\right)_p, \qquad J^\beta\left(\frac{\partial}{\partial y_j}\right)_p = -\left(\frac{\partial}{\partial x_j}\right)_p,$$

which proves the lemma.

In view of this lemma the endomorphism $J_p = J^\alpha = J^\beta$ is independent of the choice of local coordinates around p. The tensor field $J : p \to J_p$ is an almost complex structure on M, which we call the *canonical* almost complex structure associated to the complex structure on M.

Let (M, J) and (M', J') be almost complex manifolds and Φ a differentiable mapping of M into M'; the mapping Φ is called *almost complex* if

$$d\Phi_p \circ J_p = J'_{\Phi(p)} \circ d\Phi_p \qquad \text{for } p \in M. \tag{2}$$

Suppose now M and M' are complex manifolds and J and J' their corresponding almost complex structures. A mapping of M into M' is called *holomorphic* if its expression in terms of complex local coordinates is given by holomorphic functions. It is obvious from the Cauchy-Riemann equations that a holomorphic mapping is almost complex. On the other hand, suppose a mapping $\Phi : M \to M'$ satisfies (2). Let $\{z_1, ..., z_m\}$ and $\{w_1, ..., w_n\}$ be complex local coordinates in a neighborhood of p in M and of $\Phi(p)$ in M'. Put $z_j = x_j + iy_j$ $(1 \leqslant j \leqslant m)$, $w_k = u_k + iv_k$ $(1 \leqslant k \leqslant n)$. Then

$$u_k = \varphi_k(x_1, y_1, \cdots, x_m, y_m),$$
$$v_k = \psi_k(x_1, y_1, \cdots, x_m, y_m),$$

where φ_k and ψ_k are differentiable functions $(1 \leqslant k \leqslant n)$. Condition (2) implies that

$$\frac{\partial\varphi_k}{\partial x_j} = \frac{\partial\psi_k}{\partial y_j}, \qquad \frac{\partial\varphi_k}{\partial y_j} = -\frac{\partial\psi_k}{\partial x_j},$$

which shows that w_k is a holomorphic function of each variable z_j, and therefore, by a classical theorem on holomorphic functions, (see, e.g., Bochner and Martin [1], p. 33) w_k is a holomorphic function of $(z_1, ..., z_m)$. This shows that an almost complex mapping of a complex manifold into another is holomorphic.

Let M be a complex manifold and let J be the associated canonical almost complex structure. The tensor field J satisfies the integrability condition

$$S(X, Y) = 0, \qquad X, Y \in \mathfrak{D}^1(M), \tag{3}$$

where S is defined by (1). In fact, since S is $C^\infty(M)$-bilinear it suffices to check that (3) is satisfied in each coordinate neighborhood and there it obviously holds for the vector fields $\partial/\partial x_i$, $\partial/\partial y_j$. Thus the canonical almost complex structure associated with a complex structure is integrable. The converse is contained in the following theorem, first proved in full generality by Newlander and Nirenberg [1].

Theorem 1.2. *Let (M, J) be an almost complex manifold which satisfies the integrability condition* (3). *Then there exists a unique complex structure on M such that J is the associated almost complex structure.*

For the proof of this theorem which is too long to be given here, we refer to the cited article. However, we shall only use Theorem 1.2 in the case when M and J are assumed analytic.† Here much simpler proofs are available, see, e.g., Frölicher [1].

§ 2. Complex Tensor Fields. The Ricci Curvature

Let M be a C^∞ manifold. The set $C^\infty(M) + iC^\infty(M)$ of all complex-valued differentiable functions on M is an algebra over C, denoted $\mathfrak{C}_0$. A *complex vector field* on M is, by definition, a derivation of the algebra $\mathfrak{C}_0$. Let $\mathfrak{C}^1$ denote the set of complex vector fields on M. Then $\mathfrak{C}^1$ is a module over $\mathfrak{C}_0$; also $\mathfrak{C}^1$ is closed under the bracket operation $[X, Y] = XY - YX$, $(X, Y \in \mathfrak{C}^1)$. If s is an integer, $s \geqslant 1$, we consider the $\mathfrak{C}_0$ module

$$\mathfrak{C}^1 \times \dots \times \mathfrak{C}^1 \qquad (s \text{ times})$$

and let $\mathfrak{C}_s$ denote the $\mathfrak{C}_0$-module of all $\mathfrak{C}_0$-multilinear mappings of $\mathfrak{C}^1 \times \dots \times \mathfrak{C}^1$ into $\mathfrak{C}_0$. The elements of $\mathfrak{C}_1$ are called *complex* 1*-forms* on M. It follows from Lemma 2.3, Chapter I, that $\mathfrak{C}^1$ and $\mathfrak{C}_1$ are dual modules. In analogy with Chapter I, the $\mathfrak{C}_0$-multilinear mappings of the module

$$\mathfrak{C}_1 \times \dots \times \mathfrak{C}_1 \times \mathfrak{C}^1 \times \dots \times \mathfrak{C}^1 \qquad (\mathfrak{C}_1\ r \text{ times},\ \mathfrak{C}^1\ s \text{ times})$$

into $\mathfrak{C}_0$, are called *complex tensor fields*, contravariant of degree r, covariant of degree s. The set of these is denoted by $\mathfrak{C}^r_s$ (or $\mathfrak{C}^r_s(M)$).

The operation of conjugation in $\mathfrak{C}_0$ induces a similar operation in each $\mathfrak{C}^r_s$. If $Z \in \mathfrak{C}^1$, the complex vector field $\bar{Z}$ is defined by $\bar{Z}f = (Z\bar{f})^-$ for all $f \in \mathfrak{C}_0$. If $\omega \in \mathfrak{C}_1$, the complex 1-form $\bar{\omega}$ is defined by $\bar{\omega}(Z) = (\omega(\bar{Z}))^-$. Finally, if $\Omega \in \mathfrak{C}^r_s$, the tensor field $\bar{\Omega}$ is defined by

$$\bar{\Omega}(\omega_1, \dots, \omega_r, Z_1, \dots, Z_s) = (\Omega(\bar{\omega}_1, \dots, \bar{\omega}_r, \bar{Z}_1, \dots, \bar{Z}_s))^-$$

for $\omega_i \in \mathfrak{C}_1$, $Z_j \in \mathfrak{C}^1$. Each $X \in \mathfrak{D}^1$ can be regarded as a complex vector field on M by defining

$$XF = X\left(\frac{1}{2}(F + \bar{F})\right) + iX\left(\frac{1}{2i}(F - \bar{F})\right) \qquad (F \in \mathfrak{C}_0).$$

† Strictly speaking, Theorem 1.2 is not even necessary for our purposes. It will only be used to prove Prop. 4.2 of which an alternative proof is indicated in an exercise following Chapter VIII.

Similarly, $\mathfrak{D}_1$ can be regarded as a subset of $\mathfrak{C}_1$ and more generally, we shall regard the members of $\mathfrak{D}^r_s$ as complex tensor fields on M whenever this is called for by the context.

Let $p \in M$ and let M_p^C denote the complexification of the tangent space M_p. According to Chapter III, §6, M_p^C is a vector space over $\boldsymbol{C}$ consisting of all symbols $X + iY$ where $X, Y \in M_p$ with the vector space operations

$$(X_1 + iY_1) + (X_2 + iY_2) = (X_1 + X_2) + i(Y_1 + Y_2),$$
$$(a + ib)(X + iY) = aX - bY + i(bX + aY), \qquad a, b \in \boldsymbol{R}.$$

The elements of M_p^C are called *complex tangent vectors* at p. Each $X + iY \in M_p^C$ defines a complex linear function on $\mathfrak{C}_0$ given by

$$(X + iY)(f + ig) = Xf - Yg + i(Xg + Yf)$$

for $f, g \in C^\infty(M)$. Then

$$Z(FG) = F(p)\, ZG + G(p)\, ZF$$

for $Z \in M_p^C$ and $F, G \in \mathfrak{C}_0$. If Z is a complex vector field then the linear function $F \to (ZF)(p)$ on $\mathfrak{C}_0$ arises in this way from a complex tangent vector $Z_p \in M_p^C$. Thus, a complex vector field Z on M can be identified with a collection Z_p $(p \in M)$ of complex tangent vectors to M varying differentiably with p. The elements of $\mathfrak{C}^r_s$ can be described similarly.

Suppose now J is an almost complex structure on M. For each $p \in M$, the endomorphism J_p can be extended uniquely to a complex linear mapping of M_p^C onto itself. The extension, also denoted by J_p, then satisfies $(J_p)^2 = -I$. Now, since $J \in \mathfrak{D}^1_1 \subset \mathfrak{C}^1_1$, JZ is a complex vector field for each $Z \in \mathfrak{C}^1$. It is clear that $(JZ)_p = J_p Z_p$ for each $p \in M$.

Definition. Let (M, J) be an almost complex manifold and let Z be a complex vector field on M. Then Z is said to be of type $(1, 0)$ if $JZ = iZ$ and of type $(0, 1)$ if $JZ = -iZ$.

Every complex vector field Z on an almost complex manifold can be written as a sum

$$Z = Z_{1,0} + Z_{0,1},$$

where $Z_{1,0}$ and $Z_{0,1}$ are complex vector fields of type $(1, 0)$ and $(0, 1)$, respectively. In fact, it suffices to put $Z_{1,0} = \frac{1}{2}(Z - iJZ)$, $Z_{0,1} = \frac{1}{2}(Z + iJZ)$. In Example I in §1, the vector fields $\partial/\partial x - i\partial/\partial y$ and $\partial/\partial x + i\partial/\partial y$ are of type $(1, 0)$ and $(0, 1)$, respectively. They are usually denoted by $2\partial/\partial z$ and $2\partial/\partial \bar{z}$ because, if $f(z) = F(x, y)$ is a holomorphic function, then

$$2\frac{\partial f(z)}{\partial z} = \frac{\partial F(x, y)}{\partial x} - i\frac{\partial F(x, y)}{\partial y}$$

due to the Cauchy-Riemann equations.

Definition. Let M be a connected manifold with almost complex structure J. A Riemannian structure g on M is said to be a *Hermitian structure* if

$$g(JX, JY) = g(X, Y) \qquad \text{for } X, Y \in \mathfrak{D}^1 \tag{1}$$

and a *Kählerian structure* if in addition

$$\nabla_X \cdot J = 0 \qquad \text{for } X \in \mathfrak{D}^1. \tag{2}$$

In other words, the Hermitian condition means that J_p is an isometry of M_p for each $p \in M$. The Kählerian condition means that in addition the tensor field J is invariant under parallelism.

Let g be any Riemannian structure on a connected manifold M. Let $X \to \nabla_X$, ($X \in \mathfrak{D}^1$), be the corresponding Riemannian connection. Regarding now g as a complex tensor field, the covariant derivative ∇_Z can be defined by relation (2) in Chapter I, §9, for all $Z \in \mathfrak{C}^1$. Then

$$R(X, Y) = \nabla_X\nabla_Y - \nabla_Y\nabla_X - \nabla_{[X,Y]}, \qquad X, Y \in \mathfrak{C}^1,$$

because both sides are $\mathfrak{C}_0$-bilinear and coincide for $X, Y \in \mathfrak{D}^1$.

Lemma 2.1. *Let M be a connected manifold with almost complex structure J and Riemannian structure g.*

(i) *If g is Hermitian, then $g(X, Y) = 0$ if X and Y are both of type* $(1, 0)$, (*or both of type* $(0, 1)$).

(ii) *If g is Kählerian and R denotes the curvature tensor, then $R(X, Y) = 0$ if X and Y are both of type* $(1, 0)$, (*or both of type* $(0, 1)$).

Proof. Let X and Y be complex vector fields of type $(1, 0)$. Then, if g is Hermitian,

$$g(X, Y) = g(JX, JY) = g(iX, iY) = -g(X, Y)$$

so $g(X, Y) = 0$. Now let Z, T be arbitrary complex vector fields on M. The Kählerian condition (2) implies $\nabla_Z(JX) = J\nabla_Z(X)$. It follows that $R(Z, T)X$ is of type $(1, 0)$ as well as X. Hence by (i)

$$g(R(Z, T)X, Y) = 0. \tag{3}$$

The quadrilinear form $g(R(Z, T)\, U, V)$ on $\mathfrak{E}^1 \times \mathfrak{E}^1 \times \mathfrak{E}^1 \times \mathfrak{E}^1$ satisfies conditions (a), (b), (c) of Lemma 12.4, Chapter I. Owing to this lemma we have

$$g(R(U, V)\, Z, T) = g(R(Z, T)\, U, V)$$

and (3) implies $g(R\,(X, Y)\, Z, T) = 0$. Since Z and T are arbitrary and g_p is nondegenerate for each $p \in M$, the lemma follows.

Let M be a connected complex manifold of dimension m. Let p be a point in M and $\{z_1, ..., z_m\}$ local coordinates in an open neighborhood U of p. A complex-valued function f on M is said to be *holomorphic* at p if there exists a neighborhood of p where f is given by a convergent power series in the local coordinates $z_1 - z_1(p), ..., z_m - z_m(p)$. If f is holomorphic at each point of a set V, then f is said to be holomorphic on V. If we write $x_j = \frac{1}{2}(z_j + \bar{z}_j)$, $y_j = 1/(2i)\,(z_j - \bar{z}_j)$, then $\{x_1, y_1, ..., x_m, y_m\}$ is a coordinate system on the underlying analytic manifold U. The vector fields given by

$$\frac{\partial}{\partial z_j} = \frac{1}{2}\left(\frac{\partial}{\partial x_j} - i\frac{\partial}{\partial y_j}\right), \qquad \frac{\partial}{\partial \bar{z}_j} = \frac{1}{2}\left(\frac{\partial}{\partial x_j} + i\frac{\partial}{\partial y_j}\right)$$

$(1 \leqslant j \leqslant m)$ are complex vector fields on U of type $(1, 0)$ and $(0, 1)$, respectively. A function f which is holomorphic on U satisfies

$$\frac{\partial}{\partial \bar{z}_j} f = 0, \qquad \frac{\partial}{\partial z_j} \bar{f} = 0.$$

The differential forms

$$dz_j = dx_j + i\, dy_j, \qquad d\bar{z}_j = dx_j - i\, dy_j \qquad (1 \leqslant j \leqslant m)$$

are complex 1-forms on U. It is easily seen that

$$\overline{\left(\frac{\partial}{\partial z_j}\right)} = \frac{\partial}{\partial \bar{z}_j}, \qquad \overline{(dz_j)} = d\bar{z}_j \qquad (1 \leqslant j \leqslant m)$$

and

$$dz_i\left(\frac{\partial}{\partial z_j}\right) = d\bar{z}_i\left(\frac{\partial}{\partial \bar{z}_j}\right) = \delta^i{}_j,$$

$$dz_i\left(\frac{\partial}{\partial \bar{z}_j}\right) = d\bar{z}_i\left(\frac{\partial}{\partial z_j}\right) = 0$$

for $1 \leqslant i, j \leqslant m$. Let T be a complex tensor field on M of type $(1, 2)$. The coefficients $T^k{}_{ij}$, $T^{k^*}{}_{ij^*}$, etc., are defined by

$$T\left(dz_k, \frac{\partial}{\partial z_i}, \frac{\partial}{\partial z_j}\right) = T^k{}_{ij}, \qquad T\left(d\bar{z}_k, \frac{\partial}{\partial z_i}, \frac{\partial}{\partial \bar{z}_j}\right) = T^{k^*}{}_{ij^*},$$

and similarly for tensor fields of other types. We also write for simplicity $Z_i = \partial/\partial z_i$, $Z_{j^*} = \partial/\partial \bar{z}_j$. If g is a Hermitian structure on M we have from Lemma 2.1

$$g_{ij} = g_{i^*j^*} = 0 \tag{4}$$

and if g is Kählerian, we have by the same lemma

$$R^{\alpha}{}_{\beta ij} = R^{\alpha}{}_{\beta i^* j^*} = 0, \tag{5}$$

where α, β are arbitrary indices, starred or not. If $X \to \nabla_X$ is an affine connection on M, the functions $\Gamma_{ij}{}^k$, $\Gamma_{i^*j}{}^k$, ..., $\Gamma_{i^*j^*}{}^{k^*}$ are defined by

$$\nabla_{Z_i}(Z_j) = \sum_k \Gamma_{ij}{}^k Z_k + \sum_k \Gamma_{ij}{}^{k^*} Z_{k^*}$$

and the similar equations for $\nabla_{Z_{i^*}}(Z_j)$, $\nabla_{Z_i}(Z_{j^*})$ and $\nabla_{Z_{i^*}}(Z_{j^*})$.

Lemma 2.2. *A Hermitian structure g on a connected complex manifold is Kählerian if and only if*

$$\Gamma_{jk}{}^{l^*} = \Gamma_{j^*k}{}^{l} = \Gamma_{jk^*}{}^{l^*} = \Gamma_{j^*k^*}{}^{l} = 0 \tag{6}$$

in each coordinate neighborhood.

Proof. If g is Kählerian, then $J\nabla_{Z_i}(Z_j) = \nabla_{Z_i}(JZ_j) = i\nabla_{Z_i}(Z_j)$ so $\Gamma_{jk}{}^{l^*} = 0$ and conversely. The other relations are proved similarly.

Let M be a manifold with an affine connection having curvature tensor R. Let $p \in M$ and X, $Y \in \mathfrak{D}^1(M)$. The mapping

$$L \to R_p(Y_p, L) \cdot X_p, \qquad L \in M_p,$$

is an endomorphism of M_p whose trace will be denoted by $r_p(X_p, Y_p)$. The tensor field r given by

$$(r(X, Y))(p) = r_p(X_p, Y_p)$$

is called the *Ricci curvature* of the affine connection.

Lemma 2.3. *On a Riemannian manifold M the Ricci curvature is a symmetric tensor, that is,*

$$r(X, Y) = r(Y, X), \qquad X, Y \in \mathfrak{D}^1(M).$$

Proof. Let $p \in M$ and let $X_1, ..., X_m$ be a basis of the vector fields on an open neighborhood U of p such that $g(X_i, X_j) = \delta^i{}_j$ on U, g being the Riemannian structure. Then

$$R(X_i, X_j) \cdot X_l = \sum_k R^k{}_{lij} X_k$$

so

$$r(X_l, X_i) = \sum_k R^k{}_{lik}.$$

On the other hand, if X, Y, S, T are any vector fields on M,

$$g(R(X, Y)\, S, T) = g(R(S, T)\, X, Y)$$

so

$$R^m{}_{lij} = R^j{}_{ilm}$$

and $r(X_l, X_i) = r(X_i, X_l)$ follows.

If M is a complex manifold, we consider r as a complex tensor field on M with coefficients r_{ij}, r_{i^*j}, r_{ij^*}, $r_{i^*j^*}$ defined as above.

We recall that a manifold M is said to be *orientable* if there exists a collection $(U_\alpha, \psi_\alpha)_{\alpha \in A}$ of local charts such that $\{U_\alpha\}_{\alpha \in A}$ is a covering of M and such that for any $\alpha, \beta \in A$, the mapping $\psi_\beta \circ \psi_\alpha^{-1}$ has strictly positive Jacobian determinant in its domain of definition $\psi_\alpha(U_\alpha \cap U_\beta)$. The manifold M is said to be *oriented* if such a collection $(U_\alpha, \psi_\alpha)_{\alpha \in A}$ has been chosen.

Let M be a complex manifold of dimension m. Let $(V_\alpha, \varphi_\alpha)_{\alpha \in A}$ be a collection of local charts covering M such that $\varphi_\beta \circ \varphi_\alpha^{-1}$ is holomorphic for each pair $\alpha, \beta \in A$. Let

$$\varphi_\alpha(p) = (z_1, ..., z_m), \qquad \varphi_\beta(q) = (w_1, ..., w_m)$$

for $p \in V_\alpha$, $q \in V_\beta$ and let $z_j = x_j + iy_j$, $w_j = u_j + iv_j$ $(1 \leqslant j \leqslant m)$. From the Cauchy-Riemann equations follows easily by induction

$$\frac{\partial(u_1, v_1, ..., u_m, v_m)}{\partial(x_1, y_1, ..., x_m, y_m)} = \left| \frac{\partial(w_1, ..., w_m)}{\partial(z_1, ..., z_m)} \right|^2. \tag{7}$$

If we define ψ_α by

$$\psi_\alpha : p \to (x_1, y_1, ..., x_m, y_m),$$

then M, with the local charts $(V_\alpha, \psi_\alpha)_{\alpha \in A}$, is an oriented manifold.

Let M be an oriented manifold with a Riemannian structure g. Let $\{x_1, ..., x_m\}$ be a coordinate system valid on an open subset U of M. Let

$$g_{ij} = g\left(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}\right), \qquad \bar{g} = \det(g_{ij}).$$

Then $\bar{g} > 0$ and we can consider the m-form

$$\sqrt{\bar{g}}\, dx_1 \wedge ... \wedge dx_m \tag{8}$$

on U. If $\{y_1, ..., y_m\}$ is another coordinate system on U then it is easy to see that

$$\bar{g}(y_1, ..., y_m)^{1/2}\, dy_1 \wedge ... \wedge dy_m = \bar{g}(x_1, ..., x_m)^{1/2}\, dx_1 \wedge ... \wedge dx_m,$$

since the Jacobian determinant $(\partial y_j/\partial x_i)$ is positive. It follows that there exists an m-form ω on M, which on an arbitrary coordinate neighborhood has the expression (8). This form ω is called the *volume element* corresponding to the Riemannian structure g on the oriented manifold M.

Proposition 2.4. *Let ω be the volume element on an oriented Riemannian manifold M. Then*

$$\nabla_X \omega = 0$$

for each vector field X on M.

Proof. Let $p \in M$ and let $\{x_1, ..., x_m\}$ be a coordinate system valid in a neighborhood of p such that the tangent vectors $(\partial/\partial x_i)_p$ form an orthonormal basis of M_p. Let $X_1, ..., X_m$ be the vector fields on a normal neighborhood N_p of p adapted to this basis and let the forms $\omega^1, ..., \omega^m$ on N_p be determined by $\omega^i(X_j) = \delta^i{}_j$. Then

$$\omega = \omega^1 \wedge ... \wedge \omega^m$$

and since $\nabla_X \omega^i = 0$ at p, the relation $\nabla_X \omega = 0$ holds at p. The point p being arbitrary, the proposition follows.

Proposition 2.5. *Let M be a connected complex manifold and g a Riemannian structure on M. Let ω denote the corresponding volume element. In a local coordinate system $\{z_1, ..., z_m\}$, ω has an expression*

$$\omega = G\, dz_1 \wedge d\bar{z}_1 \wedge ... \wedge dz_m \wedge d\bar{z}_m,$$

where the function G is given by

$$(-2i)^m\, G = \bar{g}(x_1, y_1 ..., x_m, y_m)^{1/2}.$$

If g is Kählerian, the Ricci curvature satisfies

$$r_{ij^*} = Z_i Z_{j^*} \log |\, G\,|, \qquad r_{ij} = r_{i^*j^*} = 0.$$

Proof. The expression for ω is obvious since

$$dz_j \wedge d\bar{z}_j = (-2i)\, dx_j \wedge dy_j.$$

Now assume g Kählerian. The curvature tensor R satisfies

$$g(R(Z_i, Z_{j^*}) \cdot Z_l, Z_m) = g(R(Z_l, Z_m) \cdot Z_i, Z_{j^*}) \tag{9}$$

and this expression vanishes due to Lemma 2.1. Since g_p is nondegenerate for $p \in M$, it follows that

$$R(Z_i, Z_{j^*}) \cdot Z_l = \sum_k R^k{}_{lij^*} Z_k, \tag{10}$$

$$R(Z_i, Z_{j^*}) \cdot Z_{l^*} = \sum_k R^{k^*}{}_{l^*ij^*} Z_{k^*}. \tag{11}$$

Now the vector fields Z_i, Z_{j^*} make up a basis of the complex tangent space at each point. The trace of an endomorphism of a vector space is the same as the trace of the extension of the endomorphism to the complexified vector space. Therefore $r(Z_i, Z_{j^*})$ equals the trace of the complex endomorphism given by

$$Z_m \to R(Z_{j^*}, Z_m) \cdot Z_i,$$
$$Z_{m^*} \to R(Z_{j^*}, Z_{m^*}) \cdot Z_i.$$

It follows that

$$r_{ij^*} = \sum_m R^m{}_{ij^*m}. \tag{12}$$

Similarly, we find $r_{ij} = r_{i^*j^*} = 0$. Now, from Lemma 2.2 and the fact that covariant differentiation commutes with contractions, it follows that

$$\nabla_{Z_l}(dz_i) = -\sum_j \Gamma_{lj}{}^i \, dz_j, \qquad \nabla_{Z_l}(d\bar{z}_i) = 0.$$

Combining these equations with

$$\nabla_{Z_l}(\omega) = 0$$

one finds that

$$Z_l G = G \sum_i \Gamma_{li}{}^i. \tag{13}$$

On the other hand, we have from Lemma 2.2

$$\begin{aligned} R(Z_{j^*}, Z_k) \cdot Z_l &= (\nabla_{Z_{j^*}} \nabla_{Z_k} - \nabla_{Z_k} \nabla_{Z_{j^*}}) \cdot Z_l \\ &= \nabla_{Z_{j^*}} \Big(\sum_s \Gamma_{kl}{}^s Z_s \Big) = \sum_s (Z_{j^*} \Gamma_{kl}{}^s) \, Z_s. \end{aligned}$$

Comparing with (10) we find

$$R^s{}_{lj^*k} = Z_{j^*} \Gamma_{kl}{}^s. \tag{14}$$

Since $\Gamma_{kl}{}^s = \Gamma_{lk}{}^s$ we find from (12)-(14)

$$r_{ij^*} = Z_{j^*}\left(\frac{1}{G} Z_i\, G\right),$$

and since G is a constant multiple of $|G|$ the desired expression for r_{ij^*} follows.

§ 3. Bounded Domains. The Kernel Function

In this chapter, a *bounded domain* shall mean a bounded, open connected subset of the product $\boldsymbol{C}^N$, N being an integer > 0.

Let D be a bounded domain in $\boldsymbol{C}^N$. Let $L^2(D)$ denote the Hilbert space of complex functions on D for which $\int_D |f|^2\, d\mu < \infty$, the measure $d\mu$ denoting the Lebesgue measure on $\boldsymbol{R}^{2N}$. The inner product on $L^2(D)$ is

$$(f, g) = \int_D f(\zeta)\,\overline{g(\zeta)}\, d\mu(\zeta),$$

and as usual the norm is defined by $\|f\| = (f, f)^{1/2}$. Functions which coincide except on a set of measure 0 are considered as the same member of $L^2(D)$. Let $\mathfrak{H}(D)$ denote the set of functions in $L^2(D)$ which are holomorphic in D.

Proposition 3.1. *Let A be a compact subset of D. Then there exists a number N_A such that*

$$|f(z)| \leqslant N_A \|f\|$$

for all $z \in A$ and all $f \in \mathfrak{H}(D)$.

Proof. Let $\zeta = (\zeta_1, \ldots, \zeta_N) \in A$ and let $C(\zeta, \epsilon)$ be any polycylinder $|z_1 - \zeta_1| < \epsilon_1, \ldots, |z_N - \zeta_N| < \epsilon_N$ contained in D. Then the power series expansion for f,

$$f(z_1, \ldots, z_N) = \sum_{r_i \geqslant 0} a_{r_1 \ldots r_N} (z_1 - \zeta_1)^{r_1} \ldots (z_N - \zeta_N)^{r_N}$$

is absolutely convergent in $C(\zeta, \epsilon)$ (cf. Bochner and Martin [1], p. 33). Now the terms in the series are mutually orthogonal with respect to the inner product

$$(g, h)_\epsilon = \int_{C(\zeta, \epsilon)} g(z)\,\overline{h(z)}\, d\mu(z).$$

Consequently

$$\int_D |f(z)|^2\, d\mu(z) \geq \int_{C(\zeta,\varepsilon)} |f(z)|^2\, d\mu(z) \geq \int_{C(\zeta,\varepsilon)} |a_{0\ldots 0}|^2\, d\mu(z), \qquad (1)$$

so

$$\mu(C(\zeta, \epsilon))^{1/2}\, |f(\zeta)| \leqslant ||f||. \qquad (2)$$

Since A is compact, there exists a number $\epsilon > 0$ such that for each $(\zeta_1, \ldots, \zeta_N) \in A$, the polycylinder $|z_i - \zeta_i| < \epsilon$ $(1 \leqslant i \leqslant N)$ belongs to D. If the volume of this polycylinder is denoted by $(1/N_A)^2$, Prop. 3.1 follows from (2).

Corollary 3.2. *The set $\mathfrak{H}(D)$ is a closed linear subspace of $L^2(D)$, hence a Hilbert space with the inner product $\int_D f(z)\,(g(z))^-\, d\mu(z)$.*

In fact, let (f_n) be a sequence in $\mathfrak{H}(D)$ which converges to an element $f \in L^2(D)$. Then by Prop. 3.1

$$|f_n(\zeta) - f_m(\zeta)| \leqslant N_A\, ||f_m - f_n|| \qquad (3)$$

for all $\zeta \in A$. It follows that there exists a function g on D such that $f_n \to g$ uniformly on each compact subset of D. Hence g is holomorphic on D. By (3) we have for $\zeta \in A$

$$|f_n(\zeta) - g(\zeta)| \leqslant N_A\, ||f_n - f||. \qquad (4)$$

Given A, there exists an integer K such that the right-hand side of (4) is $\leqslant 1$ for $n \geqslant K$ and such that $||f_K|| \leqslant ||f|| + 1$. Then

$$\Big[\int_A |g(\zeta)|^2\, d\mu\Big]^{1/2} \leqslant \Big[\int_A |f_K(\zeta) - g(\zeta)|^2\, d\mu\Big]^{1/2} + \Big[\int_A |f_K(\zeta)|^2\, d\mu\Big]^{1/2}$$

$$\leqslant \mu(D)^{1/2} + ||f|| + 1.$$

Hence $g \in \mathfrak{H}(D)$. Finally, since

$$||f - g||^2 = \lim ||f_n - g||^2 = \int_D \lim |f_n(z) - g(z)|^2\, d\mu = 0$$

it follows that $f = g$ almost everywhere.

Theorem 3.3. *Let $\varphi_0, \varphi_1, \ldots$ be any orthonormal basis of the Hilbert space $\mathfrak{H}(D)$. Then the series*

$$\sum_0^\infty \varphi_n(z)\, \overline{\varphi_n(\zeta)}$$

converges uniformly on each compact subset of $D \times D$. *The sum, denoted* $K(z, \bar{\zeta})$, *is independent of the choice of orthonormal basis and*

$$F(z) = \int_D K(z, \bar{\zeta}) F(\zeta) \, d\mu(\zeta)$$

for each $F \in \mathfrak{H}(D)$.

Proof. Let $z \in D$. In view of Prop. 3.1, the linear functional $F \to F(z)$ on $\mathfrak{H}(D)$ is continuous. Now every continuous linear functional on a Hilbert space is representable as the inner product with some fixed vector in the space. Hence there exists a function $K_z \in \mathfrak{H}(D)$ such that

$$F(z) = \int_D F(\zeta) \overline{K_z(\zeta)} \, d\mu(\zeta) \qquad (F \in \mathfrak{H}(D)). \tag{5}$$

We now put $K(z, \bar{\zeta}) = \overline{K_z(\zeta)}$ for $z, \zeta \in D$. The vector K_z can be expressed by means of the basis $\varphi_0, \varphi_1, \ldots$,

$$K_z(\zeta) = \sum_0^\infty a_n \varphi_n(\zeta), \tag{6}$$

where the series converges in the L_2-norm and

$$a_n = \int_D K_z(\zeta) \overline{\varphi_n(\zeta)} \, d\mu(\zeta) \qquad (n = 0, 1, \ldots).$$

In view of Prop. 3.1, the series (6) converges uniformly on each compact subset of D. From (5) we have $a_n = \overline{\varphi_n(z)}$ and therefore

$$K(z, \bar{\zeta}) = \sum_0^\infty \varphi_n(z) \overline{\varphi_n(\zeta)}. \tag{7}$$

This shows that $\overline{K(\zeta, \bar{z})} = K(z, \bar{\zeta})$; consequently, for a given ζ, the function $z \to K(z, \bar{\zeta})$ belongs to $\mathfrak{H}(D)$. In order to prove that the series (7) converges uniformly on each compact subset of $D \times D$, it suffices, due to the inequality

$$2 \mid \varphi_n(z) \overline{\varphi_n(\zeta)} \mid \leqslant \mid \varphi_n(z) \mid^2 + \mid \varphi_n(\zeta) \mid^2,$$

to prove that the series $\Sigma_0^\infty \mid \varphi_n(z) \mid^2$ converges to $K(z, \bar{z})$, uniformly on each compact subset A of D. Let $\epsilon > 0$. There exists a number $\delta > 0$ such that for each $(\zeta_1, \ldots, \zeta_N) \in A$ the closed polycylinder $\mid z_1 - \zeta_1 \mid \leqslant \delta, \ldots, \mid z_N - \zeta_N \mid \leqslant \delta$ belongs to D. Let A_δ denote the union of these polycylinders. Then by (1) there exists a constant M_δ such that

$$|f(z)|^2 \leqslant M_\delta \int_{A_\delta} |f(\zeta)|^2 \, d\mu$$

for all $z \in A$ and all $f \in \mathfrak{H}(D)$. Since

$$\sum_0^\infty \int_{A_\delta} |\varphi_n(\zeta)|^2 \, d\mu = \int_{A_\delta} K(\zeta, \bar{\zeta}) \, d\mu < \infty,$$

there exists an integer P such that

$$\sum_{P+1}^\infty |\varphi_n(z)|^2 \leqslant M_\delta \sum_{P+1}^\infty \int_{A_\delta} |\varphi_n(\zeta)|^2 \, d\mu < \epsilon$$

for all $z \in A$. This concludes the proof.

Definition. The function K is called the *Bergman kernel function* for D.

Let $(z_1, \ldots, z_N)$ denote the components of $z \in \boldsymbol{C}^N$ and consider the complex tensor field H on D given by

$$H = \sum_{1 \leqslant i,j \leqslant N} Z_i Z_{j^*} \log K(z, \bar{z}) \, dz_i \otimes d\bar{z}_j.$$

Here $dz_i \otimes d\bar{z}_j$ denotes the complex tensor field

$$(X, Y) \to dz_i(X) \, d\bar{z}_j(Y), \qquad X, Y \in \mathfrak{C}^1(D),$$

which is covariant of degree 2. Let g denote the real part of the restriction of H to $\mathfrak{D}^1(D) \times \mathfrak{D}^1(D)$.

Proposition 3.4. *The tensor field g is a Riemannian structure on D which is Kählerian.*

Proof. If f is a holomorphic function on D, then

$$Z_{i^*} f = Z_{j^*} \bar{f} = 0, \qquad 1 \leqslant i \leqslant N,$$

due to the Cauchy-Riemann equations. We use this on the series

$$K(z, \bar{\zeta}) = \sum_0^\infty \varphi_n(z) \overline{\varphi_n(\zeta)},$$

which by Theorem 3.3 can be differentiated term by term. We obtain

$$Z_{j^*} \log K(z, \zeta) = \frac{1}{K} \sum_n \varphi_n(z) (Z_{j^*}\overline{\varphi_n}) (\zeta)$$

and writing $\log K$ for $\log K(z, \bar{z})$ we obtain, since

$$\left\{ \frac{\partial}{\partial z_i} \frac{\partial}{\partial \bar{\zeta}_j} \log K(z, \zeta) \right\}_{\zeta = z} = \frac{\partial}{\partial z_i} \frac{\partial}{\partial \bar{z}_j} \log K(z, \bar{z}),$$

$$Z_i Z_{j^*} \log K = \frac{1}{K^2} \left\{ \sum_n \varphi_n \overline{\varphi_n} \sum_m (Z_i \varphi_m) (Z_{j^*} \overline{\varphi_m}) - \sum_m \overline{\varphi_m} (Z_i \varphi_m) \sum_n \varphi_n (Z_{j^*} \overline{\varphi_n}) \right\},$$

so

$$Z_i Z_{j^*} \log K = \frac{1}{K^2} \sum_{n > m} (\varphi_n (Z_i \varphi_m) - \varphi_m (Z_i \varphi_n)) (\varphi_n (Z_j \varphi_m) - \varphi_m (Z_j \varphi_n))^-, \tag{8}$$

since $\overline{Z_j f} = Z_{j^*} \bar{f}$. Let $X, Y \in \mathfrak{E}^1(D)$ and put $\xi_i = dz_i(X)$, $\eta_j = d\bar{z}_j(Y)$. Then, by (8),

$$H(X, \bar{X}) = \sum_{i,j} (Z_i Z_{j^*} \log K) \xi_i \bar{\xi}_j \geqslant 0, \tag{9}$$

$$H(X, Y) = \sum_{i,j} (Z_i Z_{j^*} \log K) \xi_i \eta_j = H(\bar{Y}, \bar{X})^-. \tag{10}$$

For $X, Y \in \mathfrak{D}^1(D)$ this implies that $g(X, X) \geqslant 0$ and $2g(X, Y) = H(X, Y) + \overline{H(X, Y)} = H(X, Y) + H(Y, X)$ so $g(X, Y) = g(Y, X)$. Suppose now that $g(X, X) = 0$ at some point $p \in D$. Then, by (8), we have at the point p

$$\sum \xi_i (\varphi_n (Z_i \varphi_m) - \varphi_m (Z_i \varphi_n)) = 0 \tag{11}$$

for all n, m. Now, since D is bounded, the functions $1, z_1, \ldots, z_N$ all belong to $\mathfrak{H}(D)$. We can choose the system $\varphi_0, \varphi_1, \ldots$ in such a way that the functions $\varphi_0, \varphi_1, \ldots, \varphi_N$ are obtained from $1, z_1, \ldots, z_N$ by the usual orthonormalization process. Then the matrix (b_{ij}) given by

$$b_{ij} = Z_i \varphi_j \qquad (1 \leqslant i, j \leqslant N)$$

is an upper triangular matrix whose diagonal elements are constants $\neq 0$. Hence $\det (b_{ij}) \neq 0$. On the other hand, (11) implies, φ_0 being constant, that

$$\sum_i \xi_i (Z_i \varphi_j) = 0, \qquad 1 \leqslant j \leqslant N,$$

at the point p. It follows that all ξ_i vanish at p, so $X_p = 0$. It has now been shown that g is a Riemannian structure on D. The relations $dz_i(JX) = i dz_i(X)$ and $d\bar{z}_j(JY) = -i\, d\bar{z}_j(Y)$ imply that $g(JX, JY) = g(X, Y)$ so g is Hermitian. Now suppose g is extended to a complex tensor field. Then we have

$$2g(X, Y) = H(X, Y) + H(Y, X), \qquad X, Y \in \mathfrak{C}^1(D),$$

because both sides of this equation are complex tensor fields which coincide for $X, Y \in \mathfrak{D}^1(D)$. It follows from (9) and (10) that

$$g_{ij} = g_{i^*j^*} = 0, \qquad g_{ij^*} = \tfrac{1}{2} Z_i Z_{j^*} \log K. \tag{12}$$

If α, β, γ, δ are any indices, starred or not, we have from formula (2), Chapter I, §9,

$$2 \sum_\delta g_{\alpha\delta} \Gamma_{\beta\gamma}{}^\delta = Z_\beta g_{\alpha\gamma} + Z_\gamma g_{\alpha\beta} - Z_\alpha g_{\gamma\beta}. \tag{13}$$

Using the fact that g is nondegenerate one derives from (12) and (13)

$$\Gamma_{jk}{}^{l^*} = \Gamma_{j^*k}{}^l = \Gamma_{jk^*}{}^{l^*} = \Gamma_{j^*k^*}{}^l = 0. \tag{14}$$

By Lemma 2.2, g is Kählerian.

Definition. The metric induced by the Riemannian structure g is called the *Bergman metric* on D.

Let φ be a holomorphic diffeomorphism of a bounded domain $D \subset C^N$ onto a bounded domain $D' \subset C^N$; expressed in coordinates, we have

$$\varphi(z_1, \ldots, z_N) = (w_1(z_1, \ldots, z_N), \ldots, w_N(z_1, \ldots, z_N)).$$

Then the Jacobian determinant

$$J_\varphi = \frac{\partial(w_1, \ldots, w_N)}{\partial(z_1, \ldots, z_N)}$$

is a holomorphic function on D. For the real coordinates given by $z_j = x_j + iy_j$, $w_j = u_j + iv_j$ $(1 \leqslant j \leqslant N)$, we have

$$|J_\varphi|^2 = \frac{\partial(u_1, v_1, \ldots, u_N, v_N)}{\partial(x_1, y_1, \ldots, x_N, y_N)} \tag{15}$$

as noted earlier. Let μ and μ' denote the Euclidean measures on D and D', respectively; then

$$\mu'(\varphi(M)) = \int_M |J_\varphi|^2 \, d\mu$$

for each Borel subset M of D. Consequently, the mapping $f \to (f \circ \varphi) J_\varphi$ is an isometry of $\mathfrak{H}(D')$ onto $\mathfrak{H}(D)$. It follows that the kernel functions K and K' are related by

$$K(z, \bar{z}) = K'(\varphi(z), \overline{\varphi(z)}) \mid J_\varphi \mid^2 \qquad (z \in D). \tag{16}$$

Proposition 3.5. *Let D and D' be bounded domains in C^N and let g and g' denote the Riemannian structures on D and D' induced by the kernel functions. Then each holomorphic diffeomorphism φ of D onto D' is an isometry.*

Proof. Using the notation above, we have

$$dw_j = du_j + idv_j, \; d\bar{w}_j = du_j - idv_j.$$

Furthermore,

$$\varphi^*(du_j) = \sum_k \left(\frac{\partial u_j}{\partial x_k} dx_k + \frac{\partial u_j}{\partial y_k} dy_k\right)$$

and similarly for $\varphi^*(dv_j)$. Since

$$dw_j(d\varphi \cdot X) = \varphi^*(du_j)(X) + i\varphi^*(dv_j)(X), \qquad X \in \mathfrak{D}^1(D),$$

it follows from the Cauchy-Riemann equations that

$$dw_j(d\varphi \cdot X) = \sum_k \frac{\partial w_j}{\partial z_k} dz_k(X),$$

$$d\bar{w}_j(d\varphi \cdot X) = \sum_k \left(\frac{\partial w_j}{\partial z_k}\right)^{-} d\overline{z_k}(X).$$

Hence

$$g'(d\varphi \cdot X, d\varphi \cdot X) = \sum_{i,j} \frac{\partial^2 \log K'}{\partial w_i \partial \overline{w_j}} dw_i(d\varphi X)\, d\overline{w_j}(d\varphi X)$$

$$= \sum_{k,l} Z_k Z_{l^*} \log K'(\varphi(z), \overline{\varphi(z)})\, dz_k(X)\, d\overline{z_l}(X)$$

$$= \sum_{k,l} Z_k Z_{l^*} \log K(z, \bar{z})\, dz_k(X)\, d\overline{z_l}(X) = g(X, X),$$

where we have used (16) and the relation

$$\frac{\partial^2 \log |f|^2}{\partial z_k \partial \overline{z_l}} = 0,$$

which is valid for an arbitrary holomorphic function f in D without zeros. In fact, we have in a neighborhood of each point in D

$$\log |f|^2 = \log f + \log \bar{f} = \log f + (\log f)^-$$

so

$$Z_k Z_{l^*} \log |f|^2 = Z_k Z_{l^*} \log f + Z_{l^*} Z_k (\log f)^- = 0.$$

Let M be a complex manifold and let φ and ψ be holomorphic diffeomorphisms of M onto M. Then the diffeomorphisms $\varphi \circ \psi$ and φ^{-1} are almost complex, hence holomorphic. Consequently, the set of holomorphic diffeomorphisms of M onto itself forms a group, denoted $H(M)$, the group operation being composition of mappings. If $H(M)$ is transitive on M, M is said to be *homogeneous.*

Let D be a bounded domain with Riemannian structure given by the kernel function. According to Prop. 3.5 we have

$$H(D) \subset I(D).$$

If, as usual, $I(D)$ is taken with the compact open topology (Chapter IV, §2), it is clear that $H(D)$ is a closed subgroup of $I(D)$.

Proposition 3.6. *Let D be a bounded domain with Riemannian structure g given by the kernel function K. If D is homogeneous, then*

$$K(z, \bar{z}) = c\bar{g}(x_1, y_1, ..., x_N, y_N)^{1/2} \qquad (c = constant) \tag{17}$$

and

$$r = 2g,$$

r being the Ricci curvature.

Proof. Let $\varphi \in H(D)$. Then (16) implies

$$K(z, \bar{z}) = K(\varphi(z), \overline{\varphi(z)}) \mid J_\varphi \mid^2.$$

Furthermore, relation (15) shows that

$$\bar{g}(x_1, y_1, ..., x_N, y_N) = \bar{g}(u_1, v_1, ..., u_N, v_N) \mid J_\varphi \mid^4.$$

Formula (17) now follows from the homogeneity assumption; the formula $r = 2g$ follows from (12) and Prop. 2.5.

§ 4. Hermitian Symmetric Spaces of the Compact Type and the Noncompact Type

Let M be a connected complex manifold with a Hermitian structure. The set of holomorphic isometries of M forms a group, denoted $A(M)$, the group operation being composition of mappings. We have obviously

$$A(M) = H(M) \cap I(M).$$

Definition. Let M be a connected complex manifold with a Hermitian structure; M is said to be a *Hermitian symmetric space* if each point $p \in M$ is an isolated fixed point of an involutive holomorphic isometry s_p of M.

A Hermitian symmetric space M is of course a Riemannian symmetric space of even dimension. Hence the group $I(M)$ has a Lie group structure compatible with the compact open topology (Chapter IV, Lemma 3.2) and is a Lie transformation group of M. The group $A(M)$ is a closed subgroup of $I(M)$ and is therefore also a Lie transformation group of M. It is transitive on M since it contains all the symmetries. The identity component $A_0(M)$ of $A(M)$ is also transitive on M (Chapter II, Prop. 4.3(b)). Let $o \in M$ and let K be the subgroup of $G = A_0(M)$ leaving o fixed. With the automorphism $g \to s_o g s_o$ of G, the pair (G, K) is a Riemannian symmetric pair and M is diffeomorphic to G/K.

Let $X \in M_o$ and let $\gamma_X(t)$, $(t \in R)$, denote the geodesic in M having tangent vector X for $t = 0$. Let s_t denote the geodesic symmetry (extended to M) with respect to the point $\gamma_X(t)$. We have seen (Chapter IV, §3) that if $T_t = s_{t/2} s_o$ then $t \to T_t$ is a one-parameter subgroup of G, $T_t \cdot o = \gamma_X(t)$ and $(dT_t)_o$ is the parallel translation along γ_X. Since the elements of G by definition leave the complex structure of M invariant, the same is true of the parallel translation $(dT_t)_o$. This proves

Proposition 4.1. *The Hermitian structure of a Hermitian symmetric space is Kählerian.*

Next we consider the problem of constructing a complex structure on a given coset space. Let G be a connected Lie group, H a closed subgroup of G. Suppose J is an almost complex structure on the coset space $M = G/H$, invariant under the action of G. Let o denote the point $\{H\}$ in G/H. Then J_o is an endomorphism of the tangent space M_o satisfying the following conditions:

(i) $J_o^2 = -I$.

(ii) J_o commutes with each element in the linear isotropy group H^*.

On the other hand, if J_o is an endomorphism of M_o satisfying (i) and (ii), then the coset space $M = G/H$ has a unique almost complex structure which coincides with J_o at o and is invariant under the action of G.

Proposition 4.2. *Let (G, K) be a Riemannian symmetric pair. Let π denote the natural mapping of G onto $M = G/K$ and put $o = \pi(e)$. Let Q be any G-invariant Riemannian structure on M. Suppose A is an endomorphism of the tangent space M_o such that:*

(a) $A^2 = -I$.

(b) $Q_o(AX, AY) = Q_o(X, Y)$ for $X, Y \in M_o$.

(c) *A commutes with each element of the linear isotropy group K^*.*

Then M has a unique G-invariant almost complex structure J such that $J_o = A$. The structure Q is Hermitian, J is integrable, and with the corresponding[†] *complex structure, M is a Hermitian symmetric space.*

Proof. The existence and uniqueness of J is already mentioned above. That Q is Hermitian is clear from (b) since Q and J are G-invariant. We shall now verify that J is invariant under the symmetry s_o (and therefore under each symmetry s_p, $p \in M$). Let σ be any involutive analytic automorphism of G such that $(K_\sigma)_0 \subset K \subset K_\sigma$. Then, according to Prop. 3.4, Chapter IV, $s_o \circ \pi = \pi \circ \sigma$. Let $p \in M$ and $Z \in M_p$. Select $g \in G$ such that $\tau(g) \cdot o = p$ and put $Z_0 = d\tau(g^{-1})\, Z$. Then using the invariance of J under G and the relations $s_o \circ \tau(g) = \tau(\sigma(g)) \circ s_o$, $ds_o J_o = J_o\, ds_o$ it follows that

$$ds_o(J_pZ) = ds_o d\tau(g)\, J_oZ_0 = d\tau(\sigma(g)) \circ J_o(ds_oZ_0)$$
$$= J_{s_o \cdot p}(ds_o d\tau(g)\, Z_o) = J_{s_o \cdot p}(ds_oZ);$$

hence J is invariant under s_o. Next we verify that J satisfies the integrability condition

$$[X, Y] + J[JX, Y] + J[X, JY] - [JX, JY] = 0 \tag{1}$$

for arbitrary vector fields X, Y on M. Owing to the homogeneity of M it suffices to verify (1) at the point o. Moreover, since the left-hand side of (1) is $C^\infty(M)$-bilinear, we can assume that the vector fields X, Y are (in a neighborhood of o) adapted to their values X_o, Y_o at o. Since J is invariant under G, in particular under parallelism, it follows that the vector fields JX, JY are adapted to their values at o. But since the torsion is 0 we have

$$[U, V]_o = (\nabla_U(V))_o - (\nabla_V(U))_o = 0$$

† See Theorem 1.2.

for any vector fields U, V adapted to their values at o, so (1) follows. The complex structure on M corresponding to J (Theorem 1.2) is, due to its uniqueness, invariant under each s_p, so M is a Hermitian symmetric space.

Remark. The proof above uses the deep Theorem 1.2 which is unproved in this book. In an exercise following this chapter we outline a direct proof of Prop. 4.2 (under a mild restriction), which does not make use of Theorem 1.2.

The example C^2 shows that sometimes the groups $A_0(M)$ and $I_0(M)$ are different. This however, is somewhat exceptional as the following lemma shows.

Lemma 4.3. *Let M be a Hermitian symmetric space. Then $I_0(M)$ is semisimple if and only if $A_0(M)$ is semisimple. In this case $A_0(M) = I_0(M)$.*

Proof. Let $G = I_0(M)$ and let $\mathfrak{g}$ denote the Lie algebra of G. Let s denote the automorphism of $\mathfrak{g}$ which corresponds to the automorphism $g \to s_o g s_o$ of G and let $\mathfrak{p}$ denote the set of vectors $X \in \mathfrak{g}$ such that $sX = -X$. Since $A(M)$ contains the symmetries with respect to all points in M, it is clear that $A(M)$ contains all one-parameter subgroups $\exp tX$, $X \in \mathfrak{p}$. Thus the Lie algebra of $A(M)$ contains $\mathfrak{p}$ and $[\mathfrak{p}, \mathfrak{p}]$. If $I(M)$ is semisimple, then $[\mathfrak{p}, \mathfrak{p}] + \mathfrak{p} = \mathfrak{g}$ so $A_0(M) = I_0(M)$. On the other hand, if $A_0(M)$ is semisimple, then $A_0(M) = I_0(M)$ by Theorem 4.1, Chapter V.

Definition. Let M be a Hermitian symmetric space; M is said to be of the *compact type* or the *noncompact type* according to the type of the Riemannian symmetric pair $(A_0(M), K)$, K being the isotropy subgroup of $A_0(M)$ at some point $o \in M$.

Proposition 4.4. *Let M be a simply connected Hermitian symmetric space. Then M is a product*

$$M = M_o \times M_- \times M_+,$$

where all the factors are simply connected Hermitian symmetric spaces and $M_0 = C \times ... \times C$, M_- and M_+ are of the compact type and noncompact type, respectively.

Proof. Let $G = A_0(M)$, let o be a point in M and let K denote the isotropy subgroup of G at o. Let $(\tilde{G}, \varphi)$ be the universal covering group of G and let $\tilde{K}$ denote the identity component of $\varphi^{-1}(K)$. Then, if ψ denotes the mapping $g\tilde{K} \to \varphi(g) K$ of $\tilde{G}/\tilde{K}$ onto G/K, the pair $(\tilde{G}/\tilde{K}, \psi)$ is a covering space of G/K (Lemma 13.4, Chapter I); consequently $M = \tilde{G}/\tilde{K}$. Let $\mathfrak{g}$ denote the Lie algebra of G and let s denote the automorphism of $\mathfrak{g}$ which corresponds to the automorphism $g \to s_o g s_o$ of G. Then the pair $(\mathfrak{g}, s)$ is an effective orthogonal symmetric Lie algebra. We can decompose $\mathfrak{g}$ and $\mathfrak{k}$, the Lie algebra of $\tilde{K}$, according to Theorem 1.1, Chapter V,

$$\mathfrak{g} = \mathfrak{g}_0 + \mathfrak{g}_- + \mathfrak{g}_+, \qquad \mathfrak{k} = \mathfrak{k}_0 + \mathfrak{k}_- + \mathfrak{k}_+.$$

The groups $\tilde{G}$ and $\tilde{K}$ decompose accordingly

$$\tilde{G} = G_0 \times G_- \times G_+, \qquad \tilde{K} = K_0 \times K_- \times K_+$$

and the spaces $M_0 = G_0/K_0$, $M_- = G_-/K_-$, $M_+ = G_+/K_+$ are simply connected Riemannian globally symmetric spaces whose product is M. Let $\mathfrak{p}$, $\mathfrak{p}_0$, $\mathfrak{p}_-$, $\mathfrak{p}_+$ denote the eigenspaces for the eigenvalue -1 of s. As usual, these can be identified with tangent spaces to M, M_0, M_-, M_+ and

$$\mathfrak{p} = \mathfrak{p}_0 + \mathfrak{p}_- + \mathfrak{p}_+. \tag{2}$$

Let J and Q denote the almost complex structure and Riemannian structure, respectively, on M. Since $\mathfrak{p}$ is identified with the tangent space to M at o, Q_o is a bilinear form on $\mathfrak{p} \times \mathfrak{p}$ and J_o is an endomorphism of $\mathfrak{p}$. Let $Y_0 \in \mathfrak{p}_0$. Then $J_o Y_0$ can be decomposed according to (2),

$$J_o Y_0 = X_0 + X_- + X_+. \tag{3}$$

Let Ad denote the adjoint representation of $\tilde{G}$. Then

$$\operatorname{Ad}(k) Z_0 = Z_0 \qquad \text{for } k \in K_- \times K_+, \qquad Z_0 \in \mathfrak{p}_0, \tag{4}$$

by Lemma 1.5, Chapter V. Since $\operatorname{Ad}(k)$ and J_o commute, (3) and (4) imply

$$\operatorname{Ad}(k)(X_- + X_+) = X_- + X_+ \qquad \text{for } k \in K_- \times K_+.$$

This last relation, however, holds for all $k \in \tilde{K}$ since for $k_0 \in K_0$, $\operatorname{Ad}(k_0)$ keeps every vector in $\mathfrak{p}_- + \mathfrak{p}_+$ fixed. From Cor. 1.7, Chapter V, we deduce therefore that $X_- + X_+ = 0$ so $J_o\mathfrak{p}_0 \subset \mathfrak{p}_0$. Since J_o leaves Q_o invariant, it is clear that $J_o(\mathfrak{p}_- + \mathfrak{p}_+) \subset \mathfrak{p}_- + \mathfrak{p}_+$. Repeating the argument above, we find that $\mathfrak{p}_-$ and $\mathfrak{p}_+$ are invariant under J_o. Now Prop. 4.2 implies that M_- and M_+ are Hermitian symmetric.

Theorem 4.5. *Let M be a Hermitian symmetric space for which $A_0(M)$ is semisimple. Let $o \in M$ and let K denote the isotropy subgroup of $A_0(M)$*

at o. The corresponding linear isotropy group and its Lie algebra are denoted by K^* and $\mathfrak{k}^*$, respectively. Then:

(i) *The complex structure J_o of M_o belongs to the center $\mathfrak{c}$ of $\mathfrak{k}^*$.*

(ii) *The symmetry s_o is contained in the identity component of the center Z_K of K.*

Proof. Let Q and R, respectively, denote the Riemannian structure and curvature tensor of M. Then, according to Theorem 4.1, Chapter V, the Lie algebra $\mathfrak{k}^*$ consists of those endomorphisms of M_o which, when extended to the mixed tensor algebra over M_o as derivations commuting with contractions, annihilate Q_o and R_o. Now, if $X, Y \in M_o$ we have by (1) and (2), Chapter IV, §5,

$$(J_o \cdot Q_o)(X, Y) = -Q_o(X, J_oY) - Q_o(J_oX, Y), \tag{5}$$

$$(J_o \cdot R_o)(X, Y) = [J_o, R_o(X, Y)] - R_o(J_oX, Y) - R_o(X, J_oY). \tag{6}$$

The right-hand side of (5) vanishes since Q is Hermitian. The first term on the right-hand side of (6) vanishes since J_o commutes elementwise with K^* and $R_o(X, Y) \in \mathfrak{k}^*$. Finally, considering R as a complex tensor field we have by Lemma 2.1,

$$R_o(X - iJ_oX, Y - iJ_oY) = 0. \tag{7}$$

Considering the imaginary part in (7) we find that the right-hand side of (6) vanishes. Hence $J_o \in \mathfrak{k}^*$ and therefore $J_o \in \mathfrak{c}$. Identifying the Lie algebras $\mathfrak{k}$ and $\mathfrak{k}^*$ we have

$$\exp(tJ_o) \in Z_K \qquad \text{for each } t \in \boldsymbol{R}. \tag{8}$$

But on the space M_0 $(= \mathfrak{p})$ we have

$$e^{\pi J_o} = -I,$$

so $\exp(\pi J_o) = s_o$. This finishes the proof.

A compact semisimple Lie group U is a Riemannian globally symmetric space in each two-sided invariant Riemannian structure. However, this can never make U Hermitian symmetric as Theorem 4.5 shows.

Theorem 4.6. *Let M be a Hermitian symmetric space of the compact type or the noncompact type. Then M is simply connected.*

Proof. Since every Riemannian globally symmetric space of the noncompact type is simply connected we can assume that M is of the compact type. Let $U = I_0(M)$ $(= A_0(M))$; in the notation of Theorem 4.5 we have $M = U/K$. The Lie algebra $\mathfrak{u}$ of U decomposes $\mathfrak{u} = \mathfrak{k}_0 + \mathfrak{p}_*$

where $\mathfrak{k}_0$ is the Lie algebra of K and $\mathfrak{p}_*$ is the orthogonal complement of $\mathfrak{k}_0$ in $\mathfrak{u}$ with respect to the Killing form of $\mathfrak{u}$. Since we can consider $\mathfrak{k}_0$ as the Lie algebra of the linear isotropy group, corresponding to K, we have by Prop. 4.5, $J_o \in \mathfrak{k}_0$. Let S denote the closure in K of the one-parameter subgroup $\exp tJ_o$, $(t \in \boldsymbol{R})$. Since J_o annihilates no vector in $\mathfrak{p}_*$, it follows that $\mathfrak{k}_0$ is the centralizer of J_o in $\mathfrak{u}$. The centralizer Z_S of S in U therefore has Lie algebra $\mathfrak{k}_0$. From (8) we conclude $K \subset Z_S$. Being the centralizer of a torus, Z_S is connected so K is connected.

Let $(\tilde{U}, \varphi)$ be the universal covering group of U. The mapping $\sigma : u \to s_o u s_o$ is an automorphism of U. Let $\tilde{\sigma}$ denote the automorphism of $\tilde{U}$ such that $d\tilde{\sigma} = d\sigma$. Let $\tilde{K}$ denote the set of fixed points of $\tilde{\sigma}$. By Theorem 7.2, Chapter VII, the group $\tilde{K}$ is connected. Hence $\varphi(\tilde{K}) = K$ and there exists by Theorem 4.5 an element $z \in \tilde{K}$ such that $\varphi(z) = s_0$. The automorphism $\Sigma : u \to zuz^{-1}$ of $\tilde{U}$ satisfies $\varphi \circ \Sigma = \sigma \circ \varphi$ so $d\Sigma = d\sigma$. It follows that $\tilde{\sigma} = \Sigma$ and $\tilde{K}$ is the centralizer of z in $\tilde{U}$. In particular $\tilde{K}$ contains the center of $\tilde{U}$ so $\tilde{K} = \varphi^{-1}(K)$. Consequently, $U/K = \tilde{U}/\tilde{K}$ which is simply connected.

Remark. It will be proved in the next chapter that the class of symmetric bounded domains coincides with the class of Hermitian symmetric spaces of the noncompact type.

Example. It is possible for two Riemannian globally symmetric spaces M_1 and M_2 to be associated with the same orthogonal symmetric Lie algebra such that M_1 is Hermitian symmetric while M_2 is not. As an example take $M_1 = \boldsymbol{S}^2$ (two-dimensional sphere) and $M_2 = \boldsymbol{P}^2$ (two-dimensional projective space). Both are Riemannian globally symmetric (Chapter VII, Prop. 1.2) and $\boldsymbol{S}^2$ is Hermitian symmetric. However, $\boldsymbol{P}^2$ is not Hermitian symmetric since it is not even orientable.

§ 5. Irreducible Orthogonal Symmetric Lie Algebras

It is now convenient to make a further decomposition of the orthogonal symmetric Lie algebras of compact type and noncompact type.

Definition. Let $(\mathfrak{l}, s)$ be an orthogonal symmetric Lie algebra, $\mathfrak{u}$ and $\mathfrak{e}$ the eigenspaces of s for the eigenvalues $+1$ and -1, respectively; $(\mathfrak{l}, s)$ is said to be *irreducible* if the two following conditions are satisfied:

(i) $\mathfrak{l}$ is semisimple and $\mathfrak{u}$ contains no ideal $\neq \{0\}$ of $\mathfrak{l}$.

(ii) The algebra $\mathrm{ad}_{\mathfrak{l}}(\mathfrak{u})$ acts irreducibly on $\mathfrak{e}$.

Let (L, U) be a pair associated with $(\mathfrak{l}, s)$; then (L, U) is said to be irreducible if $(\mathfrak{l}, s)$ is irreducible. A Riemannian globally symmetric space M is called irreducible if the pair $(I_0(M), K)$ is irreducible, K being the isotropy subgroup of $I_0(M)$ at some point in M.

Let (L, U) be an irreducible Riemannian symmetric pair. Then all L-invariant Riemannian structures on L/U coincide except for a constant factor. In fact, $\mathrm{Ad}_L(U)$ is a compact linear group acting irreducibly on $\mathfrak{e}$ and the endomorphism $b : \mathfrak{e} \to \mathfrak{e}$ (from the proof of Lemma 1.2, Chapter V) commutes with each element of $\mathrm{Ad}_L(U)$. Hence b can only have one eigenvalue so the forms $Q(X, X)$ and $B(X, X)$ in the cited lemma are proportional. Thus L/U has an essentially unique L-invariant Riemannian structure. We can therefore always assume that this Riemannian structure is induced by $\pm B$ where B is the Killing form of $\mathfrak{l}$.

It is obvious that $(\mathfrak{l}, s)$ is irreducible if and only if the dual $(\mathfrak{l}^*, s^*)$ is irreducible.

The condition (ii) above can be described in different terms.

Proposition 5.1. *In the notations above suppose that the condition (i) is satisfied. Then (ii) holds if and only if $\mathfrak{u}$ is a maximal proper subalgebra of $\mathfrak{l}$.*

Proof. Assume first that (ii) holds. If $\mathfrak{u}$ were not maximal, there would exist a subalgebra $\mathfrak{u}^*$ of $\mathfrak{l}$ satisfying the *proper* inclusions $\mathfrak{u} \subset \mathfrak{u}^* \subset \mathfrak{l}$. Put $\mathfrak{e}^* = \mathfrak{u}^* \cap \mathfrak{e}$. Then $[\mathfrak{u}, \mathfrak{e}^*] \subset \mathfrak{u}^* \cap \mathfrak{e} = \mathfrak{e}^*$ so, due to the irreducibility, $\mathfrak{e}^* = \{0\}$ or $\mathfrak{e}^* = \mathfrak{e}$. Now the identity $\mathfrak{e}^* = \mathfrak{e}$ implies $\mathfrak{u}^* = \mathfrak{l}$ which is impossible. The identity $\mathfrak{e}^* = \{0\}$ is also impossible because if $Z \in \mathfrak{u}^*$, $Z \notin \mathfrak{u}$, then $Z = T + X$, where $T \in \mathfrak{u}$, $X \in \mathfrak{e}$, $X \neq 0$. It follows that $X = Z - T \in \mathfrak{u}^* \cap \mathfrak{e} = \mathfrak{e}^*$, which is a contradiction. The converse is trivial because if $\mathfrak{e}'$ were a proper invariant subspace of $\mathfrak{e}$, then $\mathfrak{u} + \mathfrak{e}'$ would be a proper subalgebra of $\mathfrak{l}$, properly containing $\mathfrak{u}$.

Proposition 5.2. *Let $(\mathfrak{l}, s)$ be an orthogonal symmetric Lie algebra. Let $\mathfrak{l} = \mathfrak{u} + \mathfrak{e}$ be the decomposition of $\mathfrak{l}$ into the eigenspaces of s for the eigenvalue $+1$ and -1, respectively. Assume $\mathfrak{l}$ is semisimple and that $\mathfrak{u}$ contains no ideal $\neq \{0\}$ of $\mathfrak{l}$. Then there exists ideals $\mathfrak{l}_i$ in $\mathfrak{l}$ such that:*

(a) $\mathfrak{l} = \sum_i \mathfrak{l}_i$ *(direct sum).*

(b) *The ideals $\mathfrak{l}_i$ are mutually orthogonal with respect to the Killing form B of $\mathfrak{l}$ and they are invariant under s.*

(c) *Denoting by s_i the restriction of s to $\mathfrak{l}_i$, each $(\mathfrak{l}_i, s_i)$ is an irreducible orthogonal symmetric Lie algebra.*

Proof. The proof proceeds along the same lines as that of Theorem 1.1, Chapter V. Let Q and b be as in Lemma 1.2, Chapter V. Then b is an endomorphism of $\mathfrak{e}$ which is symmetric with respect to Q, i.e.,

$$Q(bX, Y) = Q(X, bY), \qquad X, Y \in \mathfrak{e}.$$

Let

$$\mathfrak{e} = \sum_j \mathfrak{f}_j$$

be the decomposition of $\mathfrak{e}$ into the eigenspaces of b. The spaces $\mathfrak{f}_j$ are mutually orthogonal with respect to Q and B. Each $\mathfrak{f}_j$ is invariant under $\mathrm{ad}_{\mathfrak{l}}(\mathfrak{u})$ and can be decomposed into irreducible subspaces which are mutually orthogonal with respect to Q and B. Thus we get a direct decomposition

$$\mathfrak{e} = \sum_i \mathfrak{e}_i,$$

orthogonal with respect to Q and B, where the spaces $\mathfrak{e}_i$ are invariant and irreducible under $\mathrm{ad}_{\mathfrak{l}}(\mathfrak{u})$. We put $\mathfrak{u}_i = [\mathfrak{e}_i, \mathfrak{e}_i]$ and $\mathfrak{l}_i = \mathfrak{u}_i + \mathfrak{e}_i$. It can be proved just as in Chapter V, §1, that the spaces $\mathfrak{l}_i$ have the properties of Prop. 5.2.

The next theorems give an important description of the irreducible orthogonal symmetric Lie algebras.

Theorem 5.3. *The irreducible orthogonal symmetric Lie algebras of the compact type are:*

I. $(\mathfrak{l}, s)$ *where* $\mathfrak{l}$ *is a compact simple Lie algebra and* s *any involutive automorphism of* $\mathfrak{l}$.

II. $(\mathfrak{l}, s)$ *where the compact algebra* $\mathfrak{l}$ *is the direct sum* $\mathfrak{l} = \mathfrak{l}_1 + \mathfrak{l}_2$ *of simple ideals which are interchanged by an involutive automorphism* s *of* $\mathfrak{l}$.

Theorem 5.4. *The irreducible, orthogonal symmetric Lie algebras of the noncompact type are*

III. $(\mathfrak{l}, s)$ *where* $\mathfrak{l}$ *is a simple, noncompact Lie algebra over* $\boldsymbol{R}$, *the complexification* $\mathfrak{l}^C$ *is a simple Lie algebra over* C *and* s *is an involutive automorphism of* $\mathfrak{l}$ *such that the fixed points form a compactly imbedded subalgebra.*

IV. $(\mathfrak{l}, s)$ *where* $\mathfrak{l} = \mathfrak{g}^R$, $\mathfrak{g}$ *being a simple Lie algebra over* C. *Here* s *is the conjugation of* $\mathfrak{l}$ *with respect to a maximal compactly imbedded subalgebra.*

Furthermore, if $(\mathfrak{l}^*, s^*)$ *denotes the dual of* $(\mathfrak{l}, s)$,

$$(\mathfrak{l}, s) \text{ is of type III} \Leftrightarrow (\mathfrak{l}^*, s^*) \text{ is of type I},$$

$$(\mathfrak{l}, s) \text{ is of type IV} \Leftrightarrow (\mathfrak{l}^*, s^*) \text{ is of type II}.$$

Proof of Theorem 5.3. It is obvious from Prop. 5.2 that each $(\mathfrak{l}, s)$ of type I is irreducible. Next, let $(\mathfrak{l}, s)$ be of type II. Then according to

Cor. 6.3, Chapter II, the only ideals in $\mathfrak{l}$ are $\{0\}$, $\mathfrak{l}_1$, $\mathfrak{l}_2$, and $\mathfrak{l}$. Again, Prop. 5.2 shows that $(\mathfrak{l}, s)$ is irreducible.

On the other hand, suppose $(\mathfrak{l}, s)$ is irreducible and $\mathfrak{l} = \mathfrak{u} + \mathfrak{e}$ compact. Let

$$\mathfrak{l} = \mathfrak{a}_1 + \ldots + \mathfrak{a}_n$$

be the decomposition into the simple ideals of $\mathfrak{l}$, (Cor. 6.3, Chapter II). Then s permutes the ideals $\mathfrak{a}_i$. If $s\mathfrak{a}_i = \mathfrak{a}_i$, put $\mathfrak{l}_i = \mathfrak{a}_i$. If $s\mathfrak{a}_i \neq \mathfrak{a}_i$, put $\mathfrak{l}_i = \mathfrak{a}_i + s\mathfrak{a}_i$. We have then a direct decomposition

$$\mathfrak{l} = \sum_i \mathfrak{l}_i,$$

where each $\mathfrak{l}_i$ is an ideal in $\mathfrak{l}$, invariant under s, and can therefore be decomposed into eigenspaces, $\mathfrak{l}_i = \mathfrak{u}_i + \mathfrak{e}_i$. Since $\mathfrak{u} = \Sigma_i \mathfrak{u}_i$, the irreducibility of $(\mathfrak{l}, s)$ implies that all $\mathfrak{e}_i$ vanish except one, say $\mathfrak{e}_1$. But then condition (i) for irreducibility shows that $\mathfrak{u}_i = \{0\}$ for $i \neq 1$. Thus $\mathfrak{l}_i = \{0\}$ for $i \neq 1$, and this proves the theorem.

