MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Symmetric spaces and their compactifications"

verfasst von / submitted by

Paul Zellhofer, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2021 / Vienna, 2021

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 066 821
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Mathematik UG2002
Betreut von / Supervisor:	Univ.-Prof. Mag. Dr. Andreas Čap

Abstract

The aim of this thesis is to provide an introduction to the theory of Riemannian symmetric spaces with a particular view towards non-compact spaces. Starting from their classical theory in the setting of Riemannian geometry and Lie theory, we develop a correspondence between symmetric spaces and a certain type of real Lie algebras. This method leads to a complete classification of symmetric spaces and gives rise to the distinguished class of symmetric spaces of the non-compact type consisting of complete, simply connected Riemannian manifolds of non-positive sectional curvature. The question how such spaces can be compactified arises naturally in this context. In general, there are many approaches to this problem, which raises the issue of relating compactifications obtained by different methods. The core of this thesis is devoted to describing various methods of compactification of symmetric spaces of the non-compact type and to compare them in the concrete examples of hyperbolic space and open orbits in Grassmannian manifolds. These spaces provide an important class of symmetric spaces and can be studied with a wide range of tools from Riemannian geometry, Lie theory and linear algebra.

Zusammenfassung

Die vorliegende Arbeit hat zum Ziel, eine Einführung in die Theorie Riemannscher symmetrischer Räume mit besonderem Augenmerk auf nicht-kompakte Räume zu geben. Ausgehend von ihrer klassischen Theorie im Rahmen von Riemannscher Geometrie und Lie-Theorie wird eine Korrespondenz zwischen symmetrischen Räumen und einer bestimmten Art reeller Lie-Algebren entwickelt. Diese Methode ermöglicht eine vollständige Klassifizierung symmetrischer Räume und zeigt die Existenz der ausgezeichneten Klasse der symmetrischen Räume vom nicht-kompakten Typ, die aus vollständigen, einfach zusammenhängenden Riemannschen Mannigfaltigkeiten mit nicht-positiver Schnittkrümmung besteht. In diesem Kontext ergibt sich die Frage, wie solche Räume kompaktifiziert werden können. Im Allgemeinen gibt es viele Ansätze zu diesem Problem, was die Frage aufwirft, wie unterschiedliche Kompaktifizierungen miteinander verglichen werden können. Der Hauptteil der Arbeit ist der Beschreibung verschiedener Methoden der Kompaktifizierung symmetrischer Räume vom nicht-kompakten Typ gewidmet und vergleicht diese in den konkreten Beispielen des hyperbolischen Raumes sowie offenen Bahnen in Graßmann-Mannigfaltigkeiten. Diese Räume stellen eine wichtige Familie symmetrischer Räume dar und können mit einer Vielzahl an Werkzeugen der Riemannschen Geometrie, Lie-Theorie und linearen Algebra untersucht werden.

Contents

Introduction

Symmetric spaces lie at the crossroads between two viewpoints towards geometry. On the one hand, they can be approached in the framework of classical Riemannian geometry. In this setting, they are characterized by a rather simple behaviour of their geodesics. More precisely, symmetric spaces can be defined as Riemannian manifolds with the property that the reflection of geodesics at any given point extends to a globally well-defined isometry. This condition already heavily constrains the geometry of the manifold and it turns out that this property is locally equivalent to the Riemann curvature tensor being parallel.

On the other hand, a modern approach to geometry is Felix Klein's *Erlangen program* that characterizes geometries by their symmetries. In this framework, one studies the geometry of a space through an associated group of symmetries which typically turns out to be a Lie group acting smoothly on the space in question. Geometric objects on the space are then characterized by being invariant under the group action. The definition of symmetric spaces in the language of Riemannian geometry directly implies that they are endowed with a particularly large and yet very simple set of isometries. In particular, for every pair of points in a symmetric space there exists an isometry mapping one point to the other. From the viewpoint of the Erlangen program, this opens up the possibility of describing a symmetric space equivalently as a homogeneous space of its isometry group. The symmetric structure then in turn imposes strong restrictions on the resulting space, which is particularly visible in the Lie algebra structure that is naturally associated with a symmetric homogeneous space and encodes many of its geometric properties. On the one hand, this allows for a geometric study of symmetric spaces through the theory of Lie algebras. On the other hand, it is visible in this setting that in contrast to general homogeneous spaces, symmetric spaces can be completely classified.