Proof of Theorem 5.4. Let $(\mathfrak{l}, s)$ be an orthogonal symmetric Lie algebra and $(\mathfrak{l}^*, s^*)$ its dual. Since irreducibility is preserved under the duality, it suffices to prove the last two statements of the theorem. Suppose first that $(\mathfrak{l}^*, s^*)$ is of type I. Then the decomposition $\mathfrak{l} = \mathfrak{u} + \mathfrak{e}$ into eigenspaces of s is a Cartan decomposition of $\mathfrak{l}$. If $\mathfrak{l}$ were not simple we would have $\mathfrak{l} = \mathfrak{a}_1 + \mathfrak{a}_2$ where $\mathfrak{a}_1$ and $\mathfrak{a}_2$ are nonzero ideals. Let $\mathfrak{a}_1 = \mathfrak{k}_1 + \mathfrak{p}_1$, $\mathfrak{a}_2 = \mathfrak{k}_2 + \mathfrak{p}_2$ be Cartan decompositions of $\mathfrak{a}_1$ and $\mathfrak{a}_2$. (If $\mathfrak{a}_1$ or $\mathfrak{a}_2$ is compact, then $\mathfrak{p}_1$ or $\mathfrak{p}_2$ is $\{0\}$.) Since the Cartan decompositions $\mathfrak{l} = \mathfrak{u} + \mathfrak{e}$ and $\mathfrak{l} = (\mathfrak{k}_1 + \mathfrak{k}_2) + (\mathfrak{p}_1 + \mathfrak{p}_2)$ are conjugate we may assume $\mathfrak{k}_1 + \mathfrak{k}_2 = \mathfrak{u}$, $\mathfrak{p}_1 + \mathfrak{p}_2 = \mathfrak{e}$. But then $(\mathfrak{k}_1 + i\mathfrak{p}_1) + (\mathfrak{k}_2 + i\mathfrak{p}_2)$ is a decomposition of $\mathfrak{l}^*$ into nonzero ideals. Hence $\mathfrak{l}$ must be simple. The complex algebra $\mathfrak{l}^C$ is also simple because otherwise $\mathfrak{l}^C$ is a direct sum $\mathfrak{l}^C = \mathfrak{n}_1 + \mathfrak{n}_2$ where $\mathfrak{n}_1$, $\mathfrak{n}_2$ are nonzero ideals. These being semisimple, let $\mathfrak{k}_1$, $\mathfrak{k}_2$ be compact real forms of $\mathfrak{n}_1$ and $\mathfrak{n}_2$, respectively. Then $\mathfrak{l}^*$ is isomorphic to $\mathfrak{k}_1 + \mathfrak{k}_2$ which is not simple. Thus $(\mathfrak{l}, s)$ is of type III. On the other hand, let $(\mathfrak{l}, s)$ be of type III. Then $(\mathfrak{l}, s)$ and therefore $(\mathfrak{l}^*, s^*)$ is irreducible. If $(\mathfrak{l}^*, s^*)$ were of type II, the complexification $(\mathfrak{l}^*)^C = \mathfrak{l}^C$ would not be simple. Hence $(\mathfrak{l}^*, s^*)$ is of type I. Next, suppose $(\mathfrak{l}^*, s^*)$ is of type II. Then $\mathfrak{l}^* = \mathfrak{l}_1^* + \mathfrak{l}_2^*$ where $\mathfrak{l}_1^*$ and $\mathfrak{l}_2^*$ are simple ideals interchanged by s^*. Then we know from Theorem 2.4, Chapter V, and the subsequent remark that $\mathfrak{l}$ has a complex structure J, and if $\mathfrak{l} = \mathfrak{k} + J\mathfrak{k}$ is a Cartan decomposition of $\mathfrak{l}$, then $\mathfrak{k}$ and $\mathfrak{l}_1^*$ are isomorphic. Hence $\mathfrak{k}$ is simple and $\mathfrak{l}$ is simple (as a Lie algebra over $\boldsymbol{C}$). Thus $(\mathfrak{l}, s)$ is of type IV. Reversing these arguments and using Theorem 2.4, Chapter V, again we find that if $(\mathfrak{l}, s)$ is of type IV then $(\mathfrak{l}^*, s^*)$ is of type II.

Proposition 5.5. *Let M be a simply connected Riemannian globally symmetric space of the compact type or the noncompact type. Then M is a product*

$$M = M_1 \times \ldots \times M_r,$$

where the factors M_i are irreducible. If M is Hermitian, then each M_i is Hermitian.

Proof. As in the proof of Prop. 4.2, Chapter V, let $G = I_0(M)$ and let K denote the isotropy subgroup of G at some point o in M. If $(\tilde{G}, \varphi)$ is the universal covering group of G and $\tilde{K}$ is the identity component of $\varphi^{-1}(K)$, then $M = \tilde{G}/\tilde{K}$. Using Prop. 5.2 we get a decomposition

$$\begin{aligned} \tilde{G} &= G_1 \times \ldots \times G_r, \\ \tilde{K} &= K_1 \times \ldots \times K_r, \end{aligned} \tag{1}$$

where each pair (G_i, K_i) is irreducible. If we put $M_i = G_i/K_i$ we have

$$M = M_1 \times \ldots \times M_r.$$

Moreover, G_i is semisimple and the Lie algebra of K_i contains no ideal $\neq \{0\}$ of the Lie algebra of G_i. In view of Theorem 4.1, Chapter V, G_i and $I(M_i)$ have the same Lie algebra. Thus M_i is irreducible. Finally suppose M is Hermitian symmetric and let J denote the corresponding almost complex structure on M. As proved in §4, the groups $A_0(M)$ and $I_0(M)$ coincide and J_o lies in the center of the Lie algebra of K. According to (1), J_o is decomposed $J_o = J_1 \times \ldots \times J_r$ where each J_i is an endomorphism of square $-I$ of the tangent space to M_i at $\{K_i\}$ and J_i lies in the center of the Lie algebra of K_i. Since each K_i is connected, the group $\mathrm{Ad}_{G_i}(K_i)$ commutes elementwise with J_i. Proposition 4.2 now shows that M_i is Hermitian.

§ 6. Irreducible Hermitian Symmetric Spaces

Theorem 6.1.

(i) *The noncompact irreducible Hermitian symmetric spaces are exactly the manifolds G/K where G is a connected noncompact simple Lie group with center $\{e\}$ and K has nondiscrete center and is a maximal compact subgroup of G*

(ii) *The compact irreducible Hermitian symmetric spaces are exactly the manifolds U/K where U is a connected compact simple Lie group with center $\{e\}$ and K has nondiscrete center and is a maximal connected proper subgroup of U.*

Proof. Let M be an irreducible Hermitian symmetric space. Then $M = \tilde{G}/\tilde{K}$ where $\tilde{G}$ is the simply connected covering group of $I_0(M)$ and $\tilde{K}$ is connected and contains the center of $\tilde{G}$, (Theorem 1.1, Chapter VI, and Theorem 4.6). Hence $M = \mathrm{Ad}(\tilde{G})/\mathrm{Ad}(\tilde{K})$, where $\mathrm{Ad} = \mathrm{Ad}_{\tilde{G}}$. This representation of M has the properties stated in (i) and (ii) as a glance at Theorems 4.5, 5.3, and 5.4 and Prop. 5.1 shows.

On the other hand, suppose U and K have the properties in (ii). Then the center of K contains an element j of order 4. Let $s = j^2$, and let Z_s denote the centralizer of s in U. Since $s \neq e$, we have $Z_s \neq U$ so K coincides with the identity component of Z_s. The automorphism $\sigma : u \to sus^{-1}$ $(u \in U)$ turns (U, K) into a Riemannian symmetric pair. Let $\mathfrak{p}_*$ denote the eigenspace for the eigenvalue -1 of the automorphism $d\sigma$, and let J denote the restriction of $\mathrm{Ad}_U(j)$ to $\mathfrak{p}_*$. Then $J^2 = -I$ so J gives rise to a U-invariant almost complex structure on U/K and, according to Prop. 4.2, U/K is Hermitian symmetric. The statement (i) now follows by use of duality.

§ 7. Bounded Symmetric Domains

Definition. A bounded domain D is called *symmetric* if each $p \in D$ is an isolated fixed point of an involutive holomorphic diffeomorphism of D onto itself.

Theorem 7.1.

(i) *Each bounded symmetric domain D is, when equipped with the Bergman metric, a Hermitian symmetric space of the noncompact type. In particular, a bounded symmetric domain is necessarily simply connected.*

(ii) *Let M be a Hermitian symmetric space of the noncompact type. Then there exists a bounded symmetric domain D and a holomorphic diffeomorphism of M onto D.*

Proof of (i). Let D be a bounded symmetric domain. Let Q be the Riemannian structure on D corresponding to the Bergman metric. Prop. 3.5 shows that D is a Hermitian symmetric space. Let o be a fixed point in D, let σ denote the automorphism $g \to s_o g s_o$ of $A_0(D)$ and let $\mathfrak{l}$ denote the Lie algebra of $A_0(D)$. If we put $s = d\sigma$, then $(\mathfrak{l}, s)$

is an effective orthogonal symmetric Lie algebra. From Theorem 1.1. Chapter V we have the direct decompositions

$$\mathfrak{l} = \mathfrak{u} + \mathfrak{e}, \qquad \mathfrak{e} = \mathfrak{e}_0 + \mathfrak{e}_- + \mathfrak{e}_+,$$

and in order to prove (i) above, it suffices to show that $\mathfrak{e}_0 = \mathfrak{e}_- = \{0\}$. Let $X \in \mathfrak{e}$. As before, let T_X denote the restriction of $(\mathrm{ad}_{\mathfrak{l}} X)^2$ to $\mathfrak{e}$. Then the curvature tensor R of D satisfies

$$R_o(X, Y)X = T_X Y, \qquad X, Y \in \mathfrak{e}, \tag{1}$$

so by Prop. 3.6

$$2Q_o(X, X) = \mathrm{Trace}\,(T_X). \tag{2}$$

Now $[\mathfrak{e}_0, \mathfrak{e}] = \{0\}$ by Lemma 1.3, Chapter V, so (2) implies that $\mathfrak{e}_0 = \{0\}$. Next, suppose $X \in \mathfrak{e}_-$. Then $T_X \mathfrak{e}_+ = \{0\}$, $T_X \mathfrak{e}_- \subset \mathfrak{e}_-$ and

$$Q_o(T_X Y, Y) = Q_o(R_o(X, Y)X, Y) \leqslant 0$$

for $Y \in \mathfrak{e}_-$, since the curvature along two-dimensional subspaces of $\mathfrak{e}_-$ is $\geqslant 0$. Thus (2) implies that $X = 0$, so $\mathfrak{e}_- = \{0\}$. This proves (i).

Proof of Theorem 7.1 (ii) (Algebraic part). In view of Theorem 4.6 and Prop. 5.5 it can be assumed that M is irreducible. Then by Theorem 6.1, the group $I_0(M)$ is simple. Let $\mathfrak{g}_0$ denote its Lie algebra, let θ denote the involutive automorphism of $\mathfrak{g}_0$ which arises from the symmetry with respect to some point in M and let $\mathfrak{g}_0 = \mathfrak{k}_0 + \mathfrak{p}_0$ be the decomposition of $\mathfrak{g}_0$ into eigenspaces of θ for the eigenvalues $+1$ and -1, respectively. Let $\mathfrak{c}_0$ be the center of $\mathfrak{k}_0$ and let $\mathfrak{h}_0$ be some maximal abelian subalgebra of $\mathfrak{k}_0$. Then $\mathfrak{c}_0 \subset \mathfrak{h}_0$ and $\mathfrak{h}_0$ is a maximal abelian subalgebra of $\mathfrak{g}_0$. In fact, the centralizer of $\mathfrak{c}_0$ in $\mathfrak{g}_0$ contains $\mathfrak{k}_0$ but differs from $\mathfrak{g}_0$ so by Prop. 5.1, it must coincide with $\mathfrak{k}_0$.

This maximality of $\mathfrak{h}_0$ carries with it important relationship between the Cartan decomposition $\mathfrak{g}_0 = \mathfrak{k}_0 + \mathfrak{p}_0$ and the root space decomposition of the complexification $\mathfrak{g}$ of $\mathfrak{g}_0$. Let $\mathfrak{c}$, $\mathfrak{h}$, $\mathfrak{k}$, $\mathfrak{p}$ be the subspaces of $\mathfrak{g}$ spanned by $\mathfrak{c}_0$, $\mathfrak{h}_0$, $\mathfrak{k}_0$, $\mathfrak{p}_0$. Then $\mathfrak{u} = \mathfrak{k}_0 + i\mathfrak{p}_0$ is a compact real form of $\mathfrak{g}$. Let σ and τ denote the conjugations of $\mathfrak{g}$ with respect to $\mathfrak{g}_0$ and $\mathfrak{u}$, respectively, and let B denote the Killing form of $\mathfrak{g}$. The Hermitian form B_τ on $\mathfrak{g} \times \mathfrak{g}$ given by $B_\tau(X, Y) = -B(X, \tau Y)$ is strictly positive definite and

$$B_\tau([Z, X], Y) + B_\tau(X, [Z, Y]) = 0$$

for $Z \in \mathfrak{u}$, $X, Y \in \mathfrak{g}$. It follows that the endomorphism $\mathrm{ad}\,H$ of $\mathfrak{g}$ is semisimple for each $H \in \mathfrak{h}_0 \cup i\mathfrak{h}_0$. Since all $\mathrm{ad}\,H (H \in \mathfrak{h})$ commute, they are semisimple endomorphisms of $\mathfrak{g}$ so $\mathfrak{h}$ is a Cartan subalgebra of $\mathfrak{g}$. Let Δ denote the set of nonzero roots of $\mathfrak{g}$ with respect to $\mathfrak{h}$. Let $\alpha \in \Delta$. Since $[\mathfrak{h}, \mathfrak{k}] \subset \mathfrak{k}$ and $[\mathfrak{h}, \mathfrak{p}] \subset \mathfrak{p}$, it is clear that either $\mathfrak{g}^\alpha \subset \mathfrak{k}$ or $\mathfrak{g}^\alpha \subset \mathfrak{p}$. In the first case the root α is called *compact*, in the second case *noncompact* and we have the direct decompositions

$$\mathfrak{k} = \mathfrak{h} + \sum_\alpha \mathfrak{g}^\alpha, \qquad \mathfrak{p} = \sum_\beta \mathfrak{g}^\beta, \tag{3}$$

where α runs over all the compact roots, and β runs over all the noncompact roots. Let $\Delta_\mathfrak{c}$ denote the set of roots in Δ which do not vanish identically on $\mathfrak{c}$. In view of Lemma 3.1, Chapter VI, each root α is real valued on $i\mathfrak{h}_0$. We introduce compatible orderings in the duals of the real vector spaces $i\mathfrak{h}_0$ and $i\mathfrak{c}_0$. This gives an ordering of Δ which will be used in the rest of the proof. Let Δ^+ denote the set of positive roots in Δ, put $Q_+ = \Delta^+ \cap \Delta_\mathfrak{c}$ and

$$\mathfrak{p}_+ = \sum_{\beta \in Q_+} \mathfrak{g}^\beta, \qquad \mathfrak{p}_- = \sum_{-\beta \in Q_+} \mathfrak{g}^\beta.$$

Proposition 7.2. *The spaces $\mathfrak{p}_+$ and $\mathfrak{p}_-$ are abelian subalgebras of $\mathfrak{g}$ and*

$$[\mathfrak{k}, \mathfrak{p}_-] \subset \mathfrak{p}_-, \qquad [\mathfrak{k}, \mathfrak{p}_+] \subset \mathfrak{p}_+, \qquad \mathfrak{p} = \mathfrak{p}_- + \mathfrak{p}_+.$$

Proof. Let $\alpha \in \Delta$ be compact. Then $[\mathfrak{c}, \mathfrak{g}^\alpha] = \{0\}$ so α vanishes identically on $\mathfrak{c}$. Consequently $\mathfrak{p}_- + \mathfrak{p}_+ \subset \mathfrak{p}$. Using in addition the compatibility of the orderings we have $[\mathfrak{g}^\alpha, \mathfrak{p}_+] \subset \mathfrak{p}_+$, $[\mathfrak{g}^\alpha, \mathfrak{p}_-] \subset \mathfrak{p}_-$. The relations $[\mathfrak{h}, \mathfrak{p}_-] \subset \mathfrak{p}_-$, $[\mathfrak{h}, \mathfrak{p}_+] \subset \mathfrak{p}_+$ being obvious, we derive $[\mathfrak{k}, \mathfrak{p}_-] \subset \mathfrak{p}_-$, $[\mathfrak{k}, \mathfrak{p}_+] \subset \mathfrak{p}_+$ from (3).

Next, let $\beta, \gamma \in Q_+$. Then $[\mathfrak{g}^\beta, \mathfrak{g}^\gamma] \subset \mathfrak{g}^{\beta+\gamma}$ and if $\beta + \gamma$ is a root, then $\beta + \gamma \in Q_+$. But on the other hand $[\mathfrak{p}_+, \mathfrak{p}_+] \subset \mathfrak{k}$, so $[\mathfrak{p}_+, \mathfrak{p}_+] = \{0\}$. Also $\mathfrak{p}_-$ is abelian because $\tau \cdot \mathfrak{g}^\delta = \mathfrak{g}^{-\delta}$ for any $\delta \in \Delta$ (Lemma 3.1, Chapter VI).

Finally, in order to prove $\mathfrak{p} = \mathfrak{p}_- + \mathfrak{p}_+$, let $\mathfrak{q}$ denote the orthogonal complement of $\mathfrak{p}_- + \mathfrak{p}_+$ in $\mathfrak{p}$ with respect to B_τ and put

$$\mathfrak{g}_+ = \mathfrak{p}_+ + \mathfrak{p}_- + [\mathfrak{p}_+, \mathfrak{p}_-].$$

We shall prove that $\mathfrak{g}_+$ is an ideal in $\mathfrak{g}$; for this purpose it suffices to prove that

$$[\mathfrak{p}_+, \mathfrak{q}] = [\mathfrak{p}_-, \mathfrak{q}] = \{0\}. \tag{4}$$

Let $T \in \mathfrak{k}$, $X \in \mathfrak{p}_+$, $Y \in \mathfrak{q}$. Since $\tau \cdot T \in \mathfrak{k}$, and $\tau \cdot [X, \tau \cdot T] \in \mathfrak{p}_-$ we have

$$B_\tau([X, Y], T) = -B([X, Y], \tau \cdot T) = -B_\tau(Y, \tau \cdot [X, \tau \cdot T]) = 0$$

so $[\mathfrak{p}_+, \mathfrak{q}] = \{0\}$ and similarly $[\mathfrak{p}_-, \mathfrak{q}] = \{0\}$. Now, the orthogonal symmetric Lie algebra $(\mathfrak{g}_0, \theta)$ is of type III so by Theorem 5.4, $\mathfrak{g}$ is simple. We have therefore either $\mathfrak{g}_+ = \{0\}$ or $\mathfrak{g}_+ = \mathfrak{g}$. The first case implies that all the roots in Δ vanish identically on $\mathfrak{c}$ which is impossible. Thus $\mathfrak{g}_+ = \mathfrak{g}$, so $\mathfrak{p} = \mathfrak{p}_- + \mathfrak{p}_+$ and the proposition is proved.

Corollary 7.3. *A root $\alpha \in \Delta$ is compact if and only if it vanishes identically on $\mathfrak{c}$.*

In Chapter VI and VII much use was made of maximal abelian subspaces $\mathfrak{a}_0$ of $\mathfrak{p}_0$. Whereas all of these are conjugate under the linear isotropy group it is possible in the special situation here to select $\mathfrak{a}_0$ with particular reference to Δ. For each $\alpha \in \Delta$ we select a nonzero vector $X_\alpha \in \mathfrak{g}^\alpha$.

Proposition 7.4. *There exists a subset $\gamma_1, \ldots, \gamma_s$ of Q_+ such that the subspace*

$$\mathfrak{a} = \sum_{i=1}^{s} C(X_{\gamma_i} + X_{-\gamma_i}) \tag{5}$$

is a maximal abelian subspace of $\mathfrak{p}$.

The proof requires some preparation. If Q is any subset of Q_+, let

$$\mathfrak{p}_Q = \sum_{\gamma \in Q} (\mathfrak{g}^\gamma + \mathfrak{g}^{-\gamma}).$$

Let β be the lowest root in Q and let $Q(\beta)$ denote the set of all $\gamma \in Q$ such that $\gamma \neq \beta$ and neither $\gamma + \beta$ nor $\gamma - \beta$ is a root. Then the centralizer of $\mathfrak{g}^{-\beta} + \mathfrak{g}^\beta$ in $\mathfrak{p}_Q$ coincides with $\mathfrak{p}_{Q(\beta)}$.

Lemma 7.5. *The centralizer of $X_\beta + X_{-\beta}$ in $\mathfrak{p}_Q$ is $C(X_\beta + X_{-\beta}) + \mathfrak{p}_{Q(\beta)}$.*

Proof. Let $X \in \mathfrak{p}_Q$ and let Q' denote the complement of $\{\beta\}$ in Q. Then

$$X = c_\beta X_\beta + c_{-\beta} X_{-\beta} + \sum_{\gamma \in Q'} (c_\gamma X_\gamma + c_{-\gamma} X_{-\gamma}),$$

where the coefficients are complex numbers. Now $\mathfrak{g} = \mathfrak{h} + \sum_{\alpha \in \Delta} \mathfrak{g}^\alpha$ and the component of $[X, X_\beta + X_{-\beta}]$ in $\mathfrak{h}$ is $(c_\beta - c_{-\beta})\,[X_\beta, X_{-\beta}]$. Suppose X and $X_\beta + X_{-\beta}$ commute. We have then $c_\beta = c_{-\beta}$ and the vector

$$Y = \sum_{\gamma \in Q'} (c_\gamma X_\gamma + c_{-\gamma} X_{-\gamma})$$

commutes with $X_\beta + X_{-\beta}$. Since $\mathfrak{p}_+$ and $\mathfrak{p}_-$ are abelian we obtain

$$[Y, X_\beta + X_{-\beta}] = \sum_{\gamma \in Q'} (c_\gamma [X_\gamma, X_{-\beta}] + c_{-\gamma}[X_{-\gamma}, X_\beta]) = 0.$$

Here

$$c_\gamma [X_\gamma, X_{-\beta}] = c_{-\gamma}[X_{-\gamma}, X_\beta] = 0 \tag{6}$$

for each $\gamma \in Q'$. Otherwise, suppose, for example, $c_\gamma[X_\gamma, X_{-\beta}] \neq 0$. Then $\mathfrak{g}^{\gamma-\beta}$ and $\mathfrak{g}^{\beta-\gamma}$ are $\neq \{0\}$ and there exists a $\delta \in Q'$ such that

$$c_\gamma[X_\gamma, X_{-\beta}] + c_{-\delta}[X_{-\delta}, X_\beta] = 0.$$

This implies that $\alpha = \gamma - \beta = -\delta + \beta$ is a root $\neq 0$ but the relations $\gamma = \alpha + \beta$, $\delta = \beta - \alpha$ contradict the fact that β is the lowest root in Q. It follows from (6) that $Y \in \mathfrak{p}_{Q(\beta)}$ and the lemma is proved.

Proof of Proposition 7.4. We define a sequence of spaces $\mathfrak{p} = \mathfrak{p}_1 \supset \mathfrak{p}_2 \supset \ldots \supset \mathfrak{p}_{s+1} = \{0\}$, each of which has the form $\mathfrak{p}_i = \mathfrak{p}_{Q_i}$, $(Q_i \subset Q_+)$, as follows: Let $Q_1 = Q_+$ and let γ_1 be the lowest positive root in Q_1. Let $\mathfrak{p}_2$ denote the centralizer of $\mathfrak{g}^{-\gamma_1} + \mathfrak{g}^{\gamma_1}$ in $\mathfrak{p}_1 = \mathfrak{p}_{Q_1}$; then $\mathfrak{p}_2 = \mathfrak{p}_{Q_2}$ where $Q_2 = Q_1(\gamma_1)$. Denoting by γ_2 the lowest positive root in Q_2, let $\mathfrak{p}_3$ denote the centralizer of $\mathfrak{g}^{-\gamma_2} + \mathfrak{g}^{\gamma_2}$ in $\mathfrak{p}_2$ etc. Then the roots $\gamma_1, \ldots, \gamma_s$ are all different and form the desired subset of Q_+. In fact, it is clear from the construction that the space $\mathfrak{a}$ in (5) is abelian. Moreover, suppose $X \in \mathfrak{p}$ commutes with each element in $\mathfrak{a}$. We wish to prove $X \in \mathfrak{a}$. Suppose this were false. Then there exists an integer r $(1 \leqslant r \leqslant s)$, such that $X \in \mathfrak{p}_r + \mathfrak{a}$ but $X \notin \mathfrak{p}_{r+1} + \mathfrak{a}$. Let $X = Y + Z$ $(Y \in \mathfrak{p}_r, Z \in \mathfrak{a})$. Since X and Z commute with $X_{\gamma_r} + X_{-\gamma_r}$ the same is true of Y. Thus Lemma 7.5 implies that

$$Y = c(X_{\gamma_r} + X_{-\gamma_r}) + Y_1,$$

where $Y_1 \in \mathfrak{p}_{r+1}$ and $c \in \boldsymbol{C}$. Now, $Z_1 = Z + c(X_{\gamma_r} + X_{-\gamma_r})$ lies in $\mathfrak{a}$ and therefore

$$X = Y_1 + Z_1 \in \mathfrak{p}_{r+1} + \mathfrak{a},$$

which contradicts the definition of r. This proves Prop. 7.4.

Corollary 7.6. *In accordance with Lemma* 3.1, *Chapter VI, let the vectors $X_\alpha \in \mathfrak{g}^\alpha$ be chosen such that for each $\alpha \in \Delta$*

$$(X_\alpha - X_{-\alpha}), \qquad i(X_\alpha + X_{-\alpha}) \in \mathfrak{u},$$

$$[X_\alpha, X_{-\alpha}] = (2/\alpha(H_\alpha))\, H_\alpha.$$

Then the space

$$\mathfrak{a}_0 = \sum_{i=1}^{s} \boldsymbol{R}(X_{\gamma_i} + X_{-\gamma_i})$$

equals $\mathfrak{a} \cap \mathfrak{p}_0$ *and is therefore a maximal abelian subspace of* $\mathfrak{p}_0$.

In fact, owing to the choice of X_α we have $X_{\gamma_i} + X_{-\gamma_i} \in i\mathfrak{u} \cap \mathfrak{p} = \mathfrak{p}_0$ so $\mathfrak{a}_0 \subset \mathfrak{a} \cap \mathfrak{p}_0$. On the other hand, if

$$X = \sum_{i=1}^{s} c_i(X_{\gamma_i} + X_{-\gamma_i}) \in \mathfrak{p}_0 \qquad (c_i \in \boldsymbol{C}),$$

then $\tau \cdot X = -X$ so $c_i \in \boldsymbol{R}$.

Lemma 7.7. *Let* $\mathfrak{l}$ *be the three-dimensional Lie algebra over* $\boldsymbol{C}$ *given by the vector space* $\boldsymbol{C}H + \boldsymbol{C}X + \boldsymbol{C}Y$ *with the bracket defined by*

$$[X, Y] = H, \qquad [H, X] = 2X, \qquad [H, Y] = -2Y.$$

If L *is any Lie group with Lie algebra* $\mathfrak{l}^{\boldsymbol{R}}$, *then*

$$\exp t(X + Y) = \exp(\tanh t)Y \exp(\log(\cosh t))H \exp(\tanh t)X \tag{7}$$

for $t \in \boldsymbol{R}$.

Proof. Consider the group $\boldsymbol{SL}(2, \boldsymbol{C})$ of all complex 2×2 matrices of determinant 1. The Lie algebra $\mathfrak{sl}(2, \boldsymbol{C})$ of this group consists of all complex 2×2 matrices of trace 0. It is isomorphic to $\mathfrak{l}$ under the mapping

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \to H, \qquad \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \to X, \qquad \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \to Y.$$

The group $\boldsymbol{SL}(2, \boldsymbol{C})$ contains $\boldsymbol{SU}(2)$ as a maximal compact subgroup. This group consists of all matrices

$$\begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix}$$

of determinant 1. Thus $\boldsymbol{SU}(2)$ is homeomorphic to the three dimensional sphere S^3, in particular $\boldsymbol{SU}(2)$ is simply connected. In view of Theorem 2.2, Chapter VI, the group $\boldsymbol{SL}(2, \boldsymbol{C})$ is also simply connected. It suffices therefore to prove (7) for the group $L = \boldsymbol{SL}(2, \boldsymbol{C})$. This can be done by a direct computation, which is left to the reader.

Proof of Theorem 7.1 (ii) (Geometric part). The Lie algebra $\mathfrak{g}$ is the (vector space) direct sum of the Lie algebras $\mathfrak{k}$, $\mathfrak{p}_-$, and $\mathfrak{p}_+$. We shall now study the corresponding global situation.

Let G denote the simply connected Lie group with Lie algebra $\mathfrak{g}^R$. Let U, K, P_-, P_+, A^* denote the analytic subgroups of G corresponding to the subalgebras $\mathfrak{u}$, $\mathfrak{k}$, $\mathfrak{p}_-$, $\mathfrak{p}_+$, and $i\mathfrak{h}_0$, respectively, considered as real subalgebras of $\mathfrak{g}^R$. As in Chapter VI, §6, let

$$\mathfrak{n}_+ = \sum_{\alpha \in \Delta^+} \mathfrak{g}^\alpha, \qquad \mathfrak{n}_- = \sum_{-\alpha \in \Delta^+} \mathfrak{g}^\alpha,$$

considered as real subalgebras of $\mathfrak{g}^R$ and let N_+, N_-, G_0, K_0 denote the analytic subgroups of G corresponding to $\mathfrak{n}_+$, $\mathfrak{n}_-$, $\mathfrak{g}_0$, and $\mathfrak{k}_0$, respectively. Let exp denote the usual exponential mapping of $\mathfrak{g}^R$ into G, and let ad and Ad denote the adjoint representations of $\mathfrak{g}^R$ and G, respectively. Let θ, σ, τ denote the automorphisms of G which correspond to the automorphisms θ, σ, τ of $\mathfrak{g}^R$.

Lemma 7.8. *The mapping exp induces a diffeomorphism of $\mathfrak{p}_-$ onto P_- and of $\mathfrak{p}_+$ onto P_+.*

Proof. According to Lemma 3.5, Chapter VI, there exists a basis of $\mathfrak{g}$ with respect to which the matrices expressing ad $(\mathfrak{n}_+)$ are lower triangular with zeros in the diagonal. In view of Cor. 4.4 and Lemma 4.5, Chapter VI, the mapping $\operatorname{ad} X \to e^{\operatorname{ad} X} = \operatorname{Ad}(\exp X)$ is a diffeomorphism of ad $(\mathfrak{n}_+)$ onto Ad (N_+). Since ad is one-to-one and Ad is one-to-one on N_+, the mapping $\exp : \mathfrak{n}_+ \to N_+$ is a diffeomorphism of $\mathfrak{n}_+$ onto N_+. Using the fact that $\mathfrak{p}_+ \subset \mathfrak{n}_+$ and $\tau \cdot \mathfrak{p}_+ = \mathfrak{p}_-$, the lemma follows.

Lemma 7.9. *The mapping $(q, k, p) \to qkp$ is diffeomorphism of $P_- \times K \times P_+$ onto an open submanifold of G, containing G_0.*

Proof. We prove first that $P_-K \cap P_+ = \{e\}$. Suppose to the contrary that $y \in P_-K \cap P_+$, $y \neq e$. Select $Y \in \mathfrak{p}_+$ such that $\exp Y = y$. Since $[\mathfrak{k}, \mathfrak{p}_-] \subset \mathfrak{p}_-$ we have $\operatorname{Ad}(y)\, \mathfrak{p}_- \subset \mathfrak{p}_-$. Writing $Y = \Sigma_{\alpha \in Q_+} c_\alpha X_\alpha$ $(c_\alpha \in C)$, let β denote the lowest root in Q_+ such that $c_\beta \neq 0$. Then $[Y, X_{-\beta}] \equiv c_\beta[X_\beta, X_{-\beta}]$ (mod $\mathfrak{n}_+$) and it follows that

$$\operatorname{Ad}(y)\, X_{-\beta} \equiv X_{-\beta} + c_\beta[X_\beta, X_{-\beta}] \qquad (\operatorname{mod} \mathfrak{n}_+).$$

Reading this relation mod $(\mathfrak{n}_- + \mathfrak{n}_+)$, we find that $\operatorname{Ad}(y)\, X_{-\beta} \notin \mathfrak{p}_-$ which contradicts $\operatorname{Ad}(y)\, \mathfrak{p}_- \subset \mathfrak{p}_-$. This shows that $P_-K \cap P_+ = \{e\}$. Applying the mapping $x \to \tau(x^{-1})$, $(x \in G)$, it follows that $P_- \cap KP_+ = \{e\}$.

In order to prove that the mapping in the lemma is one-to-one suppose $q_1k_1p_1 = q_2k_2p_2$. This implies

$$(k_2^{-1}q_2^{-1}q_1k_2)\, k_2^{-1}k_1 = p_2p_1^{-1},$$

$$q_1^{-1}q_2 = k_1k_2^{-1}(k_2p_1p_2^{-1}k_2^{-1}),$$

which shows that $p_1 = p_2$, $q_1 = q_2$, $k_1 = k_2$. The regularity of the mapping follows by using Lemma 5.2, Chapter VI, twice, first on the algebra $\mathfrak{p}_- + \mathfrak{k}$ and then on the algebra $(\mathfrak{p}_- + \mathfrak{k}) + \mathfrak{p}_+$. The image is therefore a submanifold of G of dimension

$$\dim_R \mathfrak{p}_- + \dim_R \mathfrak{k} + \dim_R \mathfrak{p}_+ = \dim \mathfrak{g}^R$$

and is therefore an open submanifold of G. Finally, we know from Theorem 1.1, Chapter VI, that $G_0 = P_0K_0 = K_0P_0$ where $P_0 = \exp \mathfrak{p}_0$. Let $X \in \mathfrak{p}_0$ and $p = \exp \frac{1}{2} X$. From Theorem 6.3, Chapter VI,

$$p = uan, \qquad u \in U,\ a \in A^*,\ n \in N_+$$

and applying τ

$$\tau(p) = p^{-1} = ua^{-1}\tau(n),$$

so

$$\exp X = p^2 = \tau(n^{-1})\, a^2n \in N_-A^*N_+.$$

Moreover, $\mathfrak{n}_+ + \mathfrak{h} \subset \mathfrak{p}_+ + \mathfrak{k}$, $\mathfrak{n}_- \subset \mathfrak{p}_- + \mathfrak{k}$ so

$$N_-A^*N_+ \subset P_-KP_+,$$

and the lemma follows.

Remark. In general $P_-KP_+ \neq G$ as can be seen by considering the example $G = \boldsymbol{SL}(2, \boldsymbol{C})$, $G_0 = \boldsymbol{SL}(2, \boldsymbol{R})$, $K_0 = \boldsymbol{SO}(2)$. Here $K = \boldsymbol{SO}(2,\boldsymbol{C})$, the group of complex orthogonal 2×2 matrices of determinant 1. As a result of Theorem 1.1, Chapter VI, each complex orthogonal matrix can be written uniquely as $\alpha e^{i\beta}$ where α is a real orthogonal matrix and β is a real skew symmetric matrix. Hence $K = K_0 \times \boldsymbol{R}$ topologically. Thus P_-KP_+ has fundamental group $\boldsymbol{Z}$ whereas G is simply connected.

Lemma 7.10. *The set G_0KP_+ is open in P_-KP_+ and $G_0 \cap KP_+ = K_0$.*

Proof. Suppose $p \in P_0$ has the form $p = kp_+$ where $k \in K$ and $p_+ \in P_+$. Applying the automorphism $\theta = \sigma\tau$ we get $p^{-1} = k(p_+)^{-1}$ so

$p^2 = (p_+)^2$. Applying τ we get $p^{-2} \in \tau(P_+) \subset P_-$, so $p_+^2 \in P_-$. Hence $p = k = p_+ = e$. Since $G_0 = K_0 P_0$ this shows that $G_0 \cap KP_+ = K_0$.

Consider now the group KP_+. The group P_+ is closed in N_+, hence closed in G. The group K is closed in G since it is the identity component of the set of fixed points of θ. Let $(k_n p_n)$ be a sequence in KP_+ which converges in G. Applying θ we see that the sequence (p_n^2) and therefore the sequences (k_n) and (p_n) are convergent. Thus KP_+ is closed in G and due to Lemma 7.9, its Lie algebra is $\mathfrak{k} + \mathfrak{p}_+$. Consider the mapping $\psi : (g, x) \to gx$ of $G_0 \times (KP_+)$ into G. Let $Y \in \mathfrak{g}_0$, $Z \in \mathfrak{k} + \mathfrak{p}_+$. Then

$$\psi(g \exp tY, x) = gx \exp (t \operatorname{Ad}(x^{-1}) Y),$$

$$\psi(g, x \exp tZ) = gx \exp tZ$$

and consequently

$$d\psi_{(g,x)}(dL_g Y, dL_x Z) = dL_{gx}(\operatorname{Ad}(x^{-1}) Y + Z).$$

It follows that

$$\begin{aligned} d\psi(\{G_0 \times (KP_+)\}_{(g,x)}) &= dL_{gx} \circ \operatorname{Ad}(x^{-1}) (\mathfrak{g}_0 + \operatorname{Ad}(x) (\mathfrak{k} + \mathfrak{p}_+)) \\ &= dL_{gx} \circ \operatorname{Ad}(x^{-1}) (\mathfrak{g}_0 + \mathfrak{k} + \mathfrak{p}_+), \end{aligned}$$

and this image under $d\psi$ covers the whole tangent space G_{gx} due to the fact that

$$\mathfrak{g} = \mathfrak{p}_0 + \mathfrak{k} + \mathfrak{p}_+ = \mathfrak{g}_0 + \mathfrak{k} + \mathfrak{p}_+.$$

The lemma now follows from Lemma 7.9.

Lemma 7.11. *Let* $Z \in \mathfrak{a}_0$, i.e.,

$$Z = \sum_{i=1}^{s} t_i (X_{\gamma_i} + X_{-\gamma_i}), \qquad t_i \in \mathbf{R}. \tag{8}$$

Then

$$\exp Z = \exp Y \exp H \exp X,$$

where

$$Y = \sum_{i=1}^{s} (\tanh t_i)\, X_{-\gamma_i}, \qquad X = \sum_{i=1}^{s} (\tanh t_i)\, X_{\gamma_i},$$

$$H = \sum_{i=1}^{s} \log (\cosh t_i)\, [X_{\gamma_i}, X_{-\gamma_i}].$$

This follows from Lemma 7.7, combined with the fact that X_{γ_i} and $X_{-\gamma_j}$ commute if $i \neq j$.

The complex vector space $\mathfrak{g}$ becomes a finite-dimensional Hilbert space under the inner product B_τ. Let $\| X \| = B_\tau(X, X)^{1/2}$ for $X \in \mathfrak{g}$.

According to Lemma 7.8, exp induces a one-to-one mapping of $\mathfrak{p}_-$ onto P_-. Let log denote the inverse mapping. For $x \in G_0$, let $\zeta(x)$ denote the unique element in P_- such that $x \in \zeta(x)\, KP_+$ (Lemma 7.9).

Lemma 7.12. *The norm $\| \log \zeta(x) \|$ is bounded as x varies through G_0.*

Proof. Since $[\mathfrak{k}, \mathfrak{p}_-] \subset \mathfrak{p}_-$, we have $\zeta(kxk') = k\zeta(x)\, k^{-1}$ for $x \in G_0$, $k, k' \in K_0$. From Lemma 6.3, Chapter V, and $G_0 = K_0P_0$, it follows that $G_0 = K_0A_0K_0$ where A_0 is the analytic subgroup of G_0 with Lie algebra $\mathfrak{a}_0$. Writing an arbitrary $x \in G_0$ as $x = kak'$ $(k, k' \in K_0, a \in A_0)$, we get $\zeta(x) = k\zeta(a)\, k^{-1}$ and

$$\| \log \zeta(x) \| = \| \mathrm{Ad}\,(k) \log \zeta(a) \| = \| \log \zeta(a) \|.$$

Now $a = \exp Z$ where Z has the form (8), and from Lemma 7.11 follows that

$$\log \zeta(a) = \sum_{i=1}^{s} (\tanh t_i)\, X_{-\gamma_i},$$

and since $| \tanh t | \leqslant 1$ for $t \in \boldsymbol{R}$,

$$\| \log \zeta(x) \| \leqslant \sum_{i=1}^{s} \| X_{-\gamma_i} \|$$

for all $x \in G_0$, which proves the lemma.

Completion of the proof. The Hermitian manifold M is diffeomorphic to G_0/K_0 (Theorem 1.1, Chapter VI) and the complex structure on M corresponds to an endomorphism J_0 on $\mathfrak{p}_0$ which commutes with all $\mathrm{Ad}\,(k)$ $(k \in K_0)$ and satisfies $J_0^2 = -I$.

For any coset space X/Y let $\tau(x)$ as usual denote the mapping $\xi Y \to x\xi Y$ of X/Y onto itself.

As remarked earlier, the group KP_+ is closed in G and the coset space G/KP_+ contains P_-KP_+/KP_+ as an open subset which in turn contains G_0KP_+/KP_+ as an open subset. Let these imbeddings be

$$\begin{array}{ccccc}
G_0KP_+/KP_+ & \xrightarrow{I_1} & P_-KP_+/KP_+ & \xrightarrow{I_2} & G/KP_+ \\
\psi_1 \big\downarrow & & \psi_2 \big\downarrow & & \\
G_0/K_0 & \xrightarrow{\psi_0} & P_- & \xrightarrow{\log} & \mathfrak{p}_-
\end{array}$$

denoted by I_2 and I_1, respectively (see diagram). In this diagram ψ_0, ψ_1, and ψ_2 denote the mappings

$$\psi_1 : gkpKP_+ \to gK_0 \qquad (g \in G_0, k \in K, p \in P_+),$$

$$\psi_2 : qkpKP_+ \to q \qquad (q \in P_-, k \in K, p \in P_+),$$

$$\psi_0 = \psi_2 \circ I_1 \circ \psi_1^{-1}.$$

The mapping ψ_1 is a diffeomorphism of G_0KP_+/KP_+ onto G_0/K_0 (Lemma 7.10). The mapping ψ_2 is a diffeomorphism of P_-KP_+/KP_+ onto P_- (Lemma 7.9). Thus it follows from Lemma 7.10 that ψ_0 is a diffeomorphism of G_0/K_0 onto an open subset of P_-. Combining this with Lemma 7.12 it is clear that the mapping $\psi = \log \circ \psi_0$ is a diffeomorphism of G_0/K_0 onto a bounded domain D in the complex vector space $\mathfrak{p}_-$. Moreover,

$$\psi(xK_0) = \log \zeta(x), \qquad x \in G_0,$$

so

$$\psi \circ \tau(k) = \mathrm{Ad}\,(k) \circ \psi, \qquad k \in K_0, \tag{9}$$

and

$$d\psi\,(\mathrm{Ad}\,(k)\,X) = \mathrm{Ad}\,(k)\,d\psi(X) \tag{10}$$

for $k \in K_0$, $X \in \mathfrak{p}_0$. Let J^* denote the endomorphism of $\mathfrak{p}_-$ defined by $d\psi(J_0X) = J^*d\psi(X)$ for all $X \in \mathfrak{p}_0$. Then $(J^*)^2 = -I$ and J^* commutes with all $\mathrm{Ad}\,(k)$, $(k \in K_0)$. Since $\mathrm{Ad}\,(K_0)$ is irreducible on $\mathfrak{p}_-$, it follows from Schur's lemma (see e.g. Chevalley [2], Prop. 2, p. 183) that the endomorphism J^* is a scalar multiple of the identity. Hence $J^* = \pm iI$ and ψ (or the mapping ψ followed by complex conjugation) is a holomorphic diffeomorphism.

Theorem 7.13. *In the notation above, the mapping $f : uK_0 \to uKP_+$ is an analytic diffeomorphism of U/K_0 onto G/KP_+. Thus the compact Hermitian symmetric space U/K_0 contains the dual G_0/K_0 as an open submanifold.*

Proof. Let $u \in U \cap KP_+$. It is clear that $u^{-1}\theta(u) \in P_+$. Applying τ we find that $\theta(u) = u$. Since U is simply connected, it follows from Theorem 7.2, Chapter VII, that $u \in K_0$. Thus $U \cap KP_+ = K_0$ and consequently the mapping f is one-to-one. Since f is regular and $\dim U/K_0 = \dim G/KP_+$, the image $f(U/K_0)$ is an open submanifold of G/KP_+. Being compact this submanifold must coincide with G/KP_+. The mapping $\varphi = I_2 \circ I_1 \circ \psi_1^{-1}$ is a diffeomorphism of G_0/K_0 onto an open subset of U/K_0 satisfying $\varphi \circ \tau(k) = \tau(k) \circ \varphi$ for $k \in K_0$. It follows that

$$d\varphi \operatorname{Ad}(k) X = \operatorname{Ad}(k)\, d\varphi X$$

for $X \in \mathfrak{p}_0$, $k \in K_0$. If we define the endomorphism J^* by $d\varphi J_0 X = J^* d\varphi X$, $(X \in \mathfrak{p}_0)$, then J^* has square $-I$ and commutes with each $\operatorname{Ad}(k)$ $(k \in K_0)$. Thus U/K_0 is Hermitian symmetric and the mapping φ is almost complex, hence holomorphic (§1).

Remark. In view of the fact that the Lie algebra $\mathfrak{g}$ is a Lie algebra over C, one can show that G is a *complex Lie group*, that is, has a complex structure in which the group operations are holomorphic. The groups K and P_+ are also complex Lie groups and the coset space G/KP_+ has a complex structure invariant under the action of G. It is not hard to see, using the irreducibility, that (up to a sign) this coincides with the U-invariant complex structure of U/K_0. *In particular, each mapping $\tau(g)$ $(g \in G_0)$ of G_0/K_0 extends to a holomorphic diffeomorphism of U/K_0.*

Example. Let $G_0 = \boldsymbol{SL}(2, \boldsymbol{R})$, $K_0 = \boldsymbol{SO}(2)$. Then G_0/K_0 is a Hermitian symmetric space of the noncompact type. If we to each

$$g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in G_0$$

associate the mapping $T_g : z \to (az + b)/(cz + d)$ of the upper half plane, then the mapping $gK_0 \to T_g(i)$ is a holomorphic diffeomorphism of G_0/K_0 onto the upper half-plane. The representation of G_0/K_0 as a bounded domain can be realized by mapping the upper half-plane onto the open disk $| z - i | < 1$ by inversion and reflection in the x-axis. In this case $K = \boldsymbol{SO}(2, \boldsymbol{C})$, $U = \boldsymbol{SU}(2)$, and $G = \boldsymbol{SL}(2, \boldsymbol{C})$. The dual space U/K_0 is the Riemann sphere and the space G/KP_+ is identified with the complex plane with the point ∞ adjoined.

EXERCISES

A. Complex Structures

1. Let g be a Hermitian structure on an almost complex manifold (M, J). Show that the tensor field ω given by

$$\omega(X, Y) = g(X, JY), \qquad X, Y \in \mathfrak{D}^1(M),$$

is a 2-form on M. Show that the following conditions are equivalent:

(i) g is Kählerian.

(ii) $\nabla_X(\omega) = 0$ for $X \in \mathfrak{D}^1(M)$.

(iii) $d\omega = 0$.

2. Let G be a Lie group with Lie algebra $\mathfrak{g}$. Suppose the Lie algebra $\mathfrak{g}$ has a complex structure. Show that G has a complex structure in which it is a complex Lie group. (Hint: The complex structure on $\mathfrak{g}$ induces a left invariant almost complex structure on G. As a result of Theorem 1.7, Chapter II, exp is an almost complex mapping of $\mathfrak{g}$ into G.)

3. Show that the Riemannian structure corresponding to the kernel function for the unit disk $|z| < 1$ turns the disk into a Riemannian manifold of constant negative curvature.

4. Let D be a bounded symmetric domain. Representing D as $H(D)/K$ ($H(D)$ as in §3, K compact), D acquires a natural metric in two different ways: Firstly from the kernel function for D and secondly from the Killing form B of the Lie algebra of $H(D)$. Show that these two metrics coincide. (In view of (2), §7, it suffices to prove $B(X, X) = 2 \operatorname{Trace}(T_X)$ for $X \in \mathfrak{e}$.)

B. Siegel's Generalized Upper Half-Plane

Let I denote the $n \times n$ unit matrix and let

$$J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}, \qquad z = \frac{1}{\sqrt{2}} \begin{pmatrix} I & I \\ -I & I \end{pmatrix}.$$

Let $\boldsymbol{Sp}(n, \boldsymbol{R})$ denote the group of all real $2n \times 2n$ matrices g satisfying

$${}^t gJg = J.$$

1. Show that the group $G = \boldsymbol{Sp}(n, \boldsymbol{R})$ is semisimple, and that the group $K = \boldsymbol{Sp}(n, \boldsymbol{R}) \cap \boldsymbol{SO}(2n)$ is a maximal compact subgroup.

2. Let $\mathfrak{g}_0 = \mathfrak{k}_0 + \mathfrak{p}_0$ be a Cartan decomposition of the Lie algebra $\mathfrak{g}_0$ of G, $\mathfrak{k}_0$ denoting the Lie algebra of K. Then $\operatorname{Ad}_G(J)$ restricted to $\mathfrak{p}_0$ is -1; also $z^2 = -J$ and z lies in the center of K.

3. The restriction of $\operatorname{Ad}_G(z)$ to $\mathfrak{p}_0$ gives rise to an invariant complex structure on the space

$$M = \boldsymbol{Sp}(n, \boldsymbol{R})/(\boldsymbol{SO}(2n) \cap \boldsymbol{Sp}(n, \boldsymbol{R}))$$

turning M into a Hermitian symmetric space of the noncompact type.

4. The mapping $p \to g\sigma(g^{-1})$ from Exercise 5, Chapter VI, is a diffeomorphism of M onto the submanifold S of $\boldsymbol{Sp}(n, \boldsymbol{R})$ formed by the matrices in $\boldsymbol{Sp}(n, \boldsymbol{R})$ which are symmetric and strictly positive definite.

5. In the complex manifold of $n \times n$ complex symmetric matrices consider the open submanifold $\mathscr{S}_n$ of complex symmetric matrices $Z = X + iY$ where X and Y are real and Y is strictly positive definite. Each $g \in \boldsymbol{Sp}(n, \boldsymbol{R})$ gives rise to a holomorphic diffeomorphism

$$T_g: Z \to (AZ + B)(CZ + D)^{-1}$$

of $\mathscr{S}_n$, the matrix g being written

$$g = \begin{pmatrix} A & B \\ C & D \end{pmatrix}.$$

Prove that the mapping $gK \to T_g \cdot (iI)$ is a well-defined holomorphic diffeomorphism of M onto $\mathscr{S}_n$.

6. Prove that the mapping

$$Z \to (I + iZ)(I - iZ)^{-1}$$

is a holomorphic diffeomorphism of $\mathscr{S}_n$ onto the bounded domain in $\boldsymbol{C}^{\frac{1}{2}n(n+1)}$ consisting of all complex symmetric $n \times n$ matrices W for which $I - \overline{W}W$ is strictly positive definite (the generalized unit disk).

C. An Alternative Proof of Prop. 4.2

Let the assumptions be as in Prop. 4.2 but suppose in addition that the identity component K_0 of K leaves no $X \neq 0$ in M_o fixed. Show through the following steps† that $M = G/K$ is a Hermitian symmetric space.

1. Let $\mathfrak{g}_0 = \mathfrak{k}_0 + \mathfrak{p}_0$ be the decomposition of the Lie algebra of G into eigenspaces of $d\sigma$. Complexify $\mathfrak{g}_0$, $\mathfrak{k}_0$, $\mathfrak{p}_0$ to $\mathfrak{g}$, $\mathfrak{k}$, and $\mathfrak{p}$, respectively. Then $\mathfrak{g} = \mathfrak{k} + \mathfrak{p}$ and J_o extends to an endomorphism of $\mathfrak{p}$ of square -1. Let $\mathfrak{p} = \mathfrak{p}_+ + \mathfrak{p}_-$ be the decomposition of $\mathfrak{p}$ into the eigenspaces of J_o for the eigenvalues $+i$ and $-i$, respectively. As in Chapter IV, §5, show that there exists a simply connected Lie group G^c with Lie algebra $\mathfrak{g}$.

2. Let L denote the analytic subgroup of G^c with Lie algebra $\mathfrak{k} + \mathfrak{p}_-$. Then L is closed in G^c (consider the normalizer of $\mathfrak{k} + \mathfrak{p}_-$ in G^c).

3. The Lie groups L and G^c being complex (Exercise A.2 above), show that G^c/L has a complex structure invariant under the action of G^c.

4. The identity mapping $\mathfrak{g}_0 \to \mathfrak{g}$ induces a mapping $g \to \varphi(g)$ of a neighborhood of e in G into G^c. Let ψ denote the induced mapping

† This approach follows a suggestion of I. Singer, cf. Frölicher [1], § 20.

$gK \to \varphi(g) L$ of a neighborhood of $\{K\}$ in G/K into G^c/L. Show that ψ is regular at $\{K\}$ and thus is a diffeomorphism of a neighborhood of $\{K\}$ in G/K onto a neighborhood of $\{L\}$ in G^c/L.

5. Show that

$$d\psi_K(J_o X) = i d\psi_K(X), \qquad X \in \mathfrak{p}_0,$$
$$\tau(g) \circ \psi = \psi \circ \tau(g)$$

for all g in a suitable neighborhood of e in G. Deduce that the mapping ψ is almost complex and therefore G/K has a complex structure corresponding to J.

NOTES

§1-§3. The torsion tensor S was introduced by Eckmann and Frölicher, see Frölicher [1]. The equivalence of "complex structure" and "integrable almost complex structure" was first proved by Newlander and Nirenberg [1]. The analytic case had been settled by several authors, see Libermann [1]. In [1] Kähler first studied the class of Hermitian structures named after him. Lemma 2.1 and Prop. 2.5 are proved in Bochner [1]. The kernel function was introduced by Bergman [1].

§5. The decomposition of a symmetric space into irreducible ones is due to É. Cartan [16]. Theorem 5.3 is also given there.

§4, §6, §7. Theorems 4.5, 4.6, and 6.1 are proved in Borel and Lichnerowicz [1]. The theory of bounded symmetric domains was developed by É. Cartan [19]. He proved Theorem 7.1 (i), namely, that a bounded symmetric domain is a symmetric space of the noncompact type. His proof uses the Liouville theorem that a bounded holomorphic function $f(z)$ in $|z| < \infty$ is constant; the proof differs from the one given here. The second part of Theorem 7.1 states that the bounded symmetric domains exhaust the class of Hermitian symmetric spaces of the noncompact type. This fact was verified in É. Cartan [19] by means of an explicit construction for the irreducible Hermitian symmetric spaces for which $A_0(M)$ is a classical group. This leaves out two exceptional Hermitian symmetric spaces for which Cartan stated the result without proof ([19], p. 151). The first *a priori* proof was given by Harish-Chandra [6], p. 591; this proof is reproduced here in the text. In [19] Cartan raised the question whether a bounded homogeneous domain is necessarily symmetric. This question was answered in the negative by Pyatetzki-Shapiro [1] (in 4 and 5 dimensions). Although É. Cartan had answered the question affirmatively for the dimensions 1,2,3, he refrained in [19] from guessing what the result would be in higher dimensions. Nevertheless, the question is sometimes referred to as "Cartan's conjecture" in the recent literature. The imbedding theorem (Theorem 7.13) is due to Borel [2, 3].

CHAPTER IX

ON THE CLASSIFICATION OF SYMMETRIC SPACES

In his papers from the years 1926 and 1927 E. Cartan accomplished a complete classification of irreducible Riemannian globally symmetric spaces. Locally, the question amounts to a classification of all simple Lie algebras over $\boldsymbol{R}$, a difficult problem which É. Cartan had solved already in 1914. In this chapter we state the results of this classification and develop some of the general methods which lead to these results. For the actual task of carrying out the classification in detail we refer to the literature, specified later in due course.

§ 1. Reduction of the Problem

We recall that two orthogonal symmetric Lie algebras $(\mathfrak{l}_1, s_1)$ and $(\mathfrak{l}_2, s_2)$ are called isomorphic if there exists an isomorphism φ of $\mathfrak{l}_1$ onto $\mathfrak{l}_2$ such that $\varphi \circ s_1 = s_2 \circ \varphi$.

The next lemma shows that the classification of simply connected, irreducible Riemannian globally symmetric spaces up to isometry is equivalent to the classification of irreducible orthogonal symmetric Lie algebras up to isomorphism. Here it is assumed as usual that the Riemannian structure is that induced by the Killing form.

Lemma 1.1.

(i) *Let M_1 and M_2 be two irreducible Riemannian globally symmetric spaces and Φ an isometry of M_1 onto M_2. Let $p_1 \in M_1$, $p_2 \in M_2$ such that $\Phi(p_1) = p_2$. Let σ_i denote the automorphism of $I_0(M_i)$ given by $\sigma_i(g) = s_{p_i} \circ g \circ s_{p_i}$ $(i = 1, 2)$. Let s_i denote the corresponding automorphism of the Lie algebra $\mathfrak{l}_i$ of $I_0(M_i)$. Then the orthogonal symmetric Lie algebras $(\mathfrak{l}_1, s_1)$ and $(\mathfrak{l}_2, s_2)$ are isomorphic under the differential of the isomorphism $g \to \Phi \circ g \circ \Phi^{-1}$ of $I_0(M_1)$ onto $I_0(M_2)$.*

(ii) *Let $(\mathfrak{l}_1, s_1)$ and $(\mathfrak{l}_2, s_2)$ be two irreducible orthogonal symmetric Lie algebras. Let (L_1, U_1) and (L_2, U_2) be the corresponding Riemannian symmetric pairs, L_1 and L_2 simply connected, U_1 and U_2 connected. Let φ be an isomorphism of $(\mathfrak{l}_1, s_1)$ onto $(\mathfrak{l}_2, s_2)$. Then there exists an isometry Φ*

of L_1/U_1 onto L_2/U_2 such that φ is the differential of the isomorphism $g \to \Phi \circ g \circ \Phi^{-1}$ of $I_0(L_1/U_1)$ onto $I_0(L_2/U_2)$.

The proof, which is quite canonical, can be omitted.

We have seen that given an irreducible orthogonal symmetric Lie algebra of the noncompact type, there is associated with it exactly one Riemannian globally symmetric space and this space is simply connected. Owing to the duality for symmetric spaces (Chapter V, §1), it suffices therefore to classify the irreducible compact Riemannian symmetric spaces.

Definition. Let $(\mathfrak{l}, s)$ be an irreducible orthogonal symmetric Lie algebra and let M be a Riemannian globally symmetric space associated with $(\mathfrak{l}, s)$. The space M is said to be of type i (i = I, II, III, IV) if $(\mathfrak{l}, s)$ is of type i in the notation of Theorems 5.3 and 5.4 in Chapter VIII.

As mentioned above it suffices to consider the types I and II. Let us first consider type II.

Proposition 1.2. *The Riemannian globally symmetric spaces of type II are precisely the compact, connected simple Lie groups provided with a Riemannian structure invariant under left and right translations.*

Proof. It is clear from §6, Chapter IV, that a compact, connected, simple Lie group with a bi-invariant Riemannian structure is a Riemannian globally symmetric space of type II.

On the other hand, let $(\mathfrak{l}, s)$ be an orthogonal symmetric Lie algebra of type II. Then $\mathfrak{l} = \mathfrak{l}_1 + \mathfrak{l}_2$ (direct sum) where the ideals $\mathfrak{l}_1$ and $\mathfrak{l}_2$ are interchanged by s. Let $\mathfrak{l}_0$ be a Lie algebra isomorphic to both $\mathfrak{l}_1$ and $\mathfrak{l}_2$ and let I_i denote the isomorphism of $\mathfrak{l}_i$ onto $\mathfrak{l}_0$, $(i = 1, 2)$. Then the mapping

$$I_0 : X + Y \to (I_1X, I_2Y), \qquad X \in \mathfrak{l}_1,\ Y \in \mathfrak{l}_2,$$

is an isomorphism of $\mathfrak{l}$ onto the product algebra $\bar{\mathfrak{l}} = \mathfrak{l}_0 \times \mathfrak{l}_0$. Consider the automorphisms $\bar{s}$ and σ of $\bar{\mathfrak{l}}$ given by

$$\bar{s}(I_1X, I_2Y) = (I_2Y, I_1X),$$
$$\sigma(I_1X, I_2Y) = (I_1X, I_1sY)$$

for $X \in \mathfrak{l}_1$, $Y \in \mathfrak{l}_2$. Then $(\mathfrak{l}, s)$ and $(\bar{\mathfrak{l}}, \bar{s})$ are isomorphic under the mapping $\sigma \circ I_0 : \mathfrak{l} \to \bar{\mathfrak{l}}$. Let (L, H) and $(\bar{L}, \bar{H})$ be corresponding Riemannian symmetric pairs, L and $\bar{L}$ simply connected, H and $\bar{H}$ connected. Then $\bar{L}$ is the product $L_0 \times L_0$ where L_0 is a simply connected Lie group with Lie algebra $\mathfrak{l}_0$ and $\bar{H} = \{(x, x) : x \in L_0\}$. Let Q and $\bar{Q}$ be the Riemannian structures on L/H and $\bar{L}/\bar{H}$, respectively, and let $\psi : L/H \to \bar{L}/\bar{H}$ be the isometry from Lemma 1.1, induced by the isomorphism $\sigma \circ I_0 : \mathfrak{l} \to \bar{\mathfrak{l}}$. Now $\bar{L}/\bar{H}$ is a group $\bar{G}$ with the multiplication

$$(x_1, x_2)\,\bar{H} \cdot (y_1, y_2)\,\bar{H} = (x_1 x_2^{-1} y_1 y_2^{-1}, e)\,\bar{H}$$

(§6, Chapter IV). Note that $\bar{H}$ is *not* a normal subgroup of $\bar{L}$. The Riemannian structure $\bar{Q}$ is invariant under left and right translations on $\bar{G}$. The mapping ψ turns L/H into a group G isomorphic to $\bar{G}$ and Q is invariant under left and right translations on G.

In order to conclude the proof of Prop. 1.2 we need a simple lemma.

Lemma 1.3. *Let N be a subgroup of L such that H is a normal subgroup of N. Then the product in L/H satisfies.*

$$(xH)\,(nH) = xnH, \qquad x \in L,\ n \in N.$$

In particular, the factor group N/H is a subgroup of L/H (note that H is not normal in L).

Proof. Let $\bar{N}$ denote the subgroup of $\bar{L}$ which corresponds to N under the isomorphism $\sigma \circ I_0 : \mathfrak{l} \to \bar{\mathfrak{l}}$. Then $\bar{H}$ is a normal subgroup of $\bar{N}$ and it suffices to prove the lemma for $\bar{N}$, $\bar{L}$, and $\bar{H}$ instead of N, L, and H. Consider two arbitrary elements $(n_1, n_2)\,\bar{H} \in \bar{N}/\bar{H}$ and $(x_1, x_2) \in \bar{L}/\bar{H}$. Since $(n_1, n_2)\,(x, x)\,(n_1, n_2)^{-1} \in \bar{H}$ for each $x \in L_0$, it follows that $n_1^{-1}n_2$ and (therefore) $n_1 n_2^{-1}$ belong to the center of L_0. Hence the product in $\bar{G} = \bar{L}/\bar{H}$ is

$$\begin{aligned}(x_1, x_2)\,\bar{H}\,(n_1, n_2)\,\bar{H} &= (x_1 x_2^{-1} n_1 n_2^{-1}, e)\,\bar{H} = (x_1 n_1 n_2^{-1} x_2^{-1}, e)\,\bar{H} \\ &= (x_1 n_1, x_2 n_2)\,\bar{H}\end{aligned}$$

and the lemma is proved.

Turning now to Prop. 1.2, let M be an arbitrary Riemannian globally symmetric space associated with $(\mathfrak{l}, s)$. Then there exists a symmetric pair (L_1, H_1) associated with $(\mathfrak{l}, s)$ such that $L_1 = I_0(M)$ and $M = L_1/H_1$. Let π denote the homomorphism of L onto L_1 such that $d\pi$ is the identity mapping $\mathfrak{l} \to \mathfrak{l}$. Then $\pi(H) \subset H_1$. Let φ denote the mapping $xH \to \pi(x)\,H_1$ of the group $G = L/H$ onto the manifold $M = L_1/H_1$ (see diagram). Then (G, φ) is a covering manifold of M (Lemma 13.4, Chapter I), and if $o = \varphi(e)$, the geodesic symmetries s_e and s_o of G and M, respectively, are related by $\varphi \circ s_e = s_o \circ \varphi$. Consider the closed subset $\Gamma = \varphi^{-1}(o)$ of G. We shall prove that Γ is a normal subgroup of G. The set $\tilde{H} = \pi^{-1}(H_1)$ is a closed subgroup of L; its identity component is H and $\Gamma = \tilde{H}/H$ (as subsets of G). Since H is a normal subgroup of $\tilde{H}$, Lemma 1.3 shows that $\varphi(g\gamma) = \varphi(g)$ for $g \in G$, $\gamma \in \Gamma$. Thus Γ is a subgroup of G and we can define a mapping $\beta : g\Gamma \to \varphi(g)$ of the coset space G/Γ onto M.

$$
\begin{tikzcd}
G = L/H \arrow[d, "\varphi"'] & \\
M & G/\Gamma \arrow[l, "\beta"'] \arrow[d, "\eta"] \\
M \arrow[from=u, "s_o"'] & G/\Gamma \arrow[l, "\beta"']
\end{tikzcd}
$$

If $\varphi(x_1H) = \varphi(x_2H)$, then $x_2^{-1}x_1 = h \in \tilde{H}$ so $(x_2H)(hH) = x_1H$. This shows that β is one-to-one. Finally consider the mapping $\eta : G/\Gamma \to G/\Gamma$ which corresponds to s_o under β. Then since $s_e(g) = g^{-1}$ and $\varphi \circ s_e = s_o \circ \varphi$, we find that $\eta(g\Gamma) = g^{-1}\Gamma$, $(g \in G)$. This requires that $(g\gamma)^{-1}\Gamma = g^{-1}\Gamma$ for each $g \in G$, $\gamma \in \Gamma$, so Γ is a normal subgroup of G and we can turn $M = \beta(G/\Gamma)$ into a group by requiring β to be an isomorphism. Finally since β is an isometry, the metric on M is invariant under left and right translations. This finishes the proof of Prop. 1.2.

We turn now to the type I. In view of Lemma 1.1, Theorem 5.3, Chapter VIII, and Theorem 8.1, Chapter VII, the classification problem for type I reduces to the following three problems.

A. *Find all compact simple Lie algebras, isomorphic Lie algebras not distinguished.*

B. *For each compact simple Lie algebra* $\mathfrak{u}$, *find all involutive automorphisms of* $\mathfrak{u}$, *not distinguishing automorphisms which are conjugate within the group Aut* $(\mathfrak{u})$.

C. *Find the centers of all compact, simple, simply connected Lie groups.*

Now every complex semisimple Lie algebra $\mathfrak{g}$ has a compact real form $\mathfrak{u}$ which is unique up to an inner automorphism (Cor. 7.3, Chapter III). It is clear that $\mathfrak{g}$ is simple if and only if $\mathfrak{u}$ is simple. Problem A is therefore equivalent to

A'. *Find all simple Lie algebras over* C, *isomorphic Lie algebras not distinguished.*

If $\mathfrak{u}$ runs through all compact real forms of a complex semisimple Lie algebra $\mathfrak{g}$, and s runs through all involutive automorphisms of $\mathfrak{u}$, then $\mathfrak{g}_0$, in the dual $(\mathfrak{g}_0, s^*)$ to $(\mathfrak{u}, s)$, runs through all noncompact real forms of $\mathfrak{g}$. We have also seen (Prop. 2.2, Chapter V) that conjugate automorphisms s correspond to isomorphic real forms $\mathfrak{g}_0$. Hence, problem B is equivalent to:

B′. *For each simple Lie algebra* $\mathfrak{g}$ *over* C, *find all noncompact real forms of* $\mathfrak{g}$ *up to isomorphism.*

The simplest solution of problem A′ takes Theorem 5.4, Chapter III as a point of departure. According to this theorem and Theorem 6.5, Chapter V, it suffices to classify the possible systems Γ of vectors H_α ($\alpha \in \Delta$) in the Euclidean space $\mathfrak{h}^*$ with the inner product B. The system Γ is far from arbitrary, because it satisfies the following conditions, established in Chapter III, §4.

(a) If $\lambda \in \Gamma$, then an integral multiple $n\lambda$ belongs to Γ if and only if $n = \pm 1$.

(b) Let λ, μ be two nonproportional vectors in Γ and $-p \geqslant 0$, $q \geqslant 0$ the largest integers such that $\lambda + n\mu \in \Gamma$ for the integers n satisfying $p \leqslant n \leqslant q$. Then

$$p + q = -2\frac{B(\lambda, \mu)}{B(\mu, \mu)}.$$

For a simple Lie algebra $\mathfrak{g}$ over C the following additional property holds.

(c) The system Γ is indecomposable, that is, Γ cannot be written as a union $\Gamma = \Gamma_1 \cup \Gamma_2$ such that each vector in Γ_1 is perpendicular to each vector in Γ_2.

In fact, suppose $\Gamma = \Gamma_1 \cup \Gamma_2$ were such a decomposition of Γ. Let $H_\alpha \in \Gamma_1$, $H_\beta \in \Gamma_2$. Then $\beta(H_\alpha) = B(H_\alpha, H_\beta) = 0$, and $B(H_\alpha, H_\alpha \pm H_\beta) \neq 0$ and $B(H_\beta, H_\alpha \pm H_\beta) \neq 0$. This shows that neither $\alpha + \beta$ nor $\alpha - \beta$ is a root because neither $H_{\alpha+\beta}$ nor $H_{\alpha-\beta}$ can belong to $\Gamma_1 \cup \Gamma_2$. Consequently, the subspaces $\mathfrak{g}_1$ and $\mathfrak{g}_2$ of $\mathfrak{g}$ determined by

$$\mathfrak{g}_1 = \sum_{H_\alpha \in \Gamma_1} CH_\alpha + \sum_{H_\alpha \in \Gamma_1} \mathfrak{g}^\alpha, \qquad \mathfrak{g}_2 = \sum_{H_\beta \in \Gamma_2} CH_\beta + \sum_{H_\beta \in \Gamma_2} \mathfrak{g}^\beta$$

are ideals in $\mathfrak{g}$ and we have a contradiction to the assumed simplicity of $\mathfrak{g}$.

It is possible by elementary geometric methods to classify all systems Γ satisfying the conditions (a), (b), and (c) (see, for example, Pontrjagin [1], §66). It turns out that the possibilities for Γ are so restricted that it can be verified that each Γ arises from a simple Lie algebra. For this verification problem see É. Cartan [1], pp. 138-139 (also Harish-Chandra [3], Theorem 1). The list of the compact forms of all simple Lie algebras over C is given in §4.

§ 2. Automorphisms

Let $\mathfrak{g}$ be a semisimple Lie algebra over C and let $\mathfrak{h}$ be a Cartan subalgebra. Let Δ be the set of nonzero roots of $\mathfrak{g}$ with respect to $\mathfrak{h}$ and suppose Δ is ordered in some way. As usual, let $\mathfrak{h}^*$ denote the real vector space $\sum_{\alpha\in\Delta} RH_\alpha$.

Definition. A root $\alpha > 0$ is called *simple* if it cannot be written as a sum $\alpha = \beta + \gamma$ where β and γ are positive roots.

Lemma 2.1. *Let α, β be simple roots, $\alpha \neq \beta$. Then $\beta - \alpha$ is not a root and $B(H_\alpha, H_\beta) \leqslant 0$.*

Proof. If $\beta - \alpha$ were a root, say γ, then $\gamma \in \Delta$. Writing $\beta = \alpha + \gamma$ if $\gamma > 0$ and $\alpha = \beta - \gamma$ if $\gamma < 0$ we get a contradiction to the simplicity of α and β. In the notation of Theorem 4.3, Chapter III, we have $p = 0$, $q \geqslant 0$. Since

$$- 2B(H_\beta, H_\alpha) = B(H_\alpha, H_\alpha)\,(p + q)$$

the lemma follows.

Theorem 2.2. *Let $\alpha_1, \ldots, \alpha_r$ be the set of all simple roots. Then $r = \dim \mathfrak{h}^*$ and each $\beta \in \Delta$ has the form $\beta = \sum_{i=1}^r n_i\alpha_i$ where the n_i are integers which are either all positive or all negative.*

Proof. The simple roots are linearly independent. Otherwise there would be a relation

$$\sum_i a_i\alpha_i = \sum_j b_j\alpha_j$$

with nonnegative real numbers a_i, b_j, not all zero. If $\gamma = \sum a_i\alpha_i$ and $H_\gamma \in \mathfrak{h}^*$ is determined by $B(H, H_\gamma) = \gamma(H)$ $(H \in \mathfrak{h})$, then

$$B(H_\gamma, H_\gamma) = \sum_{i,j} a_ib_jB(H_{\alpha_i}, H_{\alpha_j}).$$

The left-hand side is $\geqslant 0$ but the right-hand side is $\leqslant 0$ due to the lemma. Hence $\gamma = 0$ which is a contradiction.

Now let β be a root > 0. If β is not simple it can be written $\beta = \gamma + \delta$ where γ, δ are roots > 0. It follows by induction that $\beta = \sum_{i=1}^r n_i\alpha_i$ where each n_i is an integer $\geqslant 0$. It is now obvious that $r = \dim \mathfrak{h}^*$ and the theorem is proved.

Let $\mathfrak{u}$ be a compact semisimple Lie algebra. The group Aut $(\mathfrak{u})$ of all automorphisms of $\mathfrak{u}$ is a closed subgroup of the group $GL(\mathfrak{u})$ of all invertible endomorphisms of $\mathfrak{u}$. Each member of $\operatorname{Aut}(\mathfrak{u})$ leaves the Killing form of $\mathfrak{u}$ invariant. It follows that $\operatorname{Aut}(\mathfrak{u})$ is compact. The adjoint group $\operatorname{Int}(\mathfrak{u})$ is the identity component of $\operatorname{Aut}(\mathfrak{u})$. Let $\mathfrak{t}_0$ be a maximal abelian subalgebra of $\mathfrak{u}$ and let $\mathfrak{h}$ denote the subalgebra generated by $\mathfrak{t}_0$ in the complexification $\mathfrak{g}$ of $\mathfrak{u}$. Then $\mathfrak{h}$ is a Cartan subalgebra and we can use the results above. An endomorphism φ of $\mathfrak{t}_0$ will be called a *rotation* if it maps the set of vectors $iH_\alpha(\alpha \in \Delta)$ onto itself. Each element in the Weyl group $W = W(\mathfrak{u})$ of $\mathfrak{u}$ is a rotation. Let W' denote the group of all rotations.

Definition. Let $\mathfrak{u}$ be a compact, semisimple Lie algebra, $\mathfrak{t}_0$ a maximal abelian subalgebra, $\mathfrak{g}$ the complexification of $\mathfrak{u}$ and $\mathfrak{h}$ the subalgebra of $\mathfrak{g}$ generated by $\mathfrak{t}_0$. A *Weyl basis* of $\mathfrak{g} \bmod \mathfrak{h}$ with respect to $\mathfrak{u}$ is a basis $\{X_\alpha : \alpha \in \Delta\}$ of $\mathfrak{g} \bmod \mathfrak{h}$ with the following properties.

(i) $X_\alpha \in \mathfrak{g}^\alpha$ and $[X_\alpha, X_{-\alpha}] = H_\alpha$ for each $\alpha \in \Delta$.

(ii) $[X_\alpha, X_\beta] = N_{\alpha,\beta} X_{\alpha+\beta}$ if $\alpha, \beta, \alpha+\beta \in \Delta$ where the constants $N_{\alpha,\beta}$ satisfy $N_{\alpha,\beta} = -N_{-\alpha,-\beta}$.

(iii) $X_\alpha - X_{-\alpha} \in \mathfrak{u}$, $i(X_\alpha + X_{-\alpha}) \in \mathfrak{u}$ for each $\alpha \in \Delta$.

To get such a basis one can put $X_\alpha = (\alpha(H_\alpha)/2)^{1/2} E_\alpha$ in Lemma 3.1, Chapter VI. In fact, (i) and (iii) are immediate and imply $\tau X_\alpha = -X_{-\alpha}$ (τ being the conjugation of $\mathfrak{g}$ with respect to $\mathfrak{u}$) from which (ii) follows.

Theorem 2.3. *Let A be an automorphism of $\mathfrak{u}$ leaving $\mathfrak{t}_0$ invariant. Then the restriction of A to $\mathfrak{t}_0$ is a rotation. On the other hand, each rotation of $\mathfrak{t}_0$ can be extended to an automorphism of $\mathfrak{u}$.*

Proof. It is clear that A extends uniquely to an automorphism of $\mathfrak{g}$. Denoting this extension by A, we have $A\mathfrak{h} \subset \mathfrak{h}$. For each $\alpha \in \Delta$ the linear function α^A on $\mathfrak{h}$ given by $\alpha^A(H) = \alpha(A^{-1}H)$ is a root. Since $A \cdot H_\alpha = H_{\alpha^A}$, the restriction of A to $\mathfrak{t}_0$ is a rotation. On the other hand, let φ be a rotation of $\mathfrak{t}_0$. The extension of φ to an endomorphism of $\mathfrak{h}$ will also be denoted φ. Let $\alpha' \in \Delta$ be defined by $H_{\alpha'} = \varphi H_\alpha$. Then we have from Theorem 4.3, Chapter III,

$$\frac{\beta'(H_{\alpha'})}{\alpha'(H_{\alpha'})} = \frac{\beta(H_\alpha)}{\alpha(H_\alpha)}, \qquad \alpha, \beta \in \Delta. \tag{1}$$

As shown in the proof of Theorem 5.4, Chapter III, relation (1) implies that

$$B(\varphi H, \varphi H') = B(H, H') \qquad \text{for } H, H' \in \mathfrak{h}. \tag{2}$$

From (2) it is easily seen that ${}^t\varphi \cdot \alpha' = \alpha$ $(\alpha \in \Delta)$ in the sense of Theorem 5.4, Chapter III. Hence φ can be extended to an automorphism

A of $\mathfrak{g}$. We shall now replace A by an automorphism which leaves $\mathfrak{u}$ invariant and coincides with A on $\mathfrak{h}$. Let $\{X_\alpha : \alpha \in \Delta\}$ be a Weyl basis of $\mathfrak{g}$ mod $\mathfrak{h}$ with respect to $\mathfrak{u}$. For each $\alpha \in \Delta$, let a_α be determined by $AX_\alpha = a_\alpha X_{\alpha'}$. Since $[X_\alpha, X_{-\alpha}] = H_\alpha$ it follows that

$$a_\alpha a_{-\alpha} = 1. \tag{3}$$

The numbers $N_{\alpha,\beta}$ determined by $[X_\alpha, X_\beta] = N_{\alpha,\beta} X_{\alpha+\beta}$ satisfy (Theorem 5.5, Chapter III)

$$N_{\alpha,\beta}^2 = \tfrac{1}{2}\,\alpha(H_\alpha)\,q(1-p)$$

and this number is determined by the root pattern. It follows that $N_{\alpha,\beta} = \pm N_{\alpha',\beta'}$ so

$$a_\alpha a_\beta = \pm\, a_{\alpha+\beta} \qquad \text{if } \alpha, \beta, \alpha+\beta \in \Delta. \tag{4}$$

Let $H' \in \mathfrak{h}$. Then the automorphism $B = e^{\operatorname{ad}(H')}$ of $\mathfrak{g}$ leaves $\mathfrak{h}$ pointwise fixed and

$$BX_\alpha = e^{\alpha(H')}\, X_\alpha, \qquad \alpha \in \Delta.$$

Let $\alpha_1, \ldots, \alpha_r$ be the system of simple roots. Since $a_\alpha \neq 0$, it is clear that $H' \in \mathfrak{h}$ can be chosen such that

$$a_{\alpha_i} = e^{-\alpha_i(H')}, \qquad 1 \leqslant i \leqslant r.$$

Then it follows from (4) by induction that

$$a_\alpha = \pm \prod (a_{\alpha_i})^{n_i} = \pm\, e^{-\alpha(H')}$$

if $\alpha = \sum_1^r n_i \alpha_i$. This implies that

$$ABX_\alpha = \pm\, X_{\alpha'} \quad \text{and} \quad ABX_{\alpha_i} = X_{\alpha_i'} \tag{5}$$

for $1 \leqslant i \leqslant r$. Since $\mathfrak{u}$ is spanned by $i\mathfrak{t}_0$ and $X_\alpha - X_{-\alpha}$, $i(X_\alpha + X_{-\alpha})$ $(\alpha \in \Delta)$, it is clear that AB leaves $\mathfrak{u}$ invariant and therefore gives the desired extension of φ.

Corollary 2.4. *Let A be an automorphism of $\mathfrak{u}$ leaving $\mathfrak{t}_0$ invariant. Then the extension of A to $\mathfrak{g}$ satisfies*

$$AX_\alpha = a_\alpha X_{\alpha^A},$$

where $a_\alpha a_{-\alpha} = 1$ and $|a_\alpha| = 1$.

In fact, $a_\alpha a_{-\alpha} = 1$ as before, but now we have in addition $A\tau = \tau A$. Since $\tau X_\alpha = -X_{-\alpha}$, this implies $a_{-\alpha} = \bar{a}_\alpha$ so $|a_\alpha| = 1$.

Proposition 2.5. *An automorphism A of* $\mathfrak{u}$ *leaves* $\mathfrak{t}_0$ *pointwise fixed if and only if it has the form*

$$A = e^{\operatorname{ad} H}$$

for a suitable element $H \in \mathfrak{t}_0$.

Proof. Let the extension of A to $\mathfrak{g}$ also be denoted by A. In the notation above we have $\alpha = \alpha'$. Hence (4) and (5) take the form $a_\alpha a_\beta = a_{\alpha+\beta}$ and $ABX_\alpha = + X_\alpha$ so AB is the identity. Thus $A = e^{\operatorname{ad} H}$ for some $H \in \mathfrak{h}$. The eigenvalues of A are 1 and $e^{\alpha(H)}$ ($\alpha \in \Delta$). Since the powers of A form a bounded set, $\alpha(H)$ must be purely imaginary for each $\alpha \in \Delta$. Hence $H \in \mathfrak{t}_0$ and the proposition is proved.

Theorem 2.6. *The factor group Aut* $(\mathfrak{u})$*/Int* $(\mathfrak{u})$ *is isomorphic to* W'/W, *W denoting the Weyl group, W' denoting the group of all rotations of* $\mathfrak{t}_0$.

Proof. Let $E \in \operatorname{Aut}(\mathfrak{u})$. Then $E\mathfrak{t}_0$ is a maximal abelian subalgebra of $\mathfrak{u}$ so there exists a $B_1 \in \operatorname{Int}(\mathfrak{u})$ such that $B_1E\mathfrak{t}_0 = \mathfrak{t}_0$. Consequently, each element in Aut $(\mathfrak{u})$/Int $(\mathfrak{u})$ contains an automorphism leaving $\mathfrak{t}_0$ invariant. On the other hand, W consists of the rotations in $\mathfrak{t}_0$ induced by members of Int $(\mathfrak{u})$ leaving $\mathfrak{t}_0$ invariant. Since W is generated by the reflections in the planes $\alpha(H) = 0$, it is easy to see that W is a normal subgroup of W'. We obtain now a well-defined mapping S of Aut $(\mathfrak{u})$/Int $(\mathfrak{u})$ into W'/W as follows: In each class of Aut $(\mathfrak{u})$ mod Int $(\mathfrak{u})$ select an element A leaving $\mathfrak{t}_0$ invariant; let A_t denote the restriction of A to $\mathfrak{t}_0$ and let

$$S : A \to A_t W.$$

Since W is a normal subgroup of W', S is a homomorphism. It is clear from Prop. 2.5 that S is one-to-one and Theorem 2.3 shows that S is onto.

§ 3. Involutive Automorphisms

We recall that the *rank* of a compact Lie algebra is defined as the dimension of any maximal abelian subalgebra.

Theorem 3.1. *Let θ be an involutive automorphism of a compact semisimple Lie algebra* $\mathfrak{u}$. *Let* $\mathfrak{k}_0$ *denote the set of fixed points of θ. Then* $\theta \in$ *Int* $(\mathfrak{u})$ *if and only if rank* $\mathfrak{u}$ = *rank* $\mathfrak{k}_0$.

Proof. Let U be any connected Lie group with Lie algebra $\mathfrak{u}$. Suppose first $\theta \in \operatorname{Int}(\mathfrak{u})$. Then $\theta = \operatorname{Ad}(u)$ for some $u \in U$. Now u lies in a maximal torus T of U and θ leaves the Lie algebra $\mathfrak{t}_0$ of T

pointwise fixed. Hence $\mathfrak{t}_0 \subset \mathfrak{k}_0$ so rank $\mathfrak{k}_0 =$ rank $\mathfrak{u}$. On the other hand, if $\mathfrak{k}_0$ and $\mathfrak{u}$ have the same rank, there exists a subalgebra $\mathfrak{t}_0 \subset \mathfrak{k}_0$ which is maximal abelian in $\mathfrak{u}$. Proposition 2.5 shows at once that $\theta \in \text{Int}(\mathfrak{u})$.

Theorem 3.2. *Let $(\mathfrak{g}_0, \theta)$ be an orthogonal symmetric Lie algebra of the noncompact type and let $\mathfrak{k}_0$ denote the set of fixed points of θ. Then $\theta \in Int(\mathfrak{g}_0)$ if and only if $\mathfrak{k}_0$ contains a maximal abelian subalgebra of $\mathfrak{g}_0$.*

Let as usual $\mathfrak{g}_0 = \mathfrak{k}_0 + \mathfrak{p}_0$ and let $\mathfrak{u}$ denote the subspace $\mathfrak{k}_0 + i\mathfrak{p}_0$ of the complexification $\mathfrak{g}$ of $\mathfrak{g}_0$. Then θ extends to an involutive automorphism of $\mathfrak{g}$, also denoted θ, leaving the compact real form $\mathfrak{u}$ invariant. Suppose $\mathfrak{t}_0$ is a maximal abelian subalgebra of $\mathfrak{g}_0$ contained in $\mathfrak{k}_0$. Then $\mathfrak{t}_0$ is also maximal abelian in $\mathfrak{u}$ so $\theta = e^{\text{ad}\,H}$ where $H \in \mathfrak{t}_0$. On the other hand, suppose $\theta \in \text{Int}(\mathfrak{g}_0)$. Since $\theta \in \text{Aut}(\mathfrak{u})$ we see that θ lies in the Lie subgroup $\text{Int}(\mathfrak{g}_0) \cap \text{Aut}(\mathfrak{u})$ of $\text{Int}(\mathfrak{g}_0)$ which has Lie algebra $\mathfrak{g}_0 \cap \mathfrak{u} = \mathfrak{k}_0$. It follows from Theorem 1.1, Chapter VI, that $\text{Int}(\mathfrak{g}_0) \cap \text{Aut}(\mathfrak{u}) = \text{Int}(\mathfrak{g}_0) \cap \text{Int}(\mathfrak{u})$. Hence $\theta \in \text{Int}(\mathfrak{u})$ and Theorem 3.2 follows from Theorem 3.1.

Corollary 3.3. *Let $M = I_0(M)/K$ be a Riemannian globally symmetric space of the noncompact type, K being the isotropy subgroup of $I_0(M)$ at some point $o \in M$. Let $I_0(M)$ and K have Lie algebras $\mathfrak{g}_0$ and $\mathfrak{k}_0$, respectively. Then the geodesic symmetry s_o belongs to $I_0(M)$ if and only if $\mathfrak{k}_0$ contains a maximal abelian subalgebra of $\mathfrak{g}_0$.*

In fact, if $s_o \in I_0(M)$ and if Ad denotes the adjoint representation of $I_0(M)$ then $\theta = \text{Ad}(s_o) \in \text{Int}(\mathfrak{g}_0)$. On the other hand, if $\theta \in \text{Int}(\mathfrak{g}_0)$, then $\theta \in \text{Int}(\mathfrak{g}_0) \cap \text{Aut}(\mathfrak{u}) = \text{Ad}(K)$ by Theorem 1.1, Chapter VI. Let $s \in K$ such that $\text{Ad}(s) = \theta$. Then $(ds)_o = -I$ so $s = s_o$.

Let θ_1 and θ_2 be two involutive automorphisms of $\mathfrak{u}$ leaving a maximal abelian subalgebra $\mathfrak{t}_0$ invariant, such that θ_1 and θ_2 are identical on $\mathfrak{t}_0$. In general θ_1 and θ_2 are not conjugate† in $\text{Aut}(\mathfrak{u})$. However, we have

Theorem 3.4. *Let θ_1 and θ_2 be two involutive automorphisms of $\mathfrak{u}$ such that $\theta_1(H) = \theta_2(H) = -H$ for§ $H \in \mathfrak{t}_0$. Then there exists a $\sigma \in Int(\mathfrak{u})$ such that $\theta_2 = \sigma\theta_1\sigma^{-1}$.*

Proof. Let $\{X_\alpha : \alpha \in \Delta\}$ be a Weyl basis of $\mathfrak{g}$ mod $\mathfrak{h}$ with respect to $\mathfrak{u}$. Then for each $\alpha \in \Delta$

$$\theta_1 X_\alpha = a_\alpha X_{-\alpha}, \tag{1}$$

† An example is given by the spaces *A III* in É. Cartan's list (§4) of Riemannian globally symmetric spaces.

§ Such automorphisms exist as a result of Theorem 2.3.

where the number a_α satisfies

$$a_\alpha a_{-\alpha} = 1, \qquad |a_\alpha| = 1, \qquad a_\alpha a_\beta = -a_{\alpha+\beta},$$

if $\alpha, \beta, \alpha+\beta \in \Delta$. Suppose now Δ ordered in some way; let Δ^+ denote the set of positive roots, $\alpha_1, \ldots, \alpha_r$ the simple roots. There exists a vector $H_1 \in \mathfrak{h}$ such that

$$a_{\alpha_j} = -e^{\alpha_j(H_1)} \qquad (1 \leqslant j \leqslant r), \tag{2}$$

and since each a_α has modulus 1, we have $H_1 \in \mathfrak{t}_0$. Now $(-a_\alpha)(-a_\beta) = (-a_{\alpha+\beta})$ so we obtain from (2) by induction

$$-a_\alpha = e^{\alpha(H_1)}, \qquad \alpha \in \Delta.$$

Thus we have for all $\alpha \in \Delta$

$$\theta_1 X_\alpha = -e^{\alpha(H_1)} X_{-\alpha}, \qquad \theta_2 X_\alpha = -e^{\alpha(H_2)} X_{-\alpha}, \tag{3}$$

where H_1 and H_2 are certain fixed vectors in $\mathfrak{t}_0$. Consider now the automorphism

$$\sigma = e^{\frac{1}{2}\operatorname{ad}(H_1-H_2)}$$

of $\mathfrak{g}$. This automorphism leaves $\mathfrak{u}$ invariant and keeps $\mathfrak{t}_0$ pointwise fixed. Since

$$\theta_2\sigma X_\alpha = \theta_2 e^{\frac{1}{2}\alpha(H_1-H_2)} X_\alpha = -e^{\frac{1}{2}\alpha(H_1+H_2)} X_{-\alpha},$$

$$\sigma\theta_1 X_\alpha = -e^{\alpha(H_1)}\sigma X_{-\alpha} = -e^{\frac{1}{2}\alpha(H_1+H_2)} X_{-\alpha},$$

the automorphism σ has the required properties.

Definition. Let $\mathfrak{g}$ be a semisimple Lie algebra over C. A real form $\mathfrak{g}_0$ of $\mathfrak{g}$ is called *normal* if in each Cartan decomposition $\mathfrak{g}_0 = \mathfrak{k}_0 + \mathfrak{p}_0$, the space $\mathfrak{p}_0$ contains a maximal abelian subalgebra of $\mathfrak{g}_0$.

Theorem 3.5. *Each semisimple Lie algebra $\mathfrak{g}$ over C has a normal real form and this is unique up to isomorphism.*

Proof. Let $\mathfrak{u}_0$ be a compact real form of $\mathfrak{g}$, $\mathfrak{t}_0$ a maximal abelian subalgebra of $\mathfrak{u}_0$, and $\mathfrak{h}$ the subspace of $\mathfrak{g}$ generated by $\mathfrak{t}_0$. Let $\{X_\alpha : \alpha \in \Delta\}$ be a Weyl basis of $\mathfrak{g}$ mod $\mathfrak{h}$ with respect to $\mathfrak{u}_0$. Then the subspace

$$\mathfrak{g}_0 = \sum_{\alpha\in\Delta} RH_\alpha + \sum_{\alpha\in\Delta} RX_\alpha$$

is a real form of $\mathfrak{g}$. The conjugation τ of $\mathfrak{g}$ with respect to $\mathfrak{u}_0$ leaves $\mathfrak{g}_0$ invariant and if $\mathfrak{k}_0 = \mathfrak{g}_0 \cap \mathfrak{u}_0$, $\mathfrak{p}_0 = \mathfrak{g}_0 \cap (i\mathfrak{u}_0)$, then

$$\mathfrak{g}_0 = \mathfrak{k}_0 + \mathfrak{p}_0$$

is a Cartan decomposition of $\mathfrak{g}_0$. Moreover, $\mathfrak{p}_0$ contains the subspace $\sum_{\alpha\in\Delta} \boldsymbol{R}H_\alpha$ which is a maximal abelian subalgebra of $\mathfrak{g}_0$. Hence $\mathfrak{g}_0$ is a normal real form. On the other hand, let $\mathfrak{g}_1$ be another normal real form of $\mathfrak{g}$, and $\mathfrak{g}_1 = \mathfrak{k}_1 + \mathfrak{p}_1$ any Cartan decomposition of $\mathfrak{g}_1$. Then $\mathfrak{u}_1 = \mathfrak{k}_1 + i\mathfrak{p}_1$ is a compact real form of $\mathfrak{g}$. The mapping $\theta_i : T + X \to T - X$ ($T \in \mathfrak{k}_i$, $X \in i\mathfrak{p}_i$) is an involutive automorphism of $\mathfrak{u}_i$ ($i = 0, 1$). There exists a maximal abelian subalgebra $\mathfrak{t}_i$ of $\mathfrak{u}_i$ such that $\theta_i(H) = -H$ for $H \in \mathfrak{t}_i$, ($i = 0, 1$). Here $\mathfrak{t}_0$ is the same as that above. Owing to previous conjugacy theorems, there exists an automorphism A of $\mathfrak{g}$ mapping $\mathfrak{u}_1$ onto $\mathfrak{u}_0$ such that $A\mathfrak{t}_1 = \mathfrak{t}_0$. According to Theorem 3.4, the automorphisms $A\theta_1 A^{-1}$ and θ_0 of $\mathfrak{u}_0$ are conjugate within $\mathrm{Int}(\mathfrak{u}_0)$. If $\mathfrak{g}_A$ denotes the real form of $\mathfrak{g}$ which corresponds to the involution $A\theta_1 A^{-1}$ of $\mathfrak{u}_0$, then Prop. 2.2, Chapter V, shows that $\mathfrak{g}_A$ and $\mathfrak{g}_0$ are isomorphic. On the other hand, A induces (by restriction) an isomorphism of $\mathfrak{g}_1$ onto $\mathfrak{g}_A$ and consequently $\mathfrak{g}_0$ and $\mathfrak{g}_1$ are isomorphic.

Let $M = I_0(M)/K$ be an irreducible Riemannian globally symmetric space of the noncompact type. The list of such spaces given in the next section shows that the Lie group K by itself does not determine M, not even locally. Nevertheless, the next theorem shows that the linear isotropy group K^* determines the curvature tensor of M at $\{K\}$ and therefore (M being simply connected) determines the space M. It is known that for an irreducible Riemannian globally symmetric space the holonomy group and the linear isotropy group have the same identity component. Therefore, the theorem below shows that problem 2 stated in the introduction to Chapter IV has an affirmative solution (after decomposition into irreducible factors).

Theorem 3.6. *Let $\mathfrak{g}_1 = \mathfrak{k}_1 + \mathfrak{p}_1$, $\mathfrak{g}_2 = \mathfrak{k}_2 + \mathfrak{p}_2$ be Cartan decompositions of two semisimple Lie algebras $\mathfrak{g}_1$ and $\mathfrak{g}_2$ over $\boldsymbol{R}$. Assume that $ad_{\mathfrak{g}_i}(\mathfrak{k}_i)$ acts irreducibly on $\mathfrak{p}_i$ ($i = 1, 2$). Let φ be a one-to-one linear mapping of $\mathfrak{g}_1$ onto $\mathfrak{g}_2$ such that*

(i) $\varphi(\mathfrak{p}_1) = \mathfrak{p}_2$.

(ii) *The restriction of φ to $\mathfrak{k}_1$ is an isomorphism of $\mathfrak{k}_1$ onto $\mathfrak{k}_2$.*

(iii) $\varphi([T, X]) = [\varphi(T), \varphi(X)]$ *for* $T \in \mathfrak{k}_1$, $X \in \mathfrak{p}_1$.

Then $\mathfrak{g}_1$ and $\mathfrak{g}_2$ are isomorphic.

Proof. Let B_1, B_2, Q_1, and Q_2 denote the Killing forms of $\mathfrak{g}_1, \mathfrak{g}_2$, $\mathfrak{k}_1$, and $\mathfrak{k}_2$, respectively.

Then

$$B_i(T, T) = Q_i(T, T) + \mathrm{Tr}_{\mathfrak{p}_i}(\mathrm{ad}_{\mathfrak{g}_i}(T)\,\mathrm{ad}_{\mathfrak{g}_i}(T)) \qquad (T \in \mathfrak{k}_i)$$

for $i = 1, 2$. Consider the bilinear form Q on $\mathfrak{p}_1 \times \mathfrak{p}_1$ given by

$$Q(X, Y) = B_2(\varphi(X), \varphi(Y)), \qquad X, Y \in \mathfrak{p}_1.$$

By (iii) Q is invariant under the action of $\mathfrak{k}_1$ on $\mathfrak{p}_1$, that is,

$$Q([T, X], Y) + Q(X, [T, Y]) = 0 \qquad X, Y \in \mathfrak{p}_1, \quad T \in \mathfrak{k}_1.$$

Since the action of $\mathfrak{k}_1$ on $\mathfrak{p}_1$ is irreducible, Q is proportional to the restriction of B_1 to $\mathfrak{p}_1 \times \mathfrak{p}_1$; hence

$$B_2(\varphi(X), \varphi(Y)) = dB_1(X, Y), \qquad X, Y \in \mathfrak{p}_1, \tag{4}$$

where d is a constant, $d > 0$. On the other hand, let us compare

$$B_2([\varphi(X), \varphi(Y)], \varphi(T)) \qquad \text{and} \qquad B_2(\varphi([X, Y]), \varphi(T))$$

for $X, Y \in \mathfrak{p}_1$, $T \in \mathfrak{k}_1$. We have

$$\begin{aligned} B_2(\varphi([X, Y]), \varphi(T)) &= Q_2(\varphi([X, Y], \varphi(T)) + \mathrm{Tr}_{\mathfrak{p}_2}(\mathrm{ad}(\varphi([X, Y]))\,\mathrm{ad}(\varphi(T))) \\ &= Q_1([X, Y], T) + \mathrm{Tr}_{\mathfrak{p}_1}(\mathrm{ad}([X, Y])\,\mathrm{ad}(T)), \end{aligned}$$

where we have used (ii) and the relation

$$\mathrm{ad}\, T(X) = (\varphi^{-1} \circ \mathrm{ad}\,(\varphi(T)) \circ \varphi)\,(X).$$

This proves that

$$B_2(\varphi([X, Y]), \varphi(T)) = B_1([X, Y], T). \tag{5}$$

On the other hand, using (iii) and (4) we have

$$\begin{aligned} B_2([\varphi(X), \varphi(Y)], \varphi(T)) &= B_2(\varphi(X), [\varphi(Y), \varphi(T)]) = B_2(\varphi(X), \varphi([Y, T]) \\ &= dB_1(X,[Y, T]) = dB_1([X, Y], T) \end{aligned}$$

so by (5)

$$B_2([\varphi(X), \varphi(Y)], \varphi(T)) = dB_2(\varphi([X, Y]), \varphi(T)). \tag{6}$$

Since B_2 is strictly negative definite on $\mathfrak{k}_2$, (6) implies

$$\varphi([X, Y]) = d^{-1}[\varphi(X), \varphi(Y)], \qquad X, Y \in \mathfrak{p}_1.$$

The desired isomorphism of $\mathfrak{g}_1$ onto $\mathfrak{g}_2$ is now obtained by defining

$$\psi(T) = \varphi(T) \quad (T \in \mathfrak{k}_1), \quad \psi(X) = d^{-1/2}\varphi(X) \qquad (X \in \mathfrak{p}_1).$$

Remark. Cartan's classification of real forms which is described in the next section shows that a simple Lie algebra over $\boldsymbol{R}$ is determined by its complexification and the structure of a maximal compactly imbedded subalgebra. In other words, Theorem 3.6 holds if (iii) is replaced by the assumption that $\mathfrak{g}_1$ and $\mathfrak{g}_2$ are real forms of the same Lie algebra over $\boldsymbol{C}$. However, no *a priori* proof of this result seems to be known.

§ 4. É. Cartan's List of Irreducible Riemannian Globally Symmetric Spaces

1. Some Matrix Groups and Their Lie Algebras

In order to describe Cartan's classification, we adopt the following (mostly standard) notation. Let $(x_1, \ldots, x_n)$ and $(z_1, \ldots, z_n)$ be variable points in $\boldsymbol{R}^n$ and $\boldsymbol{C}^n$, respectively. A matrix $A = (a_{ij})_{1 \leqslant i,j \leqslant n}$ operates on $\boldsymbol{C}^n$ by the rule

$$\begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} \to \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}.$$

As before, E_{ij} denotes the matrix $(\delta_{ai}\delta_{bj})_{1 \leqslant a,b \leqslant n}$. The transpose and conjugate of a matrix A are denoted by ${}^tA$ and $\bar{A}$, respectively; A is called skew symmetric if $A + {}^tA = 0$, Hermitian if ${}^tA = \bar{A}$, skew Hermitian if ${}^tA + \bar{A} = 0$.

If I_n denotes the unit matrix of order n, we put

$$I_{p,q} = \begin{pmatrix} -I_p & 0 \\ 0 & I_q \end{pmatrix}, \qquad J_n = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix},$$

$$K_{p,q} = \begin{pmatrix} -I_p & 0 & 0 & 0 \\ 0 & I_q & 0 & 0 \\ 0 & 0 & -I_p & 0 \\ 0 & 0 & 0 & I_q \end{pmatrix}.$$

The multiplicative group of complex numbers of modulus 1 will be denoted by $\boldsymbol{T}$.

$\boldsymbol{GL}(n, \boldsymbol{C})$, $(\boldsymbol{GL}(n, \boldsymbol{R}))$: The group of complex (real) $n \times n$ matrices of determinant $\neq 0$.

$\boldsymbol{SL}(n, \boldsymbol{C})$, $(\boldsymbol{SL}(n, \boldsymbol{R}))$: The group of complex (real) $n \times n$ matrices of determinant 1.

$U(p, q)$: The group of matrices in $GL(p + q, C)$ which leave invariant the Hermitian form

$$- z_1\bar{z}_1 - \dots - z_p\bar{z}_p + z_{p+1}\bar{z}_{p+1} + \cdots + z_{p+q}\bar{z}_{p+q}.$$

We put $U(n) = U(0, n) = U(n, 0)$ and $SU(p, q) = U(p,q) \cap SL(p+q,C)$, $SU(n) = U(n) \cap SL(n, C)$. Moreover, let $S(U_p \times U_q)$ denote the set of matrices

$$\begin{pmatrix} g_1 & 0 \\ 0 & g_2 \end{pmatrix},$$

where $g_1 \in U(p)$, $g_2 \in U(q)$ and $\det g_1 \det g_2 = 1$.

$SU^*(2n)$: The group of matrices in $SL(2n, C)$ which commute with the transformation ψ of C^{2n} given by

$$(z_1, \dots, z_n, z_{n+1}, \dots, z_{2n}) \to (\bar{z}_{n+1}, \dots, \bar{z}_{2n}, - \bar{z}_1, \dots, - \bar{z}_n).$$

$SO(n, C)$: The group of matrices in $SL(n, C)$ which leave invariant the quadratic form

$$z_1^2 + \dots + z_n^2.$$

$SO(p, q)$: The group of matrices in $SL(p + q, R)$ which leave invariant the quadratic form

$$- x_1^2 - \dots - x_p^2 + x_{p+1}^2 + \dots + x_{p+q}^2.$$

We put $SO(n) = SO(0, n) = SO(n, 0)$.

$SO^*(2n)$: The group of matrices in $SO(2n, C)$ which leave invariant the skew Hermitian form

$$- z_1\bar{z}_{n+1} + z_{n+1}\bar{z}_1 - z_2\bar{z}_{n+2} + z_{n+2}\bar{z}_2 - \dots - z_n\bar{z}_{2n} + z_{2n}\bar{z}_n.$$

$Sp(n, C)$: The group of matrices in $GL(2n, C)$ which leave invariant the exterior form

$$z_1 \wedge z_{n+1} + z_2 \wedge z_{n+2} + \dots + z_n \wedge z_{2n}.$$

$Sp(n, R)$: The group of matrices in $GL(2n, R)$ which leave invariant the exterior form

$$x_1 \wedge x_{n+1} + x_2 \wedge x_{n+2} + \dots + x_n \wedge x_{2n}.$$

$Sp(p, q)$: The group of matrices in $Sp(p + q, C)$ which leave invariant the Hermitian form

$${}^tZK_{p,q}\bar{Z}.$$

We put $Sp(n) = Sp(0, n) = Sp(n, 0)$. It is clear that $Sp(n) = Sp(n, C) \cap U(2n)$.

The groups listed above are all topological Lie subgroups of a general linear group. The Lie algebra of the general linear group $GL(n, C)$ can (as in Chapter II, §1) be identified with the Lie algebra $\mathfrak{gl}(n, C)$ of all complex $n \times n$ matrices, the bracket operation being $[A, B] = AB-BA$. The Lie algebra for each of the groups above is then canonically identified with a subalgebra of $\mathfrak{gl}(n, C)$, considered as a real Lie algebra. These Lie algebras will be denoted by the corresponding small German letters, $\mathfrak{sl}(n, R)$, $\mathfrak{su}(p, q)$, etc.

Now, if G is a Lie group with Lie algebra $\mathfrak{g}$, then the Lie algebra $\mathfrak{h}$ of a topological Lie subgroup H of G is given by

$$\mathfrak{h} = \{X \in \mathfrak{g} : \exp tX \in H \text{ for } t \in R\}. \tag{1}$$

Using this fact (Chapter II, §2) we can describe the Lie algebras of the groups above more explicitly. Since the computation is fairly similar for all the groups we shall give the details only in the cases $SU^*(2n)$ and $Sp(n, C)$.

$\mathfrak{gl}(n, C)$, $(\mathfrak{gl}(n, R))$: {all $n \times n$ complex (real) matrices},

$\mathfrak{sl}(n, C)$, $(\mathfrak{sl}(n, R))$: {all $n \times n$ complex (real) matrices of trace 0},

$$\mathfrak{u}(p, q) : \left\{\begin{pmatrix} Z_1 & Z_2 \\ {}^t\bar{Z}_2 & Z_3 \end{pmatrix} \middle| \begin{array}{l} Z_1, Z_3 \text{ skew Hermitian of order } p \text{ and } q, \\ \text{respectively, } Z_2 \text{ arbitrary} \end{array}\right\},$$

$$\mathfrak{su}(p, q) : \left\{\begin{pmatrix} Z_1 & Z_2 \\ {}^t\bar{Z}_2 & Z_3 \end{pmatrix} \middle| \begin{array}{l} Z_1, Z_3 \text{ skew Hermitian, of order } p \text{ and } q, \\ \text{respectively, } \operatorname{Tr} Z_1 + \operatorname{Tr} Z_3 = 0,\ Z_2 \text{ arbitrary} \end{array}\right\},$$

$$\mathfrak{su}^*(2n) : \left\{\begin{pmatrix} Z_1 & Z_2 \\ -\bar{Z}_2 & \bar{Z}_1 \end{pmatrix} \middle| \begin{array}{l} Z_1, Z_2\ n \times n \text{ complex matrices} \\ \operatorname{Tr} Z_1 + \operatorname{Tr} \bar{Z}_1 = 0 \end{array}\right\},$$

$\mathfrak{so}(n, C)$: {all $n \times n$ skew symmetric complex matrices},

$$\mathfrak{so}(p, q) : \left\{\begin{pmatrix} X_1 & X_2 \\ {}^tX_2 & X_3 \end{pmatrix} \middle| \begin{array}{l} \text{All } X_i \text{ real, } X_1, X_3 \text{ skew symmetric of order} \\ p \text{ and } q, \text{ respectively, } X_2 \text{ arbitrary} \end{array}\right\},$$

$$\mathfrak{so}^*(2n) : \left\{\begin{pmatrix} Z_1 & Z_2 \\ -\bar{Z}_2 & \bar{Z}_1 \end{pmatrix} \middle| \begin{array}{l} Z_1, Z_2\ n \times n \text{ complex matrices,} \\ Z_1 \text{ skew, } Z_2 \text{ Hermitian} \end{array}\right\},$$

$$\mathfrak{sp}(n, C) : \left\{\begin{pmatrix} Z_1 & Z_2 \\ Z_3 & -{}^tZ_1 \end{pmatrix} \middle| \begin{array}{l} Z_i \text{ complex } n \times n \text{ matrices,} \\ Z_2 \text{ and } Z_3 \text{ symmetric} \end{array}\right\},$$

$$\mathfrak{sp}(n, R) : \left\{\begin{pmatrix} X_1 & X_2 \\ X_3 & -{}^tX_1 \end{pmatrix} \middle| \begin{array}{l} X_1, X_2, X_3 \text{ real } n \times n \text{ matrices,} \\ X_2, X_3 \text{ symmetric} \end{array}\right\},$$

$$\mathfrak{sp}(p, q) : \left\{\begin{pmatrix} Z_{11} & Z_{12} & Z_{13} & Z_{14} \\ {}^t\bar{Z}_{12} & Z_{22} & {}^tZ_{14} & Z_{24} \\ -\bar{Z}_{13} & \bar{Z}_{14} & \bar{Z}_{11} & -\bar{Z}_{12} \\ {}^t\bar{Z}_{14} & -\bar{Z}_{24} & -{}^tZ_{12} & \bar{Z}_{22} \end{pmatrix} \middle| \begin{array}{l} Z_{ij} \text{ complex matrix; } Z_{11} \text{ and } Z_{13} \text{ of} \\ \text{order } p,\ Z_{12} \text{ and } Z_{14}\ p \times q \text{ matrices,} \\ Z_{11} \text{ and } Z_{22} \text{ are skew Hermitian,} \\ Z_{13} \text{ and } Z_{24} \text{ are symmetric} \end{array}\right\}.$$

Proof for $SU^*(2n)$. By the definition of this group, we have $g \in SU^*(2n)$ if and only if $g\psi = \psi g$ and $\det g = 1$. This shows that $A \in \mathfrak{su}^*(2n)$ if and only if $A\psi = \psi A$ and $\operatorname{Tr} A = 0$. Writing A in the form

$$A = \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix},$$

where A_i are $n \times n$ complex matrices we see that if U and V are $n \times 1$ matrices, then

$$A\psi \begin{pmatrix} U \\ V \end{pmatrix} = \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix} \begin{pmatrix} \bar{V} \\ -\bar{U} \end{pmatrix} = \begin{pmatrix} A_1\bar{V} - A_2\bar{U} \\ A_3\bar{V} - A_4\bar{U} \end{pmatrix},$$

$$\psi A \begin{pmatrix} U \\ V \end{pmatrix} = \psi \begin{pmatrix} A_1 U + A_2 V \\ A_3 U + A_4 V \end{pmatrix} = \begin{pmatrix} \bar{A}_3\bar{U} + \bar{A}_4\bar{V} \\ -\bar{A}_1\bar{U} - \bar{A}_2\bar{V} \end{pmatrix}.$$

It follows that $\bar{A}_3 = -A_2$, $A_1 = \bar{A}_4$ as desired.

Proof for $Sp(n, C)$. Writing symbolically

$$2(z_1 \wedge z_{n+1} + \dots + z_n \wedge z_{2n}) = (z_1, \dots, z_{2n}) \wedge J_n{}^t(z_1, \dots, z_{2n})$$

it is clear that $g \in Sp(n, C)$ if and only if

$${}^t g J_n g = J_n.$$

Using this for $g = \exp tZ$ $(t \in R)$, we find since $A \exp ZA^{-1} = \exp(AZA^{-1})$, ${}^t(\exp Z) = \exp {}^tZ$,

$$\exp t(J_n^{-1}\, {}^tZ J_n) = \exp(-tZ) \qquad (t \in R),$$

so $Z \in \mathfrak{sp}(n, C)$ if an only if

$${}^tZ J_n + J_n Z = 0. \tag{2}$$

Writing Z in the form

$$Z = \begin{pmatrix} Z_1 & Z_2 \\ Z_3 & Z_4 \end{pmatrix},$$

where Z_i is a complex $n \times n$ matrix, condition (2) is equivalent to ${}^tZ_1 + Z_4 = 0$, $Z_2 = {}^tZ_2$, $Z_3 = {}^tZ_3$.

Lemma 4.1. *Let $\approx$ denote topological isomorphism, and $\sim$ a homeomorphism. We then have*

(a) $SO(2n) \cap Sp(n) \approx U(n)$.
(b) $Sp(p, q) \cap U(2p + 2q) \approx Sp(p) \times Sp(q)$.
(c) $Sp(n, R) \cap U(2n) \approx U(n)$.

(d) $SO^*(2n) \cap U(2n) \approx U(n)$.
(e) $SU(p, q) \cap U(p+q) = S(U_p \times U_q) \sim Sp(p) \times T \times Sp(q)$.
(f) $SU^*(2n) \cap U(2n) = Sp(n)$.

Proof. (a) Each $g \in Sp(n)$ has determinant 1 so $g \in SO(2n) \cap Sp(n)$ is equivalent to ${}^t g g = I_{2n}$, ${}^t g J_n g = J_n$, ${}^t g \bar{g} = I_{2n}$. Writing

$$g = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

these last relations are equivalent to $A = D$, $B = -C$, $A^t B + B^t A = 0$, $A^t A + B^t B = I_n$. But the last two formulas express simply $A + iB \in U(n)$. For part (b), let

$$V = \{g \in GL(2p + 2q, C) : {}^t g K_{p,q} \bar{g} = K_{p,q}\}.$$

Then

$$g \in U(2p + 2q) \cap V \iff {}^t g \bar{g} = I_{2p+2q}, \qquad {}^t g K_{p,q} \bar{g} = K_{p,q}.$$

But the last two relations are equivalent to

$$g = \begin{pmatrix} X_{11} & 0 & X_{13} & 0 \\ 0 & X_{22} & 0 & X_{24} \\ X_{31} & 0 & X_{33} & 0 \\ 0 & X_{42} & 0 & X_{44} \end{pmatrix} \quad \text{where} \quad \begin{matrix} \begin{pmatrix} X_{11} & X_{13} \\ X_{31} & X_{33} \end{pmatrix} \in U(2p) \\ \\ \begin{pmatrix} X_{22} & X_{24} \\ X_{42} & X_{44} \end{pmatrix} \in U(2q). \end{matrix} \tag{3}$$

By definition

$$Sp(p, q) = Sp(p + q, C) \cap V$$

so

$$Sp(p, q) \cap U(2p + 2q) = Sp(p + q, C) \cap U(2p + 2q) \cap V.$$

Thus, g in (3) belongs to $Sp(p, q) \cap U(2p + 2q)$ if and only if ${}^t g J_{p+q} g = J_{p+q}$ or equivalently

$$\begin{pmatrix} X_{11} & X_{13} \\ X_{31} & X_{33} \end{pmatrix} \in U(2p) \cap Sp(p, C) = Sp(p)$$

and

$$\begin{pmatrix} X_{22} & X_{24} \\ X_{42} & X_{44} \end{pmatrix} \in U(2q) \cap Sp(q, C) = Sp(q).$$

This proves (b). For (c) we only have to note that

$$Sp(n, R) \cap U(2n) = Sp(n) \cap SO(2n),$$

which by (a) is isomorphic to $U(n)$. Part (d) is also easy; in fact, $g \in SO^*(2n)$ by definition if and only if ${}^t g g = I_{2n}$ and ${}^t g J_n \bar{g} = J_n$.

Thus

$$SO^*(2n) \cap U(2n) = SO(2n) \cap Sp(n, C) = SO(2n) \cap Sp(n) \approx U(n).$$

Part (e). We have

$$g \in SU(p, q) \cap U(p+q) \Longleftrightarrow g = \begin{pmatrix} g_1 & 0 \\ 0 & g_2 \end{pmatrix},$$

where $g_1 \in U(p)$, $g_2 \in U(q)$ and $\det g_1 \det g_2 = 1$. Such a matrix can be written

$$\begin{pmatrix} g_1 & 0 \\ 0 & g_2 \end{pmatrix} = \begin{pmatrix} \det g_1 & 0 & 0 & 0 & \\ 0 & 1 & & & \\ \vdots & & \ddots & & \\ 0 & & & 1 & 0 \\ 0 & & & 0 & \det g_2 \end{pmatrix} \cdot \begin{pmatrix} \gamma_1 & 0 \\ 0 & \gamma_2 \end{pmatrix},$$

where $\gamma_1 \in SU(p)$, $\gamma_2 \in SU(q)$. We have therefore a mapping

$$g \to (\gamma_1, \det g_1, \gamma_2)$$

of $SU(p, q) \cap U(p+q)$ into $SU(p) \times T \times SU(q)$. This mapping is not in general a homomorphism but it is continuous, one-to-one and onto; hence $SU(p, q) \cap U(p+q)$ is homeomorphic to $SU(p) \times T \times SU(q)$. Finally, $g \in SU^*(2n)$ if and only if $\bar{g}J_n = J_n g$ and $\det g = 1$. Hence $g \in SU^*(2n) \cap U(2n)$ if and only if $\bar{g}J_n = J_n g$, ${}^t g\bar{g} = I_{2n}$, $\det g = 1$. However, these conditions are equivalent to ${}^t g J_n g = J_n$, ${}^t g\bar{g} = I_{2n}$ or $g \in Sp(n)$. This finishes the proof of the lemma.

The following lemma is well known, see, e.g., Chevalley [2].

Lemma 4.2.

(a) *The groups* $GL(n, C)$, $SL(n, C)$, $SL(n, R)$, $SO(n, C)$, $SO(n)$, $SU(n)$, $U(n)$, $Sp(n, C)$, $Sp(n)$ *are all connected.*

(b) *The group* $GL(n, R)$ *has two connected components.*

In order to determine the connectivity of the remaining groups we need another lemma.

Definition. Let G be a subgroup of the general linear group $GL(n, C)$. Let $z_{ij}(\sigma)$ $(1 \leqslant i, j \leqslant n)$ denote the matrix elements of an arbitrary $\sigma \in GL(n, C)$, and let $x_{ij}(\sigma)$ and $y_{ij}(\sigma)$ be the real and imaginary part of $z_{ij}(\sigma)$. The group G is called a *pseudoalgebraic* subgroup of $GL(n, C)$ if there exists a set of polynomials P_β in $2n^2$ arguments such that $\sigma \in G$ if and only if $P_\beta(\ldots x_{ij}(\sigma), y_{ij}(\sigma), \ldots) = 0$ for all P_β.

A pseudoalgebraic subgroup of $GL(n, C)$ is a closed subgroup, hence a topological Lie subgroup.

Lemma 4.3.† *Let G be a pseudoalgebraic subgroup of $GL(n, C)$ such that the condition $g \in G$ implies ${}^t\bar{g} \in G$. Then there exists an integer $d \geqslant 0$ such that G is homeomorphic to the topological product of $G \cap U(n)$ and R^d.*

Proof. We first remark that if an exponential polynomial $Q(t) = \sum_{j=1}^{n} c_j e^{b_j t}$ ($b_j \in R$, $c_j \in C$) vanishes whenever t is an integer then $Q(t) = 0$ for all $t \in R$. Let $\mathfrak{h}(n)$ denote the vector space of all Hermitian $n \times n$ matrices. Then exp maps $\mathfrak{h}(n)$ homeomorphically onto the space $P(n)$ of all positive definite Hermitian $n \times n$ matrices (see Chevalley [2], Prop. 5, §IV, Chapter I). Let $H \in \mathfrak{h}(n)$. We shall prove

$$\text{If } \exp H \in G \cap P(n), \text{ then } \exp tH \in G \cap P(n) \text{ for } t \in R. \tag{4}$$

There exists a matrix $u \in U(n)$ such that uHu^{-1} is a diagonal matrix. Since the group uGu^{-1} is pseudoalgebraic as well as G, we may assume that H in (4) is a diagonal matrix. Let $h_1, \ldots, h_n$ be the (real) diagonal elements of H. The condition $\exp H \in G \cap P(n)$ means that the numbers $e^{h_1}, \ldots, e^{h_n}$ satisfy a certain set of algebraic equations. Since $\exp kH \in G \cap P(n)$ for each integer k, the numbers $e^{kh_1}, \ldots, e^{kh_n}$ also satisfy these algebraic equations and by the remark above the same is the case if k is any real number. This proves (4).

Each $g \in GL(n, C)$ can be decomposed uniquely $g = up$ where $u \in U(n)$, $p \in P(n)$. Here u and p depend continuously on g. If $g \in G$, then ${}^t\bar{g}g = p^2 \in G \cap P(n)$ so by (4) $p \in G \cap P(n)$ and $u \in G \cap U(n)$. The mapping $g \to (u, p)$ is a one-to-one mapping of G onto the product $(G \cap U(n)) \times (G \cap P(n))$ and since G carries the relative topology of $GL(n, C)$, this mapping is a homeomorphism.

The Lie algebra $\mathfrak{gl}(n, C)$ is a direct sum

$$\mathfrak{gl}(n, C) = \mathfrak{u}(n) + \mathfrak{h}(n).$$

Since the Lie algebra $\mathfrak{g}$ of G is invariant under the involutive automorphism $X \to -{}^t\bar{X}$ of $\mathfrak{gl}(n, C)$ we have

$$\mathfrak{g} = \mathfrak{g} \cap \mathfrak{u}(n) + \mathfrak{g} \cap \mathfrak{h}(n).$$

It is obvious that $\exp(\mathfrak{g} \cap \mathfrak{h}(n)) \subset G \cap P(n)$. On the other hand, each $p \in G \cap P(n)$ can be written uniquely $p = \exp H$ where $H \in \mathfrak{h}(n)$; by (4), $H \in \mathfrak{h}(n) \cap \mathfrak{g}$, so exp induces a homeomorphism of $\mathfrak{g} \cap \mathfrak{h}(n)$ onto $G \cap P(n)$. This proves the lemma.

† Compare Chevalley [2], p. 201.

Lemma 4.4.

(a) *The groups* $SU(p, q)$, $SU^*(2n)$, $SO^*(2n)$, $Sp(n, R)$, *and* $Sp(p, q)$ *are all connected.*

(b) *The group* $SO(p, q)$ $(0 < p < p + q)$ *has two connected components.*

Proof. All these groups are pseudoalgebraic subgroups of the corresponding general linear group and have the property that $g \in G \Rightarrow {}^t\bar{g} \in G$. Part (a) is therefore an immediate consequence of Lemma 4.3 and Lemma 4.1. For (b) we consider the intersection $SO(p, q) \cap U(p + q) = SO(p, q) \cap SO(p + q)$. This consists of all matrices of the form

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix},$$

where A and B are orthogonal matrices of order p and q respectively satisfying $\det A \det B = 1$. It follows again from Lemma 4.3 that $SO(p, q)$ has two components.

2. The Simple Lie Algebras† over **C** and Their Compact Real Forms. The Irreducible Riemannian Globally Symmetric Spaces of Type II and Type IV

In Table I $\mathfrak{g}$ runs over all simple Lie algebras over C; the subscript denotes the *rank* of $\mathfrak{g}$, that is, the dimension of a Cartan subalgebra. Moreover, G stands for a connected Lie group with Lie algebra $\mathfrak{g}^R$,

TABLE I

LIE GROUPS FOR THE SIMPLE LIE ALGEBRAS OVER C AND THEIR COMPACT REAL FORMS

$\mathfrak{g}$	G	U	$Z(\tilde{U})$	$\dim U$
$\mathfrak{a}_n (n \geqslant 1)$	$SL(n+1, C)$	$SU(n+1)$	Z_{n+1}	$n(n+2)$
$\mathfrak{b}_n (n \geqslant 2)$	$SO(2n+1, C)$	$SO(2n+1)$	Z_2	$n(2n+1)$
$\mathfrak{c}_n (n \geqslant 3)$	$Sp(n, C)$	$Sp(n)$	Z_2	$n(2n+1)$
$\mathfrak{d}_n (n \geqslant 4)$	$SO(2n, C)$	$SO(2n)$	Z_4 if $n =$ odd $Z_2 + Z_2$ if $n =$ even	$n(2n-1)$
$\mathfrak{e}_6$	$E_6^{\mathbf{C}}$	E_6	Z_3	78
$\mathfrak{e}_7$	$E_7^{\mathbf{C}}$	E_7	Z_2	133
$\mathfrak{e}_8$	$E_8^{\mathbf{C}}$	E_8	Z_1	248
$\mathfrak{f}_4$	$F_4^{\mathbf{C}}$	F_4	Z_1	52
$\mathfrak{g}_2$	$G_2^{\mathbf{C}}$	G_2	Z_1	14

† The remainder of §4 contains various results stated without proof and will not be used in Chapter X.

U is an analytic subgroup of G whose Lie algebra is a compact real form of $\mathfrak{g}$. By Chapter VI, §2, U is a maximal compact subgroup of G. Let $\tilde{U}$ denote the universal covering group of U, $Z(\tilde{U})$ the center of $\tilde{U}$, and $\boldsymbol{Z}_p$ a cyclic group of order p. The groups $\boldsymbol{SU}(n+1)$, $\boldsymbol{Sp}(n)$ are simply connected, but for $n \geqslant 3$, $\boldsymbol{SO}(n)$ has a twofold covering by the simply connected group $\boldsymbol{Spin}(n)$ (Chevalley [2], Chapter 2). The dimension of U is also listed.

The five last are called the exceptional structures. The first four classes $\mathfrak{a}_n$, $\mathfrak{b}_n$, $\mathfrak{c}_n$, $\mathfrak{d}_n$ (the classical structures) are of course defined for all $n \geqslant 1$, but then the following isomorphisms occur (see, e.g., Pontrjagin [1], §65) and $\mathfrak{d}_1$ is not semisimple.

$$\mathfrak{a}_1 = \mathfrak{b}_1 = \mathfrak{c}_1, \qquad \mathfrak{b}_2 = \mathfrak{c}_2, \qquad \mathfrak{a}_3 = \mathfrak{d}_3, \qquad \mathfrak{d}_2 = \mathfrak{a}_1 \times \mathfrak{a}_1. \tag{5}$$

With the restriction on the indices in the table each simple Lie algebra $\mathfrak{g}$ over $\boldsymbol{C}$ occurs exactly once.

Using now Theorem 5.4, Chapter VIII, and Theorem 1.1, Chapter VI, we have:

The Riemannian globally symmetric spaces of type IV are the spaces G/U where G is a connected Lie group whose Lie algebra is $\mathfrak{g}^R$ where $\mathfrak{g}$ is a simple Lie algebra over $\boldsymbol{C}$, and U is a maximal compact subgroup of G. The metric on G/U is G-invariant and is uniquely determined (up to a factor) by this condition.

Secondly, in view of Prop. 1.2:

The Riemannian globally symmetric spaces of type II are the simple, compact, connected Lie groups U. The metric on U is two-sided invariant and is uniquely determined (up to a factor) by this condition.

3. The Involutive Automorphisms of Simple Compact Lie Algebras. The Irreducible Riemannian Globally Symmetric Spaces of Type I and Type III

A. The Classical Structures

Let $\mathfrak{u}$ be a compact simple Lie algebra, θ an involutive automorphism of $\mathfrak{u}$; let $\mathfrak{u} = \mathfrak{k}_0 + \mathfrak{p}_*$ be the decomposition of $\mathfrak{u}$ into eigenspaces of θ and let $\mathfrak{g}_0 = \mathfrak{k}_0 + \mathfrak{p}_0$ (where $\mathfrak{p}_0 = i\mathfrak{p}_*$). Then $\mathfrak{g}_0$ is a real form of the complexification $\mathfrak{g} = \mathfrak{u}^C$. We reproduce below Cartan's list of all possibilities for $\mathfrak{u}$ (up to isomorphism) and all possibilities for θ (up to conjugacy). Then $\mathfrak{g}_0$ runs through all noncompact real forms of $\mathfrak{g}$ up to isomorphism. The simply connected Riemannian globally symmetric spaces corresponding to $(\mathfrak{u}, \theta)$ and $\mathfrak{g}_0$ are also listed (for $\mathfrak{u}$ classical).

As earlier, $\mathfrak{h}_{\mathfrak{p}_*}$ and $\mathfrak{h}_{\mathfrak{p}_0}$ denote maximal abelian subspaces of $\mathfrak{p}_*$ and $\mathfrak{p}_0$, respectively.

Type A I $\mathfrak{u} = \mathfrak{su}(n)$; $\theta(X) = \bar{X}$.

Here $\mathfrak{k}_0 = \mathfrak{so}(n)$ and $\mathfrak{p}_*$ consists of all symmetric purely imaginary $n \times n$ matrices of trace 0. Thus $\mathfrak{g}_0 = \mathfrak{k}_0 + \mathfrak{p}_0 = \mathfrak{sl}(n, \boldsymbol{R})$. The corresponding simply connected symmetric spaces are

$$\boldsymbol{SL}(n, \boldsymbol{R})/\boldsymbol{SO}(n), \qquad \boldsymbol{SU}(n)/\boldsymbol{SO}(n) \qquad (n > 1).$$

The diagonal matrices in $\mathfrak{p}_*$ form a maximal abelian subspace. Hence the *rank* is $n-1$. Since $\mathfrak{g} = \mathfrak{a}_{n-1}$, the algebra $\mathfrak{g}_0$ is a *normal* real form of $\mathfrak{g}$.

Type A II $\mathfrak{u} = \mathfrak{su}(2n)$; $\theta(X) = J_n \bar{X} J_n^{-1}$.

Here $\mathfrak{k}_0 = \mathfrak{sp}(n)$ and

$$\mathfrak{p}_* = \left\{ \begin{pmatrix} Z_1 & Z_2 \\ \bar{Z}_2 & -\bar{Z}_1 \end{pmatrix} \middle| \, Z_1 \in \mathfrak{su}(n),\ Z_2 = \mathfrak{so}(n, \boldsymbol{C}) \right\}.$$

Hence $\mathfrak{g}_0 = \mathfrak{k}_0 + \mathfrak{p}_0 = \mathfrak{su}^*(2n)$. The corresponding simply connected symmetric spaces are

$$\boldsymbol{SU}^*(2n)/\boldsymbol{Sp}(n), \qquad \boldsymbol{SU}(2n)/\boldsymbol{Sp}(n) \qquad (n > 1).$$

The diagonal matrices in $\mathfrak{p}_*$ form a maximal abelian subspace of $\mathfrak{p}_*$. Hence the *rank* is $n-1$.

Type A III $\mathfrak{u} = \mathfrak{su}(p+q)$; $\theta(X) = I_{p,q} X I_{p,q}$.

Here

$$\mathfrak{k}_0 = \left\{ \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \middle| \, \begin{matrix} A \in \mathfrak{u}(p),\ B \in \mathfrak{u}(q) \\ \operatorname{Tr}(A+B) = 0 \end{matrix} \right\},$$

$$\mathfrak{p}_* = \left\{ \begin{pmatrix} 0 & Z \\ -{}^t\bar{Z} & 0 \end{pmatrix} \middle| \, Z \quad p \times q \text{ complex matrix} \right\}.$$

The decomposition

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$$

$$= \begin{pmatrix} A - \frac{1}{p}(\operatorname{Tr} A) I_p & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} \frac{1}{p}(\operatorname{Tr} A) I_p & 0 \\ 0 & \frac{1}{q}(\operatorname{Tr} B) I_q \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & B - \frac{1}{q}(\operatorname{Tr} B) I_q \end{pmatrix}$$

shows that $\mathfrak{k}_0$ is isomorphic to the product

$$\mathfrak{su}(p) \times \mathfrak{c}_0 \times \mathfrak{su}(q),$$

where $\mathfrak{c}_0$ is the center of $\mathfrak{k}_0$. Also $\mathfrak{g}_0 = \mathfrak{k}_0 + \mathfrak{p}_0 = \mathfrak{su}(p, q)$. The corresponding simply connected symmetric spaces are

$$SU(p, q)/S(U_p \times U_q), \qquad SU(p + q)/S(U_p \times U_q) \qquad (p \geqslant 1, q \geqslant 1, p \geqslant q).$$

A maximal abelian subspace of $\mathfrak{p}_*$ is given by

$$\mathfrak{h}_{\mathfrak{p}_*} = \sum_{i=1}^{q} R(E_{i\,p+i} - E_{p+i\,i}). \tag{6}$$

Consequently, the *rank* is q. The spaces are *Hermitian symmetric*. For $q = 1$, these spaces are the so-called *Hermitian hyperbolic space* and the *complex projective space*.

Type BD I $\mathfrak{u} = \mathfrak{so}(p + q)$; $\theta(X) = I_{p,q}XI_{p,q}$ $(p \geqslant q)$.
Here

$$\mathfrak{k}_0 = \left\{ \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \,\middle|\, A \in \mathfrak{so}(p), B \in \mathfrak{so}(q) \right\},$$

$$\mathfrak{p}_* = \left\{ \begin{pmatrix} 0 & X \\ -{}^tX & 0 \end{pmatrix} \,\middle|\, X \text{ real } p \times q \text{ matrix} \right\}.$$

As shown in Chapter V, §2, the mapping

$$\begin{pmatrix} A & iX \\ -i{}^tX & B \end{pmatrix} \to \begin{pmatrix} A & X \\ {}^tX & B \end{pmatrix}$$

is an isomorphism of $\mathfrak{g}_0 = \mathfrak{k}_0 + \mathfrak{p}_0$ onto $\mathfrak{so}(p, q)$. The simply connected symmetric spaces associated with $\mathfrak{so}(p, q)$ and $(\mathfrak{u}, \theta)$ are

$$SO_0(p, q)/SO(p) \times SO(q), \quad SO(p + q)/SO(p) \times SO(q) \quad \begin{pmatrix} p > 1, q \geqslant 1 \\ p + q \neq 4, p \geqslant q \end{pmatrix}.$$

Here $SO_0(p, q)$ denotes the identity component of $SO(p, q)$. The compact space is the manifold of oriented p-planes of $(p + q)$-space, which is known (see, e.g., Steenrod [1], p. 134) to be simply connected. A maximal abelian subspace of $\mathfrak{p}_*$ is again given by (6), so the *rank* is q. If $p + q$ is even then $\mathfrak{g}_0$ is a *normal* real form of $\mathfrak{g}$ if and only if $p = q$. If $p + q$ is odd then $\mathfrak{g}_0$ is a *normal* real form of $\mathfrak{g}$ if and only if $p = q + 1$.

For $q = 1$, the spaces are the *real hyperbolic space* and the *sphere*. These are the simply connected Riemannian manifolds of constant sectional curvature $\neq 0$ and dimension $\neq 3$. Those of dimension 3 are $\boldsymbol{SL}(2, \boldsymbol{C})/\boldsymbol{SU}(2)$ and $\boldsymbol{SU}(2)$, i.e., $\mathfrak{a}_n$ for $n = 1$.

If $q = 2$, then $\mathfrak{k}_0$ has nonzero center and the spaces are *Hermitian symmetric*.

Type D III $\mathfrak{u} = \mathfrak{so}(2n)$; $\theta(X) = J_n X J_n^{-1}$.

Here $\mathfrak{k}_0 = \mathfrak{so}(2n) \cap \mathfrak{sp}(n)$ which by Lemma 4.1 is isomorphic to $\mathfrak{u}(n)$. Moreover,

$$\mathfrak{p}_* = \left\{ \begin{pmatrix} X_1 & X_2 \\ X_2 & -X_1 \end{pmatrix} \,\middle|\, X_1, X_2 \in \mathfrak{so}(n) \right\}.$$

Hence $\mathfrak{g}_0 = \mathfrak{k}_0 + \mathfrak{p}_0 = \mathfrak{so}^*(2n)$. The symmetric spaces are

$$\boldsymbol{SO}^*(2n)/\boldsymbol{U}(n), \qquad \boldsymbol{SO}(2n)/\boldsymbol{U}(n) \qquad (n > 2).$$

Here the imbedding of $\boldsymbol{U}(n)$ into $\boldsymbol{SO}(2n)$, (and $\boldsymbol{SO}^*(2n)$), is given by the mapping

$$A + iB \to \begin{Bmatrix} A & B \\ -B & A \end{Bmatrix}, \tag{7}$$

where $A + iB \in \boldsymbol{U}(n)$, A, B real. The spaces are *Hermitian symmetric* since $\mathfrak{k}_0$ has nonzero center. In view of Theorem 4.6, Chapter VIII, they are simply connected. A maximal abelian subspace of $\mathfrak{p}_*$ is spanned by the matrices

$$(E_{12} - E_{21}) - (E_{n+1\,n+2} - E_{n+2\,n+1}), \quad (E_{23} - E_{32}) - (E_{n+2\,n+3} - E_{n+3\,n+2}), \ldots.$$

Consequently, the *rank* is $[n/2]$.

Type C I $\mathfrak{u} = \mathfrak{sp}(n)$; $\theta(X) = \bar{X}$ $(= J_n X J_n^{-1})$.

Here $\mathfrak{k}_0 = \mathfrak{sp}(n) \cap \mathfrak{so}(2n)$ which is isomorphic to $\mathfrak{u}(n)$.

$$\mathfrak{p}_* = \left\{ \begin{pmatrix} Z_1 & Z_2 \\ Z_2 & -Z_1 \end{pmatrix} \,\middle|\, \begin{matrix} Z_1 \in \mathfrak{u}(n), \text{ purely imaginary} \\ Z_2 \text{ symmetric, purely imaginary} \end{matrix} \right\}.$$

Hence $\mathfrak{g}_0 = \mathfrak{k}_0 + \mathfrak{p}_0 = \mathfrak{sp}(n, \boldsymbol{R})$. The corresponding simply connected symmetric spaces are

$$\boldsymbol{Sp}(n, \boldsymbol{R})/\boldsymbol{U}(n), \qquad \boldsymbol{Sp}(n)/\boldsymbol{U}(n) \qquad (n \geqslant 1).$$

Here the imbedding of $\boldsymbol{U}(n)$ into $\boldsymbol{Sp}(n)$ (and $\boldsymbol{Sp}(n, \boldsymbol{R})$) is given by (7). The diagonal matrices in $\mathfrak{p}_*$ form a maximal abelian subspace. Thus the spaces have *rank* n and $\mathfrak{g}_0$ is a *normal* real form of $\mathfrak{g}$. The spaces are *Hermitian symmetric*.

Type C II $\mathfrak{u} = \mathfrak{sp}(p+q)$; $\theta(X) = K_{p,q}XK_{p,q}$.
Here

$$\mathfrak{k}_0 = \left\{ \begin{pmatrix} X_{11} & 0 & X_{13} & 0 \\ 0 & X_{22} & 0 & X_{24} \\ -\bar{X}_{13} & 0 & \bar{X}_{11} & 0 \\ 0 & -\bar{X}_{24} & 0 & \bar{X}_{22} \end{pmatrix} \middle| \begin{array}{l} X_{11} \in \mathfrak{u}(p),\ X_{22} \in \mathfrak{u}(q) \\ X_{13}\ p \times p \text{ symmetric} \\ X_{24}\ q \times q \text{ symmetric} \end{array} \right\},$$

$$\mathfrak{p}_* = \left\{ \begin{pmatrix} 0 & Y_{12} & 0 & Y_{14} \\ -{}^t\bar{Y}_{12} & 0 & {}^tY_{14} & 0 \\ 0 & -\bar{Y}_{14} & 0 & \bar{Y}_{12} \\ -{}^t\bar{Y}_{14} & 0 & -{}^tY_{12} & 0 \end{pmatrix} \middle| \begin{array}{l} Y_{12} \text{ and } Y_{14} \text{ arbitrary} \\ \text{complex } p \times q \text{ matrices} \end{array} \right\}.$$

It is clear that $\mathfrak{k}_0$ is isomorphic to the direct product $\mathfrak{sp}(p) \times \mathfrak{sp}(q)$. Moreover, $\mathfrak{g}_0 = \mathfrak{k}_0 + \mathfrak{p}_0 = \mathfrak{sp}(p, q)$. The corresponding simply connected symmetric spaces are

$$\boldsymbol{Sp}(p, q)/\boldsymbol{Sp}(p) \times \boldsymbol{Sp}(q), \qquad \boldsymbol{Sp}(p+q)/\boldsymbol{Sp}(p) \times \boldsymbol{Sp}(q) \qquad (p \geqslant q \geqslant 1).$$

Here the imbedding of $\boldsymbol{Sp}(p) \times \boldsymbol{Sp}(q)$ into $\boldsymbol{Sp}(p+q)$ (and $\boldsymbol{Sp}(p, q)$) is given by the mapping

$$\left(\begin{pmatrix} A_1 & B_1 \\ C_1 & D_1 \end{pmatrix}, \begin{pmatrix} A_2 & B_2 \\ C_2 & D_2 \end{pmatrix} \right) \to \begin{pmatrix} A_1 & 0 & B_1 & 0 \\ 0 & A_2 & 0 & B_2 \\ C_1 & 0 & D_1 & 0 \\ 0 & C_2 & 0 & D_2 \end{pmatrix}$$

A maximal abelian subspace of $\mathfrak{p}_*$ is obtained by taking $Y_{14} = 0$ and letting Y_{12} run through the space $\mathbf{R}E_{11} + \mathbf{R}E_{22} + \cdots + \mathbf{R}E_{qq}$. Consequently, the *rank* is q. For $q = 1$, the spaces are the so-called *quaternian hyperbolic spaces* and the *quaternian projective spaces*.