The thesis is organized in the following way. In Chapter 1 we begin by reviewing basic concepts from general Riemannian geometry that provides the framework in which the notion of a symmetric space is introduced. In this picture, we are directly able to discuss several of their geometric properties, such as curvature and geodesics. However, it then becomes apparent that symmetric spaces are in general best described algebraically as homogeneous spaces G/H, where G is a Lie group and $H \subset G$ a closed subgroup satisfying certain additional properties, and we motivate this shift in perspective throughout the first chapter.

Starting from this Lie-theoretic point of view, Chapter 2 is devoted to studying the Lie algebra structure that is associated to a symmetric space G/H. The Lie algebra $\mathfrak{g}$ of G is naturally equipped with an involutive Lie algebra automorphism whose eigenspace for the eigenvalue $+1$ coincides with the Lie algebra $\mathfrak{h}$ of H and whose -1-eigenspace is

canonically identified with the tangent space at a point of the symmetric space. First, we study this algebraic structure in its own right, afterwards we apply the results to obtain more refined information about the geometry and structure of symmetric spaces. In particular, we indicate how the classification of involutive automorphisms of compact, simple Lie algebras gives rise to a classification of symmetric spaces.
As these results show, there is a distinguished class of symmetric spaces, called the *non-compact type*, consisting of simply connected Riemannian manifolds of non-positive sectional curvature that are diffeomorphic to a vector space. It is then a natural question how these spaces can be compactified. In Chapter 3 we turn to a study of compactifications of symmetric spaces of the non-compact type where we present several approaches to this problem. We describe two general methods, but our focus lies on two concrete examples - hyperbolic space and open orbits in Grassmannian manifolds - where we discuss similarities and differences between various compactifications.

The text is intended to be self-contained for a reader with a working background in Riemannian geometry and basic knowledge of Lie groups and Lie algebras. Nevertheless, we review the most important notions of these subjects and provide references for further reading. In general, the books [Lee97] and [Kna96] provide thorough introductions to these fields and contain most of the relevant concepts. The material in Chapter 1 and 2 covers many aspects of the "classical" theory of symmetric spaces and the results obtained there can be found in various sources. In particular, the book [Hel01] is a standard reference. The content in Sections 1.2 and 1.3 can also be found in [Bau14] in a more general setting. Since Chapter 3 mainly focuses on specific examples, we do not intend to give a detailed account of the general theory of compactifications of symmetric spaces, which can be found in [BJ06] or [AO05], but mostly goes beyond the goals of this thesis. For an introduction to symmetric spaces of the non-compact type and a gentle approach to their more elaborate theory, we recommend the book [Ebe96].
Finally, it should be noted that there are various generalizations of symmetric spaces, such as locally or affine symmetric spaces, that we do not consider in this thesis. Also, there is an obvious adaptation of symmetric spaces to the pseudo-Riemannian setting, which is much more involved than the Riemannian case.

1 Riemannian symmetric spaces

In this chapter we introduce the notion of symmetric spaces, which are Riemannian manifolds that are equipped with a distinguished family of isometries. This structure gives rise to a purely Lie-theoretic description of these spaces, which makes it possible to study their geometry with a wide range of algebraic tools.

1.1 Riemannian manifolds and homogeneous spaces

As an introduction, we review some basic notions from Riemannian geometry that will be important for our study and motivate the shift to the Lie-theoretic perspective that we have in mind. The material covered in this section merely serves as an overview, which is why we will not go too far into details. For further reading, we recommend [Lee97] as well as [O'N83, Chapter 3 and 5] for an introduction to Riemannian geometry. Basic material about Lie groups and homogeneous spaces that will be relevant for our purposes is covered in [Bau14, Kapitel 1].
Riemannian geometry deals with the problem of transferring familiar notions from Euclidean geometry to the setting of general manifolds. Euclidean geometry is modelled on a finite-dimensional vector space $\mathbb{R}^n$, where the crucial structure that gives rise to familiar geometric notions, such as *distance*, is the Euclidean inner product