This concludes the list of involutive automorphisms of the compact classical simple Lie algebras. The restriction on the indices is made in order that the algebras should be simple, the spaces of dimension > 0, and the condition $p \geqslant q$ is required in order to avoid repetition within the same class.

B. Coincidences between Different Classes. Special Isomorphisms

Owing to the isomorphisms (5) the following overlappings occur between the different classes (É. Cartan [2], pp. 352-355).

(i) $\boldsymbol{A\,I}(n=2) = \boldsymbol{A\,III}(p=q=1) = \boldsymbol{BD\,I}(p=2, q=1) = \boldsymbol{C\,I}(n=1)$.
Corresponding isomorphisms:

$$\mathfrak{su}(2) \approx \mathfrak{so}(3) \approx \mathfrak{sp}(1),$$
$$\mathfrak{sl}(2, \boldsymbol{R}) \approx \mathfrak{su}(1,1) \approx \mathfrak{so}(2,1) \approx \mathfrak{sp}(1, \boldsymbol{R}).$$

(ii) ***BD I***$(p = 3, q = 2) =$ ***C I***$(n = 2)$.
Corresponding isomorphisms:

$$\mathfrak{so}(5) \approx \mathfrak{sp}(2),$$
$$\mathfrak{so}(3, 2) \approx \mathfrak{sp}(2, \boldsymbol{R}).$$

(iii) ***BD I***$(p = 4, q = 1) =$ ***C II***$(p = q = 1)$.
Corresponding isomorphisms:

$$\mathfrak{so}(5) \approx \mathfrak{sp}(2), \qquad \mathfrak{so}(4) \approx \mathfrak{sp}(1) \times \mathfrak{sp}(1),$$
$$\mathfrak{so}(4, 1) \approx \mathfrak{sp}(1, 1).$$

(iv) ***A I***$(n = 4) =$ ***BD I***$(p = q = 3)$.
Corresponding isomorphisms:

$$\mathfrak{su}(4) \approx \mathfrak{so}(6), \qquad \mathfrak{so}(4) \approx \mathfrak{so}(3) \times \mathfrak{so}(3),$$
$$\mathfrak{sl})4, \boldsymbol{R}) \approx \mathfrak{so}(3, 3).$$

(v) ***A II***$(n = 2) =$ ***BD I***$(p = 5, q = 1)$.
Corresponding isomorphisms:

$$\mathfrak{su}(4) \approx \mathfrak{so}(6), \qquad \mathfrak{sp}(2) \approx \mathfrak{so}(5),$$
$$\mathfrak{su}^*(4) \approx \mathfrak{so}(5, 1).$$

(vi) ***A III***$(p = q = 2) =$ ***BD I***$(p = 4, q = 2)$.
Corresponding isomorphisms:

$$\mathfrak{su}(4) \approx \mathfrak{so}(6),$$
$$\mathfrak{su}(2, 2) \approx \mathfrak{so}(4,2).$$

(vii) ***A III***$(p = 3, q = 1) =$ ***D III***$(n = 3)$.
Corresponding isomorphisms:

$$\mathfrak{su}(4) \approx \mathfrak{so}(6),$$
$$\mathfrak{su}(3, 1) \approx \mathfrak{so}^*(6).$$

(viii) ***BD I***$(p = 6, q = 2) =$ ***D III***$(n = 4)$.
Corresponding isomorphisms:

$$\mathfrak{su}(4) \approx \mathfrak{so}(6),$$
$$\mathfrak{so}^*(8) \approx \mathfrak{so}(6, 2).$$

The last isomorphism does not occur in Cartan's paper cited above. However, the isometry of the spaces $BD\ I(p = 6, q = 2)$ and $D\ III(n=4)$ is shown in Cartan [10], p. 459. The fact that they even coincide as Hermitian symmetric spaces is overlooked in Cartan [19], p. 152, but is pointed out in Morita [1], p. 195.

The spaces $BD\ I(p + q = 4)$ and $D\ III(n = 2)$ can of course be defined although $\mathfrak{so}(4)$ is not simple. The isomorphism $\mathfrak{d}_2 = \mathfrak{a}_1 \times \mathfrak{a}_1$ then yields additional isomorphisms corresponding to the three involutive automorphisms of $\mathfrak{a}_1 \times \mathfrak{a}_1$ given by: $(X, Y) \to (Y, X)$; $(X, Y) \to (\bar{X}, \bar{Y})$; $(X, Y) \to (X, \bar{Y})$.

(ix) $BD\ I(p = 3, q = 1) = \mathfrak{a}_n(n = 1)$.
Corresponding isomorphisms:

$$\mathfrak{so}(4) \approx \mathfrak{su}(2) \times \mathfrak{su}(2),$$
$$\mathfrak{so}(3, 1) \approx \mathfrak{sl}(2, \boldsymbol{C}).$$

(x) $BD\ I(p = 2, q = 2) = A\ I(n = 1) \times A\ I(n = 1)$.
Corresponding isomorphisms:

$$\mathfrak{so}(4) \approx \mathfrak{su}(2) \times \mathfrak{su}(2),$$
$$\mathfrak{so}(2, 2) \approx \mathfrak{sl}(2, \boldsymbol{R}) \times \mathfrak{sl}(2, \boldsymbol{R}).$$

(xi) $D\ III(n = 2)$ and $A\ I(n = 1)$.

$$\mathfrak{so}(4) \approx \mathfrak{su}(2) \times \mathfrak{su}(2),$$
$$\mathfrak{so}^*(4) \approx \mathfrak{su}(2) \times \mathfrak{sl}(2, \boldsymbol{R}).$$

C. The Exceptional Structures

Let $\mathfrak{g}$ be a semisimple Lie algebra over $\boldsymbol{C}$, $\mathfrak{g}_0$ a real form of $\mathfrak{g}$. The *character* of $\mathfrak{g}_0$ is defined as $\delta = \dim \mathfrak{p}_0 - \dim \mathfrak{k}_0$, $\mathfrak{g}_0 = \mathfrak{k}_0 + \mathfrak{p}_0$ being a Cartan decomposition of $\mathfrak{g}_0$. The character reaches its minimum value $\delta = -\dim_{\boldsymbol{C}} \mathfrak{g}$ if $\mathfrak{g}_0$ is a compact real form and its maximum value $\delta = \operatorname{rank} \mathfrak{g}$ if $\mathfrak{g}_0$ is a normal real form of $\mathfrak{g}$. For these extreme values of the character, the corresponding real forms are unique up to isomorphism (Cor. 7.3, Chapter III, and Theorem 3.5). In contrast, the examples $\mathfrak{so}^*(18)$ and $\mathfrak{so}(12,6)$ which are real forms of $\mathfrak{so}(18, \boldsymbol{C})$ with character -9 show that two nonisomorphic real forms of a simple Lie algebra $\mathfrak{g}$ over $\boldsymbol{C}$ may have the same character.† However, É. Cartan's classification shows that this cannot happen for the exceptional structures§ so

† This seems to be overlooked in Lardy [1], p. 195.

§ In fact only for certain $\mathfrak{d}_n$ and certain $\mathfrak{a}_n$ (e.g., $\mathfrak{su}^*(14)$ and $\mathfrak{su}(9,5)$).

we label the real forms of the exceptional complex algebras by means of their character. Thus $\mathfrak{e}_{6(\delta)}$ denotes the real form of $\mathfrak{e}_6$ with character δ. Since the involutions and real forms of the exceptional structures require descriptions of their root structure (É. Cartan [2]) we list in Table II only the Lie algebras $(\mathfrak{g}_0, \mathfrak{k}_0)$ and $(\mathfrak{u}, \mathfrak{k}_0)$ in the decompositions $\mathfrak{g}_0 = \mathfrak{k}_0 + \mathfrak{p}_0$, $\mathfrak{u} = \mathfrak{k}_0 + \mathfrak{p}_*$, when $\mathfrak{u}$ and $\mathfrak{g}_0$ are exceptional. The rank and dimension is also listed.

TABLE II

IRREDUCIBLE RIEMANNIAN GLOBALLY SYMMETRIC SPACES OF TYPE I AND TYPE III

	Noncompact	Compact	Rank	Dimension
A I	$SL(n, R)/SO(n)$	$SU(n)/SO(n)$	$n-1$	$\frac{1}{2}(n-1)(n+2)$
A II	$SU^*(2n)/Sp(n)$	$SU(2n)/Sp(n)$	$n-1$	$(n-1)(2n+1)$
A III	$SU(p, q)/S(U_p \times U_q)$	$SU(p+q)/S(U_p \times U_q)$	$\min(p, q)$	$2pq$
BD I	$SO_0(p, q)/SO(p) \times SO(q)$	$SO(p+q)/SO(p) \times SO(q)$	$\min(p, q)$	pq
D III	$SO^*(2n)/U(n)$	$SO(2n)/U(n)$	$[\frac{1}{2}n]$	$n(n-1)$
C I	$Sp(n, R)/U(n)$	$Sp(n)/U(n)$	n	$n(n+1)$
C II	$Sp(p, q)/Sp(p) \times Sp(q)$	$Sp(p+q/Sp(p) \times Sp(q)$	$\min(p, q)$	$4pq$
E I	$(\mathfrak{e}_{6(6)}, \mathfrak{sp}(4))$	$(\mathfrak{e}_{6(-78)}, \mathfrak{sp}(4))$	6	42
E II	$(\mathfrak{e}_{6(2)}, \mathfrak{su}(6) + \mathfrak{su}(2))$	$(\mathfrak{e}_{6(-78)}, \mathfrak{su}(6) + \mathfrak{su}(2))$	4	40
E III	$(\mathfrak{e}_{6(-14)}, \mathfrak{so}(10) + R)$	$(\mathfrak{e}_{6(-78)}, \mathfrak{so}(10) + R)$	2	32
E IV	$(\mathfrak{e}_{6(-26)}, \mathfrak{f}_4)$	$(\mathfrak{e}_{6(-78)}, \mathfrak{f}_4)$	2	26
E V	$(\mathfrak{e}_{7(7)}, \mathfrak{su}(8))$	$(\mathfrak{e}_{7(-133)}, \mathfrak{su}(8))$	7	70
E VI	$(\mathfrak{e}_{7(-5)}, \mathfrak{so}(12) + \mathfrak{su}(2))$	$(\mathfrak{e}_{7(-133)}, \mathfrak{so}(12) + \mathfrak{su}(2))$	4	64
E VII	$(\mathfrak{e}_{7(-25)}, \mathfrak{e}_6 + R)$	$(\mathfrak{e}_{7(-133)}, \mathfrak{e}_6 + R)$	3	54
E VIII	$(\mathfrak{e}_{8(8)}, \mathfrak{so}(16))$	$(\mathfrak{e}_{8(-248)}, \mathfrak{so}(16))$	8	128
E IX	$(\mathfrak{e}_{8(-24)}, \mathfrak{e}_7 + \mathfrak{su}(2))$	$(\mathfrak{e}_{8(-248)}, \mathfrak{e}_7 + \mathfrak{su}(2))$	4	112
F I	$(\mathfrak{f}_{4(4)}, \mathfrak{sp}(3) + \mathfrak{su}(2))$	$(\mathfrak{f}_{4(-52)}, \mathfrak{sp}(3) + \mathfrak{su}(2))$	4	28
F II	$(\mathfrak{f}_{4(-20)}, \mathfrak{so}(9))$	$(\mathfrak{f}_{4(-52)}, \mathfrak{so}(9))$	1	16
G	$(\mathfrak{g}_{2(2)}, \mathfrak{su}(2) + \mathfrak{su}(2))$	$(\mathfrak{g}_{2(-14)}, \mathfrak{su}(2) + \mathfrak{su}(2))$	2	8

4. Irreducible Hermitian Symmetric Spaces

In view of Theorem 6.1, Chapter VIII, it can be decided immediately which of the spaces in Table II are Hermitian symmetric. They are

A III, *D III*, *BD I* $(q = 2)$, *C I*, *E III*, *E VII*.

This exhausts the list of irreducible Hermitian symmetric spaces because the spaces of type II and IV cannot be Hermitian. According to Theorem 7.1, Chapter VIII, the noncompact spaces can be regarded as bounded domains in Euclidean space. In É. Cartan [19] such domains are constructed for the four large classes ***A III***, ***D III***, ***BD I***$(q = 2)$, ***C I***.

§ 5. Two-Point Homogeneous Spaces. Symmetric Spaces of Rank One. Closed Geodesics

Definition. A Riemannian manifold M is called *two-point homogeneous* if for any two point pairs $p_1, p_2 \in M$, $q_1, q_2 \in M$ satisfying $d(p_1, p_2) = d(q_1, q_2)$ there exists an isometry g of M such that $g \cdot p_1 = q_1$ and $g \cdot p_2 = q_2$.

Proposition 5.1. *Let M be a Riemannian globally symmetric space of rank one. Then M is a two-point homogeneous space.*

Proof. We can write $M = G/K$ where $G = I_0(M)$ and K is compact. The Riemannian symmetric pair (G, K) is then of the Euclidean type, compact type or noncompact type. If (G, K) is of the Euclidean type, then either $M = \boldsymbol{R}$ or $M = S^1$ (circle). If (G, K) is of the compact type or the noncompact type, then Theorem 6.2, Chapter V, shows that M is two-point homogeneous.

Two-point homogeneous spaces have been classified completely by Wang [1] in the compact case and Tits [1] in the noncompact case. The results show that the two-point homogeneous spaces are the Euclidean spaces, the circle S^1 and the symmetric spaces of rank one of the compact type and noncompact type, respectively. The proof makes use of the classification of all simple compact groups and of all groups acting transitively on spheres. The fact that a noncompact two-point homogeneous space is globally symmetric (and thus of rank one or $\boldsymbol{R}^n$) can be proved without classification (see Nagano [1], Helgason [3]). It would be desirable to have a direct proof also for the compact case.

Definition. Let $\gamma(t)$, $-\infty < t < \infty$ be a geodesic in a Riemannian manifold M. The geodesic is called *closed* if there exists a number $L > 0$ such that $\gamma(t + L) = \gamma(t)$ for all t. The geodesic is said to be *simply closed* if in addition $\gamma(t_1) \neq \gamma(t_2)$ for $0 < t_1 < t_2 \leqslant L$. If $|t|$ is the arc parameter, L is called the *length* of the simply closed geodesic.

Proposition 5.2. *Let M be a Riemannian globally symmetric space and $\gamma(t)$ $(-\infty < t < \infty)$ a geodesic in M which intersects itself. Then it is a simply closed geodesic.*

Proof. Let p_0 be some point where the geodesic intersects itself. We may assume that the parameter t is such that $\gamma(0) = \gamma(1) = p_0$ and $\gamma(t_1) \neq \gamma(t_2)$ for $0 < t_1 < t_2 \leqslant 1$. Using now the notation of Theorem 3.3, Chapter IV, we have $M = G/K$. There is a unique vector $X \in \mathfrak{p}$ such that $d\pi \cdot X = \dot{\gamma}(0)$. Then $\operatorname{Exp} d\pi X = \gamma(1) = p_0$ so $\exp X \in K$. If $Y \in M_{p_0}$, then by the theorem quoted, the vector $(d \exp X)_{p_0}(Y)$ is the parallel translate of Y along the curve segment $\gamma(t)$, $0 \leqslant t \leqslant 1$. Using (1) in Chapter IV, §3, we obtain

$$(d \exp X)_{p_0}(d\pi X) = d\pi\,(\operatorname{Ad}(\exp X)\,X) = d\pi X.$$

This shows that $\dot{\gamma}(0)$ is the parallel translate of $\dot{\gamma}(0)$ along $\gamma(t)$ $(0 \leqslant t \leqslant 1)$. Hence $\gamma(t+1) = \gamma(t)$ $(0 \leqslant t \leqslant 1)$ and the proposition follows.

Proposition 5.3. *Let M be a compact Riemannian globally symmetric space. Then M has a simply closed geodesic. If M is of rank one then all the geodesics in M are simply closed and have the same length.*

Proof. We follow again the notation of Theorem 3.3, Chapter IV. Then since M and the group K are compact, G is compact. Let $\mathfrak{a}$ be a maximal abelian subspace of $\mathfrak{p}$ and put $A = \exp \mathfrak{a}$. The closure $\bar{A}$ of A in G is a torus whose Lie algebra is contained in $\mathfrak{p}$. Using the maximality of $\mathfrak{a}$, it follows that $A = \bar{A}$. Being a torus, A contains a one-parameter subgroup $\exp tH$ $(t \in \boldsymbol{R},\ H \in \mathfrak{a})$ which intersects itself. The geodesic $\pi(\exp tH)$ $(t \in \boldsymbol{R})$ in G/K is simply closed by Prop. 5.1. If M has rank one, then all geodesics (parametrized by arc length) are congruent under an isometry of M (Theorem 6.2, Chapter V). This proves the proposition.

Definition. Let M be a compact Riemannian manifold and p a point in M. The set of points in M of maximum distance from p will be called the *antipodal set* associated to p. It will be denoted by A_p.

If M is a compact Riemannian globally symmetric space of rank one and $p \in M$, then the isotropy subgroup of $I_0(M)$ at p acts transitively on A_p. In view of Prop. 4.4, Chapter II, A_p is a compact submanifold of M. It is obvious that $\dim A_p < \dim M$.

Theorem 5.4. *Let M be a compact Riemannian globally symmetric space of rank one. Let $2L$ denote the common length of the geodesics in M. Let p be any point in M. Then Exp_p is a diffeomorphism of the open ball $\| X \| < L$ in M_p onto the complement $M - A_p$.*

Proof. We first assert that A_p coincides with the set of midpoints of the geodesics $\gamma(s)$ $(0 \leqslant s \leqslant 2L)$, starting at p, s denoting the arc length measured from p. In fact, consider one of these geodesics γ. The point

$\gamma(L)$ is clearly left fixed by the symmetry s_p. If Γ is a curve segment of length $d(p, \gamma(L))$ joining p and $\gamma(L)$ then $s_p \cdot \Gamma$ also has these properties. The curves $s_p \cdot \Gamma$ and Γ are geodesics with opposite tangent vectors at p. Using Prop. 5.2 we see that Γ followed by $s_p \cdot \Gamma$ (in opposite direction) forms a closed geodesic. Hence $L = d(p, \gamma(L))$, so γ realizes the shortest distance between any two of its points. Next, let $q \in A_p$. Then $d(p, q) \geqslant L$. We join p to q by a geodesic Γ' of shortest length. The closed geodesic in M tangent to Γ' at p must contain Γ'. Since this closed geodesic has length $2L$ we conclude that $d(p, q) \leqslant L$, hence $d(p, q) = L$. This proves the assertion above.

Next, let $\gamma_1(s)$ and $\gamma_2(s)$ $(0 \leqslant s \leqslant 2L)$ be two geodesics in M starting at p, s being arc length measured from p. Suppose they intersect at a point p' different from p and $\gamma_2(L)$. Consider the curve γ formed by the shortest part of γ_1 joining p and p' together with the shortest part of γ_2 joining p' and $\gamma_2(L)$. Then γ is a curve of length L joining p and $\gamma_2(L)$ and must be a geodesic due to Lemma 9.8, Chapter I. But this obviously implies that either $\gamma_1(s) \equiv \gamma_2(s)$ or $\gamma_1(s) \equiv \gamma_2(2L - s)$.

It has now been proved that Exp_p $(= \mathrm{Exp})$ is a one-to-one differentiable mapping of the open ball $\| X \| < L$ onto $M - A_p$. It remains to be proved that Exp is regular at $X \in M_p$ provided $0 < \| X \| < L$. For this purpose let Y be a tangent vector to M_p at X for which $(d\,\mathrm{Exp})_X (Y) = 0$. Let Q denote the Riemannian structure of M and consider the function $q(t) = Q_{\mathrm{Exp}\, tX} (d\,\mathrm{Exp}_{tX}(X),\ d\,\mathrm{Exp}_{tX}(Y))$ $(t \in \boldsymbol{R})$. Here Y is considered as a tangent vector to M_p at tX. If we decompose $Y = yX + Y_1$ where Y_1 is perpendicular to X we see from Lemma 9.7, Chapter I, that for small t,

$$q(t) = yQ_p(X, X).$$

Since $q(t)$ is an analytic function we conclude that it is a constant. But $q(1) = 0$ so $y = 0$ and Y is perpendicular to X.

Let K denote the isotropy subgroup of $G = I_0(M)$ at p. We may assume that $\dim M > 1$. Then $I_0(M)$ is semisimple and K acts transitively on each sphere $S_r(p)$ in M. In particular, the linear isotropy group K^* acts transitively on the sphere S around the origin in M_p through X. The group K^* is a compact linear group and is a Lie transformation group of S. In view of Prop. 4.3, Chapter II, there exists a vector Y_0 in the Lie algebra $\mathfrak{L}(K^*) = \mathfrak{L}(K)$ such that†

$$Y = \left\{\frac{d}{dt}\, d\tau\,(\exp tY_0) \cdot X\right\}_{t=0}. \qquad (\ )$$

† Here exp is the exponential mapping for G and $\tau(x)$ is the mapping $gK \to xgK$ of G/K onto itself.

Then, if f is a differentiable function on M,

$$0 = (d\,\mathrm{Exp})_X(Y)f = \left\{\frac{d}{dt} f(\mathrm{Exp}\,(d\tau\,(\exp tY_0)\cdot X)\right\}_{t=0}$$
$$= \left\{\frac{d}{dt} f(\exp tY_0 \cdot \mathrm{Exp}\,X)\right\}_{t=0},$$

where we have used relations (1) and (3) in Chapter IV, §3. Let $s \in \boldsymbol{R}$ and let us use the last formula on the function $f^*(q) = f(\exp sX \cdot q)$, $q \in M$. Then

$$0 = \left\{\frac{d}{dt} f^*(\exp tY_0 \cdot \mathrm{Exp}\,X)\right\}_{t=0} = \left\{\frac{d}{dt} f(\exp tY_0 \cdot \mathrm{Exp}\,X)\right\}_{t=s},$$

which shows that $f(\exp sY_0 \cdot \mathrm{Exp}\,X)$ is constant in s. Since f is arbitrary, this shows that the one-parameter subgroup $\exp sY_0$ ($s \in \boldsymbol{R}$) leaves the point $\mathrm{Exp}\,X$ fixed. The *unique* geodesic of shortest length joining p and $\mathrm{Exp}\,X$ is therefore left fixed by each $\exp sY_0$. Consequently, $d\tau(\exp sY_0)\cdot X = X$ for $s \in \boldsymbol{R}$ so $Y = 0$ by (1). Thus, Exp is regular and the theorem is proved.

EXERCISES

1. Let $\mathfrak{g}$ be a simple Lie algebra over $\boldsymbol{C}$. Prove that $\mathfrak{g}^R$ is simple and that $(\mathfrak{g}^R)^C$ is not simple. Hence the simple Lie algebras over $\boldsymbol{R}$ fall into two disjoint classes: (a) the simple Lie algebras over $\boldsymbol{C}$ considered as real Lie algebras; (b) the real forms of simple Lie algebras over $\boldsymbol{C}$.

2. Let $\mathfrak{g}_0$ be a simple Lie algebra over $\boldsymbol{R}$ and let $\mathfrak{g}_0 = \mathfrak{k}_0 + \mathfrak{p}_0$ be a Cartan decomposition of $\mathfrak{g}_0$. Show that the center of $\mathfrak{k}_0$ has dimension $\leqslant 1$. (Hint: Let (G, K) be a corresponding Riemannian symmetric pair, and take $G = \mathrm{Int}\,(\mathfrak{g}_0)$. Let C denote the center of K. If $\dim C > 0$ then $\mathfrak{p}_0$ has a complex structure commuting elementwise with $\mathrm{Ad}_G\,(K)$. Apply Schur's lemma to $\mathrm{Ad}_G\,(C)$ to get an isomorphism of C onto the unit circle.)

3. The notation being as in §4, let σ and τ denote the conjugations of $\mathfrak{g}$ with respect to $\mathfrak{g}_0$ and $\mathfrak{u}$, respectively. Show that

For ***A I***	$\sigma(X) = \bar{X}$,	$\tau(X) = -\,{}^t\bar{X}$.
A II	$\sigma(X) = J_n\bar{X}J_n^{-1}$,	$\tau(X) = -\,{}^t\bar{X}$.
A III	$\sigma(X) = -\,I_{p,q}{}^t\bar{X}I_{p,q}$,	$\tau(X) = -\,{}^t\bar{X}$.
BD I	$\sigma(X) = I_{p,q}\bar{X}I_{p,q}$,	$\tau(X) = \bar{X}$.
D III	$\sigma(X) = J_n\bar{X}J_n^{-1}$,	$\tau(X) = \bar{X}$.
C I	$\sigma(X) = \bar{X}$,	$\tau(X) = J_n\bar{X}J_n^{-1}$.
C II	$\sigma(X) = -\,K_{p,q}{}^t\bar{X}K_{p,q}$,	$\tau(X) = J_{p+q}\bar{X}J_{p+q}^{-1}$.

4. As a corollary of Theorem 2.6 (and Theorem 2.12, Chapter VII), show that for a compact semisimple Lie algebra $\mathfrak{u}$, the factor group $\mathrm{Aut}(\mathfrak{u})/\mathrm{Int}(\mathfrak{u})$ is isomorphic to the group of rotations of $\mathfrak{t}_0$ leaving a given Weyl chamber invariant.

(This fact makes it possible to determine $\mathrm{Aut}(\mathfrak{u})/\mathrm{Int}(\mathfrak{u})$ from the root pattern for simple $\mathfrak{u}$. The result is (Cartan [4]) that $\mathrm{Aut}(\mathfrak{u})/\mathrm{Int}(\mathfrak{u})$ consists of one element if $\mathfrak{u} = \mathfrak{b}_n$, $\mathfrak{c}_n$, $\mathfrak{e}_7$, $\mathfrak{e}_8$, $\mathfrak{f}_4$, $\mathfrak{g}_2$, two elements if $\mathfrak{u} = \mathfrak{a}_n$, $\mathfrak{d}_n (n > 4)$, $\mathfrak{e}_6$ and is isomorphic to the permutation group of 3 letters if $\mathfrak{u} = \mathfrak{d}_4$).

NOTES

§1. Problem A′ was attacked by Killing [1], where he set up the list of simple Lie algebras over C. However, his proofs contained errors at some important points (as observed by É. Cartan), and his treatment of the exceptional algebras was incomplete. In this Thèse [1] É. Cartan gave a rigorous solution of problem A′. Simplified treatments, based on Theorem 5.4, Chapter III, have been given by van der Waerden [1], Dynkin [1], and Freudenthal [3]. Coxeter (in [1] and Weyl [2]) and later Witt [2] classified all finite groups generated by reflections and applied the result to the classification problem. É. Cartan's long paper [2] gives the solution of problem B′. Simplified treatments were given after the equivalence with problem B had been noticed; see É. Cartan [12], Lardy [1], Gantmacher [2], and Freudenthal [4]. Problem C was solved by É. Cartan [9] by the method indicated in §6, Chapter VII.

§2. The discussion is mostly based on Gantmacher [1].

§3. In É. Cartan's paper [10] the number of components of $I(M)$ is determined for each noncompact irreducible M and Cor. 3.3 is verified in each case. The existence and uniqueness of the normal real form (Theorem 3.5) is also established by a case-by-case verification in É. Cartan [2]. This verification is particularly cumbersome for the exceptional structures. The proof of Theorem 3.6 is based on a suggestion by M. Berger. It has also been proved by B. Kostant, see Simons [1].

§4. The classification as well as a more geometric description is given in É. Cartan [10].

§5. The behavior of geodesics on compact symmetric spaces was investigated in É. Cartan [10]. The essense of Theorem 5.4 is stated there (p. 437) and the dimension of the antipodal set A_p is determined for each M of rank 1. A generalization of Prop. 5.2 is given in Kostant [2], p. 260.

CHAPTER X

FUNCTIONS ON SYMMETRIC SPACES

On Euclidean spaces $\mathbf{R}^n$ the exponential functions are characterized as the continuous nonzero solutions of the functional equation $\varphi(x + y) = \varphi(x)\varphi(y)$ or equivalently as the eigenfunctions of all differential operators with constant coefficients normalized by $\varphi(0) = 1$.

A fruitful generalization of the functional equation above to a symmetric space G/K is the equation

$$\int_K \varphi(xky)dk = \varphi(x)\ \varphi(y) \qquad x,\ y \in G, \tag{1}$$

whose solutions will be called spherical functions on G. These functions are bi-invariant under K, that is, $\varphi(kxk') = \varphi(x)$, for $x \in G$, k, $k' \in K$ and can therefore be regarded as functions on G/K, invariant under the action of K. Among all such functions the spherical functions are characterized as eigenfunctions of all differential operators on G/K which are invariant under the action of G. For each of the three types of symmetric spaces the solutions of (1) are given by a simple integral formula.

In certain special cases the spherical functions reduce to well-known classical functions (Legendre functions, Bessel functions), and one obtains thus a unified group theoretic explanation of important properties of these functions.

The continuous functions on G of compact support which are bi-invariant under K form an algebra $C^\natural(G)$ under the convolution product. It is a simple consequence of the symmetry of G/K that the algebra $C^\natural(G)$ is commutative. The theory of Banach algebras therefore gives immediately basic results for Fourier analysis of $C^\natural(G)$, that is, a decomposition of an arbitrary function $f \in C^\natural(G)$ into spherical functions. One of the results is the Plancherel formula connecting the L^2-norm of f with that of the Fourier transform. However, the abstract theory gives little information about the Plancherel measure which appears in the formula. An interesting problem is to relate this measure to the structure of G. For the compact case this is given by the Peter-Weyl theory together with Weyl's formula (Weyl [1]) for the degree of an irreducible representation of a given highest weight. For the noncompact case the problem is much more difficult and not yet completely solved, although Harish-Chandra has in [12] discovered a likely candidate for the Plancherel measure which is confirmed in certain special cases.

The methods which lead to the formula for the spherical function on G/K (§6) also lead to an isomorphism of the algebra $\mathbf{D}(G/K)$ of all G-invariant differential operators on G/K onto a polynomial ring in l generators, l being the rank of the space. In particular, $\mathbf{D}(G/K)$ is commutative, a fact which can also be established more directly.

Section 1 starts with the general measure theory on coset spaces; various integral theorems for semisimple Lie groups, related to the Iwasawa decomposition and the Cartan decomposition, are proved.

Section 2 deals with elementary facts concerning invariant differential operators on reductive coset spaces G/H and the description of these by means of the Lie algebras of G and H.

The spherical functions are studied in §3-§6. They are characterized by means of differential equations and functional equations; some examples are computed. The connection with representations of class one is established and finally in §6 an explicit formula is obtained. For the noncompact type (the most difficult case) this is done by associating to each linear function on $\mathfrak{h}_{\mathfrak{p}_0}$ an eigenfunction of the operators in $D(G/K)$ and then showing (and this is the more difficult part) that all the spherical functions occur in this manner.

In §7 we prove some mean value theorems for solutions of invariant differential equations on G/K, generalizing classical mean value theorems for certain differential equations in Euclidean space involving the Laplacian. The results are most explicit in the case of rank 1 because then the Laplace-Beltrami operator is essentially the only invariant differential operator. In this case we give a simple explicit solution of the Poisson equation $\Delta u = f$.

§ 1. Integral Formulas

1. Generalities

Let S be a locally compact Hausdorf space. The set of real-valued continuous functions on S will be denoted by $C(S)$, and $C_c(S)$ shall denote the set of functions in $C(S)$ of compact support. A *measure* on S is by definition (Bourbaki [2]) a linear mapping $\mu : C_c(S) \to \boldsymbol{R}$ with the property that for each compact subset $K \subset S$ there exists a constant M_K such that

$$|\mu(f)| \leqslant M_K \sup_{x \in S} |f(x)|$$

for all $f \in C_c(S)$ whose support is contained in K. We recall that a linear mapping $\mu : C_c(S) \to \boldsymbol{R}$ which satisfies $\mu(f) \geqslant 0$ for $f \geqslant 0$, $f \in C_c(S)$, is a measure on S. Such a measure is called a *positive measure.*

For a manifold M we put $C_c^\infty(M) = C^\infty(M) \cap C_c(M)$. Suppose M is an orientable m-dimensional manifold and let $(U_\alpha, \varphi_\alpha)_{\alpha \in A}$ be a collection of local charts on M by which M is oriented, (Chapter VIII, §2). Let ω be an m-form on M. We shall define the integral $\int_M f\omega$ for each $f \in C_c(M)$. First we assume that f has compact support contained in a coordinate neighborhood U_α and let $\varphi_\alpha(q) = (x_1(q), ..., x_m(q))$, $q \in U_\alpha$. On U_α, ω has an expression (Chapter I, §2, No. 4),

$$\omega_{U_\alpha} = F_\alpha(x_1, ..., x_m)\, dx_1 \wedge ... \wedge dx_m \tag{1}$$

and we set

$$\int_M f\,\omega = \int_{\varphi_\alpha(U_\alpha)} (f \circ \varphi_\alpha^{-1})\,(x_1, \ldots, x_m)\, F_\alpha(x_1, \ldots, x_m)\, dx_1 \ldots dx_m.$$

On using the transformation formula for multiple integrals we see that if f has compact support inside the intersection $U_\alpha \cap U_\beta$ of two coordinate neighborhoods, then the right-hand side in the formula above is

$$\int_{\varphi_\beta(U_\beta)} (f \circ \varphi_\beta^{-1})\,(y_1, \ldots, y_m)\, F_\beta(y_1, \ldots, y_m)\, dy_1 \ldots dy_m,$$

if $F_\beta dy_1 \wedge \ldots \wedge dy_m$ is the expression for ω on U_β. Thus $\int_M f\omega$ is well defined. Next, let f be an arbitrary function in $C_c(M)$. Then f vanishes outside a paracompact open submanifold of M and by Theorem 1.3, Chapter I, f can be expressed as a finite sum $f = \Sigma_i f_i$ where each f_i has compact support inside some neighborhood U_α from our covering. We put

$$\int_M f\,\omega = \sum_i \int_M f_i\,\omega.$$

Here it has to be verified (Chevalley [2], p. 163) that the right-hand side is independent of the chosen decomposition $f = \Sigma_i f_i$ of f. Let $f = \Sigma_j g_j$ be another such decomposition and select $\varphi \in C_c(M)$ such that $\varphi = 1$ on the union of the supports of all f_i and g_j. Then $\varphi = \Sigma\,\varphi_\alpha$ (finite sum) where each φ_α has support inside a coordinate neighborhood from our covering. We have

$$\sum_i f_i\varphi_\alpha = \sum_j g_j\varphi_\alpha,$$

and since each summand has support inside a fixed coordinate neighborhood

$$\sum_i \int (f_i\varphi_\alpha)\,\omega = \sum_j \int (g_j\varphi_\alpha)\,\omega.$$

For the same reason the formulas

$$f_i = \sum_\alpha f_i\varphi_\alpha, \qquad g_j = \sum_\alpha g_j\varphi_\alpha$$

imply that

$$\int f_i\omega = \sum_\alpha \int (f_i\varphi_\alpha)\,\omega, \qquad \int g_j\omega = \sum_\alpha \int (g_j\varphi_\alpha)\,\omega,$$

from which we derive the desired relation

$$\sum_i \int f_i\omega = \sum_j \int g_j\omega.$$

The integral $\int f\omega$ is now well defined and the mapping

$$f \to \int_M f\,\omega, \qquad f \in C_c(M),$$

is a measure on M. We have obviously:

Lemma 1.1. *If* $\int_M f\omega = 0$ *for all* $f \in C_c(M)$, *then* $\omega = 0$.

Definition. The m-form ω is said to be *positive* if for each $\alpha \in A$, the function F_α in (1) is > 0 on $\varphi_\alpha(U_\alpha)$.

If ω is a positive m-form on M, then it follows readily from Theorem 1.3, Chapter I, that $\int_M f\omega \geqslant 0$ for each nonnegative function $f \in C_c(M)$. Thus, a positive m-form gives rise to a positive measure.

Suppose M and N are two oriented manifolds and let Φ be a diffeomorphism of M onto N. We assume that Φ is *orientation preserving*, that is, if the collection of local charts $(U_\alpha, \varphi_\alpha)_{\alpha\in A}$ defines the orientation on M, then the collection $(\Phi(U_\alpha), \varphi_\alpha \circ \Phi^{-1})_{\alpha\in A}$ of local charts on N defines the orientation on N. Let m denote the dimension of M and N.

Let ω be an m-form on N. Then the formula

$$\int_M f\Phi^*\omega = \int_N (f \circ \Phi^{-1})\,\omega \tag{2}$$

holds for all $f \in C_c(M)$. In fact, it suffices to verify (2) in the case when f has compact support inside a coordinate neighborhood U_α. If we evaluate the left-hand side of (2) by means of the coordinate system φ_α and the right-hand side of (2) by means of the coordinate system $\varphi_\alpha \circ \Phi^{-1}$, both sides of (2) reduce to the same integral.

If M is a pseudo-Riemannian manifold, orientable or not, a measure can be defined on M as follows. Consider a local chart $(U_\alpha, \varphi_\alpha)$ on M and, as before, let $\varphi_\alpha(q) = (x_1, \ldots, x_m)$, $q \in U_\alpha$. Let g denote the pseudo-Riemannian structure and generalizing the definition in Chapter VIII, §2, put

$$g_{ij} = g\left(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}\right), \qquad \bar{g} = |\det(g_{ij})|.$$

For each nonnegative function f in $C_c(U_\alpha)$ put

$$\mu(f) = \left| \int_{\varphi_\alpha(U_\alpha)} (f \circ \varphi_\alpha^{-1})(x_1, \ldots, x_m)\sqrt{\bar{g}}\, dx_1 dx_2 \ldots dx_m \right|.$$

The expression on the right is invariant under coordinate changes. For any function $f \in C_c(U_\alpha)$ write $f = f_1 - f_2$ where f_1 and f_2 are nonnegative functions in $C_c(U_\alpha)$ and put $\mu(f) = \mu(f_1) - \mu(f_2)$. The result $\mu(f)$ is independent of the choice of decomposition $f = f_1 - f_2$. Using partition of unity as before $\mu(f)$ can be defined for all $f \in C_c(M)$. The result is a positive measure on M, which we shall refer to as the *Riemannian measure* on M.

Let M be a manifold and Φ a diffeomorphism of M onto itself. We recall that a differential form ω on M is called invariant under Φ if $\Phi^*\omega = \omega$.

Let G be a Lie group with Lie algebra $\mathfrak{g}$. A differential form ω on G is called *left invariant* if $L_x^*\omega = \omega$ for all $x \in G$, L_x (or $L(x)$) denoting the left translation $g \to xg$ on G. Also, R_x (or $R(x)$) denotes the right translation $g \to gx$ on G and *right invariant* differential forms on G can be defined. If $X \in \mathfrak{g}$, let $\tilde{X}$ denote the corresponding left invariant vector field on G. Let $X_1, ..., X_n$ be a basis of $\mathfrak{g}$. The equations $\omega^i(\tilde{X}_j) = \delta^i{}_j$ determine uniquely n 1-forms ω^i on G. These are clearly left invariant and the exterior product $\omega = \omega^1 \wedge ... \wedge \omega^n$ is a left invariant n-form on G. Each 1-form on G can be written $\Sigma_{i=1}^n f_i\omega^i$ where $f_i \in C^\infty(G)$; it follows that each n-form can be written $f\omega$ where $f \in C^\infty(G)$. Thus, except for a constant factor, ω is the only left invariant n-form on G. Let $\varphi : x \to (x_1(x), ..., x_n(x))$ be a system of canonical coordinates with respect to the basis $X_1, ..., X_n$ of $\mathfrak{g}$, valid on a connected open neighborhood U of e. On U, ω has an expression

$$\omega_U = F(x_1, ..., x_n)\, dx_1 \wedge ... \wedge dx_n$$

and $F > 0$. Now, if $g \in G$, the pair $(L_gU, \varphi \circ L_{g^{-1}})$ is a local chart on a connected neighborhood of g. We put $(\varphi \circ L_{g^{-1}})(x) = (y_1(x), ..., y_n(x))$, $(x \in L_gU)$. Since $y_i(gx) = x_i(x)$, $(x \in U \cap L_gU)$, the mapping $L_g : U \to L_gU$ has coordinate expression (Chapter I, §3(1)) given by $(y_1, ..., y_n) = (x_1, ..., x_n)$. On L_gU, ω has an expression

$$\omega_{L_gU} = G(y_1, ..., y_n)\, dy_1 \wedge ... \wedge dy_n$$

so the invariance condition $\omega_x = L_g^*\omega_{gx}$ $(x \in U \cap L_gU)$ can be written

$$G(y_1(x), ..., y_n(x))\, (dy_1 \wedge ... \wedge dy_n)_x = G(x_1(x), ..., x_n(x))\, (dx_1 \wedge ... \wedge dx_n)_x.$$

Hence $F(x_1(x), ..., x_n(x)) = G(x_1(x), ..., x_n(x))$ and

$$F(x_1(x), ..., x_n(x)) = F(y_1(x), ..., y_n(x))\, \frac{\partial(y_1(x), ..., y_n(x))}{\partial(x_1(x), ..., x_n(x))}$$

for $x \in U \cap L_g U$, which shows that the Jacobian of $(\varphi \circ L_{g^{-1}}) \circ \varphi^{-1}$ is > 0. Consequently, the collection $(L_g U, \varphi \circ L_{g^{-1}})_{g \in G}$ of local charts turns G into an oriented manifold and each left translation is orientation preserving. The orientation of G depends on the choice of basis of $\mathfrak{g}$. If $X_1', \ldots, X_n'$ is another basis, then the resulting orientation of G is the same as that before if and only if the linear transformation $X_i \to X_i'$ $(1 \leqslant i \leqslant n)$ has positive determinant.

The form ω is a positive left invariant n-form on G and except for a constant positive factor, ω is uniquely determined by these properties. We shall denote it by $d_l g$. The linear mapping of $C_c(G)$ into $\boldsymbol{R}$ given by $f \to \int f \, d_l g$ is a *measure* on G which we denote by μ_l. This measure is positive; moreover, it is *left invariant* in the sense that $\mu_l(f \circ L_x) = \mu_l(f)$ for $x \in G, f \in C_c(G)$.

Similarly, G can be turned into an oriented manifold such that each R_g $(g \in G)$ is orientation preserving. There exists a right invariant positive n-form $d_r g$ on G and this is unique except for a constant positive factor. We define the *right invariant* positive measure μ_r on G by

$$\mu_r(f) = \int f \, d_r g, \qquad f \in C_c(G).$$

The group G has been oriented in two ways. The left invariant orientation is invariant under all right translations R_x $(x \in G)$ if and only if it is invariant under all $I(x) = L_x \circ R_{x^{-1}}$ $(x \in G)$. Since the differential $dI(x)_g$ satisfies

$$dI(x)_g = dL_{xgx^{-1}} \circ \operatorname{Ad}(x) \circ dL_{g^{-1}},$$

the necessary and sufficient condition is $\det \operatorname{Ad}(x) > 0$ for all $x \in G$. This condition is always fulfilled if G is connected.

Lemma 1.2. *With the notation above we have*

$$d_r g = c \det \operatorname{Ad}(g) \, d_l g,$$

where c is a constant.

Proof. Let $\theta = \det \operatorname{Ad}(g) \, d_l g$ and let $x \in G$. Then

$$(R_{x^{-1}})^* \theta = \det \operatorname{Ad}(gx^{-1}) \, (R_{x^{-1}})^* \, d_l g = \det \operatorname{Ad}(gx^{-1}) \, I(x)^* \, d_l g.$$

At the point $g = e$ we have

$$(I(x)^* \, (d_l g))_e = \det \operatorname{Ad}(x) \, (d_l g)_e.$$

Consequently,

$$(R_{x^{-1}}{}^*\theta)_e = \det \operatorname{Ad}(e)\,(d_lg)_e = \theta_e.$$

Thus, θ is right invariant and therefore proportional to d_rg.

Remark. If G is connected it can be oriented in such a way that all left and right translations are orientation preserving. If d_rg and d_lg are defined by means of this orientation, Lemma 1.2 holds with $c > 0$.

Corollary 1.3. *Let $x, y \in G$ and put $d_l(ygx) = (L_yR_x)^*d_lg$, $d_r(xgy) = (L_xR_y)^*d_rg$. Moreover, if J denotes the mapping $g \to g^{-1}$, put $d_l(g^{-1}) = J^*(d_lg)$. Then*

$$d_l(gx) = \det \operatorname{Ad}(x^{-1})\,d_l(g), \qquad d_r(xg) = \det \operatorname{Ad}(x)\,d_rg,$$
$$d_l(g^{-1}) = (-1)^{\dim G} \det \operatorname{Ad}(g)\,d_lg.$$

In fact, the lemma implies that

$$c \det \operatorname{Ad}(g)\,d_lg = d_rg = d_r(gx) = c \det \operatorname{Ad}(gx)\,d_l(gx),$$
$$d_r(xg) = c \det \operatorname{Ad}(xg)\,d_l(xg) = c \det \operatorname{Ad}(xg)\,d_lg.$$

Finally, since $JR_x = L_{x^{-1}}J$, we have

$$(R_x)^*\,d_l(g^{-1}) = (R_x)^*\,J^*d_lg = (JR_x)^*\,d_lg = (L_{x^{-1}}J)^*\,d_lg = J^*\,d_lg,$$

so $d_l(g^{-1})$ is right invariant, hence proportional to d_rg. But obviously

$$(d_l(g^{-1}))_e = (-1)^{\dim G}\,(d_lg)_e,$$

so the corollary is verified.

Definition. A Lie group G is called *unimodular* if the left invariant measure μ_l is also right invariant.

In view of Cor. 1.3 we have by (2)

$$\mu_l(f \circ R_x) = |\det \operatorname{Ad}(x)|\,\mu_l(f). \tag{3}$$

It follows that G is unimodular if and only if $|\det \operatorname{Ad}(x)| = 1$ for all $x \in G$. If this condition is satisfied, the measures μ_l and μ_r coincide except for a constant factor.

Proposition 1.4. *The following Lie groups are unimodular:*

(i) *Lie groups G for which $Ad\,(G)$ is compact.*
(ii) *Semisimple Lie groups.*
(iii) *Connected nilpotent Lie groups.*

Proof. In the case (i), the group $\{|\det \mathrm{Ad}(x)| : x \in G\}$ is a compact subgroup of the multiplicative group of positive real numbers. This subgroup necessarily consists of one element so G is unimodular. In the case (ii), each $\mathrm{Ad}(x)$ leaves invariant a nondegenerate bilinear form (namely, the Killing form). It follows that $(\det \mathrm{Ad}(x))^2 = 1$. Finally, let N be a connected nilpotent Lie group with Lie algebra $\mathfrak{n}$. If $X \in \mathfrak{n}$, then $\mathrm{ad}\, X$ is nilpotent so $\mathrm{Tr}(\mathrm{ad}\, X) = 0$. Since

$$\det e^A = e^{\mathrm{Tr}\, A}$$

for an arbitrary linear transformation A, we obtain

$$\det \mathrm{Ad}(\exp X) = e^{\mathrm{Tr}(\mathrm{ad}\, X)} = 1.$$

This proves (iii).

Notation. In the sequel we shall mostly use the left invariant measure μ_l. For simplicity we shall write μ instead of μ_l and dg instead of d_lg.

2. Invariant Measures on Coset Spaces

Let G be a Lie group with Lie algebra $\mathfrak{g}$; let H be a closed subgroup with Lie algebra $\mathfrak{h} \subset \mathfrak{g}$. Each $x \in G$ gives rise to an analytic diffeomorphism $\tau(x)$: $gH \to xgH$ of G/H onto itself. Let π denote the natural mapping of G onto G/H and put $o = \pi(e)$. If $h \in H$, $(d\tau(h))_o$ is an endomorphism of the tangent space $(G/H)_o$. For simplicity, we shall write $d\tau(h)$ instead of $(d\tau(h))_o$ and $d\pi$ instead of $(d\pi)_e$.

Lemma 1.5.

$$\det(d\tau(h)) = \frac{\det \mathrm{Ad}_G(h)}{\det \mathrm{Ad}_H(h)} \qquad (h \in H).$$

Proof. It was shown in Chapter II, §4, that the differential $d\pi$ is a linear mapping of $\mathfrak{g}$ onto $(G/H)_o$ and has kernel $\mathfrak{h}$. Let $\mathfrak{m}$ be any subspace of $\mathfrak{g}$ such that $\mathfrak{g} = \mathfrak{h} + \mathfrak{m}$ (direct sum). Then $d\pi$ induces an isomorphism of $\mathfrak{m}$ onto $(G/H)_o$. Let $X \in \mathfrak{m}$. Then $\mathrm{Ad}_G(h)\, X = dR_{h^{-1}} \circ dL_h(X)$. Since $\pi \circ R_h = \pi$, $(h \in H)$ and $\pi \circ L_g = \tau(g) \circ \pi$, $(g \in G)$, we obtain

$$d\pi \circ \mathrm{Ad}_G(h)\, X = d\tau(h) \circ d\pi(X), \qquad h \in H, X \in \mathfrak{m}. \tag{4}$$

The vector $\mathrm{Ad}_G(h)\, X$ decomposes according to $\mathfrak{g} = \mathfrak{h} + \mathfrak{m}$,

$$\mathrm{Ad}_G(h)\, X = X(h)_{\mathfrak{h}} + X(h)_{\mathfrak{m}}.$$

The endomorphism $A_h : X \to X(h)_{\mathfrak{m}}$ of $\mathfrak{m}$ satisfies

$$d\pi \circ A_h(X) = d\tau(h) \circ d\pi(X), \qquad X \in \mathfrak{m},$$

so $\det A_h = \det (d\tau(h))$. On the other hand,

$$\exp \mathrm{Ad}_G (h)\, tT = h \exp tT\, h^{-1} = \exp \mathrm{Ad}_H (h)\, tT$$

for $t \in \boldsymbol{R}$, $T \in \mathfrak{h}$. Hence $\mathrm{Ad}_G (h)\, T = \mathrm{Ad}_H (h)\, T$ so

$$\det \mathrm{Ad}_G (h) = \det \mathrm{A}_h \det \mathrm{Ad}_H (h)$$

and the lemma is proved.

Proposition 1.6. *Let $m = \dim G/H$. The following conditions are equivalent*:

(i) *G/H has a nonzero G-invariant m-form ω*;
(ii) *$\det Ad_G (h) = \det Ad_H (h)$ for $h \in H$.*

If these conditions are satisfied, then G/H has a G-invariant orientation and the G-invariant m-form ω is unique up to a constant factor.

Proof. Let ω be a G-invariant m-form on G/H, $\omega \neq 0$. Then the relation $\tau(h)^*\omega = \omega$ at the point o implies $\det (d\tau(h)) = 1$ so (ii) holds. On the other hand, let $X_1, \ldots, X_m$ be a basis of $(G/H)_o$ and let $\omega^1, \ldots, \omega^m$ be the linear functions on $(G/H)_o$ determined by $\omega^i(X_j) = \delta_{ij}$. Consider the element $\omega^1 \wedge \ldots \wedge \omega^m$ in the Grassmann algebra of the tangent space $(G/H)_o$. Condition (ii) implies that $\det (d\tau(h)) = 1$ and the element $\omega^1 \wedge \ldots \wedge \omega^m$ is invariant under the linear transformation $d\tau(h)$. It follows that there exists a unique G-invariant m-form ω on G/H such that $\omega_o = \omega^1 \wedge \ldots \wedge \omega^m$. If ω^* is another G-invariant m-form on G/H, then $\omega^* = f\omega$ where $f \in C^\infty(G/H)$. Owing to the G-invariance, $f =$ constant.

Assuming (i), let $\varphi : p \to (x_1(p), \ldots, x_m(p))$ be a system of coordinates on an open connected neighborhood U of $o \in G/H$ on which ω has an expression

$$\omega_U = F(x_1, \ldots, x_m)\, dx_1 \wedge \ldots \wedge dx_m,$$

with $F > 0$. The pair $(\tau(g)\, U, \varphi \circ \tau(g^{-1}))$ is a local chart on a connected neighborhood of $g{\cdot}o \in G/H$. We put $(\varphi \circ \tau(g^{-1}))\, (p) = (y_1(p), \ldots, y_m(p))$ for $p \in \tau(g)\, U$. Then the mapping $\tau(g) : U \to \tau(g)\, U$ has expression (Chapter I, §3(1)) $(y_1, \ldots, y_m) = (x_1, \ldots, x_m)$. On $\tau(g)U$ ω has an expression

$$\omega_{\tau(g)U} = G(y_1, \ldots, y_m)\, dy_1 \wedge \ldots \wedge dy_m$$

and since $\omega_q = \tau(g)^*\omega_{\tau(g)q}$ we have for $q \in U \cap \tau(g)\, U$

$$\omega_q = G(y_1(q), \ldots, y_m(q))\, (dy_1 \wedge \ldots \wedge dy_m)_q = G(x_1(q), \ldots, x_m(q))\, (dx_1 \wedge \ldots \wedge dx_m)_q.$$

Hence $F(x_1(q), \ldots, x_m(q)) = G(x_1(q), \ldots, x_m(q))$ and

$$F(x_1(q), \ldots, x_m(q)) = F(y_1(q), \ldots, y_m(q)) \frac{\partial(y_1(q), \ldots, y_m(q))}{\partial(x_1(q), \ldots, x_m(q))},$$

which shows that the Jacobian of the mapping $(\varphi \circ \tau(g^{-1})) \circ \varphi^{-1}$ is > 0. Consequently, the collection $(\tau(g)\, U, \varphi \circ \tau(g^{-1}))_{g \in G}$ of local charts turns G/H into an oriented manifold and each $\tau(g)$ is orientation preserving.

The G-invariant form ω now gives rise to an integral $\int f\omega$ which is invariant in the sense that

$$\int_{G/H} f\,\omega = \int_{G/H} (f \circ \tau(g))\,\omega, \qquad g \in G.$$

However, just as the Riemannian measure did not require orientability, an invariant measure can be constructed on G/H under a condition which is slightly more general than (ii). The projective space $\boldsymbol{P}^2$ will, for example, satisfy this condition whereas it does not satisfy (ii). We recall that a measure μ on G/H is said to be invariant (or more precisely G-invariant) if $\mu(f \circ \tau(g)) = \mu(f)$ for all $g \in G$.

Theorem 1.7. *Let G be a Lie group and H a closed subgroup. The relation*

$$|\det \mathrm{Ad}_G(h)| = |\det \mathrm{Ad}_H(h)|, \qquad h \in H, \tag{5}$$

is a necessary and sufficient condition for the existence of a G-invariant measure > 0 on G/H. This measure dg_H is unique (up to a constant factor) and

$$\int_G f(g)\, dg = \int_{G/H} \left(\int_H f(gh)\, dh \right) dg_H, \qquad f \in C_c(G), \tag{6}$$

if the left invariant measures dg and dh are suitably normalized.

We begin by proving a simple lemma.

Lemma 1.8. *Let G be a Lie group and H a closed subgroup. Let dh be a left invariant measure > 0 on H and put*

$$\bar{f}(gH) = \int_H f(gh)\, dh, \qquad f \in C_c(G).$$

Then the mapping $f \to \bar{f}$ is a linear mapping of $C_c(G)$ onto $C_c(G/H)$.

Proof. Let $F \in C_c(G/H)$; we have to prove that there exists a function $f \in C_c(G)$ such that $F = \bar{f}$. Let C be a compact subset of G/H outside which F vanishes and let C' be a compact subset of G whose image is C

under the natural mapping $\pi : G \to G/H$. Let C_H be a compact subset of H of positive measure and put $\tilde{C} = C' \cdot C_H$. Then $\pi(\tilde{C}) = C$. Select $f_1 \in C_c(G)$ such that $f_1 \geqslant 0$ on G and $f_1 > 0$ on $\tilde{C}$. Then $\tilde{f}_1 > 0$ on C (since C_H has positive measure) and the function

$$f(g) = \begin{cases} f_1(g) \dfrac{F(\pi(g))}{\tilde{f}_1(\pi(g))}, & \text{if } \pi(g) \in C, \\ 0 \quad , & \text{if } \pi(g) \notin C, \end{cases}$$

belongs to $C_c(G)$ and $\tilde{f} = F$.

Now in order to prove Theorem 1.7 suppose first that the relation

$$|\det \mathrm{Ad}_G(h)| = |\det \mathrm{Ad}_H(h)|, \qquad h \in H,$$

holds. Let $\varphi \in C_c(G)$. Since we are dealing with measures rather than differential forms, we have by Cor. 1.3,

$$\int_G \varphi(g) \Big(\int_H f(gh)\, dh\Big)\, dg = \int_H dh \int_G \varphi(g) f(gh)\, dg$$

$$= \int_H dh \int_G \varphi(gh^{-1}) f(g) \,|\det \mathrm{Ad}_G(h)|\, dg = \int_G f(g)\, dg \int_H \varphi(gh^{-1}) \,|\det \mathrm{Ad}_G(h)|\, dh.$$

But the relation (5) and the last part of Cor. 1.3 shows that

$$\int_H \varphi(gh^{-1}) \,|\det \mathrm{Ad}_G(h)|\, dh = \int_H \varphi(gh)\, dh,$$

so

$$\int_G \varphi(g)\, dg \int_H f(gh)\, dh = \int_G f(g)\, dg \int_G \varphi(gh)\, dh.$$

This relation implies, φ being arbitrary, that

$$\int_G f(g)\, dg = 0 \qquad \text{if } \tilde{f} \equiv 0.$$

In view of the lemma we can therefore define a linear mapping $\mu : C_c(G/H) \to \boldsymbol{R}$ by

$$\mu(F) = \int_G f(g)\, dg \qquad \text{if } F = \tilde{f}.$$

Since $\mu(F) \geqslant 0$ if $F \geqslant 0$, μ is a positive measure on G/H; moreover,

$$\mu((\tilde{f})^{\tau(x)}) = \int_G f^{L(x)}(g)\, dg = \int_G f(g)\, dg = \mu(\tilde{f}),$$

so μ is invariant.

In order to prove the converse we shall make use of the theorem that the group G has (up to a constant factor) a unique positive left invariant measure. For this "uniqueness of Haar measure" see, e.g., Weil's book [1].

If μ is a positive invariant measure on G/H, the mapping $f \to \mu(\bar{f})$ is a positive left invariant measure on G. Owing to the uniqueness mentioned,

$$\int_G f(g)\, dg = \mu(\bar{f}).$$

In view of the lemma this proves the uniqueness of μ as well as (6). In order to derive (5), replace $f(g)$ by $f(gh_1)$ in (6). Owing to Cor. 1.3 the left-hand side is multiplied by $|\det \mathrm{Ad}_G(h_1)|$ and the right-hand side is multiplied by $|\det \mathrm{Ad}_H(h_1)|$. This finishes the proof of Theorem 1.7.

In the remainder of this section we shall often have to calculate how invariant measures transform under mappings. To a certain extent these calculations can be reduced to the following general lemma.

Lemma 1.9. *Let G and S be Lie groups and $H \subset G$ and $T \subset S$ closed subgroups. Suppose that the coset spaces G/H and S/T have the same dimension m and that they possess positive invariant differential m-forms, denoted dg_H and ds_T. Let $o = \{H\}$, $o' = \{T\}$.*

Let φ be a differentiable mapping of G/H into S/T such that $\varphi \cdot \{H\} = \{T\}$. Then

$$\varphi^*(ds_T) = D dg_H,$$

where D is a function on G/H computed as follows: Let $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_m$ be fixed (but arbitrary) bases of the tangent spaces $(G/H)_o$ and $(S/T)_{o'}$, respectively, such that

$$dg_H(X_1, \ldots, X_m) = ds_T(Y_1, \ldots, Y_m).$$

Let $g \in G$ and select $s \in \varphi(gH)$. Consider the linear mapping

$$A(g) = d\tau(s^{-1}) \circ d\varphi_{gH} \circ d\tau(g)_o$$

of $(G/H)_o$ into $(S/T)_{o'}$ and put $A(g)\, X_j = \sum_i a_{ij}(g)\, Y_i$. Then

$$D(gH) = \det(a_{ij}(g)).$$

Proof. Let $\omega^1, \ldots, \omega^m$ be the linear functions on $(G/H)_{o'}$ determined by $\omega^i(X_j) = \delta_{ij}$; let $\theta^1, \ldots, \theta^m$ be the linear functions on $(S/T)_{o'}$ determined by $\theta^i(Y_j) = \delta_{ij}$. Then the dual mapping

$$^t(A(g)) = \tau(g)^* \circ \varphi^* \circ \tau(s^{-1})^*,$$

which maps the dual of $(S/T)_{o'}$ into the dual of $(G/H)_o$ satisfies

$$ {}^t(A(g))\,\theta^i = \sum_j a_{ij}(g)\,\omega^j. $$

Now, due to the assumption about dg_H and ds_T,

$$ (dg_H)_{gH} = c\,\tau(g^{-1})^*\,(\omega^1 \wedge \dots \wedge \omega^m), $$
$$ (ds_T)_{sT} = c\,\tau(s^{-1})^*\,(\theta^1 \wedge \dots \wedge \theta^m), $$

where c is a constant. Consequently,

$$ \begin{aligned} (\varphi^*(ds_T))_{gH} &= c\,\varphi^* \circ \tau(s^{-1})^*\,(\theta^1 \wedge \dots \wedge \theta^m) \\ &= c\,\tau(g^{-1})^*\,{}^tA(g)\,(\theta^1 \wedge \dots \wedge \theta^m) \\ &= c\det(a_{ij}(g))\,\tau(g^{-1})^*\,(\omega^1 \wedge \dots \wedge \omega^m) \\ &= \det(a_{ij}(g))\,(dg_H)_{gH}. \end{aligned} $$

3. Some Integral Formulas for Semisimple Lie Groups

Lemma 1.10. *Let U be a Lie group with Lie algebra $\mathfrak{u}$. Suppose $\mathfrak{u}$ is a direct sum $\mathfrak{u} = \mathfrak{m} + \mathfrak{h}$ where $\mathfrak{m}$ and $\mathfrak{h}$ are subalgebras of $\mathfrak{u}$ (not necessarily ideals). Let M and H denote the analytic subgroups of U with Lie algebras $\mathfrak{m}$ and $\mathfrak{h}$, respectively. Suppose the mapping $\alpha : (m, h) \to mh$ is a one-to-one mapping of $M \times H$ onto U.*

Then the positive left invariant measures dh, dm, du can be normalized in such a way that

$$ \int_U f(u)\,du = \int_{M\times H} f(mh)\,\frac{\det \mathrm{Ad}_H(h)}{\det \mathrm{Ad}_U(h)}\,dm\,dh $$

for all $f \in C_c(U)$.

Proof. From Lemma 5.2, Chapter VI, we know that α is a diffeomorphism and

$$ d\alpha_{(m,h)}(dL_m Y, dL_h Z) = dL_{mh}(\mathrm{Ad}_U(h^{-1})\,Y + Z) $$

for $m \in M$, $h \in H$, $Y \in \mathfrak{m}$, and $Z \in \mathfrak{h}$. Using Lemma 1.9 we see that

$$ \alpha^*(du) = D(m, h)\,dm\,dh, $$

where $dm\,dh$ is the invariant measure on the product group $M \times H$ and $D(m, h)$ is the determinant of the linear mapping

$$ A(m, h) : (Y, Z) \to \mathrm{Ad}_U(h^{-1})\,Y + Z $$

of $\mathfrak{m} \times \mathfrak{h}$ into $\mathfrak{g}$. (Since we identify $\mathfrak{m} \times \mathfrak{h}$ with $\mathfrak{g}$ we can indeed regard $A(m, h)$ as an endomorphism of $\mathfrak{g}$ and thus can speak of its determinant.) Since

$$\mathrm{Ad}_U(h^{-1})\, Y + Z = \mathrm{Ad}_U(h^{-1})\,(Y + \mathrm{Ad}_H(h)\, Z)$$

it is clear that

$$\det A(m, h) = \frac{\det \mathrm{Ad}_H(h)}{\det \mathrm{Ad}_U(h)}$$

and the lemma follows from (2).

Let $\mathfrak{g}_0$ be a semisimple Lie algebra over $\boldsymbol{R}$ and $\mathfrak{g}$ its complexification. We recall that an Iwasawa decomposition of $\mathfrak{g}_0$ is constructed as follows: let $\mathfrak{g}_0 = \mathfrak{k}_0 + \mathfrak{p}_0$ be a Cartan decomposition of $\mathfrak{g}_0$, let $\mathfrak{h}_{\mathfrak{p}_0}$ be a maximal abelian subspace of $\mathfrak{p}_0$ and extend $\mathfrak{h}_{\mathfrak{p}_0}$ to a maximal abelian subalgebra $\mathfrak{h}_0$ of $\mathfrak{g}_0$. Let $\mathfrak{h}$ be the subalgebra of $\mathfrak{g}$ generated by $\mathfrak{h}_0$, let Δ denote the set of nonzero roots of $\mathfrak{g}$ with respect to $\mathfrak{h}$ and put $\mathfrak{h}^* = \sum_{\alpha \in \Delta} \boldsymbol{R}H_\alpha$. Then $\mathfrak{h}_{\mathfrak{p}_0} \subset \mathfrak{h}^*$ and compatible orderings can be introduced in the dual spaces of $\mathfrak{h}_{\mathfrak{p}_0}$ and $\mathfrak{h}^*$. Let P_+ denote the set of positive roots which do not vanish identically on $\mathfrak{h}_{\mathfrak{p}_0}$. Let $\rho = \frac{1}{2} \sum_{\alpha \in P_+} \alpha$, $\mathfrak{n} = \sum_{\alpha \in P_+} \mathfrak{g}^\alpha$, $\mathfrak{n}_0 = \mathfrak{g}_0 \cap \mathfrak{n}$. Then $\mathfrak{g}_0 = \mathfrak{k}_0 + \mathfrak{h}_{\mathfrak{p}_0} + \mathfrak{n}_0$ is the Iwasawa decomposition.

Let G be any connected semisimple Lie group with Lie algebra $\mathfrak{g}_0$ and let $G = KA_\mathfrak{p}N$ be the corresponding global decomposition (Theorem 5.1, Chapter VI). If $a \in A_\mathfrak{p}$, let $\log a$ denote the unique element $H \in \mathfrak{h}_{\mathfrak{p}_0}$ for which $\exp H = a$.

Proposition 1.11. *Let $G = KA_\mathfrak{p}N$ be an Iwasawa decomposition of a connected semisimple Lie group G. Let dk, da, and dn be left invariant measures on K, $A_\mathfrak{p}$, and N, respectively. Then the left invariant measure dg on G can be normalized such that*

$$\int_G f(g)\, dg = \int_{K \times A_\mathfrak{p} \times N} f(kan)\, e^{2\rho(\log a)}\, dk\, da\, dn, \qquad (f \in C_c(G)).$$

Proof. The group $S = A_\mathfrak{p}N$ is a closed subgroup of G containing $A_\mathfrak{p}$ and N as analytic subgroups. The Lie algebra $\mathfrak{s}_0$ of S is the direct sum $\mathfrak{s}_0 = \mathfrak{h}_{\mathfrak{p}_0} + \mathfrak{n}_0$. We consider now the following mappings:

$$\begin{array}{ll} \Phi : (k, a, n) \to kan & \text{of } K \times A_\mathfrak{p} \times N \text{ onto } G; \\ \alpha : (k, s) \to ks & \text{of } K \times S \text{ onto } G; \\ \beta : (k, a, n) \to (k, an) & \text{of } K \times A_\mathfrak{p} \times N \text{ onto } K \times S. \end{array}$$

These mappings are diffeomorphisms and $\Phi = \alpha \circ \beta$. Now, let the left invariant measures dk, da, and dn be arbitrarily normalized. Then according to Lemma 1.10, the left invariant measure ds on S can be normalized in such a way that

$$\beta^*(dk\,ds) = \frac{\det \mathrm{Ad}_N(n)}{\det \mathrm{Ad}_S(n)}\, dk\, da\, dn$$

and the left invariant measure dg on G can be normalized such that

$$\alpha^*(dg) = \frac{\det \mathrm{Ad}_S(s)}{\det \mathrm{Ad}_G(s)}\, dk\, ds.$$

Since $\Phi^* = \beta^* \circ \alpha^*$ it follows that

$$\begin{aligned}\Phi^*(dg) &= \frac{\det \mathrm{Ad}_S(an)}{\det \mathrm{Ad}_G(an)} \cdot \frac{\det \mathrm{Ad}_N(n)}{\det \mathrm{Ad}_S(n)}\, dk\, da\, dn \\ &= \det \mathrm{Ad}_S(a)\, dk\, da\, dn\end{aligned}$$

by Prop. 1.4. Now put $H = \log a$ for $a \in A_{\mathfrak{p}}$. Then

$$\det \mathrm{Ad}_S(a) = \exp(\mathrm{Tr}(\mathrm{ad}_{\mathfrak{s}_0}(H))).$$

A basis of $\mathfrak{n}_0$ is obtained by selecting a nonzero vector in each $\mathfrak{g}^\alpha \cap \mathfrak{g}_0$, $(\alpha \in P_+)$. Therefore

$$\mathrm{Tr}\,\mathrm{ad}_{\mathfrak{s}_0}(H) = \sum_{\alpha \in P_+} \alpha(H) = 2\rho(H).$$

This proves the proposition.

If $a \in A_{\mathfrak{p}}$, the mapping $n \to ana^{-1}$ is an automorphism of N. It follows that the mapping $\Psi : (k, n, a) \to kna$ is an analytic diffeomorphism of $K \times N \times A_{\mathfrak{p}}$ onto G. We find just as before

$$\Psi^*(dg) = \frac{\det \mathrm{Ad}_S(na)}{\det \mathrm{Ad}_G(na)} \cdot \frac{\det \mathrm{Ad}_A(a)}{\det \mathrm{Ad}_S(a)}\, dk\, dn\, da.$$

If $X \in \mathfrak{n}_0$, then $(\mathrm{ad}_{\mathfrak{s}_0} X)(H) \in \mathfrak{n}_0$ for $H \in \mathfrak{h}_{\mathfrak{p}_0}$. Since $\mathrm{ad}_{\mathfrak{s}_0} X$ has a nilpotent restriction to $\mathfrak{n}_0$, we conclude that $\mathrm{Tr}(\mathrm{ad}_{\mathfrak{s}_0} X) = 0$. It follows that $\det \mathrm{Ad}_S(n) = 1$ for $n \in N$. The formula for $\Psi^*(dg)$ therefore reduces to

$$\Psi^*(dg) = dk\, dn\, da. \tag{7}$$

Proposition 1.12. *Let G, K, $A_{\mathfrak{p}}$, N and dg, dk, da, dn be as in Prop.* 1.11; *then*

$$\int_G f(g)\, dg = \int_{K \times N \times A_{\mathfrak{p}}} f(kna)\, dk\, dn\, da \tag{8}$$

for all $f \in C_c(G)$. The invariant positive measure dg^* on $G/A_\mathfrak{p}$ can be normalized in such a way that

$$\int_{G/A_\mathfrak{p}} F(g^*)\, dg^* = \int_{K\times N} (F \circ \gamma)\,(k, n)\, dk\, dn \tag{9}$$

for all $F \in C_c(G/A_\mathfrak{p})$, γ denoting the mapping $(k, n) \to knA_\mathfrak{p}$ of $K \times N$ onto $G/A_\mathfrak{p}$.

Proof. The first statement (8) follows at once from (7). The existence of dg^* is obvious from Theorem 1.7 since G and $A_\mathfrak{p}$ are unimodular. Let $f \in C_c(G)$. From (8) and Theorem 1.7 we have

$$\int_G f(g)\, dg = \int_{K\times N} \Big(\int_{A_\mathfrak{p}} f(kna)\, da\Big)\, dk\, dn = \int_{G/A_\mathfrak{p}} \Big(\int_{A_\mathfrak{p}} f(ga)\, da\Big)\, dg^*.$$

Relation (9) will follow provided each $F \in C_c(G/A_\mathfrak{p})$ can be written in the form

$$F(gA_\mathfrak{p}) = \int_{A_\mathfrak{p}} f(ga)\, da \qquad (g \in G) \tag{10}$$

where $f \in C_c(G)$. But this is guaranteed by Lemma 1.8 so the proposition is proved.