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i y_i, \qquad x, y \in \mathbb{R}^n.$$

In order to define similar notions on a smooth manifold[1] M, it is therefore natural to require the existence of a non-degenerate symmetric bilinear form $Q_p : T_pM \times T_pM \to \mathbb{R}$ on each tangent space T_pM and - as usual in the category of smooth manifolds - it is assumed that this family depends smoothly on the base point $p \in M$ in the sense that if $X, Y \in \mathfrak{X}(M)$ are smooth vector fields on M, then $p \mapsto Q_p(X(p), Y(p))$ defines a smooth function $M \to \mathbb{R}$. Equivalently, the smoothness condition can be expressed by requiring that $p \mapsto Q_p$ defines a smooth $\binom{0}{2}$-tensor field Q on M. If Q_p is positive definite for every $p \in M$ and thus defines an inner product on every tangent space, then Q is called a *Riemannian metric* on M and the pair (M, Q) is said to be a *Riemannian manifold.*[2] The metric Q is the fundamental structure that gives rise to geometric notions on M, but

[1]We assume here that M is a second-countable Hausdorff space with a finite number of connected components and that its dimension is at least 2.

[2]In general, if Q_p is non-degenerate but not necessarily positive definite, then Q is called a *pseudo-Riemannian metric.* However, we will only consider the Riemannian case and much of the following discussion does not generalize to the pseudo-Riemannian setting.

it is important to keep in mind that their precise meaning depends heavily on the chosen metric, which is an additional structure on M and not intrinsic to the manifold. In fact, every manifold can be endowed with a Riemannian metric, which can be constructed in local coordinates using a partition of unity, but this procedure already illustrates that there is in general an enormous freedom of choice. Having chosen a Riemannian metric on M, every tangent space inherits the structure of a Euclidean vector space, so we obtain a natural notion of *length* or *angles* between vectors of the same tangent space. Since the metric depends smoothly on the base point, these notions can immediately be generalized to families of tangent vectors that also depend smoothly on the base point. A particularly important special case is the following.
Let $\gamma : I \to M$ be a smooth curve defined on some interval[3] $I = [a, b] \subset \mathbb{R}$, then at each $t \in I$ we can measure the length $||\dot{\gamma}(t)|| := Q_{\gamma(t)}(\dot{\gamma}(t), \dot{\gamma}(t))^{\frac{1}{2}}$ of the vector tangent to $\gamma(t)$, which automatically defines a continuous function $t \mapsto ||\dot{\gamma}(t)||$ that is smooth if $\dot{\gamma}(t) \neq 0$. As in the Euclidean case, this gives rise to a natural notion of *arc length* of γ defined by

$$L(\gamma) := \int_a^b ||\dot{\gamma}(t)||\, dt, \tag{1.1}$$

which is independent of the parametrization of γ. The curve is said to be *parametrized by arc length* or to have *unit-speed* if $||\dot{\gamma}(t)|| = 1$ for all $t \in I$. A reasonable notion of *distance* between two points $p, q \in M$ is then the optimal length of curves connecting them. Explicitly, if M is connected, then

$$d(p, q) := \inf\{L(\gamma) \mid \gamma : [0, 1] \to M \text{ piecewise smooth curve}, \gamma(0) = p, \gamma(1) = q\} \tag{1.2}$$

defines a distance function $d : M \times M \to \mathbb{R}$, which turns M into a metric space such that the metric topology with respect to d coincides with the natural manifold topology of M (cf. [Lee97, Lemma 6.2]). A curve in M is said to be *distance-minimizing* if its arc length coincides with the metric distance between its endpoints.
Another fundamental feature of Riemannian manifolds is that there is a natural way of differentiating vector fields. More precisely, a Riemannian metric Q on M gives rise to a unique covariant derivative $\nabla : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$ on the tangent bundle TM, called the *Levi-Civita connection* of M, that satisfies

$$\begin{aligned} \nabla_X Y - \nabla_Y X &= [X, Y] \\ X(Q(Y, Z)) &= Q(\nabla_X Y, Z) + Q(Y, \nabla_X Z) \end{aligned} \tag{1.3}$$

for all vector fields $X, Y, Z \in \mathfrak{X}(M)$. These conditions are typically phrased as *torsion-freeness* and *compatibility with the metric*. The Levi-Civita connection is uniquely determined by the *Koszul formula* (cf. [Lee97, Theorem 5.4]):

$$\begin{aligned} 2Q(\nabla_X Y, Z) = {} & X(Q(Y, Z)) + Y(Q(X, Z)) - Z(Q(X, Y)) \\ & + Q([X, Y], Z) - Q([X, Z], Y) - Q([Y, Z], X) \end{aligned} \tag{1.4}$$

[3] Smoothness on I is supposed to mean that γ can be smoothly extended to an open interval $J \supset I$.

Geometrically, the value of the vector field $\nabla_X Y$ at a point $p \in M$ expresses the directional derivative of the vector field Y in direction of $X(p)$. For a fixed vector field Y, the map $X \mapsto \nabla_X Y$ is (by definition of a covariant derivative) $C^\infty(M)$-linear, which implies that its value at p depends only on $X(p)$. This makes it possible to differentiate vector fields in the directions of individual tangent vectors and in particular, to measure the rate of change of a vector field along a curve. We refer to [Lee97, p. 55-62] for more details on the following construction.

Let $\gamma : I \to M$ again be a smooth curve, then a *vector field along* γ is a smooth map $X : I \to TM$ that satisfies $X(t) \in T_{\gamma(t)}M$ for all $t \in I$. The Levi-Civita connection gives rise to a well-defined map $X \mapsto \nabla_{\dot\gamma} X$ on the space $\mathfrak{X}_\gamma(M)$ of vector fields along γ, called the *induced Levi-Civita connection along* γ. A vector field $X \in \mathfrak{X}_\gamma(M)$ is said to be *parallel along* γ if $\nabla_{\dot\gamma} X$ vanishes identically. In particular, the derivative $\dot\gamma$ is a vector field along γ, the *velocity* of γ, so the expression $\nabla_{\dot\gamma}\dot\gamma$ is a well-defined vector field along γ. It measures the rate of change of the velocity along the curve, i.e. the *acceleration* of γ. The curve is called a *geodesic* if its velocity field is parallel, i.e. if $\nabla_{\dot\gamma}\dot\gamma = 0$. Therefore, geodesics are precisely those curves in M having no acceleration in direction of the curve. In particular, they are a natural generalization of the notion of *straight lines* from Euclidean geometry to general Riemannian manifolds.

In local coordinates, the condition $\nabla_{\dot\gamma} X = 0$ can be re-written as a first-order system of linear ordinary differential equations for the components of X. This implies that for every $t_0 \in I$ and every $v \in T_{\gamma(t_0)}M$ there exists a unique parallel vector field $X_v \in \mathfrak{X}_\gamma(M)$ with $X(t_0) = v$. Hence, for all $t_0, t_1 \in I$ there is a well-defined linear map

$$\begin{aligned} \mathcal{P}^\gamma_{t_0,t_1} : T_{\gamma(t_0)}M &\to T_{\gamma(t_1)}M \\ v &\mapsto X_v(t_1), \end{aligned}$$

which is called *parallel transport* along γ from t_0 to t_1. Similarly, the geodesic equation $\nabla_{\dot\gamma}\dot\gamma = 0$ can locally be re-written as a second-order system of nonlinear ordinary differential equations for the components of γ. In this case, it follows that for every point $p \in M$ and every tangent vector $v \in T_pM$, there is a unique maximally extended geodesic γ_v, which is defined on some maximal open interval $I \subset \mathbb{R}$ with $0 \in I$, that satisfies $\gamma_v(0) = p$ and $\dot\gamma_v(0) = v$. If we denote by $\mathcal{D}(p)$ the set of all $v \in T_pM$ such that γ_v is defined on an interval containing $[0, 1]$, then $\mathcal{D}(p)$ is an open neighbourhood of the origin in T_pM and there is a well-defined smooth map

$$\begin{aligned} \mathrm{Exp}_p : \mathcal{D}(p) &\to M \\ v &\mapsto \gamma_v(1), \end{aligned}$$