Consider now the function

$$D(H) = \prod_{\alpha \in P_+} \sinh \frac{\alpha(H)}{2}, \qquad H \in \mathfrak{h},$$

and let $\mathfrak{h}_0'$ denote the set of $H \in \mathfrak{h}_0$ for which $D(H) \neq 0$. If $H \in \mathfrak{h}_0'$ and $h = \exp H$, then as shown in Chapter VI, §4, the mapping $\xi : n \to h^{-1}\, nhn^{-1}$ is an analytic diffeomorphism of N onto itself.

Proposition 1.13. *Let $H \in \mathfrak{h}_0'$ and $h = \exp H$. Then*

$$\int_N f(n)\, dn = \prod_{\alpha \in P_+} |1 - e^{-\alpha(H)}| \int_N f(h^{-1}\, nhn^{-1})\, dn$$

for $f \in C_c(N)$.

Proof. As shown in Chapter VI, §4, ξ is composed of two diffeomorphisms, $\xi = \psi \circ \varphi$, where

$$\psi : X \to \exp(-H) \exp(H + X)$$

maps $\mathfrak{n}_0$ onto N and

$$\varphi : n \to \mathrm{Ad}\,(n)\,H - H$$

maps N onto $\mathfrak{n}_0$. The differentials of these mappings were found to be

$$d\psi_X(Y) = dL_{\psi(X)} \circ \frac{1 - e^{-\mathrm{ad}(H+X)}}{\mathrm{ad}\,(H + X)}(Y) \qquad (Y \in \mathfrak{n}_0), \tag{11}$$

$$d\varphi_n(dL_n X) = -\,\mathrm{Ad}\,(n)\,\mathrm{ad}\,H\,(X) \qquad (X \in \mathfrak{n}_0). \tag{12}$$

Let dX be the invariant measure on $\mathfrak{n}_0$ such that $(dX)_0 = (dn)_e$ (as differential forms). Then according to Lemma 1.9, the measure $\psi^*(dn)$ is

$$\psi^*(dn) = |\det A|\, dX,$$

where A is the endomorphism of $\mathfrak{n}_0$ given by

$$A : Y \to \frac{1 - e^{-\mathrm{ad}(H+X)}}{\mathrm{ad}\,(H + X)}(Y), \qquad Y \in \mathfrak{n}_0.$$

The absolute value comes from the fact that we are considering $\psi^*(dn)$ as a measure, not as a differential form. In the proof of Theorem 4.7, Chapter VI, it was established that

$$\det A = \prod_{\alpha \in P_+} \frac{1 - e^{-\alpha(H)}}{\alpha(H)}. \tag{13}$$

Next we compute the measure $\varphi^*(dX)$. From Lemma 1.9 and (12) we obtain

$$\varphi^*(dX) = |\det B|\, dn,$$

where B is the restriction of $\mathrm{Ad}_G\,(n)\,\mathrm{ad}_{\mathfrak{g}_0} H$ to $\mathfrak{n}_0$. Since $\det \mathrm{Ad}_G\,(n) = 1$ we obtain

$$\det B = \prod_{\alpha \in P_+} \alpha(H).$$

Using now $\xi^* = \varphi^* \circ \psi^*$ we find that

$$\xi^*(dn) = \prod_{\alpha \in P_+} |1 - e^{-\alpha(H)}|\, dn.$$

The lemma now follows from the formula

$$\int_N (f \circ \xi)\,(n)\, \xi^*(dn) = \int_N f(n)\, dn,$$

which is a special case of (2).

Proposition 1.14. *Let the notation be as in Prop.* 1.12, *but suppose now that G has finite center. Let*

$$D(a) = \prod_{\alpha \in P_+} \sinh \frac{\alpha(H)}{2} \qquad \text{if } H \in \mathfrak{h}_{\mathfrak{p}_0},\ a = \exp H.$$

Let $a \in A_{\mathfrak{p}}$ such that $D(a) \neq 0$. Then

$$c \mid D(a) \mid \int_{G/A_{\mathfrak{p}}} f(gag^{-1})\, dg^* = e^{\rho(\log a)} \int_{K \times N} f(kank^{-1})\, dk\, dn \tag{14}$$

for all $f \in C_c(G)$. The constant $c = 2^r$ (where r is the number of elements in P_+).

Proof. We must first verify that the integral on the left is defined. Let C denote the support of f and put

$$C^* = \{g^* \in G/A_{\mathfrak{p}} : gag^{-1} \in C\}.$$

Writing $g \in G$ as $g = kna_1$ ($k \in K$, $n \in N$, $a_1 \in A$), it is clear that $gag^{-1} \in C \Leftrightarrow knan^{-1}k^{-1} \in C \Rightarrow nan^{-1} \in KCK$. Now, since G has finite center, K is compact so if $gag^{-1} \in C$, then nan^{-1} lies in a compact subset of G. Using Cor. 4.8, Chapter VI, it follows that n lies in a compact subset of N. Hence kn lies in a compact subset of G so C^* is compact. Thus the left-hand side of (14) is finite.

Combining now Prop. 1.12 and Prop. 1.13, we find

$$\begin{aligned} \int_{G/A_{\mathfrak{p}}} f(gag^{-1})\, dg^* &= \int_{K \times N} f(knan^{-1}k^{-1})\, dk\, dn \\ &= \prod_{\alpha \in P_+} \mid 1 - e^{-\alpha(H)} \mid^{-1} \int_{K \times N} f(kank^{-1})\, dk\, dn. \end{aligned}$$

Since

$$\begin{aligned} \prod_{\alpha \in P_+} (1 - e^{-\alpha(H)}) &= \prod_{\alpha \in P_+} (e^{\frac{1}{2}\alpha(H)} - e^{-\frac{1}{2}\alpha(H)})\, e^{-\frac{1}{2}\alpha(H)} \\ &= cD(a)\, e^{-\rho(\log a)}, \end{aligned}$$

the proposition is proved.

This proposition has an interesting functional equation as a consequence.

Theorem 1.15. (Harish-Chandra) *Retaining the notation of Prop.* 1.11, *suppose G has finite center. Let W denote the Weyl group of G/K (Chapter VII,* §2), *and for* $s \in W$, $a = \exp H \in A_{\mathfrak{p}}$, *put* $a^s = \exp s(H)$. *Let* $f \in C_c(G)$ *such that* $f(kgk^{-1}) = f(g)$ *for all* $g \in G$, $k \in K$. *Then the function*

$$F_f(a) = e^{\rho(\log a)} \int_N f(an)\, dn, \qquad a \in A_{\mathfrak{p}},$$

satisfies the functional equation

$$F_f(a^s) = F_f(a), \qquad a \in A_{\mathfrak{p}},$$

for each $s \in W$.

Proof. For $g \in G$, let $a^{g^*} = gag^{-1}$. First assume $D(a) \neq 0$. Then we have by Prop. 1.14

$$F_f(a) = c\,|\,D(a)\,|\int_{G/A_{\mathfrak{p}}} f(a^{g^*})\, dg^*. \tag{15}$$

We shall now prove that the right-hand side of this relation is invariant under $a \to a^s$. For this purpose f can be arbitrary in $C_c(G)$. First, it follows from Cor. 2.11, Chapter VII, that the function

$$D(a)^2 = \prod_{\alpha, -\alpha \in P_+} \left|\,\sinh \frac{\alpha(H)}{2}\,\right|$$

is invariant under the Weyl group. Secondly, let $s \in W$ and select $u \in K$ such that s and $\mathrm{Ad}_G(u)$ coincide on $\mathfrak{h}_{\mathfrak{p}_0}$. Then $uA_{\mathfrak{p}}u^{-1} = A_{\mathfrak{p}}$ so the mapping $\varphi : gA_{\mathfrak{p}} \to ugu^{-1}A_{\mathfrak{p}}$ is a well-defined mapping of $G/A_{\mathfrak{p}}$ onto itself. Put $f^u(g) = f(ugu^{-1})$ and as in Lemma 1.8,

$$\bar{f}(g^*) = \int_{A_{\mathfrak{p}}} f(ga)\, da.$$

Then the mapping $f \to \bar{f}$ maps $C_c(G)$ onto $C_c(G/A_{\mathfrak{p}})$ and

$$\int_{G/A_{\mathfrak{p}}} \bar{f}(g^*)\, dg^* = \int_G f(g)\, dg = \int_G f^u(g)\, dg = \int_{G/A_{\mathfrak{p}}} \overline{f^u}(g^*)\, dg^*.$$

On the other hand,

$$\overline{f^u}(g^*) = \int_{A_{\mathfrak{p}}} f(ugau^{-1})\, da = \int_{A_{\mathfrak{p}}} f(ugu^{-1}a^s)\, da = \int_{A_{\mathfrak{p}}} f(ugu^{-1}a)\, da$$

so $\bar{f^u} = (\bar{f})^{\varphi^{-1}}$. This proves that $\varphi^*(dg^*) = dg^*$. Using now Prop. 1.12 we have

$$\int_{G/A_{\mathfrak{p}}} f((a^s)^{g^*})\, dg^* = \int_{G/A_{\mathfrak{p}}} f((a^s)^{(ugu^{-1})^*})\, dg^* = \int_{G/A_{\mathfrak{p}}} f(ugag^{-1}u^{-1})\, dg^*$$

$$= \int_{K\times N} f(uknan^{-1}k^{-1}u^{-1})\, dk\, dn = \int_{K\times N} f(knan^{-1}k^{-1})\, dk\, dn = \int_{G/A_{\mathfrak{p}}} f(a^{g^*})\, dg^*.$$

This proves $F_f(a^s) = F_f(a)$ for all a of a dense subset of $A_{\mathfrak{p}}$. By continuity it follows for all $a \in A$.

4. Integral Formulas for the Cartan Decomposition

Let $\mathfrak{g}_0$, $\mathfrak{k}_0$, $\mathfrak{p}_0$, $\mathfrak{h}_{\mathfrak{p}_0}$, G, K, and $A_{\mathfrak{p}}$ have the same meaning as in No. 3. We assume that G has finite center. This assumption is no restriction of generality for the space G/K, but has the convenient implication that K is now compact. As often before we identify $\mathfrak{p}_0$ with the tangent space $(G/K)_o$. Let dX and dg_K denote the volume elements of $\mathfrak{p}_0$ and G/K, respectively. Then $(dX)_0 = (dg_K)_o$, so we have from Lemma 1.9 and Theorem 4.1, Chapter IV,

$$\mathrm{Exp}^*(dg_K) = \det\left(\sum_0^\infty \frac{T_X^n}{(2n+1)!}\right) dX.$$

Since the mapping Exp is a diffeomorphism of $\mathfrak{p}_0$ onto G/K it follows from (2) that

$$\int_{G/K} f(g_K)\, dg_K = \int_{\mathfrak{p}_0} f(\mathrm{Exp}\, X) \det\left(\sum_0^\infty \frac{T_X^n}{(2n+1!)}\right) dX. \tag{16}$$

Let M denote the centralizer of $\mathfrak{h}_{\mathfrak{p}_0}$ in K and let us write Ad instead of Ad_G. Consider the mapping

$$\gamma : (kM, H) \to \mathrm{Ad}\,(k)\, H, \qquad k \in K,\ H \in \mathfrak{h}_{\mathfrak{p}_0},$$

of $(K/M) \times \mathfrak{h}_{\mathfrak{p}_0}$ onto $\mathfrak{p}_0$. Let $\mathfrak{m}_0$ denote the Lie algebra of M and let $\mathfrak{l}_0$ denote the orthogonal complement of $\mathfrak{m}_0$ in $\mathfrak{k}_0$ (with respect to the Killing form B of $\mathfrak{g}_0$). We identify $\mathfrak{l}_0$ with the tangent space $(K/M)_0$. Let $k_0 \in K$ and $H_0 \in \mathfrak{h}_{\mathfrak{p}_0}$. Then, as shown in Chapter VII, §3,

$$d\gamma_{(k_0M, H_0)}(d\tau(k_0)\, L, d\tau(H_0)\, H) = d\tau(\mathrm{Ad}\,(k_0)\, H_0)\, \mathrm{Ad}\,(k_0)\, ([L, H_0] + H) \tag{17}$$

for $L \in \mathfrak{l}_0$, $H \in \mathfrak{h}_{\mathfrak{p}_0}$. Now, $-B$ induces a positive definite metric on $\mathfrak{l}_0$ which gives rise to a K-invariant metric on K/M. Also B gives rise to a Euclidean metric on $\mathfrak{h}_{\mathfrak{p}_0}$. Let dk_M and dH denote the Riemannian measures on K/M and $\mathfrak{h}_{\mathfrak{p}_0}$ induced by these metrics. Then we derive from Lemma 1.9

$$\gamma^*(dX) = |\det A|\, dH\, dk_M,$$

where A denotes the linear mapping

$$A : (L, H) \to -\operatorname{ad} H_0(L) + H$$

of $\mathfrak{l}_0 \times \mathfrak{h}_{\mathfrak{p}_0}$ into $\mathfrak{p}_0$ and $|\det A|$ is evaluated by means of orthonormal bases. In order to compute this determinant we use Lemma 3.6 of Chapter VI and the notation adopted there. Since

$$\operatorname{ad} H_0(X_\alpha + \theta X_\alpha) = \alpha(H_0)\,(X_\alpha - \theta X_\alpha), \tag{18}$$

it follows that ad H_0 maps $\mathfrak{l}$ into $\mathfrak{q}$, $\mathfrak{l}_0$ into $\mathfrak{q}_0$, and, in terms of the bases $X_\alpha + \theta X_\alpha$ $(\alpha \in P_+)$ and $X_\alpha - \theta X_\alpha$ $(\alpha \in P_+)$ of $\mathfrak{l}$ and $\mathfrak{q}$, respectively, the determinant of ad H_0 is $\Pi_{\alpha\in P_+} \alpha(H_0)$. Now, let $B_\tau(X, Y) = -B(X, \tau Y)$ and $\|X\|^2 = B_\tau(X, X)$. Using the orthogonality of X_α and X_β for $\alpha + \beta \neq 0$ and Lemma 3.3, Chapter VI it is easily verified that

$$B_\tau(X_\alpha \pm \theta X_\alpha, X_\beta \pm \theta X_\beta) = 0 \qquad \text{for } \alpha \neq \beta \text{ in } P_+,$$
$$\|X_\alpha + \theta X_\alpha\| = \|X_\alpha - \theta X_\alpha\|.$$

Therefore the mapping ad $H_0 : \mathfrak{l} \to \mathfrak{q}$ has determinant of absolute value

$$\Big|\prod_{\alpha\in P_+} \alpha(H_0)\Big|$$

when expressed by means of *any* bases which are orthonormal with respect to B_τ. In particular we can use orthonormal bases of $\mathfrak{l}_0$ and $\mathfrak{q}_0$ with respect to $-B$ and B respectively. This shows that

$$\gamma^*(dX) = \Big|\prod_{\alpha\in P_+} \alpha(H)\Big|\, dk_M\, dH.$$

The set of points in $\mathfrak{h}_{\mathfrak{p}_0}$ where $\Pi_{\alpha\in P_+} \alpha(H)$ vanishes is a union of finitely many hyperplanes. Let $\mathfrak{h}'_{\mathfrak{p}_0}$ denote the complement of this set and let $\mathfrak{p}'_0 = \gamma(K/M \times \mathfrak{h}'_{\mathfrak{p}_0})$. Let M' denote the normalizer of $\mathfrak{h}_{\mathfrak{p}_0}$ in K. Then the factor group M'/M is the Weyl group W (Chapter VII, §2). Let w denote the order of W.

Lemma 1.16. *The mapping γ is a regular w-to-one mapping of $K/M \times \mathfrak{h}'_{\mathfrak{p}_0}$ onto $\mathfrak{p}'_0$.*

Proof. The regularity of γ is clear from (17) and (18). Now suppose $k_i \in K$, $H_i \in \mathfrak{h}'_{\mathfrak{p}_0}$ $(i = 1, 2)$ such that $\mathrm{Ad}(k_1) H_1 = \mathrm{Ad}(k_2) H_2$. Let $k = k_2^{-1}k_1$ so $\mathrm{Ad}(k) H_1 = H_2$. Then $\mathrm{Ad}(k)$ maps the centralizer of H_1 in $\mathfrak{p}_0$ onto the centralizer of H_2 in $\mathfrak{p}_0$. Owing to Cor. 3.7, Chapter VI, these centralizers coincide with $\mathfrak{h}_{\mathfrak{p}_0}$ and consequently $k \in M'$. Thus, if k runs through a complete set of representatives of M' mod M then $(k_1k^{-1}M, \mathrm{Ad}(k) H_1)$ runs through the subset $\gamma^{-1}(\mathrm{Ad}(k_1) H_1)$. This proves the lemma.

Remark. If C is a Weyl chamber in $\mathfrak{h}_{\mathfrak{p}_0}$, then γ is a diffeomorphism of $(K/M) \times C$ onto $\mathfrak{p}'_0$.

From Lemma 9.3, Chapter VII it is clear that the complement of $\mathfrak{p}'_0$ in $\mathfrak{p}_0$ has dimension $\leqslant \dim K/M + \dim \mathfrak{h}_{\mathfrak{p}_0} - 1 = \dim \mathfrak{p}_0 - 1$ and therefore has measure 0 with respect to dX. (Actually the dimension is $\leqslant \dim \mathfrak{p}_0 - 2$; this can be proved by the method of Theorem 3.3, Chapter VII.) We have therefore

$$\int_{\mathfrak{p}_0} F(X)\, dX = \frac{1}{w} \int_{\mathfrak{h}_{\mathfrak{p}_0}} \Big|\prod_{\alpha \in P_+} \alpha(H)\Big| \int_{K/M} F(\mathrm{Ad}(k) H)\, dk_M\, dH$$

for $F \in C_c(\mathfrak{p}_0)$. We normalize the measures dk on K and dg on G by

$$\int_K dk = 1, \int_G f(g)\, dg = \int_{G/K} \Big(\int_K f(gk)\, dk\Big)\, dg_K, \qquad f \in C_c(G). \quad (19)$$

Then

$$\int_{K/M} F(\mathrm{Ad}(k) H)\, dk_M = \mathrm{vol}\,(K/M) \int_K F(\mathrm{Ad}(k) H)\, dk,$$

so

$$\int_G f(g)\, dg = \int_{\mathfrak{p}_0} \Big(\int_K f(\exp X\, k)\, dk\Big) \det\Big(\sum_0^\infty \frac{T_X^n}{(2n+1)!}\Big)\, dX$$

$$= \frac{1}{w} \mathrm{vol}\,(K/M) \int_{\mathfrak{h}_{\mathfrak{p}_0}} \Big|\prod_{\alpha \in P_+} \alpha(H)\Big| \prod_{\alpha \in P_+} \frac{\sinh \alpha(H)}{\alpha(H)} \iint_{K \times K} f(k_1 \exp H k_1^{-1} k)\, dk_1\, dk\, dH.$$

This proves the following proposition.

Proposition 1.17. *Let (G, K) be a Riemannian symmetric pair of the noncompact type and assume G has finite center. Let dg_K denote the volume* *element on G/K and let the measures dk and dg be normalized by* (19). *Then if $f \in C_c(G)$*

$$\int_G f(g)\,dg = \frac{1}{w}\,\mathrm{vol}\,(K/M)\int_{\mathfrak{h}_{\mathfrak{p}_0}}\Big|\prod_{\alpha\in P_+}\sinh\alpha(H)\Big|\,dH\int\!\!\int_{K\times K} f(k_1\exp Hk_2)\,dk_1dk_2.$$

Here vol (K/M) denotes the volume of K/M in the K-invariant metric induced by the restriction of $-B$ to $\mathfrak{l}_0$. Moreover, w denotes the order of the Weyl group.

5. The Compact Case

Let $\mathfrak{g}$ denote the complexification of $\mathfrak{g}_0$ and as in Chapter VII, let $\mathfrak{u}$ denote the compact real form $\mathfrak{k}_0 + i\mathfrak{p}_0$. In general, if $\mathfrak{e}_0$ is a subspace of $\mathfrak{p}_0$ we write $\mathfrak{e}_*$ for the subspace $i\mathfrak{e}_0$ of $i\mathfrak{p}_0$. Thus

$$\mathfrak{u} = \mathfrak{k}_0 + \mathfrak{p}_*,$$

and $\mathfrak{h}_{\mathfrak{p}_*} = i\mathfrak{h}_{\mathfrak{p}_0}$ is a maximal abelian subspace of $\mathfrak{p}_*$. Let (U, K_1) be a symmetric pair such that U and K_1 have Lie algebras $\mathfrak{u}$ and $\mathfrak{k}_0$, respectively. Let M_1 denote the centralizer of $\mathfrak{h}_{\mathfrak{p}_*}$ in K_1. Writing Ad for Ad_U we consider, as in Chapter VII, §3, the mapping

$$\beta : (kM_1, H) \to \mathrm{Ad}\,(k)\,H$$

of $(K_1/M_1) \times \mathfrak{h}_{\mathfrak{p}_*}$ onto $\mathfrak{p}_*$ and the mapping $\Phi = \mathrm{Exp} \circ \beta$ of $(K_1/M_1) \times \mathfrak{h}_{\mathfrak{p}_*}$ onto U/K_1. Let du, dk, and dm denote the invariant measures on K_1, U, and M_1, respectively, normalized by

$$\int_{K_1} dk = \int_U du = \int_{M_1} dm = 1. \tag{20}$$

Now U/K_1, K_1/M_1, $\mathfrak{p}_*$, and $\mathfrak{h}_{\mathfrak{p}_*}$ have invariant metrics induced by the Killing form of $\mathfrak{u}$. Let the corresponding Riemannian measures be denoted by du_K, dk_M, dX, and dH. Then we have, just as for the noncompact case,

$$\mathrm{Exp}^*(du_K) = \Big|\det\Big(\sum_0^\infty \frac{T_X^n}{(2n+1)!}\Big)\Big|\,dX,$$

$$\beta^*(dX) = \Big|\prod_{\alpha\in P_+}\alpha(H)\Big|\,dk_M dH.$$

It follows that

$$\Phi^*(du_K) = \left| \prod_{\alpha \in P_+} \sin \alpha(iH) \right| dk_M dH.$$

The group $A_* = \exp \mathfrak{h}_{\mathfrak{p}_*}$ is a closed subgroup of U (see proof of Lemma 6.3, Chapter V), hence compact. Let da denote the invariant measure on A_* normalized by $\int_{A_*} da = 1$. Then

$$dH = c \exp^*(da) \tag{21}$$

where c is a constant. If the element $H \in \mathfrak{h}_{\mathfrak{p}_*}$ satisfies $\exp H = e$, then $\alpha(H) \in 2\pi i \boldsymbol{Z}$ for all $\alpha \in P_+$. We can therefore define a function $D_*(a)$ on A_* by the condition

$$D_*(\exp H) = \prod_{\alpha \in P_+} \sin \alpha(iH). \tag{22}$$

The mapping

$$\psi : (kM_1, a) \rightarrow kaK_1 \qquad (k \in K_1, a \in A_*),$$

of $((K_1/M_1) \times A_*)$ onto U/K_1 then satisfies

$$\psi^*(du_K) = c \mid D_*(a) \mid dk_M da. \tag{23}$$

Let A'_* denote the set of points $a \in A_*$ for which $D_*(a) \neq 0$ and let $(U/K_1)_r$ be defined as in Chapter VII, §3. The space $J = K_1 \cap A_*$ is compact and discrete, hence finite. Let j denote the number of elements in J.

Lemma 1.18. *The mapping ψ is a regular wj-to-one mapping of $(K_1/M_1) \times A'_*$ onto $(U/K_1)_r$.*

Proof. In view of Lemma 7.1, Chapter VII, it only remains to prove that each point in $(U/K_1)_r$ has exactly wj pre-images. Suppose $a_1, a_2 \in A'_*$, $k_1, k_2 \in K_1$ such that $\psi(k_1M_1, a_1) = \psi(k_2M_1, a_2)$. On putting $k = k_2^{-1}k_1$, we have $ka_1 = a_2k'$ for a suitable $k' \in K_1$. We apply the involutive automorphism θ, eliminate k' and obtain $ka_1^2k^{-1} = a_2^2$. This relation implies that $\operatorname{Ad}(k) Z_1 = Z_2$ if

$$Z_i = \{X \in \mathfrak{p}_* : \operatorname{Ad}(a_i^2) X = X\}$$

for $i = 1, 2$. Let $H \in \mathfrak{h}_{\mathfrak{p}_*}$, $X \in \mathfrak{p}_*$. Writing

$$X = H_1 + \sum_{\alpha \in P_+} a_\alpha(X_\alpha - \theta X_\alpha), \qquad H_1 \in \mathfrak{h}_{\mathfrak{p}_*}, \quad a_\alpha \in \boldsymbol{C},$$

according to Lemma 3.6, Chapter VI, we have

$$\begin{aligned}\operatorname{Ad}(\exp H)X &= e^{\operatorname{ad}H}X\\ &= H_1 + \sum_{\alpha\in P_+} a_\alpha \cos\alpha(iH)\,(X_\alpha - \theta X_\alpha) + \sum_{\alpha\in P_+} a_\alpha \sin\alpha(iH)\,(X_\alpha + \theta X_\alpha).\end{aligned}$$

In particular $\operatorname{Ad}(a_i^2)X = X$ if and only if $X \in \mathfrak{h}_{\mathfrak{p}_*}$, so

$$Z_i = \mathfrak{h}_{\mathfrak{p}_*} \qquad (i = 1, 2).$$

This proves that k lies in the normalizer M_1' of $\mathfrak{h}_{\mathfrak{p}_*}$ in K_1. We write

$$a_2 = (ka_1k^{-1})\,kk'^{-1}. \tag{24}$$

Now let $s \in W$ and $a \in A_*$. If m is any element in M_1' such that s and $\operatorname{Ad}(m)$ coincide on $\mathfrak{h}_{\mathfrak{p}_*}$ we write $a^s = mam^{-1}$. It is clear that $|D_*(a^s)| = |D_*(a)|$. Let $b \in J$ and select $H \in \mathfrak{h}_{\mathfrak{p}_*}$ such that $b = \exp H$. Since $b^2 = e$ it is clear that $\alpha(H) \in \pi i\boldsymbol{Z}$ for all $\alpha \in P_+$ and consequently $|D_*(ab)| = |D_*(a)|$ for $a \in A_*$. Formula (24) shows that $a_2 = a_1{}^s b$ where $b \in J$, $s \in W$.

On the other hand, if b runs through the set J and k runs through a complete set of representatives of M_1' mod M_1 then the elements

$$(k_1k^{-1}M_1, ka_1k^{-1}b) \in K_1/M_1 \times A_*'$$

are all different. Hence they make up the complete inverse image $\psi^{-1}(k_1a_1K_1)$. This proves the lemma.

Now, A_*' fills up A_* except for a set of measure 0 and $(U/K_1)_r$ fills up U/K_1 except for a set of measure 0. We derive therefore from Lemma 1.18 and formulas (20)-(23),

$$\int_{U/K_1} f(u_K)\,du_K = c\,\frac{1}{wj}\int_{A_*} |D_*(a)|\,da \int_{K_1/M_1} f(kaK_1)\,dk_M, \tag{25}$$

$$\operatorname{vol}(U/K_1)\int_U F(u)\,du = \int_{U/K_1}\Big(\int_{K_1} F(uk)\,dk\Big)\,du_K,$$

$$\operatorname{vol}(K_1/M_1)\int_{K_1} g(k)\,dk = \int_{K_1/M_1}\Big(\int_{M_1} g(km)\,dm\Big)\,dk_M,$$

if f, F, and g are continuous. Hence

$$\operatorname{vol}(U/K_1)\int_U F(u)\,du = \int_{U/K_1}\Big(\int_{K_1} F(uk_1)\,dk_1\Big)\,du_K$$

$$= c\,\frac{1}{wj}\int_{A_*} |\,D_*(a)\,|\,da\int_{K_1/M_1}\Big(\int_{K_1} F(kak_1)\,dk_1\Big)\,dk_M$$

$$= c\,\frac{1}{wj}\operatorname{vol}(K_1/M_1)\int_{A_*} |\,D_*(a)\,|\,da\iint_{K\times K} F(k_1ak_2)\,dk_1dk_2.$$

Putting here $F \equiv 1$ we obtain the relation

$$c\operatorname{vol}(K_1/M_1) = wj\operatorname{vol}(U/K_1)$$

and the following proposition.

Proposition 1.19. *Let (U, K_1) be a Riemannian symmetric pair of the compact type.* Let *du, dk, and da denote the invariant measures on U, K_1, and A_*, respectively, normalized by*

$$\int_U du = \int_{K_1} dk = \int_{A_*} da = 1.$$

Then

$$\int_{A_*} |\,D_*(a)\,|\,da\int_U F(u)\,du = \int_{A_*} |\,D_*(a)\,|\,da\int_K\int_K F(k_1ak_2)\,dk_1dk_2$$

for all $F \in C(U)$.

§ 2. Invariant Differential Operators

1 Generalities. The Laplace-Beltrami Operator

Let S be a topological space and $f : S \to \boldsymbol{R}$ a continuous function on S, i.e. $f \in C(S)$. Let Φ be a homeomorphism of S onto itself. As in Chapter I, §3, No. 2, we shall often write f^Φ for the composite function $f \circ \Phi^{-1}$ and if A is a mapping of $C(S)$ into itself, A^Φ shall denote the mapping $f \to (Af^{\Phi^{-1}})^\Phi$ of $C(S)$ into itself. The function f (the mapping A) is said to be invariant under Φ if $f^\Phi = f$, ($A^\Phi = A$). It is easy to verify $f^{\Phi\Psi} = (f^\Psi)^\Phi$, $A^{\Phi\Psi} = (A^\Psi)^\Phi$ if Φ and Ψ are two arbitrary homeomorphisms of S. The value of the function Af at $p \in S$ will often be denoted by $[Af]\,(p)$.

Let M be a manifold of dimension m. If (φ, U) is a local chart on M and $f \in C^\infty(M)$ we shall sometimes write f^* for the composite function $f \circ \varphi^{-1}$, defined on $\varphi(U)$. Let ∂_i stand for the partial differentiation $\partial/\partial x_i$ $(1 \leqslant i \leqslant m)$ and if $\alpha = (\alpha_1, \ldots, \alpha_m)$ is an m-tuple of indices $\alpha_i \geqslant 0$ we put $D^\alpha = \partial_1^{\alpha_1} \ldots \partial_m^{\alpha_m}$.

Definition. A linear transformation $D : C_c^\infty(M) \to C_c^\infty(M)$ is called a *differential operator* on M if the following condition is satisfied: For each $p \in M$ and each local chart (φ, U) around p there exists a finite set of functions $a_\alpha \in C^\infty(U)$ such that for each $f \in C_c^\infty(M)$ with support contained in U,

$$[Df](q) = \sum_\alpha a_\alpha(q) [D^\alpha f^*](\varphi(q)) \qquad \text{if } q \in U,$$

$$[Df](q) = 0 \qquad \text{if } q \notin U.$$

A vector field on M is clearly a differential operator on M; a function $f \in C^\infty(M)$ can also be regarded as a differential operator $F \to fF$ $(F \in C_c^\infty(M))$ on M.

A differential operator D on M has *local character* in the sense that

$$\text{support } Df \subset \text{support } f \tag{1}$$

for each $f \in C_c^\infty(M)$. Secondly, D is continuous with respect to the topology of $C_c^\infty(M)$ based on uniform convergence of sequences of functions and their successive derivatives† (L. Schwartz [1], I, p. 67). A differential operator D can be extended to a linear mapping $D : C^\infty(M) \to C^\infty(M)$ as follows: Let $f \in C^\infty(M)$ and $p \in M$; select $f_c \in C_c^\infty(M)$ such that $f = f_c$ in a neighborhood of p and define $[Df](p) = [Df_c](p)$. Then (1) holds for all $f \in C^\infty(M)$. If Φ is a diffeomorphism of M then D^Φ is a differential operator as well as D and the mapping $D \to D^\Phi$ is an automorphism of the algebra of all differential operators on M.

A pseudo-Riemannian manifold M always possesses a differential operator of particular interest, the so-called *Laplace-Beltrami operator* which we shall now define. Let g denote the pseudo-Riemannian structure on M and let $\varphi : q \to (x_1(q), \ldots, x_m(q))$ be a coordinate system valid on an open set $U \subset M$. As customary we define the functions g_{ij}, g^{ij}, $\bar{g}$ on U by

$$g_{ij} = g\left(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}\right), \qquad \sum_j g_{ij} g^{jk} = \delta_{ik},$$

$$\bar{g} = |\det(g_{ij})|.$$

† It can be shown (see, e.g., Helgason [3], p. 242) that these two properties characterize differential operators.

Each function $f \in C^\infty(M)$ gives rise to a vector field *grad* f (gradient of f) on M whose restriction to U is given by

$$\operatorname{grad} f = \sum_{i,j} g^{ij} \frac{\partial f}{\partial x_i} \frac{\partial}{\partial x_j}. \tag{2}$$

It is in fact obvious that the expression on the right-hand side is independent of the particular coordinate system used. On the other hand, if X is a vector field on M the *divergence* of X is the function on M which on U is given by

$$\operatorname{div} X = \frac{1}{\sqrt{\bar{g}}} \sum_i \frac{\partial}{\partial x_i} (\sqrt{\bar{g}}\, X_i) \tag{3}$$

if $X = \Sigma X_i \, \partial/\partial x_i$ on U. Again the right-hand side can be shown to be invariant under coordinate changes. The Laplace-Beltrami operator Δ is now defined by

$$\Delta f = \operatorname{div} \operatorname{grad} f, \qquad f \in C^\infty(M).$$

In terms of local coordinates we have

$$\Delta f = \frac{1}{\sqrt{\bar{g}}} \sum_k \partial_k \Big(\sum_i g^{ik} \sqrt{\bar{g}}\, \partial_i f\Big), \tag{4}$$

which shows that Δ is a differential operator on M.

For a coordinate-free expression for grad f and div X see the exercises following this chapter.

Proposition 2.1. *Let M be a pseudo-Riemannian manifold, Δ the Laplace-Beltrami operator on M.*

(i) *Δ is symmetric, that is*

$$\int_M u(x)\,[\Delta v]\,(x)\,dx = \int_M [\Delta u]\,(x)\,v(x)\,dx, \qquad u \in C_c^\infty(M),\ v \in C^\infty(M),$$

if dx is the Riemannian measure on M.

(ii) *Let Φ be a diffeomorphism of M. Then Φ leaves Δ invariant if and only if it is an isometry.*

Proof. (i) Let X be any vector field on M. Then we have from the definitions (2), (3),

$$\operatorname{div}(uX) = u(\operatorname{div} X) + Xu,$$

$$\operatorname{grad} u\,(v) = \operatorname{grad} v\,(u).$$

Consequently

$$u\Delta v - v\Delta u = \operatorname{div}(u \operatorname{grad} v) - \operatorname{grad} v(u) \\ - \operatorname{div}(v \operatorname{grad} u) + \operatorname{grad} u(v),$$

so

$$u\Delta v - v\Delta u = \operatorname{div}(u \operatorname{grad} v - v \operatorname{grad} u).$$

It suffices therefore to prove that

$$\int_M (\operatorname{div} X)\, dx = 0$$

for any vector field X on M vanishing outside a compact subset. Using partition of unity we may even assume that X vanishes outside a coordinate neighborhood U. Writing $X = \sum_i X_i\, \partial/\partial x_i$ on U we have

$$\int_M \operatorname{div} X\, dx = \int \sum_{i=1}^m \frac{\partial}{\partial x_i}(\sqrt{\bar{g}}\, X_i)\, dx_1 \ldots dx_m = 0,$$

which proves (i).

For (ii) let $p \in M$ and let (ψ, V) be a local chart around p. Then $(\psi \circ \Phi^{-1}, \Phi(V))$ is a local chart around $\Phi(p)$. For $x \in V$, let $y = \Phi(x)$ and

$$\psi(x) = (x_1, \ldots, x_m), \qquad x \in V,$$

$$(\psi \circ \Phi^{-1})(y) = (y_1, \ldots, y_m), \qquad y \in \Phi(V).$$

Then

$$x_i(x) = y_i(\Phi(x)), \qquad d\Phi_x \left(\frac{\partial}{\partial x_i}\right)_x = \left(\frac{\partial}{\partial y_i}\right)_{\Phi(x)}, \qquad 1 \leqslant i \leqslant m.$$

For each function $f \in C^\infty(M)$

$$[(\Delta f)^{\Phi^{-1}}](x) = [\Delta f](\Phi(x)) = \frac{1}{\sqrt{\bar{g}(y)}} \sum_k \frac{\partial}{\partial y_k} \left(\sum_i g^{ik}(y) \sqrt{\bar{g}}(y) \frac{\partial f^*}{\partial y_i}\right), \tag{5}$$

$$[\Delta f^{\Phi^{-1}}](x) = \frac{1}{\sqrt{\bar{g}(x)}} \sum_k \frac{\partial}{\partial x_k} \left(\sum_i g^{ik}(x) \sqrt{\bar{g}}(x) \frac{\partial (f \circ \Phi)^*}{\partial x_i}\right). \tag{6}$$

Owing to the choice of coordinates we have

$$\frac{\partial f^*}{\partial y_i} = \frac{\partial (f \circ \Phi)^*}{\partial x_i}, \qquad \frac{\partial^2 f^*}{\partial y_i \partial y_j} = \frac{\partial^2 (f \circ \Phi)^*}{\partial x_i \partial x_j} \qquad (1 \leqslant i, j \leqslant n).$$

Now, if Φ is an isometry, then $g_{ij}(x) = g_{ij}(y)$ for all $1 \leqslant i, j \leqslant n$. Thus the right-hand sides of (5) and (6) coincide and $\Delta^{\Phi} = \Delta$. On the other hand, if (5) and (6) agree then we obtain by equating coefficients, $g^{ij}(x) = g^{ij}(y)$, which shows that Φ is an isometry.

2. Invariant Differential Operators on Reductive Coset Spaces

Let G be a connected Lie group and H a closed subgroup. Let $\mathfrak{g}$ and $\mathfrak{h}$ denote the corresponding Lie algebras, $\mathfrak{h} \subset \mathfrak{g}$. We describe in Lie algebra terms the set $\boldsymbol{D}(G/H)$ of differential operators on G/H which are invariant under the action of G. To be precise, a differential operator D on G/H belongs to $\boldsymbol{D}(G/H)$ if and only if $D^{\tau(g)} = D$ for all $g \in G$. It will be of great convenience to restrict somewhat the generality of the coset space G/H, and consider only the so-called *reductive coset spaces* which we shall now define.

Definition. The coset space G/H is called *reductive* if there exists a subspace $\mathfrak{m} \subset \mathfrak{g}$ such that $\mathfrak{h} + \mathfrak{m} = \mathfrak{g}$ (direct sum) and $\mathrm{Ad}_G(H)\, \mathfrak{m} \subset \mathfrak{m}$.

If (G, H) is a symmetric pair, the coset space G/H is reductive because the eigenspace for the eigenvalue -1 of the involutive automorphism of $\mathfrak{g}$ can serve as the subspace $\mathfrak{m}$.

If H is compact, or more generally, if $\mathrm{Ad}_G(H)$ is compact, then G/H is reductive. This is seen by considering a strictly positive definite bilinear form on $\mathfrak{g} \times \mathfrak{g}$ invariant under $\mathrm{Ad}_G(H)$ and taking for $\mathfrak{m}$ the orthogonal complement of $\mathfrak{h}$ in $\mathfrak{g}$.

Reductive coset spaces have been studied by Nomizu [2, 3]. From a differential geometric point of view they form a natural class of spaces because they can essentially be characterized by the fact that they admit an invariant affine connection. For an example of a nonreductive coset space see the exercises following this chapter.

Let G/H be a reductive coset space and as usual, let π denote the natural mapping of G onto G/H and put $o = \pi(e)$. Let $C_0^{\infty}(G)$ denote set of C^{∞} functions on G which are constant on each coset gH. Then the mapping $f \to \tilde{f}$ where $\tilde{f} = f \circ \pi$ is an isomorphism of the algebra $C^{\infty}(G/H)$ onto $C_0^{\infty}(G)$. Let us write $\boldsymbol{D}(G)$ instead of $\boldsymbol{D}(G/e)$ and let $\boldsymbol{D}_0(G)$ denote the set of operators in $\boldsymbol{D}(G)$ which are invariant under all right translations from H; in other words a differential operator D on G belongs to $\boldsymbol{D}_0(G)$ if and only if $D^{L(g)} = D$ and $D^{R(h)} = D$ for all $g \in G$ and all $h \in H$. Each $D \in \boldsymbol{D}_0(G)$ leaves the subspace $C_0^{\infty}(G)$ invariant; let D_0 denote its restriction to $C_0^{\infty}(G)$.

Lemma 2.2. *The algebra $D(G/H)$ is isomorphic to the algebra of restrictions $\{D_0 : D \in D_0(G)\}$.*

Proof. To each $D \in D_0(G)$ we associate the linear transformation E of $C^\infty(G/H)$ determined by the requirement $(Ef)^\sim = D_0\tilde{f}$ for $f \in C^\infty(G/H)$. Naturally E only depends on the restriction D_0. We shall prove that this mapping $\mu : D_0 \to E$ is an isomorphism onto $D(G/H)$. Firstly, if $g \in G$,

$$(E^{\tau(g)}f)^\sim = ((Ef^{\tau(g^{-1})})^{\tau(g)})^\sim = ((Ef^{\tau(g^{-1})})^\sim)^{L(g)} = (D_0\tilde{f}^{L(g^{-1})})^{L(g)},$$

so $E^{\tau(g)} = E$. To see that E is a differential operator, let $U_{\mathfrak{m}}$ and $U_{\mathfrak{h}}$ be open neighborhoods of 0 in $\mathfrak{m}$ and $\mathfrak{h}$, respectively, such that the mapping $(X, T) \to \exp X \exp T$ is a diffeomorphism of $U_{\mathfrak{m}} \times U_{\mathfrak{h}}$ onto an open neighborhood of e in G and such that π is one-to-one on the set $\exp U_{\mathfrak{m}}$ (Lemmas 2.4 and 4.1, Chapter II). Let $X_1, ..., X_r$ be a basis of $\mathfrak{m}$, $X_{r+1}, ..., X_n$ a basis of $\mathfrak{h}$. Then the mapping

$$\exp(x_1X_1 + ... + x_rX_r)\exp(x_{r+1}X_{r+1} + ... + x_nX_n) \to (x_1, ..., x_n) \qquad (7)$$

is a coordinate system on the neighborhood $\exp U_{\mathfrak{m}} \exp U_{\mathfrak{h}}$ of e in G and the mapping

$$\pi(\exp(x_1X_1 + ... + x_rX_r)) \to (x_1, ..., x_r) \qquad (8)$$

is a coordinate system on the neighborhood $\pi(\exp U_{\mathfrak{m}})$ of o in G/H. If D is expressed by means of the coordinates (7), it is clear that $\tilde{f}$ and $D\tilde{f}$ are independent of $(x_{r+1}, ..., x_n)$. Since $(Ef)^\sim = D\tilde{f}$, it follows that E can be expressed (on $\pi(\exp U_{\mathfrak{m}})$) by means of partial derivatives with respect to the variables $(x_1, ..., x_r)$. Thus $E \in D(G/H)$. It is obvious that the mapping μ is one-to-one, linear, and preserves multiplication. To see that the image under μ is all of $D(G/H)$, let $E \in D(G/H)$. Then there exists a polynomial P in r variables such that

$$[Ef](o) = [P(\partial_1, ..., \partial_r)\tilde{f}(\exp(x_1X_1 + ... + x_rX_r))](0) \qquad (9)$$

for $f \in C^\infty(G/H)$. Since $E^{\tau(g)} = E$ for $g \in G$ we have

$$[Ef](p) = [P(\partial_1, ..., \partial_r)\tilde{f}(g\exp(x_1X_1 + ... + x_rX_r))](0) \qquad (10)$$

if $p = g \cdot o$, and if $h \in H$

$$\begin{aligned}[Ef](o) &= [Ef^{\tau(h^{-1})}](o)\\ &= [P({}_1\partial, ..., \partial_r)\tilde{f}(h\exp(x_1X_1 + ... + x_rX_r)h^{-1})](0)\\ &= [P(\partial_1, ..., \partial_r)\tilde{f}(\exp(\mathrm{Ad}(h)(x_1X_1 + ... + x_rX_r)))](0).\end{aligned}$$

Comparing with (9) we find

$$P(X_1, ..., X_r) = P(\mathrm{Ad}\,(h)\, X_1, ..., \mathrm{Ad}\,(h)\, X_r), \qquad h \in H, \qquad X_1, ..., X_r \in \mathfrak{m}, \quad (11)$$

that is, P is invariant under $\mathrm{Ad}_G\,(H)$. Now, for $g \in G$, the mapping

$$g \exp\,(x_1X_1 + ... + x_nX_n) \to (x_1, ..., x_n)$$

is a coordinate system valid in a neighborhood of $g \in G$ and the operator D defined by

$$[DF]\,(g) = [P(\partial_1, ..., \partial_r)\, F(g \exp\,(x_1X_1 + ... + x_nX_n))]\,(0) \qquad (12)$$

for $F \in C^\infty(G)$, is a differential operator on G. If $x \in G$ then

$$[D^{L(x)}F]\,(g) = [DF^{L(x^{-1})}]\,(x^{-1}g)$$

$$= [P(\partial_1, ..., \partial_r)\, F^{L(x^{-1})}(x^{-1}g \exp\,(x_1X_1 + ... + x_nX_n))]\,(0)$$

so $D^{L(x)} = D$. Using (11), one obtains in the same way, $D^{R(h)} = D$ for $h \in H$. Finally, we have from (10) and (12), $(Ef)^\sim = D\tilde{f}$ which finishes the proof of the lemma.

Now, each $X \in \mathfrak{g}$ determines uniquely a left invariant vector field on G whose value at e is X. This vector field will also be denoted by X. Then if $g \in G$, $\mathrm{Ad}\,(g)\, X$ as a vector field is $(X^{L(g^{-1})})^{R(g)}$. Also

$$[Xf]\,(g) = \left[\frac{d}{dt} f(g \exp tX)\right](0). \qquad (13)$$

More generally, formula (10) shows (for $H = e$) that each $D \in \boldsymbol{D}(G)$ can be written

$$[Df]\,(g) = [P(\partial_1, ..., \partial_n) f(g \exp\,(x_1X_1 + ... + x_nX_n))]\,(0). \qquad (14)$$

Using (12), Chapter II, §1, it follows that D is a linear combination $\Sigma_M\, a_M X(M)$ $(a_M \in \boldsymbol{R})$ and therefore the present definition of $\boldsymbol{D}(G)$ coincides with the one given in Chapter II, §1.

Let V be a vector space of finite dimension over a field K. Let $T(V)$ denote the tensor algebra over V and let I denote the two-sided ideal in $T(V)$ generated by the set of all elements of the form $X \otimes Y - Y \otimes X$, $X, Y \in V$. The factor algebra $S(V) = T(V)/I$ is called the *symmetric algebra* over V. If a basis $X_1, ..., X_n$ of V is chosen, $S(V)$ can be identified with the algebra of polynomials in these basis elements. Each endomorphism A of V induces a homomorphism $P \to A \cdot P$ of $S(V)$ given by $(A \cdot P)\,(X_1, ..., X_n) = P(AX_1, ..., AX_n)$. This homomorphism is independent of the choice of basis of V as is easily seen by considering P as a function on the dual space $V^\wedge$ of V. We have also $A_1A_2 \cdot P = A_1 \cdot (A_2 \cdot P)$ if A_1 and A_2 are two endomorphisms of V. The polynomial P is said to be invariant under A if $A \cdot P = P$.

If X and Y are two elements in $\mathfrak{g}$, let XY denote their product in $S(\mathfrak{g})$ and let $X \cdot Y$ denote their product in $\boldsymbol{D}(G)$, (or $U(\mathfrak{g})$, the universal enveloping algebra of $\mathfrak{g}$). According to Prop. 1.9 (a) Chapter II, there exists a linear one-to-one mapping (cf. Harish-Chandra [4], p. 192) λ of $S(\mathfrak{g})$ onto $\boldsymbol{D}(G)$ such that

$$\lambda(X_1^{m_1} \dots X_n^{m_n}) = m_1! \dots m_n! \, X(M),$$

where M denotes the n-tuple $(m_1, \dots, m_n)$ and $X(M)$ is the coefficient of $t_1^{m_1} \dots t_n^{m_n}$ in

$$((m_1 + \dots + m_n)!)^{-1}(t_1X_1 + \dots + t_nX_n)^{m_1+\dots+m_n}.$$

Then

$$\lambda(Y_1Y_2 \dots Y_p) = \frac{1}{p!} \sum_\sigma Y_{\sigma(1)} \cdot Y_{\sigma(2)} \dots Y_{\sigma(p)}, \tag{15}$$

where $Y_1, \dots, Y_p$ are arbitrary elements from the basis $X_1, \dots, X_n$ of $\mathfrak{g}$ and σ runs over the group of all permutations of the set $\{1, 2, \dots, p\}$. It is clear that (15) remains true if we replace Y_j by any linear combination $\sum_i a_{ij}X_i$. Thus (15) holds for arbitrary elements $Y_1, \dots, Y_p$ from $\mathfrak{g}$; in particular, λ is independent of the choice of basis of $\mathfrak{g}$. Somewhat loosely, we shall refer to the mapping λ as *symmetrization.* Comparing (14) above with (12) in Chapter II, §1, we obtain

$$[\lambda(P)f](g) = [P(\partial_1, \dots, \partial_n)f(g \exp(x_1X_1 + \dots + x_nX_n))](0) \tag{16}$$

for each $P \in S(\mathfrak{g})$, $f \in C^\infty(G)$.

For each $g \in G$, the automorphism $\operatorname{Ad}(g)$ of $\mathfrak{g}$ extends uniquely to an automorphism of $\boldsymbol{D}(G)$. Let this extension also be denoted by $\operatorname{Ad}(g)$. Since $\operatorname{Ad}(g)\, X = (X^{L(g)})^{R(g^{-1})} = X^{R(g^{-1})}$, we have

$$\operatorname{Ad}(g)\, D = D^{R(g^{-1})}, \qquad D \in \boldsymbol{D}(G). \tag{17}$$

For each $X \in \mathfrak{g}$, the mapping $D \to X \cdot D - D \cdot X$ is a derivation of $\boldsymbol{D}(G)$ which will be denoted $\operatorname{ad} X$. In terms of a basis $X_1, \dots, X_n$ of $\mathfrak{g}$, D can be written as a linear combination $\sum a_{e_1 \dots e_n} X_1^{e_1} \cdot X_2^{e_2} \dots X_n^{e_n}$ with constant coefficients. The largest integer $e_1 + \dots + e_n$ for which $a_{e_1 \dots e_n} \neq 0$ is called the *order* (or *degree*) of D. The order is independent of the choice of basis. Clearly the order of ad $X(D)$ is $\leqslant$ order of D. Consequently all the operators $(\operatorname{ad} X)^n(D)$, $n = 0, 1, \ldots$, lie in a finite-dimensional subspace of $\boldsymbol{D}(G)$ and $e^{\operatorname{ad} X}(D)$ can be defined by means of the series

$$e^{\operatorname{ad} X}(D) = \sum_{n=0}^{\infty} \frac{1}{n!} (\operatorname{ad} X)^n (D).$$

Since ad X is a derivation we have for $D_1, D_2 \in \boldsymbol{D}(G)$

$$(\operatorname{ad} X)^k (D_1 \cdot D_2) = \sum_{0 \leqslant i, j, i+j=k} \frac{k!}{i!j!} (\operatorname{ad} X)^i (D_1) \cdot (\operatorname{ad} X)^j (D_2),$$

so

$$e^{\operatorname{ad} X} (D_1 \cdot D_2) = e^{\operatorname{ad} X} (D_1) \cdot e^{\operatorname{ad} X} (D_2).$$

The automorphisms $\operatorname{Ad}(\exp X)$ and $e^{\operatorname{ad} X}$ of $\boldsymbol{D}(G)$ coincide on $\mathfrak{g}$. From the uniqueness mentioned above

$$\operatorname{Ad}(\exp X) D = e^{\operatorname{ad} X} (D), \qquad X \in \mathfrak{g},\ D \in \boldsymbol{D}(G). \tag{18}$$

Lemma 2.3. *Let $X \in \mathfrak{g}$, $D \in \boldsymbol{D}(G)$. Then $D \cdot X = X \cdot D$ if and only if $D^{R(\exp tX)} = D$ for all $t \in \boldsymbol{R}$.*

Proof. The operators $e^{\operatorname{ad}(tX)}D$ $(t \in \boldsymbol{R})$ lie in a finite-dimensional subspace of $\boldsymbol{D}(G)$ and

$$\lim_{t \to 0} \frac{1}{t} (e^{\operatorname{ad}(tX)} D - D) = \operatorname{ad} X(D).$$

Combining this relation with (17) and (18), the lemma follows immediately.

Definition. The center of $\boldsymbol{D}(G)$ will be denoted by $\boldsymbol{Z}(G)$; moreover let $I(\mathfrak{g})$ and $S_0(\mathfrak{g})$ denote the set of polynomials $P \in S(\mathfrak{g})$ which are invariant under $\operatorname{Ad}(G)$ and $\operatorname{Ad}_G(H)$, respectively.

Lemma 2.4. *Let λ denote the symmetrization $S(\mathfrak{g}) \to \boldsymbol{D}(G)$. Then*

$$\boldsymbol{Z}(G) = \{D \in \boldsymbol{D}(G) : \operatorname{Ad}(g) D = D \text{ for } g \in G\},$$

$$\lambda(I(\mathfrak{g})) = \boldsymbol{Z}(G), \qquad \lambda(S_0(\mathfrak{g})) = \boldsymbol{D}_0(G).$$

Proof. The first relation is obvious from Lemma 2.3 and (17). For the others, we just have to note that $\lambda \circ \operatorname{Ad}(g) = \operatorname{Ad}(g) \circ \lambda$ for $g \in G$.

Lemma 2.5. *Let $\boldsymbol{D}(G)\,\mathfrak{h}$ denote the set of all real linear combinations of elements of the form DT where $D \in \boldsymbol{D}(G)$, $T \in \mathfrak{h}$. Then*

$$\boldsymbol{D}(G) = \boldsymbol{D}(G)\,\mathfrak{h} + \lambda(S(\mathfrak{m})) \qquad (\textit{direct sum}).$$

Proof. We first prove by induction that for each $P \in S(\mathfrak{g})$ there exists a $Q \in S(\mathfrak{m})$ such that $\lambda(P - Q) \in \boldsymbol{D}(G)\,\mathfrak{h}$. The statement is obvious if P has degree 1 and we assume it is true for all $P \in S(\mathfrak{g})$ of degree $< d$. To prove it for P of degree d we can assume that $P = X_1^{e_1} \dots X_n^{e_n}$ where $X_1, \dots, X_n$ is a basis of $\mathfrak{g}$ such that $X_i \in \mathfrak{m}$ for $1 \leqslant i \leqslant r$ and $X_i \in \mathfrak{h}$ for $r + 1 \leqslant i \leqslant n$. If $e_{r+1} + \dots + e_n = 0$ it suffices to take $Q = P$; if $e_{r+1} + \dots + e_n > 0$, then $\lambda(P)$ is a linear combination of terms of the form $X_{\alpha_1} \cdot X_{\alpha_2} \dots X_{\alpha_d}$ where for some i, $X_{\alpha_i} \in \mathfrak{h}$. Let $S^e(\mathfrak{g})$ denote the set of homogeneous polynomials in $S(\mathfrak{g})$ of degree e and put $\boldsymbol{D}^d(G) = \lambda(\sum_{e=0}^d S^e(\mathfrak{g}))$. Then

$$(X_{\alpha_1} \dots X_{\alpha_d}) - (X_{\alpha_1} \dots X_{\alpha_{i-1}} \cdot X_{\alpha_{i+1}} \dots X_{\alpha_d} \cdot X_{\alpha_i}) \in \boldsymbol{D}^{d-1}(G)$$

so

$$\lambda(P) - D \in \boldsymbol{D}(G)\,\mathfrak{h}$$

for a suitable $D \in \boldsymbol{D}^{d-1}(G)$. Due to the induction assumption there exists a $Q \in S(\mathfrak{m})$ such that $\lambda(Q) - D \in \boldsymbol{D}(G)\,\mathfrak{h}$. Hence $\lambda(P - Q) \in \boldsymbol{D}(G)\mathfrak{h}$ as desired.

In order to prove that the sum is direct let $P \in S(\mathfrak{m})$, $P \neq 0$. There exists a function $f^*(x_1, \dots, x_r)$ of class C^∞ such that

$$[P(\partial_1, \dots, \partial_r) f^*]\,(0) \neq 0,$$

and using the coordinate system (8) it is clear that there exists a function $f \in C^\infty(G/H)$ such that $f(\pi(\exp(x_1X_1 + \dots + x_rX_r))) = f^*(x_1, \dots, x_r)$ for all sufficiently small x_i. From (16) follows

$$[\lambda(P)\,(f \circ \pi)]\,(e) = [P(\partial_1, \dots, \partial_r) f^*]\,(0) \neq 0,$$

but since $(f \circ \pi) \in C_0^\infty(G)$ it is clear that

$$[D(f \circ \pi)]\,(e) = 0 \quad \text{for each} \quad D \in \boldsymbol{D}(G)\,\mathfrak{h}.$$

It follows that $\lambda(S(\mathfrak{m})) \cap \boldsymbol{D}(G)\,\mathfrak{h} = \{0\}$ and the lemma is proved.

Corollary 2.6. *Let $I(\mathfrak{m})$ denote the set of polynomials in $S(\mathfrak{m})$ which are invariant under $Ad_G(H)$. Then*

$$\boldsymbol{D}_0(G) = (\boldsymbol{D}(G)\,\mathfrak{h}) \cap \boldsymbol{D}_0(G) + \lambda(I(\mathfrak{m})) \qquad (\textit{direct sum}).$$

In fact, for each $h \in H$, the automorphism $\mathrm{Ad}_G(h)$ of $\boldsymbol{D}(G)$ leaves the subspaces $\boldsymbol{D}(G)\,\mathfrak{h}, \lambda(S(\mathfrak{m}))$ and $\boldsymbol{D}_0(G)$ invariant. Since $\boldsymbol{D}_0(G) \cap \lambda(S(\mathfrak{m})) = \lambda(I(\mathfrak{m}))$ by Lemma 2.4, the corollary follows.

We now define a mapping of $I(\mathfrak{m})$ into $\boldsymbol{D}(G/H)$ as follows. If $P \in I(\mathfrak{m})$ then $\lambda(P) \in \boldsymbol{D}_0(G)$ and the restriction of $\lambda(P)$ to $C_0^\infty(G)$ gives by Lemma 2.2 rise to a well defined operator $D_P \in \boldsymbol{D}(G/H)$. This mapping $P \to D_P$ is linear; it is one-to-one because it was shown during the proof of Lemma 2.5 that if $P \neq 0$ then there exists a function $F \in C_0^\infty(G)$ such that $[\lambda(P)\, F]\,(e) \neq 0$. Finally, the mapping maps $I(\mathfrak{m})$ onto $\boldsymbol{D}(G/H)$ because Cor. 2.6 shows that for each $D \in \boldsymbol{D}_0(G)$ there exists a $P \in I(\mathfrak{m})$ such that D and $\lambda(P)$ have the same restriction to $C_0^\infty(G)$. We can therefore state

Theorem 2.7. *Let G/H be a reductive coset space, $\mathfrak{g} = \mathfrak{h} + \mathfrak{m}$ where $Ad_G(H)\,\mathfrak{m} \subset \mathfrak{m}$. Let $X_1, \ldots, X_r$ be a basis of $\mathfrak{m}$ and let $\tilde{f} = f \circ \pi$ for $f \in C^\infty(G/H)$. There is a one-to-one linear mapping $Q \to D_Q$ of $I(\mathfrak{m})$ onto $\boldsymbol{D}(G/H)$ such that*

$$[D_Q f]\,(p) = [Q(\partial_1, \ldots, \partial_r)\,\tilde{f}(g \exp(x_1 X_1 + \ldots + x_r X_r))]\,(0),$$

where $p = \pi(g)$. The operator D_Q is obtained from $Q(X_1, \ldots, X_r)$ by symmetrization (followed by the mapping μ from Lemma 2.2).

In the case when (G,H) is a Riemannian symmetric pair a more precise description of $\boldsymbol{D}(G/H)$ will be given in §6.

Remark. If $P = X_1^{e_1} \ldots X_n^{e_n}$, then (15) shows that

$$\lambda(P) = X_1^{e_1} \cdot X_2^{e_2} \ldots X_n^{e_n} + \lambda(Q),$$

where Q is of lower degree than P. It follows that if $P_1, P_2 \in I(\mathfrak{m})$ then

$$D_{P_1 P_2} = D_{P_1} D_{P_2} + D_Q,$$

where degree $(Q) <$ degree (P_1) + degree (P_2).

Corollary 2.8. *If $I(\mathfrak{m})$ has a finite system of generators, say $P_1, \ldots, P_l$, and we put $D_i = D_{P_i}$, then each $D \in \boldsymbol{D}(G/H)$ can be expressed*

$$D = \sum a_{n_1 \ldots n_l} D_1^{n_1} \ldots D_l^{n_l} \qquad (a_{n_1 \ldots n_l} \in \boldsymbol{R}).$$

In fact, suppose $D = D_P$ where $P \in I(\mathfrak{m})$. Then P can be written

$$P = \sum b_{n_1 \ldots n_l} P_1^{n_1} \ldots P_l^{n_l} \qquad (b_{n_1 \ldots n_l} \in \boldsymbol{R}).$$

The preceding remark shows that

$$D_P - \sum b_{n_1 \ldots n_l} D_1^{n_1} \ldots D_l^{n_l} = D_Q,$$

where $Q \in I(\mathfrak{m})$ and degree $(Q) <$ degree (P). The corollary now follows by induction.

3. The Case of a Symmetric Space

Theorem 2.9. *Let (G, K) be a Riemannian symmetric pair. Then the algebra $\boldsymbol{D}(G/K)$ is commutative.*

Proof. We recall the fact that if two differential operators act on different arguments of a function of two variables then the operators commute. The proof below reduces the commutativity of $\boldsymbol{D}(G/K)$ to this fact, by use of the symmetry of G/K.

Let σ be an involutive analytic automorphism of G such that $(K_\sigma)_0 \subset K \subset K_\sigma$, where K_σ is the set of fixed points of σ and $(K_\sigma)_0$ the identity component. The Lie algebra $\mathfrak{g}$ of G decomposes $\mathfrak{g} = \mathfrak{k} + \mathfrak{p}$ where $\mathfrak{k}$ is the Lie algebra of K and $\mathfrak{p} = \{X \in \mathfrak{g} : d\sigma(X) = -X\}$. The space G/K being complete, the mapping Exp maps $\mathfrak{p}$ onto G/K. Hence each $g \in G$ can be written $g = pk$ where $p \in \exp \mathfrak{p}$ and $k \in K$. Consequently

$$\sigma(g) = kg^{-1}k. \tag{19}$$

To begin with we shall prove that the operators in $\boldsymbol{D}(G/K)$ commute when restricted to functions in $C^\infty(G/K)$ which are invariant under all $\tau(k)$, $k \in K$. Let $f \in C^\infty(G)$ such that $f(kxk') = f(x)$ for all $k, k' \in K$ and all $x \in G$. Then by (19), $f(\sigma(g)) = f(g^{-1})$, $g \in G$. Consider the function $\varphi \in C^\infty(G \times G)$ given by $\varphi(x, y) = f(y^{-1}x)$. This function has the properties

(i) $\varphi(xk, yk') = \varphi(x, y)$ for $k, k' \in K$, $x, y \in G$;

(ii) $\varphi(gx, gy) = \varphi(x, y)$ for $g, x, y \in G$;

(iii) $\varphi(\sigma \cdot y, \sigma \cdot x) = \varphi(x, y)$ for $x, y \in G$.

Now, let $D \in \boldsymbol{D}_0(G)$ and write $D_1\varphi$ and $D_2\varphi$ when D acts on the first and second argument respectively. Let $\varphi' = D_1\varphi$ and $f' = Df$. Then

$$[D_1\varphi](x, y) = [Df^{L(y)}](x) = [(Df)^{L(y)}](x), \tag{20}$$

so $\varphi'(x, y) = f'(y^{-1}x)$. Since f' is invariant under left and right translations by K it follows that φ' satisfies (i), (ii), and (iii). Now extend σ

in the obvious way to an automorphism of $G \times G$, also denoted σ. Since $\varphi^\sigma(y, x) = \varphi(x, y)$ we have

$$[D_2\varphi^\sigma](y, x) = [D_1\varphi](x, y)$$

or

$$(D_2)^\sigma\varphi = D_1\varphi, \tag{21}$$

which connects the actions of D on the first and second argument. As remarked above, differential operators acting on different arguments commute. Thus, if T is another operator in $\boldsymbol{D}_0(G)$ we get from (21), applied to φ and φ',

$$D_1T_1\varphi = D_1(T_2)^\sigma\varphi = (T_2)^\sigma D_1\varphi = T_1D_1\varphi,$$

because T_2 as well as $T_2{}^\sigma$ are differential operators on $G \times G$ which act on the second argument. Now replace in (20) f by Tf whereby φ is replaced by $T_1\varphi$. Then

$$[D_1T_1\varphi](x, y) = [DTf](y^{-1}x)$$

and similarly

$$[T_1D_1\varphi](x, y) = [TDf](y^{-1}x),$$

so $TDf = DTf$. Now let F be an arbitrary function in $C_0^\infty(G)$ and put

$$F^*(x, g) = \int_K F(gkx)\, dk,$$

where the invariant measure dk is normalized by $\int_K dk = 1$. Then $F^* \in C^\infty(G \times G)$ and

$$[D_1F^*](x, g) = \int_K [DF](gkx)\, dk$$

so

$$[D_1F^*](e, g) = [DF](g).$$

But $F^*(kxk', g) = F^*(x, g)$ so from the commutativity already proved,

$$D_1T_1F^* = T_1D_1F^*,$$

and therefore $DT(F) = TD(F)$. In view of Lemma 2.2, this proves the commutativity of $\boldsymbol{D}(G/K)$.

Proposition 2.10. *Let M be a Euclidean space or a Riemannian globally symmetric space of rank one. The only differential operators on M that are invariant under the group $I(M)$ of all isometries of M are the* *polynomials in the Laplace-Beltrami operator Δ. If $\dim M > 1$ this statement holds with $I(M)$ replaced by the identity component $I_0(M)$.*

Proof. If $\dim M = 1$, then M is isometric to the real line or to a circle and in both cases the proposition is obvious. Hence we assume $\dim M > 1$. Let $G = I_0(M)$ and let K denote the isotropy subgroup of G at a point $o \in M$. As in the proof of Theorem 2.9, we have the direct decomposition $\mathfrak{g} = \mathfrak{k} + \mathfrak{p}$. Let P be an arbitrary polynomial in $I(\mathfrak{p})$ (Cor. 2.6). If $X_1, ..., X_r$ is an orthonormal basis of $\mathfrak{p}$, we write $P = \Sigma\, a_{n_1 \ldots n_r} X_1^{n_1} \ldots X_r^{n_r}$. The corresponding polynomial function $P^*(X) = \Sigma\, a_{n_1 \ldots n_r} x_1^{n_1} \ldots x_r^{n_r}$ $(X = \Sigma_{i=1}^r x_i X_i)$ on $\mathfrak{p}$ satisfies $P^*(\mathrm{Ad}\,(k)\,X) = P^*(X)$ for all $k \in K$. Owing to the assumption on M, the group $\mathrm{Ad}_G(K)$ acts transitively on the set of lines through the origin in $\mathfrak{p}$. This implies that $\mathrm{Ad}_G\,(K)$ acts transitively on each sphere in $\mathfrak{p}$ with center 0, these spheres being connected (since $\dim M > 1$). Hence P^* is constant on each such sphere and can therefore be written

$$P^*(X) = \sum_{k=1}^{N} a_k(x_1^2 + \ldots + x_r^2)^k.$$

It follows that $P = \Sigma_k\, a_k(X_1^2 + \ldots + X_r^2)^k$. Now, let $L = X_1^2 + \ldots + X_r^2$. Owing to Cor. 2.8, each $D \in \boldsymbol{D}(G/K)$ is a polynomial in D_L. In particular this holds for the Laplace-Beltrami operator Δ, so Δ and D_L can only differ by a constant factor. This proves the proposition.

§ 3. Spherical Functions. Definition and Examples

Let G be a connected Lie group, K a compact subgroup. Let π denote the natural mapping of G onto G/K and as in §2, we put $o = \pi(e)$ and $\tilde{f} = f \circ \pi$ if f is any function on G/K. We shall now study an important class of functions on G/K, the so-called spherical functions.

Definition. Let φ be a complex-valued function on G/K of class C^∞ which satisfies $\varphi(\pi(e)) = 1$; φ is called a *spherical function* if

(i) $\varphi^{\tau(k)} = \varphi$ for all $k \in K$,

(ii) $D\varphi = \lambda_D \varphi$ for each $D \in \boldsymbol{D}(G/K)$,

where λ_D is a complex number.

It is sometimes convenient to consider the function $\tilde{\varphi} = \varphi \circ \pi$ on G instead of φ. We say that $\tilde{\varphi}$ is a spherical function on G if and only if φ is a spherical function on G/K. Then a spherical function $\tilde{\varphi}$ on G is characterized by being an eigenfunction of each operator in $\boldsymbol{D}_0(G)$ and in addition satisfying the relations $\tilde{\varphi}(e) = 1$, $\tilde{\varphi}(kgk') = \tilde{\varphi}(g)$ for all $g \in G$ and all $k, k' \in K$. The last condition will be called *bi-invariance under* K.

We shall now see that spherical functions can also be characterized by means of an integral equation. Let dk denote the invariant measure on K, normalized by $\int_K dk = 1$.

Lemma 3.1. *Let U be an open subset of G such that $Uk \subset U$ for each $k \in K$. Let F be an analytic function on U. Then the function*

$$x \to \int_K F(xk)\, dk$$

is also analytic on U.

Proof. For simplicity, let us assume $U = G$. Let $x_0 \in G$ and let $\{x_1, ..., x_n\}$ be a system of coordinates valid in an open neighborhood V of x_0. There exists a finite set of coordinate neighborhoods $U_\alpha \subset K$ whose union equals K and a neighborhood N of x_0 in V, such that the function $F(xk)$ is given by a power series

$$F(xk) = P_\alpha(x_1, ..., x_n, k_1, ..., k_p), \qquad x \in N,\ k \in U_\alpha,$$

where $\{k_1, ..., k_p\}$ is a system of local coordinates on U_α. Consider a partition of unity $1 = \Sigma_\alpha\, \varphi_\alpha$ subordinate to the covering $\{U_\alpha\}$ of K. Then

$$F(xk) = \sum_\alpha \varphi_\alpha(k)\, P_\alpha(x_1, ..., x_n, k_1, ..., k_p), \qquad k \in K,\ x \in N,$$

which on integration over K gives a power series valid on N.

Proposition 3.2. *Let f be a complex-valued continuous function on G, not identically* 0. *Then f is a spherical function if and only if*

$$\int_K f(xky)\, dk = f(x) f(y) \tag{1}$$

for all $x, y \in G$.