called the *Riemannian exponential map* at p. Its differential at $0 \in \mathcal{D}(p)$ is the identity map, so it restricts to a diffeomorphism from an open neighbourhood V of $0 \in T_pM$ to an open neighbourhood U of $p \in M$ (cf. [Lee97, Proposition 5.7 and Lemma 5.10]). If V is star-shaped with respect to the origin, then U is called a *normal neighbourhood* of p. From this it follows easily that every point $p \in M$ possesses a normal neighbourhood U and every $q \in U$ can be connected to p by a geodesic contained in U which is unique up

to re-parametrization. The Riemannian manifold is said to be *complete* if $\mathcal{D}(p) = T_pM$ holds for all $p \in M$, which means that every maximally extended geodesic of M is defined on all of $\mathbb{R}$.
So far we have met two natural concepts of distinguished curves on a Riemannian manifold, namely distance-minimizing curves and geodesics. These two notions coincide in Euclidean geometry, but on a general Riemannian manifold the situation is more complicated. On the one hand, a distance-minimizing curve can always be re-parametrized to become a geodesic. On the other hand, geodesics are always locally distance-minimizing, i.e. when restricted to small enough intervals (cf. [Lee97, Theorem 6.6 and 6.12]). However, it can easily be seen that this need not hold globally by considering the unit sphere $S^n \subset \mathbb{R}^{n+1}$, which inherits a Riemannian metric by restricting the standard inner product of the surrounding vector space to its tangent spaces. In this example, geodesics parametrize great circles, which are clearly not distance-minimizing after passing through an antipodal point. Moreover, it is in general not clear whether two given points on a Riemannian manifold can be joined by a geodesic. A fundamental result in this direction is related to the *Hopf-Rinow theorem*, which implies that if M is a complete, connected Riemannian manifold, then every pair of points in M can be connected by a distance-minimizing geodesic (cf. [O'N83, p. 138-140]). Considering again the sphere shows that this geodesic is not necessarily unique.
Intuitively, these phenomena are consequences of the existence of some notion of *curvature* on the Riemannian manifold, but it should be noted that the different behaviour of geodesics can already be seen on a local level. For example, consider a "triangle-shaped" closed curve $\gamma : [0,1] \to S^n$ on the sphere consisting of three segments of great circles. If $\gamma(0) = \gamma(1) = p$, then parallel transport $\mathcal{P}^{\gamma}_{0,1}$ maps T_pS^n to itself, but it is in general not the identity map, meaning that vectors do not end up in the same position under parallel transport along the curve. Note that this phenomenon occurs for arbitrarily small triangles. In Euclidean geometry, however, a similar motion along a closed curve in $\mathbb{R}^n$ never changes the direction of a vector. On a general Riemannian manifold, curvature is formally introduced via the Levi-Civita connection which gives rise to the *Riemann curvature tensor* defined as[4]

$$R(X,Y)Z := \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z, \qquad X, Y, Z \in \mathfrak{X}(M). \tag{1.5}$$

It measures the extent to which the second derivatives of the Levi-Civita connection fail to commute, which turns out to be the fundamental local invariant for a Riemannian manifold. In addition, it also gives rise to a more intuitive quantity. The Riemann curvature tensor is $C^\infty(M)$-linear in all variables and therefore induces a trilinear map $R_p : (T_pM)^3 \to T_pM$ for every $p \in M$. If S is a two-dimensional plane in T_pM that is spanned by $v, w \in T_pM$, then the expression

$$K(S) := \frac{Q_p(R_p(v,w)w, v)}{Q_p(v,v)Q_p(w,w) - Q_p(v,w)^2} \equiv \frac{R_p(v,w,w,v)}{||v||^2||w||^2 - Q_p(v,w)^2} \tag{1.6}$$

[4] There are different conventions on the sign of the curvature tensor used in the literature. We follow the definition in [Hel01].