Proof. Let $D \in \boldsymbol{D}(G)$. Then all the differential operators $\mathrm{Ad}\,(k)\, D$, $(k \in K)$, lie in a finite-dimensional subspace of $\boldsymbol{D}(G)$ and we can form the integral

$$D_0 = \int_K \mathrm{Ad}\,(k)\, D\, dk, \tag{2}$$

which is an operator in $\boldsymbol{D}_0(G)$. The mapping $D \to D_0$ is a linear mapping of $\boldsymbol{D}(G)$ onto $\boldsymbol{D}_0(G)$. Let F be a function in $C^\infty(G)$, bi-invariant under K. Then $[DF](k) = [DF](e)$ so

$$[D_0F](e) = \int_K [D^{R(k^{-1})}F](e)\, dk = \int_K [(DF^{R(k)})^{R(k^{-1})}](e)\, dk,$$

so

$$[D_0F](e) = [DF](e). \tag{3}$$

Now suppose f is a spherical function on G. Then f has the form $f = \varphi \circ \pi$ where φ is a spherical function on G/K. The space G/K has a Riemannian structure invariant under the action of G. It is not difficult to see that this Riemannian structure is necessarily analytic. The Laplace-Beltrami operator with respect to this structure has analytic coefficients when expressed in terms of analytic local coordinates on G/K. In addition, this operator is elliptic and consequently, by a theorem of S. Bernstein, its eigenfunctions are analytic (see F. John [1], p. 57). In particular, the function φ, and therefore also the function f, is analytic.

Now, let x be a fixed element in G and consider the function

$$F(y) = \int_K f(xky)\, dk, \qquad (y \in G),$$

which is clearly bi-invariant under K. Let D and D_0 be as in (3). Since $D_0 f = \lambda_D f$ we have

$$[D_0F](y) = \int_K [D_0 f](xky)\, dk = \lambda_D F(y)$$

and consequently

$$[D_0(f(e)\, F - F(e) f)](e) = 0.$$

Combining this with (3), we obtain

$$[D(f(e)\, F - F(e) f)](e) = 0$$

for all $D \in \boldsymbol{D}(G)$. Since $f(e)\, F - F(e) f$ is an analytic function (Lemma 3.1), we conclude from Taylor's formula (Chapter II, §1, (6)) that

$$f(e)\, F = F(e) f.$$

Since $f(e) = 1$, this is just (1).

On the other hand, let f be a continuous function on G, not identically 0, satisfying (1). Select $x_0 \in G$ such that $f(x_0) \neq 0$. Now (1) implies that $f(xk) f(x_0) = f(x_0) f(x) = f(x_0) f(kx)$ for all $x \in G$, $k \in K$. Thus

f is bi-invariant under K. Putting $y = e$ in (1) gives $f(e) = 1$. In order to see that $f \in C^\infty(G) + iC^\infty(G)$, select $\rho \in C_c^\infty(G)$ such that $\int_G \rho(y) f(y)\,dy \neq 0$, dy denoting a left invariant measure on G. Then

$$f(x) \int_G f(y)\,\rho(y)\,dy = \int_G \rho(y) \left(\int_K f(xky)\,dk\right) dy$$
$$= \int_K \left(\int_G \rho(y) f(xky)\,dy\right) dk = \int_K \left(\int_G \rho(k^{-1}x^{-1}z) f(z)\,dz\right) dk$$
$$= \int_G \left(\int_K \rho(kx^{-1}z)\,dk\right) f(z)\,dz,$$

which shows that $f \in C^\infty(G) + iC^\infty(G)$. For each $D_0 \in \boldsymbol{D}_0(G)$ we get from (1)

$$f(x)\,[D_0 f]\,(y) = \int_K [D_0 f]\,(xky)\,dk.$$

Putting $y = e$ we get

$$[D_0 f]\,(x) = [D_0 f]\,(e)\,f(x), \tag{4}$$

which shows that f is spherical.

From (4) it follows that a spherical function is determined by its system of eigenvalues. More precisely, we have

Corollary 3.3. *Let φ and φ_1 be two spherical functions on G such that $D_0\varphi = \lambda_D\varphi$ and $D_0\varphi_1 = \lambda_D\varphi_1$ for all $D_0 \in \boldsymbol{D}_0(G)$. Then $\varphi = \varphi_1$.*

In fact, (4) and (3) imply that $[D\varphi]\,(e) = [D\varphi_1]\,(e)$ for each $D \in \boldsymbol{D}(G)$; being analytic, φ and φ_1 must coincide.

Examples. Let M be a simply connected Riemannian manifold of dimension 2 and constant curvature. On these spaces (which are the plane, the 2-sphere $\boldsymbol{S}^2$, and the hyperbolic plane $\boldsymbol{H}^2$), the only differential operators invariant under $I_0(M)$ are the polynomials in the Laplace-Beltrami operator. We shall now find the spherical functions. Let us first recall the definition of *geodesic polar coordinates* at a point o on a Riemannian manifold M. Let $\theta_1, \ldots, \theta_{m-1}$ be coordinates valid on an open subset U of the unit sphere in the tangent space M_o. Let N_o be a normal neighborhood of o in M and let $C(U)$ denote the "solid cone" consisting of the points $p \in N_o - \{o\}$ for which the ray in M_o tangent to the geodesic (op) intersects U. Let $(\theta_1, \ldots, \theta_{m-1})$ denote the coordinates of the point of intersection and let $r = d(o, p)$ where d denotes distance. Then the mapping

$$p \to (\theta_1, \ldots, \theta_{m-1}, r), \qquad p \in C(U),$$

is called a system of geodesic polar coordinates at o.

I. $M = \mathbf{R}^2$, $G = \mathbf{R}^2$, $K = \{e\}$.

Here $D(G/K)$ consists of all differential operators on $\boldsymbol{R}^2$ with constant coefficients. We have obviously

Proposition 3.4. *The spherical functions on G/K are precisely the exponential functions*

$$(x, y) \to e^{\alpha x + \beta y},$$

α, β being any complex numbers.

II. $M = \mathbf{R}^2$, $G = I_0(M)$, $K = \mathbf{SO}(2)$.

Let l be any ray in $\boldsymbol{R}^2$ from the origin. In polar coordinates, valid on $\boldsymbol{R}^2 - l$, the Riemannian structure of $\boldsymbol{R}^2$ is given by the symmetric differential form

$$dr^2 + r^2\, d\theta^2,$$

where $r(x, y) = (x^2 + y^2)^{1/2}$ and θ is the angle from l to the line from the origin o to (x, y). In terms of these coordinates, the Laplace-Beltrami operator is given by

$$\Delta = \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2}.$$

Let φ be a spherical function on G/K. Then $\varphi(x, y) = \psi(\sqrt{(x^2 + y^2)})$ where $\psi(r)$ is a function which satisfies the differential equation

$$\frac{d^2\psi}{dr^2} + \frac{1}{r}\frac{d\psi}{dr} = \alpha\,\psi, \qquad r > 0, \tag{5}$$

α being a complex constant. If $\lambda^2 = \alpha$, the function

$$\psi(r) = \frac{1}{2\pi}\int_0^{2\pi} e^{\lambda r \cos u}\, du \tag{6}$$

is a solution of (5). The function

$$\varphi(x, y) = \frac{1}{2\pi}\int_0^{2\pi} \exp\,(\lambda\,\sqrt{x^2 + y^2}\cos u)\, du, \qquad (\lambda^2 = \alpha), \tag{7}$$

is therefore a solution of the equation $\Delta\varphi = \alpha\varphi$ in the punctured plane $\boldsymbol{R}^2 - \{o\}$. On the other hand, the function φ in (7) is differentiable, also at the origin, because it can be written

$$\varphi(x, y) = \frac{1}{2\pi}\int_0^{2\pi} \exp\,(\lambda(x\cos u + y\sin u))\, du.$$

Thus φ is a spherical function on G/K. Noting Cor. 3.3, we have therefore:

Proposition 3.5. *The spherical functions on G/K are precisely the functions* (7), *λ being an arbitrary complex number.*

The functional equation (1) now can be expressed in terms of ψ. If $g, h \in G$ denote translations in the direction of the x-axis of distance r and s, respectively, and k is a rotation around the origin of angle u then the point $gkh \cdot o$ has distance $(r^2 + s^2 + 2rs \cos u)^{1/2}$ from the origin. The functional equation (1) therefore takes the form

$$\frac{1}{2\pi} \int_0^{2\pi} \psi((r^2 + s^2 + 2rs \cos u)^{1/2})\, du = \psi(r)\, \psi(s). \tag{8}$$

For $\lambda = i$ the function $\psi(r)$ is the classical Bessel function $J_0(r)$.

III. $M = S^2$, $G = SO(3)$, $K = SO(2)$.

Let o be the north pole on M and let l be a semicircle on M joining o to the south pole. If (θ, r) are ordinary polar coordinates on the tangent space M_o, then we regard (θ, r) as geodesic polar coordinates at $o \in M$. They are valid on $S^2 - l$. In terms of these coordinates the Riemannian structure on S^2 is given by the symmetric differential form

$$dr^2 + (\sin r)^2\, d\theta^2$$

and the Laplace-Beltrami operator is

$$\Delta = \frac{\partial^2}{\partial r^2} + \frac{\cos r}{\sin r}\frac{\partial}{\partial r} + \frac{1}{(\sin r)^2}\frac{\partial^2}{\partial \theta^2}. \tag{9}$$

Let φ be a spherical function on G/K. Then $\varphi(p) = \psi(d(o, p))$ where d denotes distance and $\psi(r)$ is a function which satisfies the differential equation

$$\frac{d^2\psi}{dr^2} + \frac{\cos r}{\sin r}\frac{d\psi}{dr} = \alpha\psi \qquad (0 < r < \pi), \tag{10}$$

α being a complex constant. If $\alpha = -n(n+1)$, where n is an integer $\geqslant 0$, a well-known solution of (10) is given by the Legendre polynomial

$$\psi(r) = P_n(\cos r) = \frac{1}{2\pi}\int_0^{2\pi} (\cos r + i \sin r \cos u)^n\, du. \tag{11}$$

Proposition 3.6. *The spherical functions on G/K are precisely the functions*

$$\varphi_n(p) = P_n(\cos(d(o, p))), \tag{12}$$

where P_n is the Legendre polynomial of degree n.

Proof. Since the function φ_n is differentiable on S^2 it remains to be proved that each spherical function φ can be expressed in the form (12). If this were not so, the eigenvalue λ in $\Delta\varphi = \lambda\varphi$ must be different from $-n(n+1)$ for each integer $n \geqslant 0$ (Cor. 3.3). On the other hand, the operator Δ is symmetric, that is

$$\int_{S^2} f(p)\,[\Delta g]\,(p)\,\omega = \int_{S^2} [\Delta f]\,(p)\,g(p)\,\omega \tag{13}$$

for $f, g \in C^\infty(S^2)$, where ω is the volume element, (Prop. 2.1). Using this property we conclude that

$$\int_{S^2} \varphi(p)\,\varphi_n(p)\,\omega = 0 \tag{14}$$

for all integers $n \geqslant 0$. Owing to the completeness of the system of Legendre polynomials, relation (14) implies $\varphi \equiv 0$ which is a contradiction.

The functional equation (1) can now be expressed in terms of $P_n(\cos r)$. Consider rectangular xyz-coordinate system in which S^2 is given by $x^2 + y^2 + z^2 = 1$ and o is the point $(0, 0, 1)$. In the equation

$$\int_K f(gkh)\,dk = f(g)\,f(h)$$

let g and h denote rotations around the y-axis through the angles r and s, respectively, and let k denote a rotation of angle u around the z-axis. Then the point $kh \cdot o$ is given in coordinates

$$kh \cdot o = (\sin s \cos u, \sin s \sin u, \cos s),$$

and since the rotation g has coordinate expression

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \to \begin{pmatrix} \cos r & 0 & \sin r \\ 0 & 1 & 0 \\ -\sin r & 0 & \cos r \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

the coordinates of $gkh \cdot o$ are

$$(\cos r \sin s \cos u + \sin r \cos s, \sin s \sin u, -\sin r \sin s \cos u + \cos r \cos s).$$

Equation (1) therefore takes the form

$$\frac{1}{2\pi} \int_0^{2\pi} P_n(\cos r \cos s - \sin r \sin s \cos u)\,du = P_n(\cos r)\,P_n(\cos s). \tag{15}$$

IV. $G = \mathbf{SL}(2, \mathbf{R})$, $K = \mathbf{SO}(2)$, $M = G/K$.

The group G acts transitively on the upper half-plane $\operatorname{Im} z > 0$ by means of the mappings

$$z \to g \cdot z = \frac{az + b}{cz + d} \quad \text{if} \quad g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \boldsymbol{SL}(2, \boldsymbol{R}).$$

The isotropy subgroup of G at i is K so M can be identified with the upper half-plane. Let $\mathfrak{g}_0$ and $\mathfrak{k}_0$ denote the Lie algebras of G and K, respectively, and as usual we have $\mathfrak{g}_0 = \mathfrak{k}_0 + \mathfrak{p}_0$, where $\mathfrak{p}_0$ is the orthogonal complement of $\mathfrak{k}_0$ in $\mathfrak{g}_0$ with respect to the Killing form B of $\mathfrak{g}_0$. Since $\mathfrak{g}_0$ has complexification $\mathfrak{sl}(2, \boldsymbol{C})$ we have by the example in Chapter VII, §6,

$$B(X, Y) = 4 \operatorname{Tr}(XY), \qquad X, Y \in \mathfrak{g}_0.$$

The restriction of $\frac{1}{2} B$ to $\mathfrak{p}_0 \times \mathfrak{p}_0$ gives rise to a G-invariant Riemannian structure on M with respect to which M has constant curvature (G is transitive). According to formula (1), Chapter V, §3, this curvature is

$$\tfrac{1}{2} B([X, Y], [X, Y]),$$

if X, Y is an orthonormal basis of $\mathfrak{p}_0$. For example, we can take

$$X = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad Y = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Then

$$[X, Y] = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

so the curvature of M is -1. In geodesic polar coordinates (θ, r), say at the point $i \in M$, the Riemannian structure on M is given by

$$dr^2 + (\sinh r)^2 \, d\theta^2,$$

and the Laplace-Beltrami operator is

$$\Delta = \frac{\partial^2}{\partial r^2} + \frac{\cosh r}{\sinh r} \frac{\partial}{\partial r} + \frac{1}{(\sinh r)^2} \frac{\partial^2}{\partial \theta^2}.$$

Each spherical function φ on G/K has the form $\varphi(p) = \psi(d(i, p))$ where the function $\psi(r)$ satisfies the differential equation

$$\frac{d^2\psi}{dr^2} + \frac{\cosh r}{\sinh r} \frac{d\psi}{dr} = \alpha \, \psi \qquad (r > 0), \tag{16}$$

for some complex constant α. The general solution of this equation is in analogy with Example III given by (see Erdélyi *et al.* [1], p. 156)

$$\psi(r) = P_\rho(\cosh r) = \frac{1}{2\pi}\int_0^{2\pi} (\cosh r + \sinh r \cos u)^\rho \, du, \tag{17}$$

where the exponent ρ satisfies $\rho(\rho + 1) = \alpha$.

Proposition 3.7. *The spherical functions on G/K are precisely the functions*

$$\varphi(p) = P_\rho(\cosh d(i, p)),$$

where P_ρ is given by (17) *and ρ is an arbitrary complex number.*

Proof. Since (17) is the most general solution of (16) which is 1 for $r = 0$, it remains only to be proved that each function (17) arises from a spherical function f on G. We shall do this by writing down an appropriate formula for f; this will at the same time illustrate how the formula (17) is a special case of Harish-Chandra's formula for the spherical functions on a general G/K of the noncompact type.

Consider the Iwasawa decomposition of the group $G = \boldsymbol{SL}(2, \boldsymbol{R})$, namely, $G = KAN$, where A is the group of diagonal matrices

$$a = \begin{pmatrix} d & 0 \\ 0 & d^{-1} \end{pmatrix}, \qquad d > 0,$$

and N is the group of matrices

$$n = \begin{pmatrix} 1 & z \\ 0 & 1 \end{pmatrix}, \qquad z \in \boldsymbol{R}.$$

For each $x \in G$, let $H(x)$ denote the unique element in $\mathfrak{a}$, the Lie algebra of A, such that $x = k \exp(H(x))\, n$. If we take

$$a = \begin{pmatrix} e^{r/2} & 0 \\ 0 & e^{-r/2} \end{pmatrix}$$

then $(d(K, aK))^2 = \frac{1}{2} B(H(a), H(a)) = r^2$. Moreover, if

$$k = \begin{pmatrix} \cos u & \sin u \\ -\sin u & \cos u \end{pmatrix}$$

and we decompose $ak = k_1 a_1 n_1$, then

$$a_1 = \begin{pmatrix} (\cosh r + \sinh r \cos 2u)^{1/2} & 0 \\ 0 & (\cosh r + \sinh r \cos 2u)^{-1/2} \end{pmatrix}.$$

The integral (17) can therefore be written

$$\int_K e^{\nu(H(ak))}\,dk,$$

where ν is the linear function on $\mathfrak{a}$ given by

$$\nu\begin{pmatrix} t & 0 \\ 0 & -t \end{pmatrix} = 2\rho t, \qquad t \in \boldsymbol{R}.$$

The function f on G given by

$$f(x) = \int_K e^{\nu(H(xk))}\,dk \tag{18}$$

is differentiable and has the form $f = \varphi \circ \pi$. Thus φ is differentiable and $\varphi(p) = \psi(d(i, p))$ where ψ is given by (17). This proves the proposition.

In order to find the explicit form of the functional equation (1), let

$$a_i = \begin{pmatrix} e^{\frac{1}{2}r_i} & 0 \\ 0 & e^{-\frac{1}{2}r_i} \end{pmatrix}, \qquad (i = 1, 2) \qquad k = \begin{pmatrix} \cos u & \sin u \\ -\sin u & \cos u \end{pmatrix}.$$

Then $a_1 k a_2$ can be written $k_1 a k_2$ where $k_1, k_2 \in K$ and

$$a = \begin{pmatrix} e^{r/2} & 0 \\ 0 & e^{-r/2} \end{pmatrix},$$

$$\cosh r = \cosh r_1 \cosh r_2 + \sinh r_1 \sinh r_2 \cos 2u.$$

Hence the functions (17) are characterized by the integral formula

$$\frac{1}{2\pi}\int_0^{2\pi} P(\cosh r \cosh s + \sinh r \sinh s \cos u)\,du = P(\cosh r)\,P(\cosh s). \tag{19}$$

V. Compact groups

In this last example let U be an arbitrary compact connected Lie group. Let G denote the product group $U \times U$ and let K denote the diagonal in $U \times U$. Then G/K is identified with U under the mapping $(u_1, u_2)\,K \to u_1 u_2^{-1}$. Under this identification, the mapping $\tau(u_1, u_2)$ of G/K corresponds to the mapping $u \to u_1 u u_2^{-1}$ of U. Consequently, a differential operator D on G/K belongs to $\boldsymbol{D}(G/K)$ if and only if, when considered as a differential operator on U, it is invariant under all left and right translations. Thus $\boldsymbol{D}(G/K)$ is identified with the center $\boldsymbol{Z}(U)$

of $\boldsymbol{D}(U)$, (Lemma 2.4). If we interpret functions on G/K as functions on U, the functional equation (1) becomes

$$\int_U \varphi(xuyu^{-1})\,du = \varphi(x)\,\varphi(y), \tag{20}$$

which characterizes those eigenfunctions of the operators in $\boldsymbol{Z}(U)$ which are invariant under all inner automorphisms of U and satisfy $\varphi(e) = 1$.

From the theory of representations of compact groups (see e.g. Weil [1], p. 87) it is known, that the solutions of (20) are precisely

$$\varphi = \frac{1}{d_\chi}\chi,$$

where χ is the character of an irreducible representation of U and d_χ denotes its dimension. (For the definition of these terms see the following section.)

§ 4. Elementary Properties of Spherical Functions

Let G be a Lie group and dg a left invariant measure on G. The set of complex-valued continuous functions on G of compact support can be turned into an associative algebra over C, the multiplication being the *convolution* product

$$f_1 * f_2(x) = \int_G f_1(g)\,f_2(g^{-1}x)\,dg$$

and the addition being pointwise addition of functions. This algebra is called the *group algebra* of G. If θ is an analytic automorphism of G which preserves the measure dg then the mapping $f \to f^\theta$ is an automorphism of the group algebra.

Suppose now that (G, K) is a Riemannian symmetric pair and let σ be a corresponding involutive automorphism of G. We assume that K is compact. Let $C^\natural(G)$ denote the set of functions in the group algebra which are bi-invariant under K. Then $C^\natural(G)$ is a subalgebra of the group algebra.

Theorem 4.1. *The algebra $C^\natural(G)$ is commutative.*

Proof. Since σ is involutive it will clearly preserve the left invariant measure dg on G. For each function f on G let $\check{f}$ denote the function $x \to f^\sigma(x^{-1})$. Then since each $x \in G$ can be written $x = kp$ where

$\sigma(k) = k$, $\sigma(p) = p^{-1}$ it follows that $f^\sigma = \check{f}$ for $f \in C^\natural(G)$. Now, for any two functions f, g in the group algebra we have

$$\check{f}*\check{g}(x) = \int_G f(y^{-1})\, g(x^{-1}y)\, dy = \int_G g(z)\, f(z^{-1}x^{-1})\, dz,$$

so

$$\check{f}*\check{g} = (g*f)^\vee.$$

Since

$$f^\sigma * g^\sigma = (f*g)^\sigma,$$

we obtain $f*g = g*f$ for $f, g \in C^\natural(G)$.

Lemma 4.2. *Let φ be a continuous, complex-valued function on G, bi-invariant under K. Then φ is a spherical function if and only if the mapping*

$$L : f \to \int_G f(x)\, \varphi(x)\, dx$$

is a homomorphism of $C^\natural(G)$ onto $\mathbf{C}$.

Proof. Let f, g be two functions in the group algebra and put

$$f^\natural(x) = \iint_{K\times K} f(kxk')\, dk\, dk', \qquad L(f) = L(f^\natural).$$

Then the mapping $f \to f^\natural$ maps the group algebra onto $C^\natural(G)$ and

$$(f^\natural * g)^\natural = f^\natural * g^\natural. \tag{1}$$

For a suitably normalized measure dy_K on G/K we have

$$\int_G (f*g)(x)\, \varphi(x)\, dx = \int_G \Big(\int_G f(y)\, g(y^{-1}x)\, dy\Big)\, \varphi(x)\, dx$$

$$= \int_G \int_G f(y)\, g(z)\, \varphi(yz)\, dy\, dz = \int_G g(z) \int_{G/K} \Big(\int_K f(yk)\, \varphi(ykz)\, dk\Big)\, dy_K\, dz$$

so

$$L(f^\natural * g) = \int_G \int_G f^\natural(y)\, g(z) \Big(\int_K \varphi(ykz)\, dk\Big)\, dy\, dz.$$

Since φ is bi-invariant and since dy is invariant under the mappings $y \to kyk'$ it follows that

$$L(f^\natural * g) = \int_G \int_G f(y)\, g(z) \Big(\int_K \varphi(ykz)\, dk\Big)\, dy\, dz. \tag{2}$$

Moreover,

$$L(f)\,L(g) = \int_G \int_G f(y)\,g(z)\,\varphi(y)\,\varphi(z)\,dy\,dz. \tag{3}$$

Considering (1), the lemma follows immediately.

The norm $\|f\| = \int_G |f(x)|\,dx$ turns the group algebra into a normed vector space. Owing to the additional property $\|f*g\| \leqslant \|f\|\,\|g\|$ the group algebra is a *normed algebra*. The algebra $C^\natural(G)$ is a closed subalgebra.

Theorem 4.3. *The continuous homomorphisms of the algebra $C^\natural(G)$ onto $\mathbf{C}$ are the mappings*

$$f \to \int_G f(x)\,\varphi(x)\,dx,$$

where φ is a bounded spherical function on G.

Proof. Let L be a continuous homomorphism of $C^\natural(G)$ onto $\mathbf{C}$. Then the mapping

$$f \to L(f^\natural)$$

is a continuous linear function on the group algebra. Hence there exists a bounded measurable function φ on G such that

$$L(f^\natural) = \int_G f^\natural(x)\,\varphi(x)\,dx$$

for all $f \in C_c(G)$. Here we may assume that φ is bi-invariant under K because otherwise it can be replaced by $\varphi^\natural$. Since L is a homomorphism, the relations (2) and (3) imply (by approximation in $C_c(G \times G)$) that

$$\int_K \varphi(xky)\,dk = \varphi(x)\,\varphi(y)$$

except for a set of $(x, y) \in G \times G$ of measure 0. In order to see that φ is equal to a continuous function almost everywhere select $\rho \in C_c(G)$ such that $\int_G \rho(y)\,\varphi(y)\,dy \neq 0$. Then for almost all $x \in G$,

$$\varphi(x) \int_G \varphi(y)\,\rho(y)\,dy = \int_G \rho(y) \left(\int_K \varphi(xky)\,dk\right) dy$$

$$= \int_K \left(\int_G \rho(k^{-1}x^{-1}z)\,\varphi(z)\,dz\right) dk = \int_G \left(\int_K \rho(kx^{-1}z)\,dk\right) \varphi(z)\,dz$$

and this last expression is continuous in x. This proves the theorem.

By means of Theorems 4.1 and 4.3, the stage is set for applications of the general theory of normed algebras and abstract harmonic analysis. The *Fourier transform* $\tilde{f}$ of a function $f \in C^{\natural}(G)$ is defined by

$$\tilde{f}(\varphi) = \int_G f(x)\, \varphi(x^{-1})\, dx$$

as φ varies through the set $\mathfrak{M}$ of positive definite (see definition below) spherical functions on G. The set $\mathfrak{M}$ is given the weakest topology for which all Fourier transforms $\tilde{f}$ are continuous. Then $\mathfrak{M}$ is locally compact. The Plancherel formula for the classical Fourier transform generalizes as follows (Godement [4]): There exists a unique positive measure $d\varphi$ on $\mathfrak{M}$ such that the abstract Plancherel formula

$$\int_G |f(x)|^2\, dx = \int_{\mathfrak{M}} |\tilde{f}(\varphi)|^2\, d\varphi$$

holds for all $f \in C^{\natural}(G)$; finally the mapping $f \to \tilde{f}$ extends to an isomorphism of the Hilbert space of square integrable functions on G bi-invariant under K onto $L^2(\mathfrak{M})$. In Example I, §3, the abstract Fourier transform reduces to the classical one

$$\tilde{f}(u, v) = \frac{1}{2\pi} \int_{\mathbf{R}^2} f(x, y)\, e^{-i(ux+vy)}\, dx\, dy,$$

and the Plancherel formula is

$$\int_{\mathbf{R}^2} |f(x, y)|^2\, dx\, dy = \int_{\mathbf{R}^2} |\tilde{f}(u, v)|^2\, du\, dv.$$

In Example II, §3, we obtain the Bessel transform

$$\tilde{f}(\lambda) = \int_0^{\infty} f(r)\, J_0(r\lambda)\, r dr$$

and

$$\int_0^{\infty} |f(r)|^2\, r dr = \int_0^{\infty} |\tilde{f}(\lambda)|^2\, \lambda d\lambda.$$

Example III, §3, leads to the Plancherel formula for the Laplace series for zonal functions on a sphere. In Example V, §3, the space $C^{\natural}(G)$ consists of the *central functions* on U, that is continuous functions f on U satisfying $f(xy) = f(yx)$ for all $x, y \in U$. The Fourier transform $\tilde{f}$ is

$$\tilde{f}(\chi) = \frac{1}{d\chi} \int_U f(x)\, \chi(x^{-1})\, dx,$$

and the abstract Plancherel formula becomes

$$\int_U |f(x)|^2\,dx = \sum_\chi |\hat{f}(\chi)|^2\,d\chi^2,$$

which is the Peter-Weyl formula specialized to central functions.

We shall not explore this topic further but instead establish a connection between the spherical functions and representations of the group G.

Definition. A complex-valued continuous function φ on a topological group G is called *positive definite* if

$$\sum_{i,j=1}^{n} \varphi(x_i^{-1}x_j)\,\alpha_i\bar{\alpha}_j \geqslant 0$$

for all finite sets $x_1, \ldots, x_n$ of elements in G and any complex numbers $\alpha_1, \ldots, \alpha_n$.

A positive definite function φ satisfies the conditions

$$\varphi(e) \geqslant 0, \qquad \varphi(x^{-1}) = \overline{\varphi(x)}, \qquad |\varphi(x)| \leqslant \varphi(e),$$

which are easily derived from the definition. In particular, a positive definite function is necessarily bounded.

Definition. Let G be a topological group and $\mathfrak{H}$ a Banach space over C. A *representation* of G on $\mathfrak{H}$ is a mapping π which assigns to each $x \in G$ a continuous endomorphism of $\mathfrak{H}$ such that

(i) $\pi(xy) = \pi(x)\,\pi(y)$, $(x, y \in G)$, $\pi(e) = I$.

(ii) For each $\boldsymbol{a} \in \mathfrak{H}$, the mapping $x \to \pi(x)\,\boldsymbol{a}$ is a continuous mapping of G into $\mathfrak{H}$.

A representation π is called *irreducible* if no closed subspace of $\mathfrak{H}$ (except 0 and $\mathfrak{H}$) is invariant under all $\pi(x)$. If $\mathfrak{H}$ is of finite dimension the function $x \to \mathrm{Tr}\,(\pi(x))$ is called the *character* of the representation π. If $\mathfrak{H}$ is a Hilbert space and each $\pi(x)$ is unitary the representation π is called a *unitary* representation.

Let π be a unitary representation of a topological group G on a Hilbert space $\mathfrak{H}$. Let $\langle , \rangle$ denote the scalar product in $\mathfrak{H}$. For each fixed vector $\boldsymbol{e} \in \mathfrak{H}$, the function $x \to \langle \boldsymbol{e}, \pi(x)\,\boldsymbol{e}\rangle$ is a positive definite function on G. The continuity is obvious so let $x_1, \ldots, x_n$ be any finite set of elements in G and $\alpha_1, \ldots, \alpha_n$ any complex numbers.

Then

$$\sum_{i,j=1}^{n} \langle e, \pi(x_i^{-1}x_j)\, e\rangle\, \alpha_i\bar{\alpha}_j = \sum_{i,j=1}^{n} \langle \pi(x_i)\, e, \pi(x_j)\, e\rangle\, \alpha_i\bar{\alpha}_j$$

$$= \Big\langle \sum_i \alpha_i\pi(x_i)\, e, \sum_j \alpha_j\pi(x_j)\, e\Big\rangle \geqslant 0.$$

On the other hand, a positive definite function $\varphi \not\equiv 0$ on G gives in a canonical way rise to a unitary representation of G as follows: Let V_φ denote the set of all complex linear combinations of left translates $\varphi^{L(x)}$ $(x \in G)$ of φ. We define a scalar product in V_φ by the formula

$$\langle f, g\rangle = \sum_{i,j} \alpha_i\bar{\beta}_j\varphi(x_i^{-1}y_j) \tag{4}$$

if $f = \sum_i \alpha_i\varphi^{L(x_i)}$, $g = \sum_j \beta_j\varphi^{L(y_j)}$. Now $\langle f, g\rangle = \Sigma\, \alpha_i\, \overline{g(x_i)} = \Sigma\bar{\beta}_j f(y_j)$ so it is clear that (4) depends only on f and g but not on their special expressions in terms of φ. Except for completenesss and the fact that $\langle f, f\rangle$ might be 0 without f being $\equiv 0$, the space V_φ is a Hilbert space with inner product $\langle,\rangle$. From the Schwarz inequality $|\langle f, g\rangle|^2 \leqslant \langle f, f\rangle\, \langle g, g\rangle$ it follows that the set of vectors $f \in V_\varphi$ for which $\langle f, f\rangle = 0$ is a subspace N of V_φ. The factor space V_φ/N inherits a scalar product $\langle,\rangle$ and norm $\|\ \|$ from V_φ and the completion of V_φ/N in this norm is a Hilbert space $\mathfrak{H}_\varphi$. Each $x \in G$ gives rise to an endomorphism $f \to f^{L(x)}$ of V_φ; this endomorphism preserves the inner product $\langle,\rangle$, maps the subspace N into itself and extends uniquely to a unitary operator $\pi(x)$ of $\mathfrak{H}_\varphi$. The mapping $x \to \pi(x)$ satisfies $\pi(xy) = \pi(x)\, \pi(y)$ for all $x, y \in G$. Moreover, if $f \in V_\varphi$ we have

$$\langle f^{L(x)} - f, f^{L(x)} - f\rangle = 2\langle f, f\rangle - \langle f^{L(x)}, f\rangle - \langle f, f^{L(x)}\rangle$$

$$= 2\sum_{i,j} \alpha_i\bar{\alpha}_j\varphi(x_i^{-1}x_j) - \sum_{i,j} \alpha_i\bar{\alpha}_j\varphi(x_i^{-1}x^{-1}x_j) - \sum_{i,j} \alpha_i\bar{\alpha}_j\varphi(x_i^{-1}xx_j),$$

which tends to 0 for $x \to e$. Thus for a dense set of vectors $\boldsymbol{a} \in \mathfrak{H}_\varphi$, the mapping $x \to \pi(x)\, \boldsymbol{a}$ is continuous for $x = e$, hence for all $x \in G$. From the inequality

$$\|\, \pi(x)\, \boldsymbol{b} - \boldsymbol{b}\, \| \leqslant \|\, \pi(x)\, (\boldsymbol{a} - \boldsymbol{b})\, \| + \|\, \pi(x)\, \boldsymbol{a} - \boldsymbol{a}\, \| + \|\, \boldsymbol{a} - \boldsymbol{b}\, \|$$

it follows that for each $\boldsymbol{b} \in \mathfrak{H}_\varphi$, the mapping $x \to \pi(x)\, \boldsymbol{b}$ is continuous.

Thus π is a unitary representation of G on $\mathfrak{H}_\varphi$. Finally, if $\boldsymbol{e}$ denotes the vector in $\mathfrak{H}_\varphi$ which corresponds to the vector $\varphi \in V_\varphi$ we have

$$\varphi(x) = \langle \boldsymbol{e}, \pi(x)\, \boldsymbol{e} \rangle. \tag{5}$$

We shall call π the unitary representation *associated to* φ. Summarizing the results above we have

Theorem 4.4. *Let π be a unitary representation of a topological group G on a Hilbert space $\mathfrak{H}$. For each vector $\boldsymbol{e} \in \mathfrak{H}$ the function $\langle \boldsymbol{e}, \pi(x)\, \boldsymbol{e} \rangle$ is a positive definite function on G. Conversely, to any positive definite function $\varphi \not\equiv 0$ on G corresponds a unitary representation π of G such that $\varphi(x) = \langle \boldsymbol{e}, \pi(x)\, \boldsymbol{e} \rangle$ for a suitable vector $\boldsymbol{e}$.*

Definition. Let G be a topological group and K a closed subgroup. A representation π of G on a Hilbert space $\mathfrak{H}$ is said to be of *class* 1 if it is irreducible and unitary and if there exists a vector $\boldsymbol{e} \neq 0$ in $\mathfrak{H}$ which is left fixed by each $\pi(k)$, $k \in K$.

Let (G, K) be a Riemannian symmetric pair, K compact. We shall now prove that under the correspondance in Theorem 4.4, the positive definite spherical functions on G correspond to representations of G of class 1.

Theorem 4.5. *Let (G, K) be a Riemannian symmetric pair, K compact. Let $\varphi \not\equiv 0$ be a positive definite spherical function on G and let π be the unitary representation of G associated to φ. Then π is of class 1. On the other hand, if π is a representation of G of class 1 and if $\boldsymbol{e}$ is a unit vector left fixed by all $\pi(k)$ $(k \in K)$, then the function $\langle \boldsymbol{e}, \pi(x)\, \boldsymbol{e} \rangle$ is a positive definite spherical function on G.*

Proof. Let φ be a positive definite spherical function, let V_φ, $\mathfrak{H}_\varphi$, and π be as defined above. Let $\boldsymbol{e}$ be the vector in $\mathfrak{H}_\varphi$ which corresponds to $\varphi \in V_\varphi$. Since φ is bi-invariant under K it follows that $\pi(k)\, \boldsymbol{e} = \boldsymbol{e}$ for all $k \in K$. In order to prove that π is irreducible, consider for each pair $\boldsymbol{a}, \boldsymbol{b} \in \mathfrak{H}_\varphi$ the integral

$$B(\boldsymbol{a}, \boldsymbol{b}) = \int_K \langle \pi(k)\, \boldsymbol{a}, \boldsymbol{b} \rangle \, dk.$$

Since $| B(\boldsymbol{a}, \boldsymbol{b}) | \leqslant || \boldsymbol{a} || \, || \boldsymbol{b} ||$ there exists a bounded operator P on $\mathfrak{H}_\varphi$ such that

$$\langle P\boldsymbol{a}, \boldsymbol{b} \rangle = B(\boldsymbol{a}, \boldsymbol{b})$$

for all $\boldsymbol{a}, \boldsymbol{b} \in \mathfrak{H}_\varphi$. Now, since φ is spherical it follows from the definition of the scalar product in V_φ that

$$\int_K \langle \varphi^{L(kx)}, \psi \rangle \, dk = \varphi(x^{-1}) \langle \varphi, \psi \rangle$$

for all $\psi \in V_\varphi$. Consequently we have in $\mathfrak{H}_\varphi$

$$\int_K \langle \pi(kx)\, \boldsymbol{e}, \boldsymbol{b} \rangle\, dk = \varphi(x^{-1}) \langle \boldsymbol{e}, \boldsymbol{b} \rangle$$

for all $\boldsymbol{b} \in \mathfrak{H}_\varphi$. This means that

$$P\, \pi(x)\, \boldsymbol{e} = \varphi(x^{-1})\, \boldsymbol{e}. \tag{6}$$

Since the space $\mathfrak{H}_\varphi$ is generated by the complex linear combinations of the vectors $\pi(x)\, \boldsymbol{e}$ $(x \in G)$, it follows that $P^2 = P$ and

$$P(\mathfrak{H}_\varphi) = \boldsymbol{C}\boldsymbol{e}. \tag{7}$$

Let $\mathfrak{H}'$ denote the closure of the sum of all closed subspaces $U' \subset \mathfrak{H}_\varphi$ which are invariant under π and which satisfy $PU' = \{0\}$. Then $\mathfrak{H}'$ and its orthogonal complement $\mathfrak{H}''$ in $\mathfrak{H}_\varphi$ are invariant under π. Let U be any closed subspace of $\mathfrak{H}''$ invariant under π $(U \neq \{0\})$. Then $PU \neq \{0\}$ so by (7), $\boldsymbol{e} \in U$. But then $\pi(x)\, \boldsymbol{e} \subsetneq U$ for all $x \in G$ so $U = \mathfrak{H}_\varphi$. This proves firstly that $\mathfrak{H}_\varphi = \mathfrak{H}''$ and secondly that π is irreducible.

On the other hand, let π be a representation of G on a Hilbert space $\mathfrak{H}$ such that π is of class 1. Let $\boldsymbol{e}$ be a unit vector in $\mathfrak{H}$ such that $\pi(k)\, \boldsymbol{e} = \boldsymbol{e}$ for all $k \in K$. Before proving that the function $\langle \boldsymbol{e}, \pi(x)\, \boldsymbol{e} \rangle$ is spherical we establish a few facts about π.

Let f be a continuous function on G with compact support. Then the Hermitian form

$$C(\boldsymbol{a}, \boldsymbol{b}) = \int_G f(x) \langle \pi(x)\, \boldsymbol{a}, \boldsymbol{b} \rangle\, dx \qquad (\boldsymbol{a}, \boldsymbol{b} \in \mathfrak{H}),$$

satisfies

$$| C(\boldsymbol{a}, \boldsymbol{b}) | \leqslant \left(\int_G | f(x) |\, dx \right) \| \boldsymbol{a} \|\, \| \boldsymbol{b} \|,$$

so there exists a bounded operator π_f on $\mathfrak{H}$ such that $\langle \pi_f(\boldsymbol{a}), \boldsymbol{b} \rangle = C(\boldsymbol{a}, \boldsymbol{b})$. We write symbolically

$$\pi_f = \int_G f(x)\, \pi(x)\, dx.$$

Then the mapping $f \to \pi_f$ is a homomorphism of the group algebra of G into the algebra of bounded operators on $\mathfrak{H}$, in other words we have a representation of the group algebra on $\mathfrak{H}$.

Let $\mathfrak{N}$ be the subspace of $\mathfrak{H}$ consisting of all vectors $\boldsymbol{a} \in \mathfrak{H}$ which are left fixed by each $\pi(k)$, $k \in K$.

Lemma 4.6. *The subspace* $\mathfrak{N}$ *is invariant under each operator* π_f, $f \in C^\natural(G)$.

Proof. If $\boldsymbol{a} \in \mathfrak{N}$ and $f \in C^\natural(G)$ we have for $k \in K$,

$$\pi(k)\,\pi_f \boldsymbol{a} = \int_G f(x)\,\pi(kx)\,\boldsymbol{a}\,dx = \int_G f(x)\,\pi(x)\,\boldsymbol{a}\,dx = \pi_f \boldsymbol{a}$$

where we have used $f(kx) = f(x)$.

Lemma 4.7. *The space* $\mathfrak{N}$ *is one-dimensional.*

Proof. For $f \in C^\natural(G)$ let A_f denote the restriction of π_f to the Hilbert space $\mathfrak{N}$. Let f^* denote the function

$$f^*(x) = \overline{f(x^{-1})}\,\det \mathrm{Ad}_G(x), \qquad x \in G.$$

Then $f^* \in C^\natural(G)$ and using Cor. 1.3 we have

$$\begin{aligned}\langle \boldsymbol{a}, \pi_{f^*} \boldsymbol{b}\rangle &= \int_G f(x^{-1})\,\det \mathrm{Ad}_G(x)\,\langle \boldsymbol{a}, \pi(x)\,\boldsymbol{b}\rangle\,dx\\ &= \int_G f(x^{-1})\,\langle \pi(x^{-1})\,\boldsymbol{a}, \boldsymbol{b}\rangle\,d(x^{-1}).\end{aligned}$$

Hence $\langle \boldsymbol{a}, A_{f^*}\boldsymbol{b}\rangle = \langle A_f \boldsymbol{a}, \boldsymbol{b}\rangle$ so the operator A_{f^*} is the adjoint of A_f. The operators A_f therefore constitute a commutative family of normal bounded operators. Consequently they have a common spectral resolution

$$A_f = \int p_f(\lambda)\,dE_\lambda,$$

where λ varies over some (unspecified) space, p_f is a complex-valued function, and dE_λ is a measure whose values are projection operators on $\mathfrak{N}$. These operators $E(S)$ commute with each A_f and the range $E(S)\,\mathfrak{N}$ is therefore invariant under each A_f.

Now suppose that dim $\mathfrak{N}$ were > 1. Then either all A_f are scalar multiples of I or $E(S)\,\mathfrak{N}$ is for some S different from 0 and $\mathfrak{N}$. In both of these cases $\mathfrak{N}$ can be decomposed $\mathfrak{N} = \mathfrak{N}_1 + \mathfrak{N}_2$ where $\mathfrak{N}_1$ and $\mathfrak{N}_2$ are nonzero, closed mutually orthogonal subspaces of $\mathfrak{N}$, invariant under each A_f. Select $\boldsymbol{a}_1 \neq 0$ in $\mathfrak{N}_1$ and let $\mathfrak{M}_1$ denote the set of vectors $\pi_f \boldsymbol{a}_1$ as f runs through the group algebra. Then $\mathfrak{M}_1$ is not $\{0\}$ and is invariant under each $\pi(x)$, $x \in G$. We shall now show that $\mathfrak{M}_1$ and $\mathfrak{N}_2$ are orthogonal. This would imply that the closure of $\mathfrak{M}_1$ is different from $\mathfrak{H}$ which in turn contradicts the assumed irreducibility of π.

Let $f \in C_c(G)$ and $\boldsymbol{a}_2 \in \mathfrak{N}_2$. Then

$$\langle \pi_f \boldsymbol{a}_1, \boldsymbol{a}_2 \rangle = \int_G f(x) \langle \pi(x) \boldsymbol{a}_1, \boldsymbol{a}_2 \rangle \, dx = \int_G f^\natural(x) \langle \pi(x) \boldsymbol{a}_1, \boldsymbol{a}_2 \rangle \, dx$$

so

$$\langle \pi_f \boldsymbol{a}_1, \boldsymbol{a}_2 \rangle = \langle A_{f^\natural} \boldsymbol{a}_1, \boldsymbol{a}_2 \rangle.$$

This last expression vanishes since $\mathfrak{N}_1$ is invariant under $A_{f^\natural}$. This concludes the proof.

We return now to the proof of Theorem 4.5. It remains to prove that the function $\varphi(x) = \langle \boldsymbol{e}, \pi(x) \boldsymbol{e} \rangle$ is spherical. Let $f \in C^\natural(G)$. In view of Lemma 4.6 and 4.7 the vector $\pi_f \boldsymbol{e}$ is a scalar multiple of $\boldsymbol{e}$ and since

$$\langle \pi_f \boldsymbol{e}, \boldsymbol{e} \rangle = \int_G f(x) \overline{\varphi(x)} \, dx$$

it is clear that

$$\pi_f \boldsymbol{e} = \left(\int_G f(x) \overline{\varphi(x)} \, dx \right) \boldsymbol{e}.$$

Since the mapping $f \to \pi_f$ is a representation of $C^\natural(G)$ on $\mathfrak{H}$, the mapping

$$f \to \int_G f(x) \overline{\varphi(x)} \, dx$$

is a homomorphism of $C^\natural(G)$ onto $\boldsymbol{C}$. Lemma 4.2 now shows that φ is a spherical function on G.

The connection between positive definite spherical functions and representations of class 1 established in Theorem 4.5 can be made more precise by using the standard concept of equivalence of representations which we now define.

Definition. Let G be a topological group. Two unitary representations π and π' of G on Hilbert spaces $\mathfrak{H}$ and $\mathfrak{H}'$ are called *equivalent* if there exists a linear mapping A of $\mathfrak{H}$ onto $\mathfrak{H}'$ preserving scalar products such that $\pi'(x) \circ A = A \circ \pi(x)$ for all $x \in G$.

Theorem 4.8. *Let (G, K) be a Riemannian symmetric pair, K compact. For each representation π of G of class* 1 *on a Hilbert space $\mathfrak{H}$ let $\mathfrak{N}_\pi$ denote the (one-dimensional) space of vectors in $\mathfrak{H}$ which are fixed under each $\pi(k)$, $k \in K$. Let Ω denote the set of all equivalence classes ω of representations of G of class* 1. *Then Ω is in a natural one-to-one correspondence with the set $\mathfrak{P}$ of all positive definite spherical functions φ on G satisfying $\varphi(e) = 1$. This correspondence $\omega \to \varphi$ has the properties*

(i) *If $\pi \in \omega$ and $\boldsymbol{e}$ is any unit vector in $\mathfrak{N}_\pi$, then*

$$\varphi(x) = \langle \boldsymbol{e}, \pi(x) \boldsymbol{e} \rangle.$$

(ii) *ω contains the representation associated to φ.*

Proof. First we note that $\langle e, \pi(x)\, e\rangle$ is independent of the choice of the unit vector $e \in \mathfrak{N}_\pi$ and of the choice of π in ω. Thus we have a mapping of Ω into $\mathfrak{P}$. This mapping is onto because

$$\varphi(x) = \langle e_\varphi, \pi_\varphi(x)\, e_\varphi\rangle \qquad (x \in G)$$

if $\varphi \in \mathfrak{P}$, and π_φ the representation associated to φ and e_φ any unit vector in $\mathfrak{N}_{\pi_\varphi}$. In order to prove that the mapping is one-to-one it suffices to prove that if $\omega \in \Omega$, $\pi \in \omega$ and if we put $\varphi(x) = \langle e, \pi(x)\, e\rangle$ where e is a unit vector in $\mathfrak{N}_\pi$, then $\pi_\varphi \in \omega$. But the desired mapping from the Hilbert space $\mathfrak{H}$ (on which π acts) onto $\mathfrak{H}_\varphi$ is given by

$$A : \sum_{i=1}^{r} a_i \pi(x_i)\, e \to \sum_{i=1}^{r} a_i \pi_\varphi(x_i)\, e_\varphi$$

where $x_1, \cdots, x_r$ are arbitrary in G and $a_1, \cdots, a_r$ are arbitrary complex numbers. This proves the theorem.

§ 5. Some Algebraic Tools

In this section we collect some known algebraic results which will be used in §6. The notation and treatment follow that of the book by Zariski and Samuel [1] to which we also refer concerning omitted proofs.

We recall that "field" always means commutative field of characteristic 0. In this section we only consider rings and algebras which are commutative, have an identity element and no divisors of 0. A module M over a ring A is called *finite* if there exists finitely many elements $x_1, \ldots, x_m \in M$ such that $M = Ax_1 + \ldots + Ax_m$.

Definition. Let B be a ring and A a subring of B (with the same identity). An element $x \in B$ is called *integral* over A if it satisfies an equation

$$x^n + a_1 x^{n-1} + \ldots + a_n = 0 \tag{1}$$

with leading coefficient 1, where $a_i \in A$. If each $x \in B$ is integral over A, the ring B is said to be integral over A.

Lemma 5.1. *Let A be a subring of a ring B and let $x \in B$. The following conditions are equivalent:*

(i) *x is integral over A.*

(ii) *The ring $A[x]$, (the subring of B generated by A and x), is a finite A-module.*

(iii) *There exists an intermediate ring* R, $A[x] \subset R \subset B$ *such that* R *is a finite* A*-module.*

Proof. (i)$\Rightarrow$(ii): There exists an integer $n > 0$ such that $x^n \in \sum_{i=0}^{n-1} Ax^i$. It follows that $x^{n+r} \in \sum_{i=0}^{n-1} Ax^{i+r}$ so by induction on r, $x^{n+r} \in \sum_{i=0}^{n-1} Ax^i$ for each $r > 0$.

(ii)$\Rightarrow$(iii): Take $R = A[x]$.

(iii)$\Rightarrow$(i): Select $y_1, \ldots, y_n \in R$ such that $R = \sum_{i=1}^n Ay_i$. Then we have for suitable elements $a_{ij} \in A$, $xy_i = \sum_{j=1}^n a_{ij}y_j$ $(1 \leqslant i \leqslant n)$. Writing this system of linear equations as

$$\sum_{j=i}^{n} (\delta_{ij}x - a_{ij})\, y_j = 0 \qquad (1 \leqslant i \leqslant n),$$

we conclude that the determinant $d = \det(\delta_{ij}x - a_{ij})$ satisfies $dy_i = 0$ for each i. Then $dR = 0$ so $d = 0$. But this is an equation of the form (1) so x integral over A.

By induction we conclude from Lemma 5.1,

Lemma 5.2. *Let* $x_1, \ldots, x_n$ *be elements of a ring* B *which are integral over a subring* A. *Then the ring* $A[x_1, \ldots, x_n]$ (*the subring of* B *generated by* A *and* $x_1, \ldots, x_n$) *is a finite* A*-module.*

From this lemma and Lemma 5.1 (iii) we obtain the following

Corollary 5.3. *Let* A *be a subring of a ring* B. *The set of elements in* B *which are integral over* A *form a subring* $\bar{A}$ *of* B *containing* A.

Definition. The subring $\bar{A}$ is called the *integral closure* of A in B.

Definition. A ring A is called *integrally closed* if it coincides with its integral closure in the quotient field $C(A)$ of A.

Lemma 5.4. *A unique factorization domain* A *is integrally closed.*

Proof. Let $x \in C(A)$ be integral over A. Then $x = \alpha/\beta$ where we may assume that α and β are relatively prime. Now x satisfies an equation

$$x^n + a_1x^{n-1} + \ldots + a_n = 0$$

so

$$\alpha^n + a_1\alpha^{n-1}\beta + \ldots + a_n\beta^n = 0.$$

If β has a prime factor γ, then γ must divide α^n and therefore γ divides α. This contradicts the fact that α and β are relatively prime. Hence $x \in A$.

In particular, the symmetric algebra $S(V)$ over a finite-dimensional vector space V is integrally closed.

Theorem 5.5. *Let K be an algebraically closed field and A and B finitely generated algebras (with identity) over K. Suppose that $A \subset B$ and that B is integral over A. Then*

(i) *Each homomorphism $\varphi : A \to K$ extends to a homomorphism ψ of B into K.*

(ii) *If each homomorphism $\varphi : A \to K$ extends uniquely to a homomorphism of B into K then the quotient fields $C(A)$ and $C(B)$ coincide.*

Proof of (i). If $\varphi(A) = \{0\}$ we define ψ by $\psi(B) = \{0\}$. If $\varphi(A) \neq \{0\}$ then the kernel of φ is a maximal ideal $\mathfrak{m}$ of A. We first prove the existence of a maximal ideal $\mathfrak{n}$ of B such that $\mathfrak{n} \cap A = \mathfrak{m}$. For this purpose consider the set of all ideals $\mathfrak{p}$ of B satisfying $\mathfrak{p} \cap A \subset \mathfrak{m}$. This set is partially ordered under inclusion and every (totally) ordered subset has an upper bound. By Zorn's lemma there exists a maximal element $\mathfrak{n}$ of the set. If $\mathfrak{n} \cap A$ were a proper subset of $\mathfrak{m}$ let x be an element in $\mathfrak{m}$ which does not belong to $\mathfrak{n}$. Then $\mathfrak{n}$ is a proper subset of the ideal $\mathfrak{n} + Bx$; by the choice of $\mathfrak{n}$, $(\mathfrak{n} + Bx) \cap A$ is not a part of $\mathfrak{m}$. In other words, there exists an element $z \in B$ and an element $y \in A$, not in $\mathfrak{m}$, such that $zx - y \in \mathfrak{n}$. Now z satisfies an equation $z^n + a_1 z^{n-1} + \dots + a_n = 0$, $a_i \in A$, which after multiplication by x^n and use of the congruence $zx \equiv y \pmod{\mathfrak{n}}$ gives

$$y^n + a_1 x y^{n-1} + \dots + a_n x^n \equiv 0 \pmod{\mathfrak{m}}.$$

This contradicts $x \in \mathfrak{m}$, $y \notin \mathfrak{m}$. Hence $\mathfrak{n} \cap A = \mathfrak{m}$.

If $\mathfrak{n}$ were not a maximal ideal in B, suppose $\mathfrak{n}'$ is a maximal ideal in B satisfying the proper inclusions $\mathfrak{n} \subset \mathfrak{n}' \subset B$. Then $\mathfrak{n}' \cap A$ is an ideal in A properly containing $\mathfrak{m}$ so $A \subset \mathfrak{n}'$. Let $x \in B$. Since x is integral over $\mathfrak{n}'$ we have $x^l \in \mathfrak{n}'$ for some integer l. Since $B/\mathfrak{n}'$ is a field we conclude that $x \in \mathfrak{n}'$. Thus $\mathfrak{n}' = B$ which is a contradiction. This proves that $\mathfrak{n}$ is a maximal ideal in B so $B/\mathfrak{n}$ is a field. Let 1 denote the identity of A. Then the mapping $\alpha \to \alpha \cdot 1$ $(\alpha \in K)$ is an isomorphism of K into A and the mapping $\alpha \to \alpha \cdot 1 + \mathfrak{n}$ is an isomorphism of K into $B/\mathfrak{n}$. Now select $b_1, \dots, b_n \in B$ such that $B = K[b_1, \dots, b_n]$. Then $B = A[b_1, \dots, b_n]$ so by Lemma 5.2, B is a finite A-module. Writing $B = Ax_1 + \dots + Ax_m$ $(x_i \in B)$ and using $A = K + \mathfrak{m}$ we see that $B/\mathfrak{n}$ is a finite extension of K. Since K is algebraically closed we have $B/\mathfrak{n} = K$. The natural mapping $\psi : B \to K$ gives the desired extension of φ.

Proof† of (ii). In order to prove (ii) we make use of the following theorem (the Noether normalization theorem, see, e.g., Zariski-Samuel [1], Vol. I, p. 266).

† This proof was kindly communicated to the author by A. Mattuck.

Let $R = k[y_1, ..., y_n]$ be a finitely generated algebra over a field k and let d be the transcendence degree of the quotient field $k(y_1, ..., y_n)$ over k. There exist d linear combinations $z_1, ..., z_d$ of the y_i with coefficients in k, algebraically independent over k, such that R is integral over $k[z_1, ..., z_d]$.

Combining this theorem with (i) we see that if k is algebraically closed there exists a homomorphism of $k[y_1, ..., y_n]$ onto k.

Suppose now that the quotient fields $C(A)$ and $C(B)$ were different. Let $a_1, ..., a_m$ be a set of generators of A so $A = K[a_1, ..., a_m]$. Pick any element $\alpha \in B$ which does not belong to $C(A)$. We shall find a homomorphism of $K[a_1, ..., a_m]$ into K which has more than one extension to a homomorphism of $K[a_1, ..., a_m, \alpha]$ into K. Since B is integral over $K[a_1, ..., a_m, \alpha]$, (ii) will then follow from (i). Let

$$p_\alpha(x) = x^n + f_1(a)\, x^{n-1} + ... + f_n(a) = 0$$

be the polynomial with coefficients in the field $C(A) = K(a_1, ..., a_m)$ of lowest degree having α as a zero and leading coefficient 1. (Here a stands for the m-tuple $(a_1, ..., a_m)$). Let $q(a)$ denote the product of all the denominators of all the $f_i(a)$ with the discriminant of the polynomial $p_\alpha(x)$. From the remark above we see that there exists a homomorphism φ of $K[a_1, ..., a_m, 1/q(a)]$ onto K. The image of the polynomial $p_\alpha(x)$ under φ will then be a polynomial with coefficients in K having n distinct roots, say $\alpha_1, ..., \alpha_n$. Fix one α_i. We wish to extend φ (or more precisely, the restriction of φ to $K[a_1, ..., a_m]$) to a homomorphism $\psi_i : K[a_1, ..., a_m, \alpha] \to K$ by putting $\psi_i(\alpha) = \alpha_i$. The condition for this being possible is that whenever α satisfies a polynomial equation $p(a, x) = 0$ with coefficients in $K[a_1, ..., a_m]$ then α_i satisfies the corresponding equation $p(\varphi(a), x) = 0$ with coefficients in K. Since the polynomial $p_\alpha(x)$ has minimum degree it divides $p(a, x)$:

$$p(a, x) = p_\alpha(x)\, q(a, x).$$

Here the polynomial $q(a, x)$ can be found by long division; its coefficients $g_i(a)$ are rational expressions in $a_1, ..., a_m$ whose denominators divide the product of the denominators of the $f_i(a)$. Since φ does not map this product into 0 it is clear that $p(\varphi(a), x)$ vanishes for $x = \alpha_i$. Now $n > 1$ and the homomorphisms ψ_i $(1 \leqslant i \leqslant n)$, are all different. This concludes the proof.

Theorem 5.6. *Let H be a compact group of linear transformations of a finite-dimensional vector space V over a field K, $(K = \boldsymbol{R}$ or $\boldsymbol{C})$. Then the set $I(V)$ of all polynomials $P \in S(V)$ which are invariant under all $h \in H$ is a finitely generated algebra over K.*

Proof. Let $I_+(V)$ denote the set of invariants without constant term and let J denote the ideal in $S(V)$ generated by $I_+(V)$. By Hilbert's basis theorem the ideal J has a finite basis. It follows that there exist finitely many homogeneous invariants $j_1, ..., j_s \in I_+(V)$ such that each homogeneous invariant $j \in I_+(V)$ can be written

$$j = p_1 j_1 + ... + p_s j_s, \tag{2}$$

where all $p_k \in S(V)$ are homogeneous and degree $(p_k) =$ degree (j) − degree (j_k). Applying the linear transformation h and integrating over the compact group H we get

$$j = i_1 j_1 + ... + i_s j_s,$$

where $i_k = \int_H (h \cdot p_k)\, dh$. Applying (2) to the homogeneous invariant i_k we obtain by induction on degree (j), $j \in K[j_1, ..., j_s]$. This proves the theorem.

§ 6. The Formula for the Spherical Function

1. The Euclidean Type

Let M be a Euclidean space of dimension m and let G stand for $I_0(M)$, the largest connected group of isometries of M. Let K denote the compact subgroup of G leaving the origin $o \in M$ fixed. Then $M = G/K$. The group of all translations of M will also be denoted by M. Then M is a normal subgroup of G.

Proposition 6.1. *The spherical functions on G/K are the functions*

$$\varphi_\nu(x) = \int_K \exp(\nu(kxk^{-1}))\, dk, \qquad x \in M,$$

where ν is an arbitrary complex-valued linear function on M. (The product kxk^{-1} is the product in the group G.)

Proof. First let $m > 1$. In this case $D(G/K)$ consists of all polynomials in the Laplacian Δ on M. The spherical functions on M are the eigenfunctions of Δ which have value 1 at o, and are invariant under the action $x \rightarrow kxk^{-1}$ of K on M. Let $(x_1, ..., x_m)$ be coordinates on M with respect to an orthonormal basis (X_i) and write $\nu(X) = \nu_1 x_1 + ... + \nu_m x_m$ if $X = \sum x_i X_i$. Then $\Delta e^\nu = (\nu_1^2 + ... + \nu_m^2)\, e^\nu$ and since Δ is invariant under the isometry $x \rightarrow kxk^{-1}$ of M we have

$$\Delta\varphi_\nu = (\nu_1^2 + ... + \nu_m^2)\, \varphi_\nu.$$

In addition, $\varphi_\nu(o) = 1$ and $\varphi_\nu(kxk^{-1}) = \varphi_\nu(x)$, so φ_ν is spherical. On the other hand, since a spherical function is determined by its system of eigenvalues (Cor. 3.3), the φ_ν give all the spherical functions.

If $m = 1$, then $G = M$, $K = \{e\}$ and $\boldsymbol{D}(G/K)$ consists of all polynomials in d/dx. In this case the proposition is obvious.

2. The Compact Type

Let (G, K) be a Riemannian symmetric pair, G compact. We shall express the spherical functions on G/K by means of characters of irreducible representations of G.

Let φ be a spherical function on G. Then $\varphi(e) = 1$, $\varphi(kgk') = \varphi(g)$, $(k, k' \in K, g \in G)$, and

$$D\varphi = \lambda_D \varphi \tag{1}$$

for each $D \in \boldsymbol{D}_0(G)$, λ_D being a complex number. For each $g \in G$, the left translate $\varphi^{L(g)}$ of φ is also a solution of (1). Let V_φ denote the vector space formed by all finite complex linear combinations of left translates of φ.

Lemma 6.2. *The space V_φ has finite dimension.*

An elementary proof of this lemma is indicated in Exercise D. 2 following this chapter. We shall now give a different proof making use of some facts about the Laplace-Beltrami operator Δ on a compact Riemannian manifold M (see de Rham [2]). Let $m = \dim M$ and let d denote the distance function on M. In view of Prop. 2.1, the operator Δ is symmetric, that is

$$\int_M (\Delta f)\, g\, dx = \int_M (\Delta g)\, f\, dx, \qquad f, g \in C^\infty(M),$$

if dx denotes the Riemannian measure on M. It follows that the Poisson equation $\Delta u = f$ where $f \in C^\infty(M)$ can not have a solution unless $\int f\, dx = 0$. On the other hand, if this condition is satisfied there exists a solution u. Suppose u_1 and u_2 are two solutions. Then the function

$$v = u_1 - u_2 - \frac{1}{\operatorname{vol}(M)} \int_M (u_1 - u_2)\, dx$$

satisfies

$$\int_M v\, dx = 0, \qquad \int_M v(\Delta f)\, dx = 0$$

for each $f \in C^\infty(M)$.

Hence we have for all $g \in C^\infty(M)$,

$$\int_M v\, g\, dx = 0,$$

so $v = 0$. The solution to $\Delta u = f$ is therefore unique up to an additive constant. Hence, if we consider the subspace consisting of all C^∞ functions satisfying $\int f\, dx = 0$ the Laplace-Beltrami operator Δ is a one-to-one linear mapping of this subspace onto itself. The inverse operator, denoted G, is of the form

$$[Gf](x) = \int_M g(x, y) f(y)\, dy,$$

where the kernel $g(x, y)$, (singular for $x = y$), has the following properties.

(i) If D denotes the diagonal in $M \times M$ then $g \in C^\infty(M \times M - D)$.

(ii) $d(x, y)^{m-2} g(x, y)$ is bounded. (If $m = 2$ this property should be replaced by "$(\log d(x, y))^{-1} g(x, y)$ is bounded").

(iii) $g(x, y) = g(y, x)$.

Lemma 6.3. *Let Δ be the Laplace-Beltrami operator on a compact Riemannian manifold M. The vector space of eigenfunctions $u \in C^\infty(M)$ of Δ for a given eigenvalue λ is of finite dimension.*

Proof. It suffices to consider an eigenvalue $\lambda \neq 0$. Let V denote the corresponding space of eigenfunctions. Then $\int u(x)\, dx = 0$ for $u \in V$. Let $u_1, \ldots, u_n$ be any set of functions in V, orthonormal with respect to the scalar product

$$\langle f_1, f_2 \rangle = \int_M f_1(x) f_2(x)\, dx$$

in $C^\infty(M)$. Assuming, as we may, that $\lambda \neq 0$, let

$$g^*(x, y) = \frac{1}{\lambda} \sum_{i=1}^{n} u_i(x)\, u_i(y)$$

and consider the integral

$$c = \iint (g(x, y) - g^*(x, y))^2\, dx\, dy,$$

ignoring for the moment the question of convergence. Since $Gu_i = \lambda^{-1} u_i$ $(1 \leqslant i \leqslant n)$ we obtain, using (ii) and (iii),

$$c = \iint (g(x, y))^2\, dx\, dy - \frac{n}{\lambda^2}.$$

But $c \geqslant 0$ and consequently

$$n \leqslant \lambda^2 \iint (g(x, y))^2 \, dx \, dy.$$

However, this integral is in general infinite. The difficulty can be overcome by replacing $g(x, y)$ by a suitably iterated kernel. This method is based on the following fact: Let $e_1 > 0$, $e_2 > 0$. Then the function

(iv) $d(x, y)^{m-e_1-e_2} \int_M d(x, z)^{e_1-m} \, d(z, y)^{e_2-m} \, dz$ is bounded.

This can be seen by breaking the integral up into two parts, namely, over the ball $B_{2d(x,y)}(x)$ and over the complement. (For the details, see de Rham [2], pp. 141-142.) In view of (ii) we can therefore form the iterated kernel

$$g_p(x, y) = \int_M g(x, z_1) \, dz_1 \dots \int_M g(z_{p-2}, z_{p-1}) \, dz_{p-1} \int_M g(z_{p-1}, z_p) \, g(z_p, y) \, dz_p$$

which (by (iv)) is bounded provided $2(p+1) > m$. The operator G^p has kernel $g_p(x, y)$, that is,

$$[G^p f](x) = \int_M g_p(x, y) f(y) \, dy \qquad \left(\int f(x) \, dx = 0\right)$$

and the functions $u_1, \cdots, u_n$ are eigenfunctions of G^p for the eigenvalue λ^{-p}. Repeating the computation above with $g(x, y)$ replaced by $g_p(x, y)$ and $g^*(x, y)$ replaced by

$$g_p^*(x, y) = \frac{1}{\lambda^p} \sum_{i=1}^{n} u_i(x) \, u_i(y)$$

one finds

$$n \leqslant \lambda^{2p} \iint (g_p(x, y))^2 \, dx \, dy,$$

which gives an upper bound for the dimension of V.

Returning now to the space V_φ we first observe that each vector in V_φ is an eigenfunction of each $D \in D_0(G)$ with the same eigenvalue λ_D. If we introduce a bi-invariant Riemannian structure on G the corresponding Laplace-Beltrami operator belongs to $D_0(G)$. From Lemma 6.3 it is clear that V_φ is of finite dimension.

Consider now the representation $g \rightarrow \pi(g)$ of G on V_φ defined by $\pi(g) \psi = \psi^{L(g)}$ for $g \in G$, $\psi \in V_\varphi$. Then $\pi(k) \varphi = \varphi$ for all $k \in K$. Since

G is compact, a scalar product can be defined on V_φ such that $\pi(g)$ is unitary for each $g \in G$.

Lemma 6.4. *The representation π is irreducible.*

Proof. Since V_φ is finite-dimensional the operator $P = \int_K \pi(k)\, dk$ is well defined. Moreover,

$$[P\varphi^{L(g)}](x) = \int_K [\pi(k)\, \varphi^{L(g)}](x)\, dk = \int_K \varphi(g^{-1}k^{-1}x)\, dk = \varphi(g^{-1})\, \varphi(x),$$

so $PV_\varphi = C\varphi$ and $P^2 = P$. Let $V_\varphi = U_1 + \ldots + U_n$ be a direct decomposition of V_φ into subspaces which are invariant and irreducible under π. For a suitable i, we have $PU_i \neq \{0\}$. Then $PU_i \subset PV_\varphi = C\varphi$ so $\varphi \in PU_i \subset U_i$. Hence $\pi(g)\, \varphi \in U_i$ for $g \in G$ so $U_i = V_\varphi$ and π is irreducible.

The mapping $P\pi(g)\, P$ maps V_φ onto $C\varphi$ and maps the vector φ into $\varphi(g^{-1})\, \varphi$. Hence

$$\varphi(g^{-1}) = \operatorname{Tr}(P\pi(g)\, P) = \operatorname{Tr}(\pi(g)\, P) = \operatorname{Tr} \int_K \pi(gk)\, dk = \int_K \chi(gk)\, dk$$

if χ denotes the character of the representation π.

Theorem 6.5. *The spherical functions on G are precisely the functions of the form*

$$\varphi(g) = \int_K \chi(g^{-1}k)\, dk,$$

where χ is the character of a finite-dimensional representation π of G of class 1. *Here φ is positive definitie and π is the representation associated to φ.*

Proof. It remains to prove that if π is a finite-dimensional representation of G on a Hilbert space $\mathfrak{H}$ such that π is of class 1, then the integral above is a spherical function. Let e be a unit vector in $\mathfrak{H}$ which is left fixed by each $\pi(k)$, $k \in K$. Then if we put $\psi(g) = \langle e, \pi(g)\, e\rangle$ we know from Theorem 4.5 that ψ is spherical. Now put as before

$$P = \int_K \pi(k)\, dk.$$

Then $P^2 = P$ so

$$\operatorname{Tr}(P\pi(g)\, P) = \operatorname{Tr}(\pi(g)\, P) = \int_K \operatorname{Tr} \pi(gk)\, dk = \int_K \chi(gk)\, dk \qquad (2)$$

if χ denotes the character of π. Now for each $x \in G$, the vector $P\pi(x)\,e$ is left fixed by each $\pi(k)$, $k \in K$. Using Lemma 4.7 we conclude that

$$P\mathfrak{H} = Ce.$$

Let $e_1 = e, e_2, \dots, e_n$ be an orthonormal basis of $\mathfrak{H}$. Since $Pe_i = 0$ for $i \geqslant 2$ it follows that

$$\operatorname{Tr}(P\pi(g)\,P) = \langle \pi(g)\,e, e\rangle = \psi(g^{-1}). \tag{3}$$

From (2) and (3) it now follows that the integral $\int_K \chi(g^{-1}k)\,dk$ is a spherical function.

3. The Noncompact Type

Let (G, K) be a Riemannian symmetric pair of the noncompact type. We assume that G has finite center so K is compact. The meaning of the symbols $\mathfrak{g}_0$, $\mathfrak{k}_0$, $\mathfrak{p}_0$, $\mathfrak{h}_{\mathfrak{p}_0}$, $\mathfrak{g}$, $\mathfrak{k}$, $\mathfrak{p}$, $\mathfrak{h}_{\mathfrak{p}}$, $\mathfrak{h}$, Δ, P_+, $\mathfrak{n}$, and $\mathfrak{n}_0$ will be the same as in Chapter VI. We also make the following convention: if E and F are linear subspaces of an associative algebra, then EF denotes the vector space spanned by all elements of the form ef $(e \in E, f \in F)$. Regarding each element of $\mathfrak{g}_0$ as a left invariant differential operator on G, we have $\mathfrak{g}_0 \subset D(G)$. Let $\mathfrak{H}_{\mathfrak{p}_0}$ denote the subalgebra of $D(G)$ generated by $\mathfrak{h}_{\mathfrak{p}_0}$ and the identity.

Lemma 6.6. *For any $D \in D(G)$ there exists a unique element $D_{\mathfrak{h}} \in \mathfrak{H}_{\mathfrak{p}_0}$ such that*

$$D - D_{\mathfrak{h}} \in \mathfrak{k}_0 D(G) + D(G)\,\mathfrak{n}_0.$$

Moreover, degree $(D_{\mathfrak{h}}) \leqslant$ degree (D).

Proof. Let $X_1, \dots, X_n$ be a basis of $\mathfrak{g}_0$ such that $X_i \in \mathfrak{k}_0$ for $1 \leqslant i \leqslant \dim \mathfrak{k}_0$, $X_i \in \mathfrak{h}_{\mathfrak{p}_0}$ for $\dim \mathfrak{k}_0 < i \leqslant \dim \mathfrak{k}_0 + \dim \mathfrak{h}_{\mathfrak{p}_0}$ and $X_i \in \mathfrak{n}_0$ for $\dim \mathfrak{k}_0 + \dim \mathfrak{h}_{\mathfrak{p}_0} < i \leqslant \dim \mathfrak{g}_0$. Then, due to Cor. 1.10, Chapter II, D can be expressed uniquely

$$D = \sum a_{e_1 \dots e_n} X_1^{e_1} \dots X_n^{e_n}. \tag{4}$$

The existence of $D_{\mathfrak{h}}$ follows at once. The uniqueness of $D_{\mathfrak{h}}$ means that $\mathfrak{H}_{\mathfrak{p}_0} \cap (\mathfrak{k}_0 D(G) + D(G)\,\mathfrak{n}_0) = \{0\}$ and this follows from the uniqueness of the representation (4) combined with the fact that $\mathfrak{k}_0$ and $\mathfrak{n}_0$ are subalgebras. The statement about the degrees is obvious.

Consider now the Iwasawa decomposition $G = KA_{\mathfrak{p}}N$. If $x \in G$, let $H(x)$ denote the unique element in $\mathfrak{h}_{\mathfrak{p}_0}$ for which $x = k \exp H(x)\,n$ where $k \in K$, $n \in N$. Let $D_0(G)$ and $D(G/K)$ be defined as in §2, No. 2

The algebra $\mathfrak{H}_{\mathfrak{p}_0}$ is just the symmetric algebra $S(\mathfrak{h}_{\mathfrak{p}_0})$ so each linear mapping $\nu : \mathfrak{h}_{\mathfrak{p}_0} \to C$ extends uniquely to a homomorphism $\chi_\nu : \mathfrak{H}_{\mathfrak{p}_0} \to C$ satisfying $\chi_\nu(1) = 1$.

Proposition 6.7. *For each linear function $\nu : \mathfrak{h}_{\mathfrak{p}_0} \to C$ the function*

$$\varphi(x) = \int_K e^{\nu(H(xk))}\, dk \qquad (x \in G),$$

is a spherical function on G.

Proof. It is obvious that $\varphi(e) = 1$ and since $H(kx) = H(x)$ for $k \in K$, φ is bi-invariant under K. Let $F(x) = e^{\nu(H(x))}$. Then since $H(xn) = H(x)$ for $n \in N$ it follows that

$$DF = 0 \qquad \text{for } D \in \boldsymbol{D}(G)\,\mathfrak{n}_0. \tag{5}$$

Secondly, since $H(x \exp H) = H(x) + H$ $(H \in \mathfrak{h}_{\mathfrak{p}_0})$, we have

$$DF = \chi_\nu(D)\, F \qquad \text{for } D \in \mathfrak{H}_{\mathfrak{p}_0}. \tag{6}$$

Moreover, if $f \in C^\infty(G)$, the integral $\int_K f(xkk_1)\, dk$ is independent of k_1; consequently, $\int [Tf]\,(xk)\, dk = 0$ for $T \in \mathfrak{k}_0$. In particular

$$\int_K [DF]\,(xk)\, dk = 0 \qquad \text{for } D \in \mathfrak{k}_0\boldsymbol{D}(G). \tag{7}$$

Now let $D \in \boldsymbol{D}_0(G)$. Then by Lemma 6.6, $D = D_1 + D_2 + D_{\mathfrak{h}}$ where $D_1 \in \boldsymbol{D}(G)\,\mathfrak{n}_0$, $D_2 \in \mathfrak{k}_0\boldsymbol{D}(G)$. Since $D^{R(k)} = D$, $(k \in K)$, we have $[DF^{R(k^{-1})}]\,(x) = [DF]\,(xk)$ so by (5)-(7)

$$D\varphi = \chi_\nu(D_{\mathfrak{h}})\,\varphi, \qquad D \in \boldsymbol{D}_0(G), \tag{8}$$

which proves the proposition.

As earlier, let $\rho = \frac{1}{2}\sum_{\alpha \in P_+} \alpha$ and let W denote the Weyl group of G/K, acting on $\mathfrak{h}_{\mathfrak{p}_0}$. *If $s \in W$* and ν is a linear function on $\mathfrak{h}_{\mathfrak{p}}$, we write $s\nu$ instead of ν^s.

Proposition 6.8. *For each linear function ν on $\mathfrak{h}_{\mathfrak{p}_0}$ let*

$$\varphi_\nu(x) = \int_K e^{(i\nu-\rho)(H(xk))}\, dk.$$

Then $\varphi_{s\nu} = \varphi_\nu$ for each $s \in W$.

Proof. It suffices to prove that

$$\int_G \varphi_{s\nu}(x) f(x)\, dx = \int_G \varphi_\nu(x) f(x)\, dx \tag{9}$$

for each $f \in C_c(G)$ and since φ_ν and $\varphi_{s\nu}$ are bi-invariant under K it suffices to prove (9) for all f bi-invariant under K. Using the formula $dx = e^{2\rho(\log a)}\, dk\, da\, dn$ from Prop. 1.11 we have

$$\int_G \varphi_\nu(x) f(x)\, dx = \int_K dk \int_G e^{(i\nu-\rho)(H(xk))} f(x)\, dx$$

$$= \int_G e^{(i\nu-\rho)(H(x))} f(x)\, dx = \int_{A_\mathfrak{p}} \int_N e^{(i\nu+\rho)(\log a)} f(an)\, da\, dn,$$

so

$$\int_G \varphi_\nu(x) f(x)\, dx = \int_{A_\mathfrak{p}} e^{i\nu(\log a)}\, F_f(a)\, da, \tag{10}$$

where

$$F_f(a) = e^{\rho(\log a)} \int_N f(an)\, dn.$$

Using now Theorem 1.15, formula (9) follows. Note that formula (10) connects the abstract Fourier transform $\tilde{f}$ of f with the classical Fourier transform of the function F_f.

Consider now the symmetric algebras $S(\mathfrak{p})$ and $S(\mathfrak{h}_\mathfrak{p})$. The group $\mathrm{Ad}_G(K)$ operates on $S(\mathfrak{p})$ and the Weyl group W operates on $S(\mathfrak{h}_\mathfrak{p})$. Let $I(\mathfrak{p})$ and $I(\mathfrak{h}_\mathfrak{p})$ be the corresponding sets of invariant polynomials. Let $S^*(\mathfrak{p})$ and $S^*(\mathfrak{h}_\mathfrak{p})$ denote the symmetric algebras over the dual spaces $\mathfrak{p}^\wedge$ and $(\mathfrak{h}_\mathfrak{p})^\wedge$. Since the Killing form B is nondegenerate on $\mathfrak{p} \times \mathfrak{p}$ there is a one-to-one linear mapping of $\mathfrak{p}$ onto $\mathfrak{p}^\wedge$ which sends $X \in \mathfrak{p}$ into the linear function $Y \to B(X, Y)$. This mapping extends uniquely to an isomorphism of $S(\mathfrak{p})$ onto $S^*(\mathfrak{p})$. Similarly, since B is nondegenerate on $\mathfrak{h}_\mathfrak{p} \times \mathfrak{h}_\mathfrak{p}$ we have an isomorphism of $S(\mathfrak{h}_\mathfrak{p})$ onto $S^*(\mathfrak{h}_\mathfrak{p})$. These isomorphisms being canoncial we shall identify $S(\mathfrak{p})$ and $S^*(\mathfrak{p})$, $S(\mathfrak{h}_\mathfrak{p})$ and $S^*(\mathfrak{h}_\mathfrak{p})$. In particular, $I(\mathfrak{p})$ and $I(\mathfrak{h}_\mathfrak{p})$ will be considered as polynomial functions on $\mathfrak{p}$ and $\mathfrak{h}_\mathfrak{p}$, respectively. We shall also consider the real symmetric algebras $S(\mathfrak{p}_0)$, $S(\mathfrak{h}_{\mathfrak{p}_0})$ and the corresponding sets of invariants $I(\mathfrak{p}_0)$ and $I(\mathfrak{h}_{\mathfrak{p}_0})$.

Lemma 6.9. *The ring $S(\mathfrak{h}_{\mathfrak{p}_0})$ is integral over $I(\mathfrak{h}_{\mathfrak{p}_0})$.*

Proof. Let $X \in \mathfrak{p}_0$ and let as usual T_X denote the restriction of $(\mathrm{ad}\, X)^2$ to $\mathfrak{p}_0$. In the characteristic polynomial for T_X

$$\det(\lambda I - T_X) = \lambda^r + p_1(X)\,\lambda^{r-1} + \ldots + p_{r-l}(X)\,\lambda^l$$

the coefficients p_i are invariant polynomial functions on $\mathfrak{p}_0$, i.e., $p_i \in I(\mathfrak{p}_0)$. From Lemma 2.9, Chapter VII, we know that the eigenvalues of T_H $(H \in \mathfrak{h}_{\mathfrak{p}_0})$ are 0 and $(\alpha(H))^2$ as α varies through P_+. Consequently

$$\det(\alpha(H)^2 I - T_H) = 0 \qquad \text{for all } H \in \mathfrak{h}_{\mathfrak{p}_0}.$$

This shows that for each $\alpha \in P_+$, the function $\alpha^2 \in S(\mathfrak{h}_{\mathfrak{p}_0})$ and hence α itself is integral over $I(\mathfrak{h}_{\mathfrak{p}_0})$. Since every linear function on $\mathfrak{h}_{\mathfrak{p}_0}$ is a linear combination of these α, the lemma follows (Cor. 5.3).

If f is a function on $\mathfrak{p}$ (or $\mathfrak{p}_0$), let $\bar{f}$ denote its restriction to $\mathfrak{h}_\mathfrak{p}$ (or $\mathfrak{h}_{\mathfrak{p}_0}$).

Theorem 6.10. *The mapping $p \to \bar{p}$ is an isomorphism of $I(\mathfrak{p})$ onto $I(\mathfrak{h}_\mathfrak{p})$.*

This theorem is due to Chevalley (unpublished, cf. Harish-Chandra [12], I, §3). A proof will be given at the end of this section. At present we shall use the theorem to show that all the spherical functions on G are given by the formula in Prop. 6.7.

Since the linear function ρ is real-valued on $\mathfrak{h}_{\mathfrak{p}_0}$ there exists a unique automorphism $p \to {}'p$ of $\mathfrak{H}_{\mathfrak{p}_0} = S(\mathfrak{h}_{\mathfrak{p}_0})$ such that ${}'H = H - \rho(H)$ for $H \in \mathfrak{h}_{\mathfrak{p}_0}$. We now construct a mapping γ of $\boldsymbol{D}_0(G)$ into $S(\mathfrak{h}_{\mathfrak{p}_0})$ as follows. If $D \in \boldsymbol{D}_0(G)$, let $\gamma(D) = {}'D_\mathfrak{h}$.

Lemma 6.11. *The mapping γ is a homomorphism.*

Proof. Let $D', D'' \in \boldsymbol{D}_0(G)$. Then

$$D'D'' - D'_\mathfrak{h} D''_\mathfrak{h} = D'(D'' - D''_\mathfrak{h}) + (D' - D'_\mathfrak{h})\, D''_\mathfrak{h},$$

where $D'' - D''_\mathfrak{h}$ and $D' - D'_\mathfrak{h}$ belong to $\mathfrak{k}_0\boldsymbol{D}(G) + \boldsymbol{D}(G)\,\mathfrak{n}_0$. Since $TD' = D'T$ for $T \in \mathfrak{k}_0$ we have

$$D'(D'' - D''_\mathfrak{h}) \in \mathfrak{k}_0\boldsymbol{D}(G) + \boldsymbol{D}(G)\,\mathfrak{n}_0$$

and since $[\mathfrak{n}_0, \mathfrak{h}_{\mathfrak{p}_0}] \subset \mathfrak{n}_0$,

$$(D' - D'_\mathfrak{h})\, D''_\mathfrak{h} \in \mathfrak{k}_0\boldsymbol{D}(G) + \boldsymbol{D}(G)\,\mathfrak{n}_0.$$

Hence $(D'D'')_\mathfrak{h} = D'_\mathfrak{h} D''_\mathfrak{h}$ from which the lemma follows.

Let λ denote the symmetrization mapping of $S(\mathfrak{g}_0)$ onto $\boldsymbol{D}(G)$. If $p \in I(\mathfrak{p}_0)$ then $\lambda(p) \in \boldsymbol{D}_0(G)$ by Lemma 2.4.

Lemma 6.12. *For each nonconstant $p \in I(\mathfrak{p}_0)$*

$$\textit{degree}\,(\gamma(\lambda(p)) - \bar{p}) < \textit{degree}\ p. \tag{11}$$

Proof. Let d denote the degree of p; we may assume that p is homogeneous. The restriction $\bar{p}$ can be regarded as an element of $S(\mathfrak{p}_0)$ and since $\mathfrak{g}_0 = \mathfrak{k}_0 + \mathfrak{h}_{\mathfrak{p}_0} + \mathfrak{n}_0$ it is clear that

$$p - \bar{p} \in \mathfrak{k}_0 S^{d-1}(\mathfrak{g}_0) + S^{d-1}(\mathfrak{g}_0)\, \mathfrak{n}_0,$$

where $S^e(\mathfrak{g}_0)$ denotes the set of homogeneous elements in $S(\mathfrak{g}_0)$ of degree e. Now $\lambda\bar{p} = \bar{p}$ and from the definition of λ,

$$\lambda(q_1 q_2) - \lambda(q_1)\,\lambda(q_2) \in \sum_{e < d_1 + d_2} \lambda(S^e(\mathfrak{g}_0)),$$

if $q_1 \in S^{d_1}(\mathfrak{g}_0)$, $q_2 \in S^{d_2}(\mathfrak{g}_0)$. It follows that

$$\lambda(p) - \bar{p} \in \mathfrak{k}_0 \boldsymbol{D}(G) + \boldsymbol{D}(G)\, \mathfrak{n}_0 + \sum_{e<d} \lambda(S^e(\mathfrak{g}_0)).$$

The inequality (11) now follows from Lemma 6.6 and the definition of γ.

Lemma 6.13. *The mapping γ has kernel $\boldsymbol{D}_0(G) \cap \boldsymbol{D}(G)\, \mathfrak{k}_0$.*

Proof. Let $D \in \boldsymbol{D}_0(G) \cap \boldsymbol{D}(G)\, \mathfrak{k}_0$. Then it follows from (8) that $\chi_\nu(D_{\mathfrak{h}}) = 0$ for each linear function ν on $\mathfrak{h}_{\mathfrak{p}_0}$. This implies that $D_{\mathfrak{h}} = 0$ so $\gamma(D) = 0$. On the other hand, let D be in the kernel of γ. In view of Cor. 2.6, we have $D = \lambda(p) + D'$ where $D' \in \boldsymbol{D}_0(G) \cap \boldsymbol{D}(G)\, \mathfrak{k}_0$ and $p \in I(\mathfrak{p}_0)$. Thus $\gamma(\lambda(p)) = 0$. By decomposing p into homogeneous components it is clear that the restriction $\bar{p}$ has the same degree as p. Thus Lemma 6.12 implies that $p = 0$ so $D \in \boldsymbol{D}_0(G) \cap \boldsymbol{D}(G)\, \mathfrak{k}_0$ as desired.

Lemma 6.14. *The image $\gamma(\boldsymbol{D}_0(G))$ is $I(\mathfrak{h}_{\mathfrak{p}_0})$.*

Proof. Let $D \in \boldsymbol{D}_0(G)$ and consider the function φ_ν from Prop. 6.8. From relation (8) and the definition of γ it follows that

$$D\varphi_\nu = \gamma(D)\,(i\nu)\,\varphi_\nu, \tag{12}$$

where $\gamma(D)\,(i\nu)$ is the value of the polynomial function $\gamma(D) \in S(\mathfrak{h}_{\mathfrak{p}})$ at $i\nu \in \mathfrak{h}_{\mathfrak{p}}$. Using Prop. 6.8 it follows that $\gamma(D)^s = \gamma(D)$ for $s \in W$ so $\gamma(D) \in I(\mathfrak{h}_{\mathfrak{p}_0})$. In order to see that the image under γ is all of $I(\mathfrak{h}_{\mathfrak{p}_0})$ we make use of Theorem 6.10. Let q be an arbitrary homogeneous polynomial in $I(\mathfrak{h}_{\mathfrak{p}_0})$ and let d denote its degree. We prove by induction on d that there exists a $D \in \boldsymbol{D}_0(G)$ such that $\gamma(D) = q$. This is obvious if $d = 0$ so suppose that $d > 0$. Let p be the polynomial in $I(\mathfrak{p}_0)$ satisfying $\bar{p} = q$. By Lemma 6.12, the polynomial $\gamma(\lambda(p)) - \bar{p}$ has degree $< d$ and by Lemma 6.14 it is invariant. By the inductive assumption there exists a $D_1 \in \boldsymbol{D}_0(G)$ such that $\gamma(D_1) = \gamma(\lambda(p)) - \bar{p}$. But then the operator $D = \lambda(p) - D_1$ lies in $\boldsymbol{D}_0(G)$ and satisfies $\gamma(D) = q$.

The four last lemmas show that γ is a homomorphism of $\boldsymbol{D}_0(G)$ onto $I(\mathfrak{h}_{\mathfrak{p}_0})$ with kernel $\boldsymbol{D}_0(G) \cap \boldsymbol{D}(G)\,\mathfrak{k}_0$. Passing to the quotient we get an isomorphism γ^* of $\boldsymbol{D}_0(G)/(\boldsymbol{D}_0(G) \cap \boldsymbol{D}(G)\,\mathfrak{k}_0)$ onto $I(\mathfrak{h}_{\mathfrak{p}_0})$. On the other hand, two operators in $\boldsymbol{D}_0(G)$ have the same restriction to $C_0^\infty(G)$ if and only if their difference lies in $\boldsymbol{D}_0(G) \cap \boldsymbol{D}(G)\,\mathfrak{k}_0$ (Cor. 2.6). Thus Lemma 2.2 gives an isomorphism μ of $\boldsymbol{D}_0(G)/(\boldsymbol{D}_0(G) \cap \boldsymbol{D}(G)\,\mathfrak{k}_0)$ onto $\boldsymbol{D}(G/K)$. Putting $\Gamma = \gamma^* \circ \mu^{-1}$ we have:

Theorem 6.15. *The mapping Γ is an isomorphism of $\boldsymbol{D}(G/K)$ onto $I(\mathfrak{h}_{\mathfrak{p}_0})$.*

In particular, the operators in $\boldsymbol{D}(G/K)$ all commute, as shown earlier (Theorem 2.9). A theorem of Chevalley [7] states that if a finite group W of linear transformations of an n-dimensional vector space is generated by reflections then the algebra of polynomials invariant under W is generated by n algebraically independent homogeneous polynomials and the identity. Since the Weyl group acting on $\mathfrak{h}_{\mathfrak{p}_0}$ is generated by reflections, the theorem above shows that the algebra $\boldsymbol{D}(G/K)$ is generated by l algebraically independent elements, where $l = \text{rank}\,(G/K)$.

Using the lemmas above about γ it is now possible to prove that all the spherical functions on G are given by linear functions on $\mathfrak{h}_{\mathfrak{p}_0}$.

Theorem 6.16. (Harish-Chandra) *Let $\rho = \frac{1}{2} \sum_{\alpha \in P_+} \alpha$. The functions*

$$\varphi_\nu(x) = \int_K e^{(i\nu - \rho)(H(xk))}\, dk, \qquad x \in G,$$

where ν is a complex-valued linear function on $\mathfrak{h}_{\mathfrak{p}_0}$ exhaust the class of spherical functions on G. Moreover, two such functions φ_ν and φ_λ are identical if and only if $\nu = \lambda^s$ for some s in the Weyl group.

Proof. From Prop. 6.7 we know that each φ_ν is a spherical function. Now let ψ be an arbitrary spherical function on G/K. Then for $D \in \boldsymbol{D}(G/K)$ we have $D\psi = \lambda_D \psi$ and the mapping $D \to \lambda_D$ is a homomorphism of $\boldsymbol{D}(G/K)$ into $\boldsymbol{C}$. Owing to Theorem 6.15, this corresponds to a homomorphism of the algebra $I(\mathfrak{h}_{\mathfrak{p}_0})$ into $\boldsymbol{C}$. This homomorphism extends to a homomorphism $\chi : S(\mathfrak{h}_{\mathfrak{p}_0}) \to \boldsymbol{C}$ because $S(\mathfrak{h}_{\mathfrak{p}_0})$ is integral over $I(\mathfrak{h}_{\mathfrak{p}_0})$ (Lemma 6.9). The homomorphism χ is obviously the extension of a linear function μ on $\mathfrak{h}_{\mathfrak{p}_0}$. Since a spherical function is determined by the corresponding system of eigenvalues (Cor. 3.3), the function

$$\int_K e^{(\mu - \rho)(H(xk))}\, dk$$

must coincide with $\tilde{\psi}(x)$. Next suppose $\varphi_\nu = \varphi_\lambda$. Then by (9), $\gamma(D)\,(i\nu) = \gamma(D)\,(i\lambda)$ for all $D \in \boldsymbol{D}_0(G)$. In view of Theorem 6.15 this means that each polynomial $p \in I(\mathfrak{h}_{\mathfrak{p}_0})$ takes the same value at ν and λ. If ν and λ were not conjugate under the Weyl group there exists a polynomial $p \in S(\mathfrak{h}_{\mathfrak{p}_0})$ such that $p > 0$ on the orbit ν^s $(s \in W)$ and $p < 0$ on the orbit λ^s $(s \in W)$. But then the polynomial $p^* = \Sigma_{s \in W} p^s$ belongs to $I(\mathfrak{h}_{\mathfrak{p}_0})$ but takes different values at ν and λ. This concludes the proof.

We shall now prove Theorem 6.10. The proof which follows was generously put at our disposal by Harish-Chandra.

Lemma 6.17. *If $p_1, p_2 \in I(\mathfrak{p})$ and $p_1 = p_2 q$ where $q \in S(\mathfrak{p})$, then $q \in I(\mathfrak{p})$.*

Proof. If $q_0 = \int_K q^{\mathrm{Ad}(k)}\, dk$, then $p_1 = p_2 q_0$ so $q = q_0 \in I(\mathfrak{p})$.

Lemma 6.18. *The ring $I(\mathfrak{p})$ is integrally closed.*

Proof. Let x be an element in the quotient field $C(I(\mathfrak{p}))$ of $I(\mathfrak{p})$. Suppose x is integral over $I(\mathfrak{p})$; then x, considered as an element of $C(S(\mathfrak{p}))$ is integral over $S(\mathfrak{p})$. But $S(\mathfrak{p})$ is integrally closed, being a unique factorization domain (Lemma 5.4). Hence $x \in S(\mathfrak{p})$, so by Lemma 6.17, $x \in I(\mathfrak{p})$, proving that $I(\mathfrak{p})$ is integrally closed.

Let $J \subset I(\mathfrak{h}_{\mathfrak{p}})$ denote the image of $I(\mathfrak{p})$ under the restriction mapping $p \to \bar{p}$. Since every vector $X \in \mathfrak{p}_0$ can be rotated into $\mathfrak{h}_{\mathfrak{p}_0}$ by some $\mathrm{Ad}\,(k)$ $(k \in K)$ (Lemma 6.3, Chapter V), it follows that the restriction mapping $p \to \bar{p}$ is an isomorphism of $I(\mathfrak{p})$ into $I(\mathfrak{h}_{\mathfrak{p}})$. In particular, J is integrally closed. Owing to Theorem 5.6, $I(\mathfrak{p})$ and J have a finite number of generators.

Lemma 6.19. *The ring $S(\mathfrak{h}_{\mathfrak{p}})$ is integral over J.*

This lemma is contained in the proof of Lemma 6.9. One just has to remark that for $H \in \mathfrak{h}_{\mathfrak{p}}$, T_H has characteristic polynomial

$$\det\,(\lambda^2 - T_H) = \lambda^{2r} + p_1(H)\,\lambda^{2r-2} + \ldots + p_{r-l}(H)\,\lambda^{2l} \tag{13}$$

where $p_i \in J$, $(1 \leqslant i \leqslant r - l)$.

Lemma 6.20. *Let $H_0 \in \mathfrak{h}_{\mathfrak{p}_0}$ and $H_1 \in \mathfrak{h}_{\mathfrak{p}}$ and suppose that $p(H_0) = p(H_1)$ for all $p \in J$. Then $H_1 = sH_0$ for some $s \in W$.*

Proof. It is clear from Lemma 2.9, Chapter VII, that the roots of the polynomial (13) are 0 and $\pm\, \alpha(H)$ $(\alpha \in P_+)$. Since

$$\det\,(\lambda^2 - T_{H_0}) \equiv \det\,(\lambda^2 - T_{H_1})$$

it follows that $\alpha(H_1)$ is real for each $\alpha \in P_+$ so $H_1 \in \mathfrak{h}_{\mathfrak{p}_0}$. Next we prove that H_0 and H_1 are conjugate under $\mathrm{Ad}_G\,(K)$. In fact, suppose this were not the case. Then there exists a real-valued continuous function

f on $\mathfrak{p}_0$ identically 0 on the compact orbit $\mathrm{Ad}_G(K)\,H_0$ and identically 1 on the compact orbit $\mathrm{Ad}_G(K)\,H_1$. Owing to the Weierstrass approximation theorem there exists a polynomial $p \in S(\mathfrak{p}_0)$ such that

$$|p(H) - 0| < \tfrac{1}{3} \qquad \text{for} \qquad H \in \mathrm{Ad}_G(K)\,H_0,$$
$$|p(H) - 1| < \tfrac{1}{3} \qquad \text{for} \qquad H \in \mathrm{Ad}_G(K)\,H_1.$$

Then the polynomial

$$p^* = \int_K p^{\mathrm{Ad}(k)}\,dk$$

belongs to $I(\mathfrak{p}_0)$ but has different values at H_0 and H_1. This contradiction shows that $\mathrm{Ad}(k)\,H_0 = H_1$ for some $k \in K$. Using Prop. 2.2, Chapter VII, we conclude that $sH_0 = H_1$ for some $s \in W$, proving the lemma.

Consider now the polynomial

$$F(\lambda) = \lambda^{2r} + p_1\lambda^{2r-2} + \dots + p_{r-l}\lambda^{2l}.$$

Its coefficients lie in J and its roots are 0 and $\pm\,\bar{\alpha}$ (if $\bar{\alpha}$ denotes the restriction of $\alpha \in P_+$ to $\mathfrak{h}_\mathfrak{p}$). Since the $\bar{\alpha}$ ($\alpha \in P_+$) span the space $\mathfrak{h}_\mathfrak{p}$ ($\mathfrak{h}_\mathfrak{p}$ being identified with its dual), it is clear that the quotient field $Q = C(S(\mathfrak{h}_\mathfrak{p}))$ is obtained by adjoining all the roots of $F(\lambda)$ to the field $C(J)$. Hence the field extension $Q/C(J)$ is normal.

Let σ be an automorphism of Q leaving each element of $C(J)$ fixed. Then σ permutes the roots of $F(\lambda)$ and therefore leaves $S(\mathfrak{h}_\mathfrak{p})$ invariant. Fix $H_0 \in \mathfrak{h}_{\mathfrak{p}_0}$. Then the mapping

$$\lambda : p \to p^\sigma(H_0), \qquad p \in S(\mathfrak{h}_\mathfrak{p}),$$

is a homomorphism of $S(\mathfrak{h}_\mathfrak{p})$ into C. Let H_1 denote the unique element in $\mathfrak{h}_\mathfrak{p}$ such that under the identification between $S(\mathfrak{h}_\mathfrak{p})$ and $S^*(\mathfrak{h}_\mathfrak{p})$, $\lambda(H) = B(H, H_1)$ for all $H \in \mathfrak{h}_\mathfrak{p}$. Then $\lambda(p) = p(H_1)$ for all $p \in S(\mathfrak{h}_\mathfrak{p})$. Hence if $p \in J$,

$$p(H_1) = \lambda(p) = p^\sigma(H_0) = p(H_0)$$

so by Lemma 6.20, $H_1 = sH_0$ for some $s \in W$. Now let $q \in I(\mathfrak{h}_\mathfrak{p})$. Then

$$q^\sigma(H_0) = q(H_1) = q(sH_0) = q(H_0)$$

and since H_0 as arbitrary in $\mathfrak{h}_{\mathfrak{p}_0}$, $q^\sigma = q$. Thus every automorphism of Q which is identity on $C(J)$ leaves $I(\mathfrak{h}_\mathfrak{p})$ pointwise fixed. By Galois theory $I(\mathfrak{h}_\mathfrak{p}) \subset C(J)$. Now $I(\mathfrak{h}_\mathfrak{p})$ is integral over J (Lemma 6.19) and as mentioned above J is integrally closed. Hence $I(\mathfrak{h}_\mathfrak{p}) = J$ and the proof is finished.

§ 7. Mean Value Theorems

1. The Mean Value Operators

Let G be a connected Lie group and K a compact subgroup. Let π denote the natural mapping of G onto the manifold G/K and as usual put $o = \pi(e)$. For each $x \in G$ we associate a linear transformation M^x of the space of continuous functions on G/K, defined as follows: Let $f \in C(G/K)$ and $p \in G/K$. Select any $g \in G$ such that $\pi(g) = p$. The function $M^x f$ is then defined by

$$[M^x f](p) = \int_K f(gk \cdot \pi(x))\, dk,$$

where dk is the bi-invariant measure on K, normalized by $\int_K dk = 1$.

The isotropy subgroup of G at p is gKg^{-1}. The orbit of the point $g \cdot \pi(x)$ under this group is the set $\{gk \cdot \pi(x) : k \in K\}$ so $[M^x f](p)$ is the average of the values of f on this orbit. Thus it is not necessary to assume f defined on all of G/K. If $\tilde{f} = f \circ \pi$ then

$$[M^x f](p) = \int_K \tilde{f}(gkx)\, dk, \tag{1}$$

and it is obvious that $M^{kxk'} = M^x$ for $x \in G$, $k, k' \in K$.

In view of Cor. 2.8 and Theorem 5.6, the algebra $D(G/K)$ is generated by finitely many operators, say $\Delta_1, ..., \Delta_l$, and the identity operator I. We can assume that each of the operators $\Delta_1, ..., \Delta_l$ annihilates the constants ("has no constant term"). The next lemma shows, roughly speaking, that each mean value operator M^x is a function of the differential operators $\Delta_1, ..., \Delta_l$.

Lemma 7.1. *Let $p \in G/K$ and let f be an analytic function on a neighborhood U of p. There exists a neighborhood N of e in G and a neighborhood $V \subset U$ of p with the following property: for each $x \in N$ there exists a sequence of polynomials $P_1, P_2, ...$ without constant term such that*

$$[M^x f](q) = f(q) + \sum_{n=1}^{\infty} [P_n(\Delta_1, ..., \Delta_l) f](q)$$

for $q \in V$.

Proof. Let $\mathfrak{g}$ and $\mathfrak{k}$ denote the Lie algebras of G and K, respectively, and let $\mathfrak{p}$ denote a fixed subspace of $\mathfrak{g}$ such that $\mathfrak{g} = \mathfrak{k} + \mathfrak{p}$ (direct sum) and $\mathrm{Ad}_G(K)\, \mathfrak{p} \subset \mathfrak{p}$. Select a strictly positive definite quadratic form on $\mathfrak{p}$, invariant under $\mathrm{Ad}_G(K)$. This gives rise to a G-invariant Riemannian structure on G/K. Select an orthonormal basis $X_1, \ldots, X_r$ of $\mathfrak{p}$. Let B be a submanifold of G which π maps diffeomorphically onto a neighborhood $W \subset U$ of p. For each $q \in W$ let g_q be the element in B for which $\pi(g_q) = q$. The mapping

$$(t_1, \ldots, t_r) \to \pi(g_p \exp(t_1X_1 + \ldots + t_rX_r))$$

is a diffeomorphism of an open ball in $\boldsymbol{R}^r$ onto an open neighborhood W_p of p. The inverse mapping, say φ_p, is a coordinate system on W_p. For each $q \in W$ put

$$\varphi_q = \varphi_p \circ \tau(g_p g_q^{-1}), \qquad W_q = \tau(g_q g_p^{-1})\, W_p.$$

Then the pair (φ_q, W_q) is a local chart on M containing q. It is clear that W_q is "close to" W_p if q is close to p. We can therefore select an open neighborhood V' of p so small that for each $q \in V'$, the neighborhood W_q contains V' and such that the function $f \circ \varphi_p^{-1}$ has a Taylor series expansion

$$(f \circ \varphi_p^{-1})(t_1, \ldots, t_r) = \sum a_{n_1 \ldots n_r}(p)\, t_1^{n_1} \ldots t_r^{n_r}$$

absolutely convergent on $\varphi_p(V')$. The coordinates $t_i(\varphi_p(s))$ and $t_j(\varphi_q(s))$ are connected by

$$\pi(g_p \exp(t_1(\varphi_p(s))\, X_1 + \ldots + t_r(\varphi_p(s))\, X_r))$$
$$= \pi(g_q \exp(t_1(\varphi_q(s))\, X_1 + \ldots + t_r(\varphi_q(s))\, X_r))$$

for $s \in W_q \cap W_p$. Shrinking V' further we can assume that for each $q \in V'$

$$t_i(\varphi_p(s)) = P_q^i(t_1(\varphi_q(s)), \ldots, t_r(\varphi_q(s))) \qquad (1 \leqslant i \leqslant r)$$

where $P_q^1, \ldots, P_q^r$ are power series absolutely convergent for $s \in V'$. The Taylor series for $f \circ \varphi_q^{-1}$ around the origin is obtained by substituting the expressions for $t_i(\varphi_p(s))$ into the Taylor series for $f \circ \varphi_p^{-1}$. We can therefore shrink V' once more so that for each $q \in V'$ the Taylor expansion for $f \circ \varphi_q^{-1}$,

$$f \circ \varphi_q^{-1} = \sum a_{n_1 \ldots n_r}(q)\, t_1^{n_1} \ldots t_r^{n_r},$$

is absolutely convergent on $\varphi_q(V')$. Since $\varphi_q(V') = \varphi_p(\tau(g_p g_q^{-1})\, V')$ we can select an open neighborhood V of p, $p \in V \subset V'$, such that the set $\bigcap_{q \in V} \varphi_q(V')$ contains an open ball $\sum_i t_i^2 < \delta^2$ in $\boldsymbol{R}^r$. We may assume that $K \cdot V \subset V$ and that V has compact closure contained in V'.

Now let $X \in B_\delta(0) \subset \mathfrak{p}$ and $q \in V$. Then

$$[M^{\exp X} f](q) = \int_K \tilde{f}(g_q k \exp X k^{-1})\, dk = \int_K \tilde{f}(g_q \exp \operatorname{Ad}(k) X)\, dk.$$

Putting $\operatorname{Ad}(k) X = t_1 X_1 + \dots + t_r X_r$ we have by (16), §2,

$[(\operatorname{Ad}(k) X)^m \tilde{f}](g_q)$

$$= \left[\left(t_1 \frac{\partial}{\partial x_1} + \dots + t_r \frac{\partial}{\partial x_r}\right)^m \tilde{f}(g_q \exp(x_1 X_1 + \dots + x_r X_r))\right](0)$$

$$= \left[\left(t_1 \frac{\partial}{\partial x_1} + \dots + t_r \frac{\partial}{\partial x_r}\right)^m (f \circ \varphi_q^{-1})(x_1, \dots, x_r)\right](0).$$

Since $t_1^2 + \dots + t_r^2 < \delta^2$ and since the Taylor series for $f \circ \varphi_q^{-1}$ converges in the ball $t_1^2 + \dots + t_r^2 < \delta^2$, we have

$$\int_K \tilde{f}(g_q \exp \operatorname{Ad}(k) X)\, dk = \int_K \sum_0^\infty \frac{1}{m!} [(\operatorname{Ad}(k) X)^m \tilde{f}](g_q)\, dk.$$

Due to the uniform convergence of the Taylor series, the summation and integration can be interchanged. For $m > 0$, the differential operator

$$D_X = \int_K (\operatorname{Ad}(k) X)^m dk = \int_K \operatorname{Ad}(k) \cdot X^m\, dk$$

belongs to $\boldsymbol{D}_0(G)$ (see (2), §3). This operator gives rise to an operator in $\boldsymbol{D}(G/K)$ (Lemma 2.2) which can be written $m! P_m(\Delta_1, \dots, \Delta_l)$ where P_m is a polynomial without constant term. Since

$$D_X \tilde{F} = (m! P_m(\Delta_1, \dots, \Delta_l) F)^\sim$$

for $F \in C^\infty(G/K)$, we obtain

$$[M^{\exp X} f](q) = f(q) + \sum_1^\infty [P_m)\Delta_1, \dots, \Delta_l) f](q).$$

Lemma 7.1 is therefore proved by taking

$$N = \exp B_\delta(0) \cdot K,$$

which indeed is a neighborhood of e in G.

Theorem 7.2. *Let V be an open subset of G/K. Let φ be a complex-valued function on V which is an eigenfunction of each of the operators $\Delta_1, \ldots, \Delta_l$. Then for each $q \in V$*

$$[M^x\varphi](q) = \lambda_x\varphi(q), \qquad \lambda_x \in \boldsymbol{C}, \tag{2}$$

provided x is sufficiently close to e in G. On the other hand, a continuous function on V with the property (2) *is an eigenfunction of the operators $\Delta_1, \ldots, \Delta_l$.*

Proof. Suppose first that $\Delta_i\varphi = \lambda_i\varphi$ for $1 \leqslant i \leqslant l$, where $\lambda_i \in \boldsymbol{C}$. Then φ is an eigenfunction of each $D \in \boldsymbol{D}(G/K)$. Among these D is the Laplace-Beltrami operator on G/K. This being elliptic and having analytic coefficients, the function φ is analytic (compare the proof of Prop. 3.2). Now (2) follows immediately from Lemma 7.1.

On the other hand, suppose a continuous function φ satisfies (2). Let $q \in V$ and select $g_0 \in G$ such that $\pi(g_0) = q$. Then

$$\int_K \tilde{\varphi}(gkx)\, dk = \lambda_x\tilde{\varphi}(g) \tag{3}$$

for all g in a neighborhood of g_0 and all x in a neighborhood N of e. It is clear from (3) (assuming $\tilde{\varphi} \not\equiv 0$) that λ_x depends continuously on x. Let $\rho(x)$ be a differentiable function on G with support inside N such that $\int_G \rho(x)\, \lambda_x dx = 1$. Then (3) implies

$$\tilde{\varphi}(g) = \int_G \rho(x) \int_K \tilde{\varphi}(gkx)\, dk\, dx = \int_G \varphi(y) \int_K \rho(yk^{-1}g^{-1})\, dk\, dy,$$

which shows that $\tilde{\varphi}$, and therefore φ, is differentiable. Hence $\lambda_x = \lambda(x)$ depends differentiably on x. Let D be an arbitrary operator in $\boldsymbol{D}_0(G)$. We apply D to (3) with respect to the variable x. Then

$$\int_K [D\tilde{\varphi}](gkx)\, dk = [D\lambda](x)\, \tilde{\varphi}(g).$$

Putting $x = e$ we obtain

$$[D\tilde{\varphi}](g) = [D\lambda](e)\, \tilde{\varphi}(g) \tag{4}$$

so φ is an eigenfunction of each Δ_i $(1 \leqslant i \leqslant l)$.

Corollary 7.3. *Let V be an open subset of G/K. A function u on V satisfying the equations*

$$\Delta_1 u = \ldots = \Delta_l u = 0$$

is characterized by the mean value property

$$[M^x u](q) = u(q)$$

for each $q \in V$, x *being sufficiently close to* e *in* G.

In fact, (4) shows that $[D\lambda](e) = 0$ for each $D \in \boldsymbol{D}_0(G)$; since $\lambda(x)$ is bi-invariant under K, relation (3), §3, shows that $\lambda(x)$ is constant.

Remark. Consider the special case when M is a Euclidean space or a Riemannian globally symmetric space of rank 1 (and dimension > 1) and $G = I_0(M)$. In this case $\boldsymbol{D}(G/K)$ is generated by the Laplace-Beltrami operator Δ alone and K acts transitively on each sphere $S_r(o)$. Hence M^x is just the operation of averaging over a sphere of fixed radius. Corollary 7.3 states that the solutions of the equation $\Delta u = 0$ on M are characterized by the fact that the mean value of u on each sphere in M equals the value at the center. In Euclidean space this is Gauss' mean value theorem for harmonic functions.

In Prop. 3.2 we have an integral equation characterizing those eigenfunctions of $\Delta_1, \ldots, \Delta_l$ which are invariant under the action of K. From Theorem 7.2 we obtain the following generalization.

Corollary 7.4. *Let* φ *be a function on* G/K *satisfying* $\varphi(o) = 1$. *Then* φ *is an eigenfunction of all* $\Delta_1, \ldots, \Delta_l$ *if and only if*

$$\int_K \tilde{\varphi}(gkx)dk = \tilde{\varphi}(g) \int_K \tilde{\varphi}(kx)dk, \qquad x, g \in G. \tag{5}$$

In fact, if (5) holds for all x in a neighborhood of e it must hold for all $x \in G$ due to the analyticity of $\tilde{\varphi}$ (Lemma 4.3, Chapter VI).

Remark. Let φ be a spherical function on G/K. Then $\Delta_i\varphi = \lambda_i\varphi$ $(1 \leqslant i \leqslant l)$ where $\lambda_i \in \boldsymbol{C}$ and $M^x\varphi = \tilde{\varphi}(x)\,\varphi$. From Lemma 7.1 we have therefore

$$\tilde{\varphi}(x) = 1 + \sum_{n=1}^{\infty} P_n(\lambda_1, \ldots, \lambda_l),$$

if x is sufficiently small and formally

$$M^x = I + \sum_{n=1}^{\infty} P_n(\Delta_1, \ldots, \Delta_l). \tag{6}$$

As an example take $M = \boldsymbol{R}^2$, $G = I_0(M)$. According to §3 the function

$$\varphi(x, y) = J_0(\sqrt{-\lambda}\sqrt{x^2 + y^2}) = \sum_{k=0}^{\infty} \frac{1}{2^{2k}(k!)^2} \lambda^k (x^2 + y^2)^k$$

is a spherical function on M satisfying $\Delta\varphi = \lambda\varphi$. The formula (6) becomes in this case a well-known expansion

$$M^r = J_0(\sqrt{-\Delta}\, r) = \sum_{k=0}^{\infty} \frac{1}{2^{2k}(k!)^2} \Delta^k r^{2k}$$

for the operation M^r of averaging over spheres of radius r.

2. Approximations by Analytic Functions

Since Lemma 7.1 is limited to analytic functions we must, in order to apply it to C^∞ functions, approximate these by analytic functions. For this problem we recall some facts concerning analytic vectors of representations of Lie groups.

Let M be an analytic manifold and $\mathfrak{H}$ a (complex) Banach space. A mapping $f : M \rightarrow \mathfrak{H}$ is called *analytic* at a point $p \in M$ if there exists a coordinate system $\{x_1, ..., x_m\}$ on a neighborhood U of p such that $x_1(p) = ... = x_m(p) = 0$ and

$$f(q) = \sum a_{n_1 \cdots n_m} x_1(q)^{n_1} \ldots x_m(q)^{n_m} \qquad (q \in U),$$

where the coefficients $a_{n_1 \ldots n_m}$ belong to $\mathfrak{H}$ and the series converges absolutely. (A series $\Sigma\, a_n$ where $a_n \in \mathfrak{H}$ is said to converge absolutely if $\Sigma \| a_n \|$ converges, $\| \ \|$ denoting the norm in $\mathfrak{H}$.)

The following two lemmas are obvious.

Lemma 7.5. *Consider two mappings $\varphi : M \rightarrow M$ and $f : M \rightarrow \mathfrak{H}$. If φ is analytic at $p \in M$ and f is analytic at $\varphi(p)$, then the composite mapping $f \circ \varphi : M \rightarrow \mathfrak{H}$ is analytic at p.*

Lemma 7.6. *Let the mapping $f : M \rightarrow \mathfrak{H}$ be analytic at p and let F be a continuous linear mapping of $\mathfrak{H}$ into $\mathbf{C}$. Then the function $F \circ f$ is analytic at p.*

Definition. Let G be a Lie group and let π be a representation of G on a Banach space $\mathfrak{H}$. A vector $e \in \mathfrak{H}$ is called *analytic* (under π) if the mapping $x \rightarrow \pi(x)\, e$ is an analytic mapping of G into $\mathfrak{H}$.

Let $\mathfrak{A}_\pi$ denote the set of analytic vectors under π. It is obvious that $\mathfrak{A}_\pi$ is a linear subspace of $\mathfrak{H}$. If $e \in \mathfrak{A}_\pi$ and $y \in G$, then the mapping $x \rightarrow \pi(x)\, \pi(y)\, e$ is composed of the analytic mappings $x \rightarrow xy$ and $x \rightarrow \pi(x)\, e$. Using Lemma 7.5 it follows that $\pi(y)\, e \in \mathfrak{A}_\pi$.

Lemma 7.7. *The space $\mathfrak{A}_\pi$ is invariant under each $\pi(y)$, $y \in G$.*

If $\mathfrak{H}$ is finite-dimensional, then each vector $e \in \mathfrak{H}$ is analytic. This is a special case of the theorem that a continuous homomorphism of one Lie group into another is analytic. For Banach spaces $\mathfrak{H}$ of infinite dimension the basic fact is:

Theorem 7.8. *The space $\mathfrak{A}_\pi$ of analytic vectors is dense in $\mathfrak{H}$.*

This theorem was proved by Harish-Chandra [4], p. 220, for semisimple Lie groups and generalized by E. Nelson to arbitrary Lie groups [1]. The quickest proof is due to L. Gårding [1] who, like Nelson, uses the heat equation on the group. We shall use this theorem in order to prove the following lemma on approximation by analytic functions.

Lemma 7.9. *Let $F_1, \ldots, F_n$ be a finite set of bounded, continuous functions on G, let C be a compact subset of G and ϵ a number > 0. Then there exists an analytic, integrable function φ on G such that the convolution $\varphi * F_i$ is analytic and such that*

$$|(\varphi * F_i)(x) - F_i(x)| < \epsilon \qquad (1 \leqslant i \leqslant n) \tag{7}$$

for all $x \in C$. If each F_i has compact support, φ can be chosen such that (7) holds for all $x \in G$.

Proof. Let $L^1(G)$ denote the Banach space of integrable functions on G with respect to a left invariant measure dg on G, the norm on $L^1(G)$ being

$$||f|| = \int_G |f(g)|\, dg.$$

The space $L^1(G)$ is the completion of the group algebra in this norm. For each $x \in G$ let $\pi(x)$ denote the endomorphism of $L^1(G)$ given by $[\pi(x)f](y) = f(x^{-1}y)$, $f \in L^1(G)$, $y \in G$. By the left invariance of dg, $||\pi(x)f|| = ||f||$. If $f \in C_c(G)$ and $\delta > 0$ there exists a neighborhood V of e in G such that $|f(x^{-1}y) - f(y)| < \delta$ for all $x \in V$ and all $y \in G$. This implies that the mapping $x \to \pi(x)f$ of G into $L^1(G)$ is continuous. Using the fact that the group algebra is dense in $L^1(G)$ this continuity follows for all $f \in L^1(G)$. Hence the mapping $x \to \pi(x)$ is a representation of G on $L^1(G)$. Let $h \in L^1(G)$ be an analytic vector under π. If $x \in G$, the vector $\pi(x)\, h$ is also analytic (Lemma 7.7). If F is a bounded continuous function on G then by Lemma 7.6, the function

$$x \to \int_G F(y^{-1})\, [\pi(x^{-1})\, h]\, (y)\, dy,$$

which is just the function $h*F$, is analytic. Now let U be a compact neighborhood of e in G and let γ_U be a positive continuous function on G with support inside U and satisfying $\int \gamma_U(x)\,dx = 1$. Then if U is sufficiently small we have

$$|(\gamma_U * \gamma_U * F_i)(x) - F_i(x)| < \frac{\epsilon}{2} \qquad (1 \leqslant i \leqslant n) \tag{8}$$

for all $x \in C$. According to Theorem 7.8 there exists a sequence (φ_n) of analytic vectors converging to γ_U. Since

$$\int_G |\varphi_n(y) - \gamma_U(y)|\,dy \to 0$$

as $n \to \infty$ it is clear that for each bounded continuous function F on G the sequence $\varphi_n * F$ $(n = 1, 2, \ldots)$ converges to $\gamma_U * F$ uniformly on G. In particular, if N is sufficiently large, the function φ_N satisfies

$$|(\varphi_N * \gamma_U * F_i)(x) - (\gamma_U * \gamma_U * F_i)(x)| < \frac{\epsilon}{2} \qquad x \in G \qquad (1 \leqslant i \leqslant n). \tag{9}$$

The first statement of Lemma 7.9 (with $\varphi = \varphi_N * \gamma_U$) follows at once from (8) and (9). Moreover, if each F_i has compact support, U can be chosen such that (8) holds for all $x \in G$. This completes the proof of the lemma.

3. The Darboux Equation in a Symmetric Space

Theorem 7.10. *Let (G, K) be a Riemannian symmetric pair and K compact. Then for each $x \in G$, the mean value operator M^x commutes with all the operators in $\boldsymbol{D}(G/K)$.*

Proof. First let f be an analytic function on G/K and let $D \in \boldsymbol{D}(G/K)$. Then the function Df is also analytic. Using now the commutativity of $\boldsymbol{D}(G/K)$ (Theorem 2.9) and Lemma 7.1 it follows that for each $p \in G/K$ the relation

$$[DM^xf](p) = [M^xDf](p) \tag{10}$$

holds for all x in a neighborhood of e in G. Now, in view of Lemma 3.1, both sides of (10) are analytic in the variable x. Thus (10) holds for all $x \in G$ and all $p \in G/K$.

Now fix $x \in G$ and consider the operator N^x on $C(G)$ given by

$$[N^xF](g) = \int_K F(gkx)\,dk.$$

Let D_0 denote the operator in $\boldsymbol{D}_0(G)$ which corresponds to D according to Lemma 2.2. We shall now prove (10) for each $f \in C_c^\infty(G/K)$. Lemma 7.9 shows that there exists a sequence (φ_n) of continuous functions on G such that the sequences $(\varphi_n * \tilde{f})$, $(\varphi_n * D_0 N^x \tilde{f})$ and $(\varphi_n * N^x D_0 \tilde{f})$ consist of analytic functions and converge to the functions $\tilde{f}$, $D_0 N^x \tilde{f}$ and $N^x D_0 \tilde{f}$, uniformly on G. A straightforward verification shows that

$$\varphi_n * N^x \tilde{f} = N^x(\varphi_n * \tilde{f}), \tag{11}$$

$$\varphi_n * D_0 \tilde{f} = D_0(\varphi_n * \tilde{f}). \tag{12}$$

Since (10) holds for analytic f we have

$$D_0 N^x(\varphi_n * \tilde{f}) = N^x D_0(\varphi_n * \tilde{f}),$$

so by (11) and (12)

$$\varphi_n * D_0 N^x \tilde{f} = \varphi_n * N^x D_0 \tilde{f}.$$

Taking limits we get $D_0 N^x \tilde{f} = N^x D_0 \tilde{f}$ so $DM^x f = M^x D f$. Finally, let f be an arbitrary function in $C^\infty(G/K)$. For each compact subset $C \subset G$ there exists a function $f_C \in C_c^\infty(G/K)$ which coincides with f on C. If $p \in G/K$ and $x \in G$ the numbers $[DM^x f](p)$ and $[M^x D f](p)$ only depend on the values of f on a compact set. Since (10) holds for f_C it holds also for f.

Corollary 7.11. *If $\pi(x) = q$ let M^q stand for M^x. If $f \in C^\infty(G/K)$ put*

$$F(p, q) = [M^q f](p), \qquad p, q \in G/K.$$

Then for each $D \in \boldsymbol{D}(G/K)$

$$D_1 F = D_2 F, \tag{13}$$

where the subscripts 1, 2 *indicate that D acts on the first and second argument, respectively.*

In order to verify (13) select $g \in G$ such that $\pi(g) = p$. Then, using the theorem and the invariance of D, we obtain

$$[D_1 F](p, q) = [DM^q f](p) = [M^q D f](p) = \int_K [Df](gk \cdot q)\, dk$$

$$= \int_K [(Df)^{\tau(k^{-1}g^{-1})}](q)\, dk = \left[D \int_K f^{\tau(k^{-1}g^{-1})}\, dk\right](q) = [D_2 F](p, q).$$

Consider (13) in the case of Euclidean space $\boldsymbol{R}^n$. If a function $\varphi \in C^\infty(\boldsymbol{R}^n)$ has the form

$$\varphi(x_1, \ldots, x_n) = \psi(r), \qquad r = (x_1^2 + \ldots + x_n^2)^{1/2},$$

then

$$\Delta\varphi = \frac{d^2\psi}{dr} + \frac{n-1}{r}\frac{d\psi}{dr}$$

(see Lemma 7.12 below). Writing M^r instead of M^q if $r = d(0, q)$ we see from (13) that the function $F(p, r) = [M^r f](p)$ satisfies the differential equation

$$\Delta_p F = \frac{\partial^2 F}{\partial r^2} + \frac{n-1}{r}\frac{\partial F}{\partial r},$$

which in the literature is sometimes called "the Darboux equation."

4. Poisson's Equation in a Two-Point Homogeneous Space

Let M be a Euclidean space or a Riemannian globally symmetric space of rank 1. We shall now give an explicit formula for the solution of Poisson's equation $\Delta u = f$ on M.

For each $p \in M$ let as usual $S_r(p)$ denote the sphere in M with center p and radius r. If M is noncompact, $S_r(p)$ is the image of a sphere in the tangent space M_p under the diffeomorphism Exp_p. If M is compact, let L denote its *diameter*, that is the maximum distance between any pairs of points in M. If $r < L$, $S_r(p)$ is the image of a sphere in M_p under a diffeomorphism (Theorem 5.4, Chapter IX). This of course fails to hold for $S_L(p)$ (see, e.g., $M = S^2$). However, $S_L(p)$ is a submanifold of M, being the orbit under a compact subgroup of $I(M)$ (Prop. 4.4, Chapter II).

In any case $S_r(p)$ is for each r a submanifold of M and has a Riemannian structure induced by that of M. Let $A(r)$ denote the total measure of $S_r(p)$ according to the Riemannian measure on $S_r(p)$. (If $\dim M = 1$ so $M = \boldsymbol{R}$ or $M = S^1$, $A(r)$ is to be understood as the number of points in $S_r(p)$.) Naturally, we call $A(r)$ the area of $S_r(p)$. The volume of the ball $B_r(p)$ in M will be denoted $V(r)$. Since M is homogeneous, that is, has a transitive group of isometries, $A(r)$ and $V(r)$ are independent of p. Let d denote the distance function on M.

Let $n = \dim M$ and let $\{\theta_1, \ldots, \theta_{n-1}, r\}$ be a system of geodesic polar coordinates at the point $p \in M$ (§3). Here r runs through the intervals $0 < r < \infty$ in the noncompact case, $0 < r < L$ in the compact case (Theorem 5.4, Chapter IX). The system $\{\theta_1, \ldots, \theta_{n-1}\}$ can be regarded as a system of coordinates on an open subset of $S_r(p)$.

Lemma 7.12. *In geodesic polar coordinates at p as described above, the Laplace-Beltrami operator Δ on M has the form*

$$\Delta = \frac{\partial^2}{\partial r^2} + \frac{1}{A(r)} \frac{dA}{dr} \frac{\partial}{\partial r} + \Delta',$$

where Δ' is the Laplace-Beltrami operator on $S_r(p)$. Here $0 < r < \infty$ if M is noncompact, $0 < r < L$ if M is compact.

Proof. Put $\theta_n = r$ and let

$$ds^2 = \sum_{i,j=1}^{n} g_{ij}(\theta_1, \ldots, \theta_n)\, d\theta_i\, d\theta_j \tag{14}$$

denote the Riemannian structure of M expressed in geodesic polar coordinates at p. As before, let $\bar{g} = \det(g_{ij})$ and $(g^{ij}) = (g_{ij})^{-1}$ (the inverse of the matrix (g_{ij})). As shown in Chapter I, the geodesics emanating from p intersect $S_r(p)$ orthogonally. Hence (14) reduces to

$$ds^2 = dr^2 + \sum_{i,j=1}^{n-1} g_{ij}(\theta_1, \ldots, \theta_{n-1}, r)\, d\theta_i\, d\theta_j. \tag{15}$$

If I denotes the identity mapping of $S_r(p)$ into M, we have $I^*(dr^2) = 0$ so the Riemannian structure of $S_r(p)$ is given by

$$\sum_{i,j=1}^{n-1} g_{ij}(\theta_1, \ldots, \theta_{n-1}, r)\, d\theta_i\, d\theta_j. \tag{16}$$

Using (15) we obtain

$$\Delta = \frac{\partial^2}{\partial r^2} + \frac{1}{\sqrt{\bar{g}}} \frac{\partial \sqrt{\bar{g}}}{\partial r} \frac{\partial}{\partial r} + \frac{1}{\sqrt{\bar{g}}} \sum_{i,j=1}^{n-1} \frac{\partial}{\partial \theta_j} \left(g^{ij} \sqrt{\bar{g}} \frac{\partial}{\partial \theta_i} \right). \tag{17}$$

If A denotes the matrix $(g_{ij})_{1 \leqslant i,j \leqslant n}$ and B denotes the matrix $(g_{ij})_{1 \leqslant i,j \leqslant n-1}$, we have

$$A = \begin{pmatrix} B & 0 \\ 0 & 1 \end{pmatrix}, \qquad A^{-1} = \begin{pmatrix} B^{-1} & 0 \\ 0 & 1 \end{pmatrix}, \qquad \det A = \det B.$$

Consequently, the last term in (17) is the Laplace-Beltrami operator Δ' on $S_r(p)$.

The isotropy subgroup of $I(M)$ at p acts transitively on the set of geodesics emanating from p. It follows that the function

$$\Delta r = \frac{1}{\sqrt{\bar{g}}} \frac{\partial \sqrt{\bar{g}}}{\partial r}$$

is a function of r alone so

$$\log \sqrt{\bar{g}} = \alpha(r) + \beta(\theta_1, ..., \theta_{n-1})$$

and

$$\sqrt{\bar{g}} = e^{\alpha(r)}\, e^{\beta(\theta_1, ..., \theta_{n-1})}$$

for suitable functions α and β. Note that $\alpha'(r)$ is independent of the choice of $\{\theta_1, ..., \theta_{n-1}\}$. So is $\alpha(r)$ if we, for example, fix $\alpha(L/2)$ by $\alpha(L/2) = 1$. Now

$$A(r) = \int_{S_r(p)} d\mu,$$

where $d\mu$ is the Riemannian measure on $S_r(p)$ which is locally given by $|\det B|^{1/2}\, d\theta_1 \dots d\theta_{n-1} = \sqrt{\bar{g}}\, d\theta_1 \dots d\theta_{n-1}$. Remembering that $\alpha(r)$ is independent of the choice of $\{\theta_1, ..., \theta_{n-1}\}$ we have

$$A(r) = Ce^{\alpha(r)} \qquad (C = \text{constant}),$$

so

$$\frac{1}{\sqrt{\bar{g}}} \frac{\partial \sqrt{\bar{g}}}{\partial r} = \frac{1}{A(r)} \frac{dA}{dr}.$$

This concludes the proof.

The volume $V(r)$ of $B_r(p)$ is given by

$$V(r) = \int_{B_r(p)} dq,$$

where dq is the Riemannian measure on M which is locally given by $\sqrt{\bar{g}}\, dr\, d\theta_1 \dots d\theta_{n-1}$. It follows that

$$V(r) = \int_0^r A(t)\, dt.$$

Theorem 7.13. *Let M be a Euclidean space or a Riemannian globally symmetric space of rank* 1. *Let $f \in C_c^\infty(M)$. A solution to Poisson's equation*

$$\Delta u = f$$

is given as follows:

(i) *If M is noncompact,*

$$u(p) = \int_M H(p, q) f(q)\, dq, \tag{18}$$

where

$$H(p, q) = \int_1^{d(p,q)} \frac{1}{A(r)}\, dr. \tag{19}$$

(ii) *If M is compact a solution exists if and only if $\int f(q)\,dq = 0$. If this condition is satisfied, the unique solution to $\Delta u = f$ satisfying $\int u(q)\,dq = 0$ is given by*

$$u(p) = \int_M G(p, q) f(q)\,dq,$$

where

$$G(p, q) = \frac{1}{V(L)} \int_L^{d(p,q)} \frac{V(L) - V(r)}{A(r)}\,dr.$$

Proof of (i). We have

$$u(p) = \int_0^\infty dr \int_{S_r(p)} H(p, q) f(q)\,d\mu(q) = \int_0^\infty \psi(r)\,A(r)\,[M^r f]\,(p)\,dr,$$

where

$$\psi(r) = \int_0^r \frac{1}{A(t)}\,dt,$$

and $[M^r f]\,(p)$ is the mean value of f on $S_r(p)$. Using now Cor. 7.11 and Lemma 7.12 we obtain

$$[\Delta u]\,(p) = \int_0^\infty \psi(r)\,A(r)\,[\Delta M^r f]\,(p)\,dr = \int_0^\infty \psi(r)\,A(r)\,\Delta_r [M^r f]\,(p)\,dr$$

where

$$\Delta_r = \frac{\partial^2}{\partial r^2} + \frac{1}{A(r)} \frac{dA}{dr} \frac{\partial}{dr}\,.$$

Keeping p fixed and putting $F(r) = [M^r f]\,(p)$ we have

$$[\Delta u]\,(p) = \int_0^\infty \{\psi(r)\,A(r)\,F''(r) + \psi(r)\,A'(r)\,F'(r)\}\,dr$$

$$= \lim_{\epsilon \to 0} [\psi(r)\,A(r)\,F'(r)]_\epsilon^\infty - \int_0^\infty F'(r)\,dr.$$

Now $\psi(\epsilon)\,A(\epsilon) \to 0$ as $\epsilon \to 0$ and $F(r)$ vanishes for all sufficiently large r. Hence

$$[\Delta u]\,(p) = F(0) = f(p).$$

The choice of $H(p, q) = \psi(d(p, q))$ was of course motivated by the fact that the equation

$$[\Delta_2 H]\,(p, q) = \delta_p, \tag{20}$$

where δ_p is the delta distribution $f \to f(p)$, leads to the equation

$$\frac{d^2\psi}{dr^2} + \frac{1}{A(r)} \frac{dA}{dr} \frac{d\psi}{dr} = 0.$$

Proof of (ii). In the compact case we have $\int \Delta u(q)\, dq = 0$ for each $u \in C^\infty(M)$ (§6, No. 3). Taking this fact into account we replace (20) by the equation

$$[\Delta_2 G]\,(p, q) = \frac{1}{V(L)}\,(\delta_p - 1),$$

which by $G(p, q) = \varphi(d(p, q))$ leads to the equation

$$\frac{d^2\varphi}{dr^2} + \frac{1}{A(r)} \frac{dA}{dr} \frac{d\varphi}{dr} = -\frac{1}{V(L)}, \qquad 0 < r < L, \tag{21}$$

whose coefficients are in general unbounded near $r = 0$ and $r = L$. The function

$$\varphi(r) = \frac{1}{V(L)} \int_L^r \frac{V(L) - V(t)}{A(t)}\, dt$$

is a solution of (21), unbounded near $r = 0$ (if $n > 1$) but bounded near $r = L$ since $V'(t) = A(t)$. We have

$$u(p) = \int_0^L dr \int_{S_r(p)} G(p, q)\, f(q)\, d\mu(q) = \int_0^L \varphi(r)\, A(r)\, [M^r f]\,(p)\, dr.$$

We put $F(r) = [M^r f]\,(p)$ and observe that $\varphi(r)\, A(r) \to 0$ as $r \to L$ and also as $r \to 0$ if $n > 1$. If $n = 1$ we have $F'(0) = 0$. Then we obtain as above

$$\begin{aligned}
[\Delta u]\,(p) &= \int_0^L \{\varphi(r)\, A(r)\, F''(r) + \varphi(r)\, A'(r)\, F'(r)\}\, dr \\
&= [\varphi(r)\, A(r)\, F'(r)]_0^L - \int_0^L \varphi'(r)\, A(r)\, F'(r)\, dr \\
&= -\int_0^L \left(1 - \frac{V(r)}{V(L)}\right) F'(r)\, dr \\
&= F(0) - F(L) + \left[F(r)\,\frac{V(r)}{V(L)}\right]_0^L - \frac{1}{V(L)} \int_0^L A(r)\, [M^r f]\,(p)\, dr \\
&= f(p) - F(L) + F(L) - \frac{1}{V(L)} \int_M f(q)\, dq = f(p).
\end{aligned}$$

Thus $\Delta u = f$. Moreover,

$$\int_M u(p)\,dp = \int_M f(q)\,dq \int_M G(p, q)\,dp = 0, \tag{22}$$

because $\int G(p, q)\,dp$ is independent of q and $\int f(q)\,dq = 0$. Since the only solutions of $\Delta v = 0$ are the constants (§6, No. 3), the function u is the only solution of the Poisson's equation satisfying (22).

EXERCISES

A. Measure Theory

1. Let V_n be a vector space over $\boldsymbol{R}$ of dimension $n > 0$. For $0 < h < n$ let $\mathscr{S}_h$ denote the set of h-dimensional vector subspaces of V_n. The group $\boldsymbol{SL}(V_n)$ of endomorphisms of V_n with determinant 1 acts transitively on $\mathscr{S}_h$. Let S_0 be a fixed element in $\mathscr{S}_h$ and let H denote the topological subgroup of $\boldsymbol{SL}(V_n)$ leaving S_0 invariant. Prove (using Theorem 1.7 or otherwise) that the space $\boldsymbol{SL}(V_n)/H$ has no measure > 0 invariant under $\boldsymbol{SL}(V_n)$.

2. Let G be a Lie group, H and N closed subgroups such that $H \subset N \subset G$. Assume that G/H and G/N have positive G-invariant measures dg_H and dg_N. Show that N/H has an N-invariant positive measure dn_H which (suitably normalized) satisfies

$$\int_{G/H} f(gH)dg_H = \int_{G/N} \left(\int_{N/H} f(gnH)\,dn_H\right) dg_N$$

for all $f \in C_c(G/H)$.

3. Let G be a semisimple, connected, compact Lie group and let T be a maximal torus in G. Let w denote the order of the Weyl group of G and let

$$D(t) = \prod_{\alpha \in \Delta^*} 2 \sin\left(\tfrac{1}{2}\,\alpha(iH)\right),$$

if $t = \exp H \in T$. Let dt and dg, respectively, denote the invariant measures on T and G normalized by

$$\int_T dt = \int_G dg = 1.$$

Derive Weyl's integral formula

$$\int_G f(g)\,dg = \frac{1}{w}\int_T |D(t)|\,dt \int_G f(gtg^{-1})\,dg, \qquad f \in C(G),$$

as a special case of Prop. 1.19 or directly from Lemma 1.9.

B. A Nonreductive Coset Space

Let $\boldsymbol{P}^2$ denote the 2-dimensional projective space consisting of all real 3-tuples $(x, y, z) \neq (0, 0, 0)$ where proportional 3-tuples are identified. The group $G = \boldsymbol{SL}(3, \boldsymbol{R})$ acts transitively on $\boldsymbol{P}^2$ as well as on the set of lines in $\boldsymbol{P}^2$. Let H_1 denote the subgroup of G leaving the point $y = z = 0$ fixed and let H_2 be the subgroup of G mapping the line $x = 0$ into itself.

1. Show that the mapping $g \to {}^t(g^{-1})$ is an automorphism of G mapping H_1 onto H_2. (The ensuing diffeomorphism between G/H_1 and G/H_2 is the polarity between points and lines in elliptic geometry.)

2. Show that the coset spaces G/H_1 and G/H_2 are not reductive.

C. Differential Operators

1. Let g be a pseudo-Riemannian structure on a manifold M. Let $f \in C^\infty(M)$. Then

$$Xf = g(\operatorname{grad} f, X)$$

for any vector field X on M.

2. Let M be an oriented Riemannian manifold, and let ω denote the volume element on M. Let X be a vector field on M and $\theta(X)$ the Lie derivative with respect to X (Exercises, Chapter I). Prove that

$$\theta(X)\,\omega = (\operatorname{div} X)\,\omega.$$

3. Let M be a manifold of dimension m and suppose ω is an m-form on M of maximal rank, that is, $\omega_p(X_1, \ldots, X_m) \neq 0$ if $X_1, \ldots, X_m$ are arbitrary linearly independent tangent vectors at an arbitrary point $p \in M$. Let D be a differential operator on M. Prove that there exists a unique differential operator D^* on M such that

$$\int_M (Df)\,g\,\omega = \int_M f(D^*g)\,\omega, \qquad f, g \in C_c^\infty(M).$$

The operator D^* is called the *adjoint* of D (Hint: Prove the formula first for all g with sufficiently small supports. The general case can then be settled by a partition of unity argument similar to the one used in §1, No. 1 for integration of a differential form.)

4. Let G be a Lie group and dx a left invariant measure on G. Let X be a left invariant vector field on G. Using the formula

$$[Xf](x) = \left[\frac{d}{dt} f(x \exp tX)\right](0)$$

show that the adjoint X^* is given by

$$X^* = -X + [X\delta](e),$$

where $\delta(x) = \det \mathrm{Ad}_G(x)$.

5. Let G be a connected semisimple Lie group with Lie algebra $\mathfrak{g}$. Let B denote the Killing form of $\mathfrak{g}$. Let $X_1, \ldots, X_n$ be a basis of $\mathfrak{g}$, put $g_{ij} = B(X_i, X_j)$ $(1 \leqslant i, j \leqslant n)$ and let (g^{ij}) denote the inverse of the matrix (g_{ij}). Show that

(i) The differential operator $\gamma = \sum_{i,j} g^{ij} X_i \cdot X_j$ is independent of the choice of basis (X_i) and lies in the center of the algebra $D(G)$. The operator γ is called the *Casimir operator*.