is independent of the basis $\{v, w\} \subset S$ and called the *sectional curvature* of M along S. This concept is indeed a generalization of the classical notion of *Gaussian curvature* for surfaces in $\mathbb{R}^3$ (cf. [Lee97, p. 142-146]). As suggested by the definition, it is customary to write $R(X,Y,Z,W) := Q(R(X,Y)Z,W)$ and we will employ both conventions simultaneously.
All the constructions mentioned so far are induced by the Riemannian metric, so the appropriate class of *isomorphisms* in the category of Riemannian manifolds are those maps that preserve the Riemannian structure. If (M', Q') is another Riemannian manifold, then a smooth map $f : M \to M'$ is called an *isometry* if it is a diffeomorphism and satisfies $f^*Q' = Q$, where f^*Q' denotes the pullback of the metric Q' under f. Explicitly, this condition is equivalent to the following requirement.

$$Q'_{f(p)}(T_pf(v), T_pf(w)) = Q_p(v,w) \qquad \forall p \in M, \forall v, w \in T_pM \tag{1.7}$$

This ensures that isometries are compatible with all geometric objects induced by the metric. For example, the pullback $f^*\nabla'$ of the Levi-Civita connection on M' defines a torsion-free connection on M that is compatible with the metric $Q = f^*Q'$ and thus coincides with the Levi-Civita connection of M, which also directly implies that the Riemann curvature tensors satisfy $f^*R' = R$. In particular, this shows that the Riemann curvature tensor is really an invariant of a Riemannian manifold. Moreover, isometries preserve geodesics in the following sense: If γ is the unique geodesic of M with $\gamma(0) = p$ and $\dot\gamma(0) = v$, then $\sigma := f \circ \gamma$ is the unique geodesic of M' with $\sigma(0) = f(p)$ and $\dot\sigma(0) = T_pf(v)$ (cf. [Lee97, Proposition 5.6]). This can be expressed by the following commutative diagram.

$$\begin{array}{ccc} \mathcal{D}(p) \subset T_pM & \xrightarrow{\mathrm{Exp}_p} & M \\ \downarrow{\scriptstyle T_pf} & & \downarrow{\scriptstyle f} \\ \mathcal{D}(f(p)) \subset T_{f(p)}M' & \xrightarrow{\mathrm{Exp}_{f(p)}} & M' \end{array}$$

Let f and g be two isometries that satisfy $f(p_0) = g(p_0)$ and $T_{p_0}f = T_{p_0}g$ for some point $p_0 \in M$, then the set $A := \{p \in M : f(p) = g(p) \text{ and } T_pf = T_pg\}$ is closed and non-empty. In a normal neighbourhood U of a point $p \in M$, every $q \in U$ can be connected to p by a unique unit-speed geodesic γ contained in U. If $p \in A$, then the geodesics $f \circ \gamma$ and $g \circ \gamma$ coincide, which implies $f(q) = g(q)$ and $T_qf = T_qg$ by the diagram above. Therefore, the set A is also open, so if M is connected, then $A = M$ shows that an isometry is uniquely determined by its value and tangent map at a single point.
The set of all isometries $f : M \to M$ forms a group under composition of maps which is called the *isometry group* of M and denoted by $I(M)$. Moreover, the map

$$\begin{aligned} I(M) \times M &\to M \\ (f,p) &\mapsto f \cdot p := f(p) \end{aligned} \tag{1.8}$$

defines a canonical group action of $I(M)$ on M and it is natural to ask whether this action is compatible with the manifold structure of M. In concrete examples, it is often possible to find subgroups of $I(M)$ which are indeed Lie groups acting smoothly on the manifold. For example, it is easy to see that the orthogonal group $O(n+1)$ is a Lie group that acts smoothly on the sphere by isometries, but it is more difficult to show that $I(S^n) = O(n+1)$. However, there is a powerful theorem by Myers and Steenrod which answers the above question positively for all Riemannian manifolds. The concept of isometric actions of Lie groups will be very important for our discussion and we do not want to restrict our attention to specific examples. Therefore, we will make use of the following result whose proof can be found in [KN63, Chapter VI, Theorem 3.4].

Theorem 1.1.1 (Myers, Steenrod)**.** *Let M be a Riemannian manifold and $I(M)$ its group of isometries.*

(i) $I(M)$ possesses a smooth structure such that it becomes a Lie group and acts smoothly on M via the action in (1.8)*.*

(ii) The stabilizer $I(M)_p := \{f \in I(M) : f(p) = p\}$ is compact for every $p \in M$.