(ii) Suppose G has finite center and let K be a maximal compact subgroup of G. Let Γ denote the member of $D(G/K)$ which corresponds to γ according to Lemma 2.2. Then Γ is the Laplace-Beltrami operator on G/K (assuming the Riemannian structure induced by B).

6*. Let M be a compact Riemannian globally symmetric space of rank 1. The notation being as in Theorem 7.13, let

$$K(p, q) = \int_{L/2}^{d(p,q)} \frac{1}{A(r)}\, dr,$$

$$v(p) = \int_M K(p, q) f(q)\, dq, \qquad f \in C^\infty(M).$$

Prove that the function v satisfies the following modification of Poisson's equation (Helgason [3], p. 281),

$$\Delta v = f - M^L f.$$

D. Spherical Functions on a Compact Symmetric Space

1. Let G be a compact group and π an irreducible unitary representation of G on a Hilbert space $\mathfrak{H}$. Prove the following assertions (Séminaire Sophus Lie [1], Exposé 22) which show that $\dim \mathfrak{H} < \infty$.

(i) Let T be a continuous operator on $\mathfrak{H}$. Then there exists a continuous operator $T^\natural$ on $\mathfrak{H}$ such that

$$\langle T^\natural \boldsymbol{a}, \boldsymbol{b} \rangle = \int_G \langle T\pi(x)\, \boldsymbol{a}, \pi(x)\, \boldsymbol{b} \rangle\, dx, \qquad \boldsymbol{a}, \boldsymbol{b} \in \mathfrak{H}.$$

dx being the normalized Haar measure on G.

(ii) $T^\natural$ commutes with each $\pi(x)$, $x \in G$; owing to the irreducibility, $T^\natural$ is a scalar multiple of the identity.

(iii) If T satisfies an inequality

$$\left| \sum_{i=1}^{n} \langle Te_i, e_i \rangle \right|^2 \leqslant M n \qquad (M \text{ constant})$$

for every set of orthonormal vectors $e_1, \ldots, e_n$, then the same holds for $T^\natural$.

(iv) Let e be a unit vector in $\mathfrak{H}$ and let P be the orthogonal projection of $\mathfrak{H}$ on the subspace Ce. Then $P^\natural \neq 0$; also P and $P^\natural$ satisfy the inequality above with $M = 1$. Now $P^\natural$ being a scalar, $P^\natural = \alpha I$, deduce that

$$\dim \mathfrak{H} \leqslant \frac{1}{|\alpha|^2}.$$

2. Let (G, K) be a Riemannian symmetric pair, G compact. Let φ be a spherical function on G and let V_φ denote the vector space formed by all finite complex linear combinations of left translates $\varphi^{L(g)}$ of φ. Prove the following statements (compare Godement [3], §3), which show that $\dim V_\varphi < \infty$, proving Lemma 6.2.

(i) Let $L^2(G)$ denote the Hilbert space of complex-valued square-integrable functions on G. Let $L^\natural(G)$ denote the image of $L^2(G)$ under the mapping $f \to f^\natural$ where

$$f^\natural(x) = \int_K \int_K f(kxk')\, dk\, dk', \qquad f \in L^2(G).$$

Under the convolution product, $L^2(G)$ is a Banach algebra and $L^\natural(G)$ is a commutative closed subalgebra. The kernel $\mathfrak{m}$ of the homomorphism $f \to \int f(x)\, \overline{\varphi(x)}\, dx$, $(f \in L^\natural(G))$, is a closed maximal ideal in $L^\natural(G)$.

(ii) Let

$$\mathfrak{n} = \{f \in L^2(G) : (g * f)^\natural \in \mathfrak{m} \text{ for all } g \in L^2(G)\}.$$

Using the formula

$$(f^\natural * g)^\natural = f^\natural * g^\natural, \qquad f, g \in L^2(G),$$

prove that $\mathfrak{n}$ is a maximal left ideal in $L^2(G)$ and $\mathfrak{n} \cap L^\natural(G) = \mathfrak{m}$.

(iii) For each $x \in G$, the ideal $\mathfrak{n}$ is invariant under the mapping $f \to f^{L(x)}$ $(f \in L^2(G))$. The resulting unitary representation of G on $L^2(G)/\mathfrak{n}$ is irreducible.

(iv) The vector space $L^2(G)/\mathfrak{n}$ is finite-dimensional and isomorphic to V_φ.

E. The Fourier Transform on a Symmetric Space

1. Let (G, K) be a Riemannian symmetric pair, K compact. Let $L^1(G)$ denote the space of complex-valued integrable functions on G with the norm

$$\|f\| = \int_G |f(x)|\, dx.$$

Let $L^\natural(G)$ denote the set of functions in $L^1(G)$ which are bi-invariant under K. Then under convolution product $L^\natural(G)$ is a commutative Banach algebra. The following assertions show that $L^\natural(G)$ is semisimple, that is, if $f \in L^\natural(G)$, $f \neq 0$ then there exists a continuous homomorphism $\alpha : L^\natural(G) \to C$ such that $\alpha(f) \neq 0$. Let $f^*(x) = \overline{f(x^{-1})}$.

(i) Suppose there exists $f_0 \in L^\natural(G)$, $f_0 \neq 0$ such that $\alpha(f_0) = 0$ for all continuous homomorphisms $\alpha : L^\natural(G) \to C$. Show that there exists a bounded $f_0 \in L^\natural(G)$, $f_0 \neq 0$ with this property and consequently if $f \in L^\natural(G)$, the convolution $(f_0 * f_0^* * f)(x)$ exists for *all* x.

(ii) The linear function $F : L^\natural(G) \to C$ defined by

$$F(f) = (f_0 * f_0^* * f)(e), \qquad f \in L^\natural(G),$$

satisfies

$$F(f * f^*) \geqslant 0$$

and

$$|F(f * g)|^2 \leqslant F(f * f^*)\, F(g * g^*), \qquad f, g \in L^\natural(G).$$

(iii) Deduce from (ii) that

$$|F(f)|^2 \leqslant M F(f * f^*), \qquad f \in L^\natural(G),$$

where M is a constant. Iteration of this inequality and use of the formula

$$\sup_\alpha |\alpha(f)| = \lim_{n \to \infty} \|f^n\|^{1/n}$$

(valid for every commutative Banach algebra) gives the inequality

$$|F(f)| \leqslant M \sup_\alpha |\alpha(f)|,$$

as α ranges over all homomorphisms of $L^\natural(G)$ into C.

(iv) Deduce from (iii) that $F(f_0 * f_0^*) = 0$ and as a consequence, $f_0 \equiv 0$.

2. Let (G, K) be a Riemannian symmetric pair of the noncompact type and assume that G has finite center. Let

$$F_f(a) = e^{\rho(\log a)} \int_N f(an)\, dn, \qquad f \in L^\natural(G),$$

the notation being as in §6. As proved in §6,

$$\int_G \varphi_\nu(x) f(x)\, dx = \int_{A_\mathfrak{p}} e^{i\nu(\log a)} F_f(a)\, da.$$

Verify the following assertions:

(i) $F_{f*g} = F_f * F_g, \qquad f, g \in L^\natural(G)$.

(ii) $F_{f^*} = (F_f)^*$.

(iii) The mapping $f \to F_f$ $(f \in L^\natural(G))$ is one-to-one.

3*. Let (G, K) be a Riemannian symmetric pair, G compact. Let Φ denote the set of all spherical functions on G/K. For each $\varphi \in \Phi$ let V_φ denote the set of all complex linear combinations of translates $\varphi^{\tau(g)}$, $g \in G$; let λ_φ denote the homomorphism of $\mathbf{D}(G/K)$ into $\mathbf{C}$ given by $D\varphi = \lambda_\varphi(D)\varphi$. Then

$$L^2(G/K) = \sum_{\varphi \in \Phi} V_\varphi$$

is an orthogonal decomposition of $L^2(G/K)$ into the set of all G-invariant, G-irreducible subspaces of $L^2(G/K)$. Furthermore

$$V_\varphi = \{ f \in C^\infty(G/K) : Df = \lambda_\varphi(D) f \text{ for all } D \in \mathbf{D}(G/K) \}.$$

4. Let U be a compact connected Lie group. Deduce from Exercise E. 3 and Cor. 7.4 that the functional equation

$$\varphi(e)\int \varphi(xuyu^{-1})du = \varphi(x)\int \varphi(uyu^{-1})du$$

characterizes the elements of the U-invariant, U-irreducible subspaces of $L^2(U)$.

NOTES

§1. The invariant integral on a compact Lie group was used by Hurwitz [1], Schur [1], and especially Weyl [1]. Invariant measures on coset spaces occur in special cases in the classical Integral Geometry but Theorem 1.7 (actually for arbitrary locally compact groups) was first proved in Weil [1] (see also Chern [1]). The integral formulas in No.3-No.5 are due to Harish-Chandra [4], I, p. 239,

[5], p. 507, [6], VI, §12. For Prop. 1.19 see also É. Cartan [13], §20. For the case of the complex classical groups many such integral formulas are proved in Gelfand and Naĭmark [1].

§2. The discussion of invariant differential operators on reductive coset spaces is taken from the author's paper [3]. The representation of the center $\boldsymbol{Z}(G)$ (Lemma 2.4) is given in Gelfand [2] and Harish-Chandra [4], p. 192. The commutativity of $\boldsymbol{D}(G/K)$ for symmetric G/K was shown by Gelfand [1]. The proof in the text is modelled after Selberg [1]; in this paper it is also shown that $\boldsymbol{D}(G/K)$ is finitely generated if K is compact, G/K symmetric or not (Cor. 2.8).

§3-§6. For a compact irreducible symmetric space the spherical functions were introduced by É. Cartan [13] and determined for the complex projective space $\boldsymbol{SU}(p+1)/\boldsymbol{T}\,\boldsymbol{SU}(p)$ and for the sphere $\boldsymbol{SO}(p+1)/\boldsymbol{SO}(p)$. Cartan used the spherical functions to prove that a compact irreducible symmetric space U/K can be imbedded into a Euclidean space $\boldsymbol{R}^n$ in such a way that the isometries from U correspond to rotations around 0 in $\boldsymbol{R}^n$. As shown in Mostow [4] and Palais [2] this theorem holds under much more general circumstances. In Gelfand's paper [1] the spherical functions are introduced for general symmetric spaces; here appear the functional equation and the differential equations for the spherical functions. Gelfand also discovered the commutativity of $C^\natural(G)$ (Theorem 4.1) and applied the abstract harmonic analysis to it, the spherical functions corresponding to maximal ideals (Lemma 4.2). The integral formula for the spherical function (Theorem 6.16) was found by Harish-Chandra ([4], p. 64). The proof in the text which also leads to the isomorphism between $\boldsymbol{D}(G/K)$ and $I(\mathfrak{h}_{\mathfrak{p}_0})$ is from Harish-Chandra [12], I, §4. Weyl's formula [1], Kap. IV, for the characters of irreducible representations of compact semisimple Lie groups expresses each character as a ratio of two exponential polynomials on a maximal abelian subalgebra. In view of the duality for symmetric spaces it is natural to expect a similar formula for the spherical functions on G/K (G complex semisimple, K maximal compact). This is indeed the case, see Harish-Chandra [4], p. 253. For the space $\boldsymbol{SL}(n,\boldsymbol{C})/\boldsymbol{SU}(n)$ the formula is given in Gelfand and Naĭmark [1]. A different proof, using differential equations instead of representations, was given by Berezin [2] and Harish-Chandra [12], p. 304. The theory of spherical functions was generalized substantially by Godement [3] and Harish-Chandra [13]. Here the bi-invariance of the spherical function under K is replaced by a weaker condition involving arbitrary irreducible unitary representations of K (rather than just the unit representation). Theorem 4.4 goes back to Gelfand and Raikov [1]; see also Godement [1]. Lemma 4.7 is due to Gelfand and Naĭmark [2]. The formula for the spherical function in the compact case was communicated to the author by Harish-Chandra.

§7. The material of No.1 and No.3 is based on Helgason [1, 3] where some other integral theorems are also established. For symmetric spaces Lemma 7.1 and Cor. 7.11 were also stated in Berezin and Gelfand [1], but verified only for spherical functions and their translates. The generalization of the mean value theorem for harmonic functions (Cor. 7.3) is due to Godement [2]. Theorem 7.13 (i) was proved by the author in [3]. Part (ii) was obtained in 1960 (unpublished); in [2] Allamigeon obtained the same result under the assumption that M is an analytic, complete, simply connected and completely harmonic Riemannian manifold. Laplace's equation on harmonic spaces has been studied by various authors, cf. Feller [1], Günther [1], Lichnerowicz [3], Ruse [1], and Willmore [1].

This Page Intentionally Left Blank

BIBLIOGRAPHY

ADLER, A.

1. Characteristic classes of homogeneous spaces. *Trans. Amer. Math. Soc.* **86** (1957), 348-365.

ALLAMIGEON, A. C.

1. Espaces homogènes symétriques harmoniques. *C. R. Acad. Sci., Paris* **246** (1958), 795-798.
2. Propriétés globales des espaces harmoniques. *C. R. Acad. Sci., Paris* **252** (1961), 1093-1095.

AMBROSE, W.

1. Parallel translation of Riemannian curvature. *Ann. of Math.* **64** (1956), 337-363.
2. The Cartan structural equations in classical Riemannian geometry. *J. Ind. Math. Soc.* **24** (1960), 23-76.

AMBROSE, W., and SINGER, I. M.

1. On homogeneous Riemannian manifolds. *Duke Math. J.* **25** (1958), 647-669.

ARENS, R.

1. Topologies for homeomorphism groups. *Amer. J. Math.* **68** (1946), 593-610.

BARGMANN, V.

1. Irreducible unitary representations of the Lorentz group. *Ann. of Math.* **48** (1947), 568-640.

BARRETT, W.

1. Sur la structure des groupes semi-simples réels. *Bull. Sci. École polytechn. Timisoara,* **9** (1939), 22-90.

BEREZIN, F. A.

1. Representation of complex semi-simple Lie groups in a Banach space. *Doklady Akad. Nauk. S.S.S.R.* **110** (1956), 897-900.
2. Laplace operators on semi-simple Lie groups. *Trudy Moskov. Mat. Obšč.* **6** (1957), 371-463.
3. An analog of Liouville's theorem for symmetric spaces of negative curvature. *Doklady Akad. Nauk. S.S.S.R.* **125** (1959), 1187-1189.

BEREZIN, F. A., and GELFAND, I. M.

1. Some remarks on the theory of spherical functions on symmetric Riemannian manifolds. *Trudy Moskov. Mat. Obšč.* **5** (1956), 311-356.

BEREZIN, F. A., GELFAND, I. M., GRAEV, M. I., and NAĬMARK, M. A.

1. Representation of groups. *Uspekhi Mat. Nauk.* **11** (1956), 13-40.

BEREZIN, F. A., and KARPELEVIČ, F. I.

1. Zonal spherical functions and Laplace operators on some symmetric spaces. *Doklady Akad. Nauk. S.S.S.R.* **118** (1958), 9-12.

BERGER, M.

1. Sur les groupes d'holonomie homogène des variétés à connexions affine et des variétés riemanniennes. *Bull. Soc. Math. France,* **83** (1955), 279-330.
2. Les espaces symétriques non compacts. *Ann. Sci. École Norm. Sup.*, **74** (1957), 85-177.
3. Sur quelques variétés riemanniennes suffisamment pincées. *Bull. Soc. Math. France,* **88** (1960), 57-71.

4. Les variétés riemanniennes homogènes normales simplement connexes à courbure strictement positive. *Ann. Scuola Norm. Sup. Pisa* **15** (1961), 179-246.

BERGMAN, S.

1. Über die Kernfunktion eines Bereiches und ihr Verhalten am Rande. *J. Reine Angew. Math.*, **169** (1933), 1-42.

BHANU MURTI, T. S.

1. Plancherel's measure for the factor space $SL(n,\boldsymbol{R})/SO(n,\boldsymbol{R})$. *Doklady Akad. Nauk. S.S.S.R.*, **133** (1960), 503-506.
2. The asymptotic behavior of zonal spherical functions on the Siegel upper half-plane. *Doklady Akad. Nauk. S.S.S.R.*, **135** (1960), 1027-1030.

BIRKHOFF, G.

1. Representability of Lie algebras and Lie groups by matrices. *Ann. of Math.*, **38** (1937), 526-533.
2. Analytical groups. *Trans Amer. Math. Soc.*, **43** (1938), 61-101.

BOCHNER, S.

1. Vector fields and Ricci curvature. *Bull. Amer. Math. Soc.*, **52** (1946), 776-797.
2. Curvature in Hermitian metric. *Bull. Amer. Math. Soc.*, **53** (1947), 179-195.

BOCHNER, S., and MARTIN, W. T.

1. " Several Complex Variables. " Princeton Univ. Press, Princeton, New Jersey, 1948.

BOREL, A.

1. Exposé 12-13 in Séminaire H. Cartan, Paris, 1949-1950.
2. Les fonctions automorphes de plusieurs variables complexes. *Bull. Soc. Math. France* **80** (1952), 167-182.
3. Les espaces hermitiens symétriques. Séminaire Bourbaki, 1952.
4. Sous-groupes compact maximaux des groupes de Lie. Séminaire Bourbaki, 1952.
5. Kählerian coset spaces of semi-simple Lie groups. *Proc. Nat. Acad. Sci., U.S.A.* **40** (1954), 1147-1151.
6. Topology of Lie groups and characteristic classes. *Bull. Amer. Math. Soc.* **61** (1955), 397-432.
7. On the curvature tensor of the Hermitian symmetric manifolds. *Ann. of Math.* **71** (1960), 508-521.
8. Sous-groupes commutatifs et torsion des groupes de Lie compacts connexes. *Tôhoku Math. J.* **13** (1961), 216-240.
9. Lectures on symmetric spaces. Mass. Inst. of Tech. Notes. 1958.

BOREL, A., and HIRZEBRUCH, F.

1. Characteristic classes and homogeneous spaces, Part I. *Amer. J. Math.* **80** (1958), 459-538.

BOREL, A., and LICHNEROWICZ, A.

1. Espaces riemanniens et hermitiens symétriques. *C. R. Acad. Sci., Paris* **234** (1952) 2332-2334.

BOTT, R.

1. An application of the Morse theory to the topology of Lie groups. *Bull. Soc. Math. France* **84** (1956), 251-282.
2. The stable homotopy of the classical groups. *Ann. of Math.* **70** (1959), 313-337.

BOTT, R., and SAMELSON, H.

1. Applications of the theory of Morse to symmetric spaces. *Amer. J. Math.* **70** (1958), 964-1028.

BOURBAKI, N.

1. "Éléments de mathématiques", Vol. III, "Topologie générale." Hermann, Paris, 1949.
2. "Éléments de mathématique", Vol. VI, "Intégration." Chapters I-V. Hermann, Paris, 1952, 1956.
3. "Éléments de mathématique." Groupes et algèbres de Lie, Chapter I. Hermann, Paris, 1960.

BREMERMANN, H. J.

1. Holomorphic continuation of the kernel function and the Bergman metric in several complex variables. "Lectures on functions of a complex variable," pp. 349-383. The Univ. of Mich. Press, Ann Arbor, Michigan, 1955.

BRUHAT, F.

1. Sur les représentations induites des groupes de Lie. *Bull. Soc. Math. France* **84** (1956), 97-205.
2. Lectures on Lie groups and representations of locally compact groups. Tata Institute of Fundamental Research, Bombay, 1958.

BUSEMANN, H.

1. "The Geometry of Geodesics." Academic Press, New York, 1955.

CALABI, E., and VESENTINI, E.

1. On compact locally symmetric Kähler manifolds. *Ann. of Math.* **71** (1960), 472-507.

CARTAN, É.

1. "Sur la structure des groupes de transformations finis et continus." Thèse, Paris, Nony, 1894; 2nd éd. Vuibert, 1933.
2. Les groupes réels simples finis et continus. *Ann. Sci. École Norm. Sup.* **31** (1914), 263-355.
3. Les tenseurs irréductibles et les groupes linéaires simples et semi-simples. *Bull. Sci. Math.* **49** (1925), 130-152.
4. Le principe de dualité et la théorie des groupes simples et semi-simples. *Bull. Sci. Math.* **49** (1925), 361-374.
5. Sur les espaces de Riemann dans lesquels le transport par parallélism conserve la courbure. *Rend. Acc. Lincei* 3^{I} (1926), 544-547.
6. Sur une classe remarquable d'espaces de Riemann. *Bull. Soc. Math. France* **54** (1926), 214-264.
7. Sur une classe remarquable d'espaces de Riemann. *Bull. Soc. Math. France* **55** (1927), 114-134.
8. La géométrie des groupes de transformations. *J. Math. Pures Appl.* **6** (1927), 1-119.
9. La géométrie des groupes simples. *Ann. Mat. Pura Appl.* **4** (1927), 209-256.
10. Sur certaines formes riemanniennes remarquables des géométries a groupe fondamental simple. *Ann. Sci. École Norm. Sup.* **44** (1927), 345-467.
11. Complément au mémoire "sur la géométrie des groupes simples." *Ann. Math. Pura Appl.* **5** (1928), 253-260.
12. Groupes simples clos et ouverts et géométrie riemannienne. *J. Math. Pures Appl.* **8** (1929), 1-33.
13. Sur la détermination d'un système orthogonal complet dans un espace de Riemann symétrique clos. *Rend. Circ. Mat. Palermo* **53** (1929), 217-252.
14. Sur les invariants intégraux de certains espaces homogènes clos et les propriétés topologiques de ces espaces. *Ann. Soc. Pol. Math.* **8** (1929), 181-225.
15. Sur les représentations linéaires des groupes clos. *Comm. Math. Helv.* **2** (1930), 269-283.

16. La théorie des groupes finis et continus et l'Analysis situs. *Mém. Sci. Math.* fasc. XLII (1930).
17. "Leçons sur la géométrie projective complexe." Gauthier-Villars, Paris, 1931.
18. Les espaces riemanniens symétriques. *Verh. Int. Math. Kongr. Zurich* **1** (1932), 152-161.
19. Sur les domaines bornés homogènes de l'espace de n variables complexes. *Abh. Math. Sem. Univ. Hamburg* **11** (1935), 116-162.
20. "La topologie des espaces représentatifs des groupes de Lie." *Actual. Scient. Ind.* No. 358. Hermann, Paris, 1936.
21. La topologie des espaces homogènes clos. *Mém. Sémin. Anal. Vect., Moscow,* **4** (1937), 388-394.
22. Leçons sur la géométrie des espaces de Riemann." Gauthier-Villars, Paris, 1928, 2nd ed., 1946.
23. "Œuvres complètes," Partie I, Vols. 1 and 2. Gauthier-Villars, Paris, 1952.

CARTAN, É., and SCHOUTEN, J. A.

1. On the geometry of the group-manifold of simple and semi-simple groups. *Proc. Akad. Wetensch., Amsterdam* **29** (1926), 803-815.

CARTAN, H.

1. Notions d'algèbre différentielle; applications aux groupes de Lie et aux variétés où opère un groupe de Lie. Colloque de topologie (espaces fibrés), Bruxelles, 1950.
2. La transgression dans un groupe de Lie et dans un espace fibré principal. Colloque de topologie (espaces fibrés), Bruxelles, 1950.
3. Fonctions automorphes. Séminaire 1957-1958, Paris.

CARTIER, P. and DIXMIER, J.

1. Vecteurs analytiques dans les représentations des groupes de Lie. *Amer. J. Math.* **80** (1958), 131-145.

CHAE, Y.

1. Symmetric spaces which are mapped conformally on each other. *Kyungpook Math. J.* **2** (1959), 65-72.

CHEN, K. T.

1. On the composition functions of nilpotent Lie groups. *Proc. Amer. Math. Soc.* **8** (1957), 1158-1159.

CHERN, S. S.

1. On integral geometry in Klein spaces. *Ann. of Math.* **43** (1942), 178-189.
2. On a generalization of Kähler geometry. "Algebraic Geometry and Topology" (S. Lefschetz symposium), pp. 103-121. Princeton Univ. Press, Princeton, New Jersey. 1957.
3. Topics in differential geometry. Inst. Adv. Study, Princeton, New Jersey. Notes. 1951.

CHERN, S. S., and CHEVALLEY, C.

1. Élie Cartan and his mathematical work. *Bull. Amer. Math. Soc.* **58** (1952), 217-250.

CHEVALLEY, C.

1. An algebraic proof of a property of Lie groups. *Amer. J. Math.* **63** (1941), 785-793.
2. "Theory of Lie groups," Vol. 1. Princeton Univ. Press, Princeton, New Jersey, 1946.
3. Algebraic Lie algebras. *Ann. of Math.* **48** (1947), 91-100.
4. The Betti numbers of the exceptional simple Lie groups. *Proc. Int. Congr. of Math. Harvard* ***II*** (1950), 21-24.

5. Sur certains groupes simples. *Tôhoku Math. J.* **7** (1955), 14-66.
6. Théorie des groupes de Lie," Vol. III, Paris, Hermann, 1955.
7. Invariants of finite groups generated by reflections. *Amer. J. Math.* **77** (1955), 778-782.

COHN, P. M.
1. "Lie Groups." Cambridge Univ. Press, Cambridge, 1957.

COXETER, H. S. M.
1. Discrete groups generated by reflections. *Ann. of Math.* **35** (1934), 588-621.

CRITTENDEN, R.
1. Minimum and conjugate points in symmetric spaces. *Canad. J. Math.* **14** (1962), 320-328.

DALLA VOLTA, V.
1. Geodesic manifolds in the spaces of Siegel-Hua. *Ann. Mat. Pura Appl.* **46** (1958), 19-42.

DE RHAM, G.
1. Sur la reductibilité d'un espace de Riemann. *Comment. Math. Helv.* **26** (1952), 328-344.
2. "Variétés différentiables." Hermann, Paris, 1955.

DIXMIER, J.
1. Représentations intégrables du groupe de de Sitter. *Bull. Soc. Math. France* **89** (1961), 9-41.

DYNKIN, E. B.
1. The structure of semi-simple Lie algebras. *Uspekhi Mat. Nauk.* **2** (1947), 59-127.
2. Topological characteristics of homomorphisms of compact Lie groups. *Mat. Sb.* **35** (1954), 129-177.

DYNKIN, E. B. and ONIŠČIK, A. L.
1. Compact global Lie groups. *Uspekhi Mat. Nauk.* **10** (1955), 3-74.

EHRENPREIS, L., and MAUTNER, F. I.
1. Some properties of the Fourier transform on semi-simple Lie groups, I. *Ann. of Math.* **61** (1955), 406-439; II, III. *Trans. Amer. Math. Soc.* **84** (1957), 1-55; **90** (1959), 431-484.

EHRESMANN, C.
1. Sur certains espaces homogènes de groupes de Lie. *Enseignement Math.* **35** (1936), 317-333.

EISENHART, L. P.
1. "Continuous Groups of Transformations." Princeton Univ. Press, Princeton, New Jersey, 1933.

ERDÉLYI, A., *et al.*
1. "Higher Transcendental Functions" (Bateman Manuscript Project), Vol. I. McGraw-Hill, New York, 1953.

FEDENKO, A. S.
1. Symmetric spaces with simple non-compact fundamental groups. *Doklady Akad. Nauk. S.S.S.R.* **108** (1956), 1026-1028.

FELLER, W.
1. Über die Lösungen der linearen partiellen Differentialgleichungen zweiter Ordnung vom elliptischen Typus. *Math. Ann.* **102** (1930) 633-649.

FREUDENTHAL, H.
1. Sur les invariants caractéristiques des groupes semi-simples. *Nederl. Akad. Wetensch. Proc.* Ser. A **56** (1953), 90-94.

2. Zur Berechnung der Charaktere der halbeinfachen Lieschen Gruppen, I, II. *Nederl. Akad. Wetensch. Proc.* Ser. A **57** (1954), 369-376, 487-491.
3. Zur Klassifikation der einfachen Lie-Gruppen. *Nederl. Akad. Wetensch. Proc.* Ser. A **60** (1958), 379-383.
4. Lie groups. Dittoed lecture notes, Yale University, New Haven, Connecticut, 1961, 309 p.

FRÖLICHER, A.
1. Zur Differentialgeometrie der komplexen Strukturen. *Math. Ann.* **129** (1955), 50-95.

GANTMACHER, F.
1. Canonical representation of automorphisms of a complex semi-simple Lie group. *Mat. Sb.* **5** (1939), 101-144.
2. On the classification of real simple Lie groups. *Mat. Sb.* **5** (1939), 217-249.

GÅRDING, L.
1. Vecteurs analytiques dans les représentations des groupes de Lie. *Bull. Soc. Math. France* **88** (1960), 73-93.

GELFAND, I. M.
1. Spherical functions on symmetric spaces. *Doklady Akad. Nauk. S.S.S.R.* **70** (1950), 5-8.
2. The center of an infinitesimal group algebra. *Mat. Sb.* **26** (1950), 103-112.

GELFAND, I. M. and FOMIN, S. V.
1. Geodesic flows on manifolds of constant negative curvature. *Uspekhi Mat. Nauk.* **7** (1952), 118-137.

GELFAND, I. M., and GRAEV, M. I.
1. Analogue of the Plancherel formula for real semi-simple Lie groups. *Doklady Akad. Nauk. S.S.S.R.* **92** (1953), 461-464.
2. Analogue of the Plancherel formula for the classical groups. *Trudy Moscov. Mat. Obšč.* **4** (1955), 375-404.
3. The geometry of homogeneous spaces, group representations in homogeneous spaces and questions in integral geometry related to them, I. *Trudy Moskov. Mat. Obšč.* **8** (1959), 321-390.

GELFAND, I. M., and NAĬMARK, M. A.
1. "Unitary Representations of the Classical Groups." (Trudy Mat. Inst. Steklov. **36**). Moskov-Leningrad, 1950. German transl. Akademie-Verlag, Berlin, 1957.
2. Unitary representations of a unimodular group containing the identity representation of the unitary subgroup. *Trudy Moskov. Mat. Obšč.* 1 (1952), 423-475.

GELFAND, I. M., and RAIKOV, D. A.
1. Irreducible unitary representations of locally compact groups. *Mat. Sb.*, **13** (1943), 301-316.

GODEMENT, R.
1. Les fonctions de type positif et la théorie des groupes. *Trans. Amer. Math. Soc.* **63** (1948), 1-84.
2. Une généralisation du théorème de la moyenne pour les fonctions harmoniques. *C. R. Acad. Sci., Paris* **234** (1952), 2137-2139.
3. A theory of spherical functions, I. *Trans. Amer. Math. Soc.* **73** (1952), 496-556.
4. Introduction aux travaux de A. Selberg. Séminaire Bourbaki, 1957.
5. Articles in H. Cartan [3].

GOTO, M.
1. Faithful representations of Lie groups, I. *Math. Japonica* **1** (1948), 107-119.

GUNNING, R. C.

1. Factors of automorphy and other formal cohomology groups for Lie groups. *Ann. of Math.* **69** (1959), 314-326.

GÜNTHER, P.

1. Über einige spezielle Probleme aus der Theorie der linearen partiellen Differentialgleichungen 2. Ordnung. *Ber. Verh. Sächs. Akad. Wiss., Leipzig* **102** (1957), 1-50.

HADAMARD, J.

1. Les surfaces à courbures opposées. *J. Math. Pures Appl.* **4** (1898), 27-73.
2. Sur les éléments linéaires à plusieurs dimensions. *Bull. Soc. Math. France* **25** (1901), 37-40.

HANO, J.

1. On Kählerian homogeneous spaces of unimodular groups. *Amer. J. Math.* **69** (1957). 885-900.

HANO, J. and MATSUSHIMA, Y.

1. Some studies on Kählerian homogeneous spaces. *Nagoya Math. J.* **11** (1957), 77-92.

HANO, J., and MORIMOTO, A.

1. Note on the group of affine transformations of an affinely connected manifold. *Nagoya Math. J.* **8** (1955), 71-81.

HARISH-CHANDRA

1. On representations of Lie algebras. *Ann. of Math.* **50** (1949), 900-915.
2. Lie algebras and the Tannaka duality theorem. *Ann. of Math.* **51** (1950), 299-330.
3. On some applications of the universal enveloping algebra of a semi-simple Lie algebra. *Trans. Amer. Math. Soc.* **70** (1951), 28-96.
4. Representations of semi-simple Lie groups, I, II, III, *Trans. Amer. Math. Soc.* **75** (1953), 185-243; **76** (1954), 26-65, 234-253.
5. The Plancherel formula for complex semi-simple Lie groups. *Trans. Amer. Math. Soc.* **76** (1954), 485-528.
6. Representations of semi-simple Lie groups, IV, V, VI. *Amer. J. Math.* **77** (1955), 743-777, **78** (1956), 1-41, 564-628.
7. On a lemma of F. Bruhat. *J. Math. Pures Appl.* **35** (1956), 203-210.
8. The characters of semi-simple Lie groups. *Trans. Amer. Math. Soc.* **83** (1956), 98-163.
9. Differential operators on a semi-simple Lie algebra. *Amer. J. Math.* **79** (1957), 87-120.
10. Fourier transforms on a semi-simple Lie algebra, I, II. *Amer. J. Math.* **79** (1957). 193-257, 653-686.
11. A formula for semi-simple Lie groups. *Amer. J. Math.* **79** (1957), 733-760.
12. Spherical functions on a semi-simple Lie group, I, II. *Amer. J. Math.* **80** (1958), 241-310, 553-613.
13. Some results on differential equations and their applications. *Proc. Nat. Acad. Sci. U.S.A.* **45** (1959), 1763-1764.
14. Plancherel formula for the 2×2 real unimodular group. *Proc. Nat. Acad. Sci., U.S.A.* **38** (1952), 337-342.

HELGASON, S.

1. Partial differential equations on Lie groups. Scand. Math. Congr. XIII, Helsinki, 1957, 110-115.
2. On Riemannian curvature of homogeneous spaces. *Proc. Amer. Math. Soc.* **9** (1958), 831-838.
3. Differential operators on homogeneous spaces. *Acta Math.* **102** (1959), 239-299.

4. Some remarks on the exponential mapping for an affine connection. *Math. Scand.* **9** (1961), 129-146.
5. Some results on invariant theory. *Bull. Amer. Math. Soc.* **68** (1962).

HERMANN, R.
1. Variational completeness for compact symmetric spaces. *Proc. Amer. Math. Soc.* 11 (1960), 544-546.
2. Geodesics of bounded symmetric domains. *Comment. Math. Helv.* **35** (1961), 1-8.
3. A Poisson kernel for certain homogeneous spaces. *Proc. Amer. Math. Soc.* **12** (1961), 892-899.

HICKS, N.
1. A theorem on affine connexions. *Illinois J. Math.* **3** (1959), 242-254.
2. An example concerning affine connexions. *Proc. Amer. Math. Soc.* **11** (1960), 952-956.

HIRZEBRUCH, F.
1. Automorphe Formen und der Satz von Riemann-Roch. Int. Symposium on Algebraic Topology, Mexico, 1958.

HOCHSCHILD, G.
1. The automorphism group of a Lie group. *Trans. Amer. Math. Soc.* **72** (1952), 209-216.

HOPF, H.
1. Über den Rang geschlossenen Liescher Gruppen. *Comment. Math. Helv.* **13** (1940), 119-143.
2. Maximale Toroide und singuläre Elemente in geschlossenen Lieschen Gruppen. *Comment. Math. Helv.* **15** (1943), 59-70.

HOPF, H. and RINOW, W.
1. Über den Begriff der vollständigen differentialgeometrischen Fläche. *Comment. Math. Helv.* **3** (1931), 209-225.

HUA, L. K.
1. On the Riemannian curvature of the non-Euclidean space of several complex variables. *Acta Math. Sinica*, **4** (1954), 141-168.
2. "Harmonic Analysis of Functions of Several Complex Variables in the Classical Domains." Peking, 1958. Russian Transl. Moscow, 1959.

HUNT, G. A.
1. Semi-groups of measures of Lie groups. *Trans. Amer. Math. Soc.* **81** (1956), 264-293.
2. A theorem of Élie Cartan. *Proc. Amer. Math. Soc.* **7** (1956), 307-308.

HUREWICZ, W., and WALLMAN, H.
1. "Dimension Theory." Princeton Univ. Press, Princeton, New Jersey, 1948.

HURWITZ, A.
1. Über die Erzeugung der Invarianten durch Integration. *Gött. Nachr.* (1897), 71-90.

IWAMOTO, H.
1. On integral invariants and Betti numbers of symmetric Riemannian spaces, I, II. *J. Math. Soc. Japan* **1** (1949), 91-110, 235-243.

IWASAWA, K.
1. On some types of topological groups. *Ann. of Math.* **50** (1949), 507-558.

JACOBSON, N.
1. "Lie Algebras." Interscience, New York, 1962.

JOHN, F.
1. "Plane Waves and Spherical Means, Applied to Partial Differential Equations." Interscience, New York, 1955.

Kähler, E.
1. Über eine bemerkenswerte Hermitische Metrik. *Abh. Math. Sem. Univ. Hamburg* **9** (1933), 173-186.

Karpelevič, F. I.
1. Surfaces of transitivity of a semi-simple subgroup of the group of motions of a symmetric space. *Doklady Akad. Nauk. S.S.S.R.* **93** (1953), 401-404.
2. The simple subalgebras of the real Lie algebras. *Trudy Moskov. Mat. Obšč.* **4** (1955), 3-112.
3. Geodesics and harmonic functions on symmetric spaces. *Doklady Akad. Nauk. S.S.S.R.* **124** (1959), 1199-1202.

Kelley, J. L.
1. "General Topology." Van Nostrand, Princeton, New Jersey, 1955.

Killing, W.
1. Die Zusammensetzung der stetigen endlichen Transformationsgruppen, I, II, III, IV. *Math. Ann.* **31** (1888), 252-290; **33** (1889), 1-48; **34** (1889), 57-122; **36** (1890), 161-189.

Klingen, H.
1. Diskontinuierliche Gruppen in symmetrischen Räumen, I, II. *Math. Ann.* **129** (1955), 345-369; **132** (1956), 134-144.

Kobayashi, S.
1. Une remarque sur la connexion affine symétrique. *Proc. Japan Acad.* **31** (1955), 13-14.
2. Espaces à connexions affines et riemanniens symétriques. *Nagoya Math. J.* **9** (1955), 25-37.
3. Fixed points of isometries. *Nagoya Math. J.* **13** (1958), 63-68.
4. Geometry of bounded domains. *Trans. Amer. Math. Soc.* **92** (1959), 267-290.

Koecher, M.
1. Positivitätsbereiche im $\boldsymbol{R}^n$. *Amer. J. Math.* **79** (1957), 575-596.
2. Die Geodätischen von Positivitätsbereichen. *Math. Ann.* **135** (1958), 192-202.

Kolchin, E.
1. Algebraic matric groups and the Picard-Vessiot theory of homogeneous linear ordinary differential equations. *Ann. of Math.* **49** (1948), 1-42.

Kostant, B.
1. On the conjugacy of real Cartan subalgebras, I. *Proc. Nat. Acad. Sci. U.S.A.* **41** (1955), 967-970.
2. On differential geometry and homogeneous spaces, I, II. *Proc. Nat. Acad. Sci. U.S.A.* **42** (1956), 258-261, 354-357.
3. A characterization of the classical groups. *Duke Math. J.* **25** (1958), 107-124.
4. The principal three-dimensional subgroup and the Betti-numbers of a complex simple Lie group. *Amer. J. Math.* 81 (1959), 973-1032.
5. A characterization of invariant affine connections. *Nagoya Math. J.* **16** (1960), 35-50.
6. On invariant skew-tensors. *Proc. Nat. Acad. Sci. U.S.A.* **40** (1955), 148-151.
7. Lie algebra cohomology and the generalized Borel-Weil theorem. *Ann. of Math.* **74** (1961). 329-387.

Koszul, J. L.
1. Sur la forme hermitienne canonique des espaces homogènes complexes. *Canad. J. Math.* **7** (1955), 562-576.

2. Variétés Kähleriennes. Universidade de São Paulo, 1957, 100 p.
3. Exposés sur les espaces homogènes symétriques. Soc. Math. São Paulo, 1959, 71 p.

KREIN, M. G.

1. Positive-hermitian kernels on homogeneous spaces, I, II. *Ukrain. Math. J.* **1** (1949), 64-98; **2** (1950), 10-59.

KUNZE, R. A., and STEIN, E. M.

1. Uniformly bounded representations and harmonic analysis of the 2×2 real unimodular group. *Amer. J. Math.* **82** (1960), 1-62.

LARDY, P.

1. Sur la détermination des structures réelles de groupes simples, finis et continus, au moyen des isomorphies involutives. *Comment. Math. Helv.* **8** (1935-1936), 189-234.

LEDGER, A. J.

1. Symmetric harmonic spaces. *J. London Math. Soc.* **32** (1957), 53-56.

LIBERMANN, P.

1. Sur les structures presque complexes et autres structures infinitésimales régulières. *Bull. Soc. Math. France* **83** (1955), 195-224.

LICHNEROWICZ, A.

1. Courbure, nombres de Betti, et espaces symétriques. *Proc. Int. Congr. of Math. Harvard II* (1950), 216-223.
2. Espaces homogènes kähleriens. Coll. Int. Géom. Diff., Strasbourg, (1953), 171-184.
3. Équations de Laplace et espaces harmoniques. Premier Colloq. sur les équations aux dérivées partielles, Louvain, 1953, 9-23.
4. Un théorème sur les espaces homogènes complexes. *Arch. Math.* **5** (1954), 207-215.
5. "Théorie globale des connexions et des groupes d'holonomie." Éd. Cremonese, Rome, 1955.
6. "Géométrie des groupes de transformations." Dunod, Paris, 1958.

LIE, S.

1. "Theorie der Transformationsgruppen I, II, III." Unter Mitwirkung von F. Engel. Leipzig, 1888, 1890, 1893.

LISTER, W. G.

1. A structure theory of Lie triple systems. *Trans. Amer. Math. Soc.* **72** (1952), 217-242.

LOWDENSLAGER, D. B.

1. Potential theory in bounded symmetric homogeneous complex domains. *Ann. of Math.* **67** (1958), 467-484.
2. Potential theory and a generalized Jensen-Nevanlinna formula for functions of several complex variables. *J. Math. Mech.* **7** (1958), 207-218.

MAASS, H.

1. Spherical functions and quadratic forms. *J. Indian Math. Soc.* **20** (1956), 117-162.
2. Zur Theorie der Kugelfunktionen einer Matrixvariablen. *Math. Ann.* **135** (1958). 391-416.

MALCEV, A.

1. On the theory of Lie groups in the large. *Mat. Sb.* **16** (58) (1945), 163-190.

MATSUSHIMA, Y.

1. Un théorème sur les espaces homogènes complexes. *C. R. Acad. Sci., Paris* **241** (1955), 785-787.

2. Espaces homogènes de Stein des groupes de Lie complexes. *Nagoya Math. J.* **16** (1960), 205-218.

Mautner, F.

1. Fourier analysis and symmetric spaces. *Proc. Nat. Acad. Sci. U.S.A.* **37** (1951), 529-533.
2. Geodesic flows on symmetric Riemann spaces. *Ann. of Math.* **65** (1957), 416-431.
3. Spherical functions over $\mathfrak{P}$-adic fields, I. *Amer. J. Math.* **80** (1958), 441-457.

Mayer, W., and Thomas, T. Y.

1. Foundations of the theory of Lie groups. *Ann. of Math.* **30** (1935), 770-822.

Miller, K. S. and Murray, F. J.

1. "Existence Theorems for Ordinary Equations." New York Univ. Press, New York, 1954.

Montgomery, D. and Zippin, L.

1. "Topological Transformation Groups." Interscience, New York, 1955.

Morita, K.

1. On the kernel functions of symmetric domains. *Sci. Rep. Tokyo Kyoiku Daigaku*, Sect. A **5** (1956), 190-212.

Mostow, G. D.

1. A new proof of É. Cartan's theorem on the topology of semi-simple groups. *Bull. Amer. Math. Soc.* **55** (1949), 969-980.
2. The extensibility of local Lie groups of transformations and groups on surfaces. *Ann. of Math.* **52** (1950), 606-636.
3. Some new decomposition theorems for semi-simple Lie groups. *Mem. Amer. Math. Soc.* **14** (1955), 31-54.
4. Equivariant embeddings in Euclidean space. *Ann. of Math.* **65** (1957), 432-446.

Murakami, S.

1. On the automorphisms of a real semi-simple Lie algebra. *J. Math. Soc. Japan* **4** (1952), 103-133.
2. Supplements and corrections to [1]. *J. Math. Soc. Japan* **5** (1953), 105-112.

Myers, S. B.

1. Riemannian manifolds in the large. *Duke Math. J.* **1** (1935), 39-49.

Myers, S. B. and Steenrod, N.

1. The group of isometries of a Riemannian manifold. *Ann. of Math.* **40** (1939), 400-416.

Nagano, T.

1. Homogeneous sphere bundles and the isotropic Riemannian manifolds. *Nagoya Math. J.* (1959), 29-55.

Naĭmark, M. A.

1. On the description of all unitary representations of the complex classical groups, I, II. *Mat. Sb.* **35** (1954), 317-356; **37** (1955), 121-140.
2. "Representations of the Lorentz Group." Moscow, 1958
3. "Normierte Algebren," Moscow, 1956. German Transl. V. E. B. Deutscher Verlag der Wissenschaften, Berlin, 1959.

Nelson, E.

1. Analytic vectors. *Ann. of Math.* **70** (1959), 572-615.

Newlander, A. and Nirenberg, L.

1. Complex analytic coordinates in almost complex manifolds. *Ann. of Math.* **65** (1957), 391-404.

NOMIZU, K.

1. On the group of affine transformations of an affinely-connected manifold. *Proc. Amer. Math. Soc.* **4** (1953), 816-823.
2. Invariant affine connections on homogeneous spaces. *Amer. J. Math.* **76** (1954), 33-65.
3. Reduction theorem for connections and its application to the problem of isotropy and holonomy groups of a Riemannian manifold. *Nagoya Math. J.* **9** (1955), 57-66.
4. "Lie Groups and Differential Geometry." (Math. Soc. Japan Publ., Vol. 2), 1956.
5. On infinitesimal holonomy and isotropy groups. *Nagoya Math. J.* **11** (1957), 111-114.
6. On local and global existence of Killing vector fields. *Ann. of Math.* **72** (1960), 105-120.

PALAIS, R. S.

1. On the differentiability of isometries. *Proc. Amer. Math. Soc.* **8** (1957), 805-807.
2. Imbeddings of compact differentiable transformation groups in orthogonal representations. *J. Math. Mech.* **6** (1957), 673-678.
3. A definition of the exterior derivative in terms of Lie derivatives. *Proc. Amer. Math. Soc.* **5** (1954), 902-908.

PETER, F., and WEYL, H.

1. Die Vollständigkeit der primitiven Darstellungen einer geschlossenen kontinuierlichen Gruppe. *Math. Ann.* **97** (1927), 737-755.

PFLUGER, A.

1. "Theorie der Riemannschen Flächen." Springer, Berlin, 1959.

PONTRJAGIN, L. S.

1. "Topological Groups," 2nd ed. Moscow, 1954. German transl. Teubner, Leipzig, 1957, 1958.

PYATETZKI-SHAPIRO, I.

1. On a problem proposed by É. Cartan. *Doklady Akad. Nauk. S.S.S.R.* **124** (1959), 272-273.

RAŠEVSKIĬ, P. K.

1. On the geometry of homogeneous spaces. *Trudy Sem. Vector. Tenzor. Analizu* **9** (1952), 49-74.

RAUCH, H. E.

1. Geodesics, symmetric spaces and differential geometry in the large. *Comment. Math. Helv.* **27** (1953), 294-320.
2. The global study of geodesics in symmetric and nearly symmetric Riemannian manifolds. *Comment. Math. Helv.* **35** (1961), 111-125.
3. "Geodesics and Curvature in Differential Geometry in the Large." Yeshiva Univ., New York, 1959.

RIEMANN, B.

1. "Über die Hypothesen, welche der Geometrie zu Grunde liegen." Habilitationvorlesung, 1854. Gesamm. Werke, 2nd ed., Leipzig, 1892.

RINOW, W.

1. Über die Zusammenhänge zwischen der Differentialgeometrie im grossen und im kleinen. *Math. Z.* **35** (1932), 512-538.

ROTHAUS, O. S.

1. Domains of positivity. *Abh. Math. Sem. Univ. Hamburg*, **24** (1960), 189-235.

ROZENFELD, B. A.

1. "Non-Euclidean Geometries." Moscow, 1955.

2. A geometric interpretation of symmetric spaces with simple fundamental group. *Doklady Akad. Nauk. S.S.S.R.* **110** (1956), 23-26.
3. On the theory of symmetric spaces af rank one. *Mat. Sb.* **41** (83) (1957), 373-380.

RUSE, H. S.
1. On the elementary solution of Laplace's equation. *Proc. Edinburgh Math. Soc.* **2** (1931), 135-139.

SAMELSON, H.
1. A note on Lie groups. *Bull. Amer. Math. Soc.* **52** (1946), 870-873.
2. Topology of Lie groups. *Bull. Amer. Math. Soc.* **58** (1952), 2-37.
3. On curvature and characteristic of homogeneous spaces. *Michigan Math. J.* **5** (1958), 13-18.

SATAKE, I.
1. On a theorem of É. Cartan. *J. Math. Soc. Japan* **2** (1951), 284-305.
2. A remark on bounded symmetric domains. *Sci. Papers Coll. Gen. Ed. Univ. Tokyo* **3** (1953), 131-144.
3. On representations and compactifications of symmetric Riemann spaces. *Ann. of Math.* **71** (1960), 77-110.
4. On compactifications of the quotient space for arithmetically defined discontinuous groups. *Ann. of Math.* **72** (1960), 555-580.

SCHOUTEN, J. A.
1. "Ricci Calculus," 2nd ed. Springer, Berlin, 1954.

SCHREIER, O.
1. Abstrakte kontinuierliche Gruppen. *Abh. Math. Sem. Univ. Hamburg* **4** (1925), 15-32.

SCHUR, I.
1. Neue Anwendungen der Integralrechnung auf Probleme der Invariantentheorie. *Sitzungsber. d. Preuss. Ak. d. Wiss.* (1927), 189, 297, 346.

SCHWARTZ, L.
1. "Théorie des distributions," I, II. Hermann, Paris, 1950, 1951.

SEGAL, I. E.
1. Hypermaximality of certain operators on Lie groups. *Proc. Amer. Math. Soc.* **3** (1952), 13-15.
2. A theorem on the measurability of group-invariant operators. *Duke Math. J.* **26** (1959), 549-552.

SEIFERT, H. and THRELFALL, W.
1. "Lehrbuch der Topologie." Teubner. Leipzig, 1934.

SELBERG, A.
1. Harmonic analysis and discontinuous groups in weakly symmetric Riemannian spaces with applications to Dirichlet series. *J. Ind. Math. Soc.* **20** (1956), 47-87.

SÉMINAIRE "SOPHUS LIE"
1. "Théorie des algèbres de Lie, Topologie des groupes de Lie." Paris, 1955.

SIEGEL, C. L.
1. Symplectic geometry. *Amer. J. Math.* **65** (1943), 1-86.
2. "Analytic functions of several complex variables." Inst. Adv. Study, Princeton, New Jersey, 1949.

SIMONS, J.
1. On the transitivity of holonomy systems. To appear in *Ann. of Math.*

SINGER, I. M.

1. Infinitesimally homogeneous spaces. *Comm. Pure Appl. Math.* **13** (1960), 685-697.

SINYUKOV, N. S.

1. On geodesic mappings of Riemannian spaces onto symmetric Riemannian spaces. *Doklady Akad. Nauk. S.S.S.R.* **98** (1954), 21-23.

SIROKOV, P. A.

1. On a certain type of symmetric spaces. *Mat. Sb.* **41** (83) (1957), 361-372.

STEENROD, N.

1. "Topology of Fibre Bundles." Princeton Univ. Press, Princeton, New Jersey, 1951.

STIEFEL, E.

1. Kristallographische Bestimmung der Charaktere der geschlossenen Lie'schen Gruppen. *Comment. Math. Helv.*, **17** (1944-45), 165-200.

STRUIK, D. J.

1. "Grundzüge der Mehrdimensionalen Differentialgeometrie in direkter Darstellung." Springer, Berlin, 1922.

SUGIURA, M.

1. Conjugate classes of Cartan subalgebras in real semi-simple Lie algebras. *J. Math. Soc. Japan* 11 (1959), 374-434.
2. Spherical functions and representation theory of compact Lie groups. *Sci. Papers Coll. Gen. Ed. Univ. Tokyo* **10** (1960), 187-193.

SUMITOMO, T.

1. Projective and conformal transformations in compact Riemannian manifolds. *Tensor (N.S.)* **9** (1959), 113-135.

TAKAHASHI, R.

1. Sur les groupes de Lorentz généralisés. *C. R. Acad. Sci., Paris* **252** (1961), 835-837, 980-982.

TAMAGAWA, T.

1. On Selberg's trace formula. *J. Fac. Sci. Univ. Tokyo.* Sect. I, **8** (1960), 363-386.

TELEMAN, C.

1. Sur une classe d'espaces riemanniens symétriques. *Rev. Math. Pures Appl.* **2** (1957), 445-470.

TITS, J.

1. Sur certains classes d'espaces homogènes de groupes de Lie. *Acad. Roy. Belg. Cl. Sci. Mém. Coll.* **29** (1955), No. 3.

VAN DANTZIG, D., and VAN DER WAERDEN, B. L.

1. Über metrische homogene Räume. *Abh. Math. Sem. Univ. Hamburg* **6** (1928), 367-376.

VAN DER WAERDEN, B. L.

1. Die Klassifizierung der einfachen Lie'schen Gruppen. *Math. Z.* **37** (1933), 446-462.

VAN EST, W. T.

1. Dense imbeddings of Lie groups. *Nederl. Akad. Wetensch., Proc.* Ser. A, **54** (1951), 321-328.

VILENKIN, N. YA.

1. Bessel functions and representations of the group of Euclidean motions. *Uspekhi Mat. Nauk.* **11** (1956), 69-112.

VON NEUMANN, J.

1. Über die analytischen Eigenschaften von Gruppen linearen Transformationen und ihrer Darstellungen. *Math. Z.* **30** (1929), 3-42.

VRĂNCEANU, G.

1. Sur une classe d'espaces symétriques. "Bericht von der Riemann-Tagung des Forschungsinstituts für Mathematik," pp. 112-123. Akademie-Verlag, Berlin, 1957.

WANG, H. C.

1. Two-point homogeneous spaces. *Ann. of Math.* **55** (1952), 177-191.

WEIL, A.

1. "L'intégration dans les groupes topologiques et ses applications." Hermann, Paris, 1940.
2. "Variétés Kähleriennes." Hermann, Paris, 1958.

WEYL, H.

1. Theorie der Darstellung kontinuierlicher halbeinfacher Gruppen durch lineare Transformationen, I, II, III, und Nachtrag. *Math. Z.* **23** (1925), 271-309; **24** (1926), 328-376, 377-395, 789-791.
2. The structure and representations of continuous groups. Inst. Adv. Study, Princeton, New Jersey. Notes. 1935.
3. "The Classical Groups." Princeton Univ. Press, Princeton, New Jersey, 1939.

WHITEHEAD, J. H. C.

1. Affine spaces of paths which are symmetric about each point. *Math. Z.* **35** (1932), 644-659.
2. Convex regions in the geometry of paths. *Quart. J. Math.* 3 (1932), 33-42.
3. On the covering of a complete space by the geodesics through a point. *Ann. of Math.* **36** (1935), 679-704.

WILLMORE, T. J.

1. Mean value theorems in harmonic Riemannian spaces. *J. London Math. Soc.* **25** (1950), 54-57.
2. "An Introduction to Differential Geometry." Oxford Univ. Press, London, 1959.

WITT, E.

1. Treue Darstellung Lie'scher Ringe. *J. Reine Angew. Math.* **177** (1937), 152-166.
2. Spiegelungsgruppen und Aufzählung halb-einfacher Lie'scher Ringe. *Abh. Math. Sem. Univ. Hamburg* **14** (1941), 289-322.

WOLF, J. A.

1. Homogeneous manifolds of constant curvature. *Comment. Math. Helv.* **36** (1961), 112-147.

YANO, K.

1. "The Theory of Lie Derivatives and Its Applications." North-Holland, Amsterdam, and Interscience, New York, 1957.

YEN, C. T.

1. Sur les représentations linéaires de certains groupes et les nombres de Betti des espaces homogènes symétriques. *C. R. Acad. Sci., Paris* **228** (1949), 1367-1369.

ZARISKI, O., and SAMUEL, P.

1. "Commutative Algebra," Vols I and II. Van Nostrand, Princeton, New Jersey, 1958, 1960.

This Page Intentionally Left Blank

LIST OF NOTATIONAL CONVENTIONS

I. Set theory. Let A and B be sets. The symbol $A \subset B$ means that A is a subset of B. If $A \subset B$ and $A \neq B$, then A is called a *proper* subset of B. The empty set is denoted by $\emptyset$. The set $A - B$ is the set of elements in A not in B. The symbols $\cap, \cup$, respectively, denote intersection and union of sets. The symbol $x \in A$ ($x \notin A$) means that x is (x is not) an element of the set A. The subset of A consisting of x alone is denoted $\{x\}$. If M and N are sets the symbol f: $M \to N$ means a mapping of M into N. If $M \subset N$ and $f(m) = m$ for all $m \in M$, f is called the *identity mapping* of M into N and is denoted by I or 1. If $f: M \to N$ and $g: N \to P$, then the mapping which assigns to every $m \in M$ the element $g(f(m)) \in P$ is denoted $g \circ f$. If $f: M \to N$ and $A \subset N$, then $f^{-1}(A)$ denotes the set of elements $m \in M$ for which $f(m) \in A$. If $\mathscr{P}$ is a property and M a set then $\{x \in M: x \text{ has property } \mathscr{P}\}$ denotes the set of $x \in M$ with property $\mathscr{P}$. Thus $f^{-1}(A) = \{m \in M: f(m) \in A\}$. The sign $\Rightarrow$ means "implies." In order to save parentheses, the image $f(m)$ of m under a mapping f will sometimes be denoted $f \cdot m$. A mapping $f: M \to N$ is said to be *one-to-one* if $m_1 \neq m_2 \Rightarrow f(m_1) \neq f(m_2)$. If $f(M) = N$, f is said to map M *onto* N ("f is onto").

II. Algebra. The identity element of a group will usually be denoted by e. If K is a subgroup of a group G, the symbol G/K denotes the set of left cosets gK, $g \in G$. When K is considered as an element in G/K it will sometimes be denoted by $\{K\}$. If $x \in G$, the mapping $gK \to xgK$ of G/K onto itself will be denoted by $\tau(x)$.

By a *field* we shall always mean a commutative field of characteristic 0. Let V be a vector space over a field K. The *dual space*, consisting of all linear mappings of V into K, is denoted by V^* or $V^\wedge$. The dimension of the vector space V is denoted by $\dim V$ or $\dim_K V$. If $\dim V < \infty$, then $(V^*)^*$ can be identified with V. If $e_1, \ldots, e_n$ is a basis of V and $f_1, \ldots, f_n$ are linear mappings of V into K such that $f_i(e_j) = \delta_{ij}$ (Kronecker symbol), then $f_1, \ldots, f_n$ is called the *basis* of V^* *dual* to $e_1, \ldots, e_n$. Let W be a subspace of V. A *basis of* V (mod W) is a set of elements in V which together with a basis of W constitute a basis of V. The number of elements in a basis of V (mod W) is called the *codimension* of W. If A and B are subspaces of a vector space such that each $v \in V$ can be written $v = a + b$ where $a \in A$, $b \in B$, then we write $V = A + B$. If, in addition $A \cap B = \{0\}$, V is called the *direct sum* of A and B and the subspace B is said to be *complementary* to A. If V and W are vector spaces over the same field K, then the ***product*** of V and W, denoted $V \times W$, is the set of all pairs (v, w) where $v \in V$, $w \in W$, turned into a vector space over K by the rules

$$(v, w) + (v', w') = (v + v', w + w'), \quad \alpha(v, w) = (\alpha v, \alpha w), \quad \alpha \in K.$$

The subsets $\{(v, 0): v \in V\}$ and $\{(0, w): w \in W\}$ are subspaces of $V \times W$, isomorphic to V and W, respectively, and $V \times W$ is the direct sum of those subspaces.

If $v \in V$, $v^* \in V^*$, then the value $v^*(v)$ will sometimes be denoted by $\langle v, v^* \rangle$. Let A be a linear mapping of a vector space V into a vector space W over the same field. The ***transpose*** of A (the ***dual*** of A), denoted ${}^tA$, is the linear map $W^* \to V^*$ determined by $\langle Av, w^* \rangle = \langle v, {}^tAw^* \rangle$. A linear map $A: V \to V$ will often be called ***endomorphism*** of V. If V has finite dimension, the ***determinant*** and ***trace*** of A will be denoted by $\det(A)$ and $\mathrm{Tr}\,(A)$, respectively.

Let V and W be vector spaces over the same field K. A ***bilinear form*** on $V \times W$ is a mapping $B: V \times W \to K$ such that for each $v \in V$, the mapping $B_v: w \to B(v, w)$ belongs to W^* and such that for each $w \in W$, the mapping $B^w: v \to B(v, w)$ belongs to V^*. Thus a bilinear form on $V \times W$ gives rise to linear mappings $V \to W^*$ and $W \to V^*$. The bilinear form B is called ***nondegenerate*** if $v \neq 0$ implies $B_v \not\equiv 0$ and if $w \neq 0$ implies $B^w \not\equiv 0$. The set of all bilinear forms on $V \times W$ is a vector space over K whose dual is denoted $V \otimes W$ and called the ***tensor product*** of V and W. Each element $(v, w) \in V \times W$ gives rise to an element $v \otimes w$ in $V \otimes W$ determined by $(v \otimes w)(B) = B(v, w)$. The direct sum $K + V + V \otimes V + V \otimes V \otimes V + \dots$ is an associative algebra, the ***tensor algebra*** $T(V)$ over V, the multiplication being $\otimes$.

Let $\boldsymbol{R}$ and $\boldsymbol{C}$, respectively, denote the fields of real and complex numbers. Let $\boldsymbol{Z}$ denote the ring of integers. Let V be a vector space over $\boldsymbol{R}$. A bilinear form B on $V \times V$ is called ***symmetric*** if $B(v, v') = B(v', v)$ for $v, v' \in V$, ***positive definite*** if $B(v, v) \geqslant 0$ for $v \in V$, ***strictly positive definite*** if $B(v, v) > 0$ for $v \neq 0$ in V. Let W be a vector space over $\boldsymbol{C}$. A mapping $B: W \times W \to \boldsymbol{C}$ is called a ***Hermitian form*** if for each $w_0 \in W$ the mapping $w \to B(w, w_0)$ is linear and if for each pair $(w', w'') \in W \times W$ the numbers $B(w', w'')$ and $B(w'', w')$ are conjugate complex numbers.

By a ***ring*** we shall always mean a commutative ring with an identity element. Let A be a ring. A commutative group M is called a ***module*** over A (or an A-module) if for each $a \in A$ and $m \in M$ an element am is defined such that

$$a(m_1 + m_2) = am_1 + am_2, \quad (a_1 + a_2)m = a_1m + a_2m,$$
$$(a_1a_2)m = a_1(a_2m), \quad 1m = m.$$

A subset $N \subset M$ such that $n_1, n_2 \in N$ implies $n_1 + n_2 \in N$, $an_1 \in N$ for each $a \in A$ is called a *submodule* of M.

A vector space V over a field K is called an (associative) *algebra* (over K) if there exists a multiplication in V with the properties: $\alpha(v_1v_2) = (\alpha v_1)v_2 = v_1(\alpha v_2)$, $1v = v$, $(v_1v_2)v_3 = v_1(v_2v_3)$, $v_1(v_2 + v_3) = v_1v_2 + v_1v_3$, $(v_1 + v_2)v_3 = v_1v_3 + v_2v_3$ for $\alpha \in K$, $v, v_1, v_2, v_3 \in V$. Let A be in algebra (with identity 1) over a field K and V a vector space over K. A *representation* of A on V is a homomorphism ρ of A into the algebra of all endomorphisms of V such that $\rho(1) = \mathrm{I}$.

III. Topology. A topological space shall always mean a topological space in which the Hausdorf separation axiom holds. Let M be a topological space. A collection $\{U_\alpha\}$, $(\alpha \in A)$ of open subsets of M is called a *basis for the open sets* if each open set can be written as a union of some U_α. If $p \in M$, a *neighborhood* of p is a subset of M containing an open subset of M containing p. A *fundamental system of neighborhoods* of p is a system $\{N_\alpha\}_{\alpha \in A}$ of neighborhoods of p such that each neighborhood of p contains some N_α. A topological space is called *separable* if it has a countable dense subset. For metric spaces, separability is equivalent to the existence of a countable basis for the open sets. A topological space is called *compact* if each open covering has a finite subcovering. A subset of a topological space is called *relatively compact* if its closure is compact. A mapping $f: M \to N$ of a topological space M onto a topological space N is called a *local homeomorphism* of M onto N if each point $m \in M$ has an open neighborhood which f maps homeomorphically onto an open neighborhood of $f(m)$ in N. A *domain* in a topological space is an open connected subset. A *path* (or a *continuous curve*) in a topological space is a continuous mapping of a closed interval $[a, b]$ into the space. A space is called *pathwise connected* if any two points in the space can be joined by means of a path. A topological space is said to be *locally connected* (*locally pathwise connected*) if each neighborhood of any point p in the space contains a connected (pathwise connected) neighborhood of p.

SYMBOLS FREQUENTLY USED

In addition to the preceding conventions the list below contains many of the symbols whose meaning is usually fixed throughout the book. The symbols from Chapter V, § 5 have not been relisted.

ad: adjoint representation of a Lie algebra, 89
Ad: adjoint representation of a Lie group, 117
$A[x_1, \ldots, x_n]$: ring generated by $x_1, \ldots, x_n$ over A, 419
$A(M)$, $A_0(M)$: group of holomorphic isometries, and its identity component, 301
Aut($\mathfrak{a}$): group of automorphisms of $\mathfrak{a}$, 116
$\mathfrak{A}(M)$: Grassmann algebra, 19
$\mathfrak{A}_s(M)$: set of s-forms on M, 17
$B_r(p)$: open ball with center p, radius r, 51
$B(X, Y)$: Killing form, 121
$\mathbf{C}^n$: complex n-space, 4
$C(M)$, $C_c(M)$: set of continuous functions, set of continuous functions of compact support, 361
C^∞: indefinitely differentiable, 4, 6
$C^\infty(M)$, $C_c^\infty(M)$: set of differentiable functions, set of differentiable functions of compact support, 6, 361
$G^\natural(G)$: set of bi-invariant functions in the group algebra, 408
γ_X: maximal geodesic determined by X, 31
d: exterior differentiation, 19
$d\Phi_p$: differential of Φ at p, 22
div: divergence, 387
$D(U, K)$, $D(\mathfrak{u}, \theta)$: diagram of (U, K), 252
$D(U)$, $D(\mathfrak{u})$: diagram of U, 256
$\mathbf{D}(G)$, $\mathbf{D}_0(G)$, $\mathbf{D}(G/H)$: algebras of invariant differential operators, 389
$\mathfrak{D}^1(M)$, $\mathfrak{D}_1(M)$: set of vector fields (1-forms), 9, 11
$\mathfrak{D}(M)$, $\mathfrak{D}(p)$, $\mathfrak{D}^*(M)$, $\mathfrak{D}^*(p)$, $\mathfrak{D}_*(M)$, $\mathfrak{D}_*(p)$: tensor algebras, 15, 16
$\mathfrak{D}_s^r(M)$: set of tensor fields of type (r, s), 13
Δ: set of nonzero roots, 141; Laplace-Beltrami operator, 387
Exp: Exponential mapping for an affine connection, 33
exp: exponential mapping of a Lie group, 94
E^R: complex vector space with scalars restricted to $\mathbf{R}$, 152
$\mathfrak{E}_0$: set of complex-valued differentiable functions on M, 285
$\mathfrak{E}_s^r(M)$: set of complex tensor fields of type (r, s), 285
$f * g$: convolution product, 408
$\mathfrak{F}$: set of real-valued differentiable functions, 5
grad: gradient, 387
G/H: space of left cosets gH, $g \in G$, 110
$\mathbf{GL}(n, \mathbf{R})$: group of nonsingular $n \times n$ real matrices, 100

$\mathfrak{gl}(n, \mathbf{R})$: Lie algebra of all $n \times n$ real matrices, 100
$\mathfrak{gl}(V)$: Lie algebra of all endomorphisms of V, 89
$\mathfrak{g}^\alpha$: root subspace, 140
$\mathfrak{h}_0$, $\mathfrak{h}$, $\mathfrak{h}_{\mathfrak{p}_0}$, $\mathfrak{h}_{\mathfrak{p}}$, $\mathfrak{h}_{\mathfrak{p}_*}$: maximal abelian subalgebras, 140, 221, 224
$\mathfrak{h}^*$: real space spanned by the root vectors, 145
$\mathrm{Int}(\mathfrak{a})$: adjoint group of $\mathfrak{a}$, 116
$I(M)$, $I_0(M)$: group of isometries and its identity component, 166, 167
J: almost complex structure, 281
$K(z, \zeta)$: kernel function, 295
L_ρ, $L(\rho)$: left translation by ρ, 89, 364
$\mathfrak{L}(G)$: Lie algebra of G, 89
λ: symmetrization mapping of $S(\mathfrak{g})$ onto $\mathbf{D}(G)$, 392
$\mathfrak{m}_0$, $\mathfrak{l}_0$: centralizer of $\mathfrak{h}_{\mathfrak{p}_0}$ and its orthogonal complement in $\mathfrak{k}_0$, 223
M, M': centralizer and normalizer of $\mathfrak{h}_{\mathfrak{p}_*}$, 244
M_p: tangent space at p, 10
N^x: mean value operator, 435
$\mathfrak{n}$, $\mathfrak{n}_0$: nilpotent Lie algebras, 222
∇: affine connection, 26
∇_X: covariant differentiation, 26, 41
P_+, P_-: sets of positive roots, 222
π: natural projection of G onto G/H, 111; representation, 412
X^Φ, A^Φ: transform of a vector field X (operator A), 24, 25
Φ^*T: transform of a covariant tensor field T, 25
R: curvature tensor, 43
$\mathbf{R}^n$: Euclidean n-space, 2
R_ρ, $R(\rho)$: right translation by ρ, 364
$S_r(p)$: sphere with center p, radius r, 51
s_p: geodesic symmetry with respect to p, 163
$S(V)$: symmetric algebra over V, 391
T_X: restriction of $(ad\, X)^2$, 179
$T(V)$: tensor algebra over V, 90, 391, 474
$\tau(x)$: mapping $gH \to xgH$, 110
$W(U, K)$, $W(\mathfrak{u}, \theta)$: Weyl group of (U, K), 244
$W(U)$, $W(\mathfrak{u})$: Weyl group of U, 256
W^C : complexification of W, 153
X^*: vector field adapted to X, 36
$U(\mathfrak{g})$: universal enveloping algebra, 90
Z_x: centralizer of x, 258
$[\,,\,]$: bracket, 9, 89
$\wedge$: exterior multiplication, 19
$\langle\,,\,\rangle$: scalar (inner) product, 412
$\|\ \|$: length of a vector, norm, 48, 413

This Page Intentionally Left Blank

AUTHOR INDEX

Numbers in italic indicate the page on which the full reference appears.

SUBJECT INDEX

H

I

J

K

L

M

S

T

U

V

W