Trivially, the only isometry of M that fixes every point is the identity, which means that the action of $I(M)$ is *effective*. If we write $G = I(M)$ and $G \cdot p := \{g(p) : g \in G\}$ for the orbit of a point $p \in M$ under the action of G, then we obtain a bijection between the coset space G/G_p and the orbit $G \cdot p$. Since the stabilizer G_p is closed in G, there is a natural manifold structure on G/G_p, which can be used to define a manifold structure on $G \cdot p \subset M$ and turns the orbit into an immersed submanifold of M. If we have $G \cdot p = M$ for some $p \in M$, then this condition in fact holds for all $p \in M$. In this case, the action of G on M is *transitive* and the manifold M is said to be *homogeneous*. In particular, setting $H = G_p$ yields a bijection

$$\begin{aligned} \Phi : G/H &\to M \\ gH &\mapsto g(p) \end{aligned}$$

which can be shown to be a diffeomorphism (cf. [Bau14, Satz 1.25]). Under this identification, the point $p \in M$ corresponds to the coset $eH = H$, where $e \in G$ denotes the neutral element, and the action of G on M is identified with the action

$$\begin{aligned} l : G \times G/H &\to G/H \\ (g, g'H) &\mapsto l_g(g'H) \equiv g \cdot (g'H) := gg'H \end{aligned}$$

of left-multiplication on G/H. Moreover, we can equip G/H with the pullback $Q^{G/H} := \Phi^* Q$ of the Riemannian metric of M. This turns G/H into a Riemannian manifold that is isometric to M such that the metric is *G-invariant* in the sense that l_g is an isometry for every $g \in G$. The canonical quotient map

$$\begin{aligned} \pi : G &\to G/H \\ g &\mapsto gH = l_g(eH) \end{aligned}$$

is a surjective submersion, so its differential $T_e\pi : \mathfrak{g} \to T_{eH}(G/H)$ is surjective and its kernel is precisely the Lie algebra $\mathfrak{h}$ of H. Under the diffeomorphism Φ, it corresponds to the map

$$\begin{aligned} \tau : G &\to M \\ g &\mapsto g(p). \end{aligned}$$

Therefore, τ is a submersion as well and its differential $T_e\tau : \mathfrak{g} \to T_pM$ is surjective with kernel $\mathfrak{h}$. If $\mathfrak{p}$ is any linear subspace of $\mathfrak{g}$ that is complementary to $\mathfrak{h}$, then $T_e\tau$ induces a linear isomorphism $\mathfrak{p} \cong T_pM$. However, in general there is no canonical choice of such a complementary space.
Every isometry $h \in H$ fixes the point p, so its tangent map T_ph is a linear automorphism of T_pM. This induces a well-defined map

$$\begin{aligned} \lambda : H &\to GL(T_pM) \\ h &\mapsto T_ph, \end{aligned}$$

which is a Lie group homomorphism, called the *isotropy representation* of H on T_pM. The fact that H consists of isometries implies that λ actually takes values in the orthogonal group $O(T_pM)$. If M is connected, which we will from now on always assume, then every isometry in H is uniquely determined by its tangent map at p, in which case the isotropy representation is injective. Moreover, as for every Lie group, there is a natural representation $\mathrm{Ad} : G \to GL(\mathfrak{g})$ of G on its Lie algebra $\mathfrak{g}$, called the *adjoint representation* of G, where $\mathrm{Ad}(g)$ is defined to be the differential at e of the conjugation map $\mathrm{conj}_g : G \to G$, $h \mapsto ghg^{-1}$. In particular, we can restrict Ad to $H \subset G$ in the situation above to obtain a representation of H on $\mathfrak{g}$. The relation between these representations is clarified by the following lemma.

Lemma 1.1.2. *There exists a complementary subspace $\mathfrak{p}$ to $\mathfrak{h}$ in $\mathfrak{g}$ that is $Ad(H)$-invariant, i.e. $Ad(h)\mathfrak{p} \subset \mathfrak{p}$ for every $h \in H$. Under the isomorphism $\mathfrak{p} \cong T_pM$, the isotropy representation of H on T_pM corresponds to the restriction of the adjoint representation of $H \subset G$ to $\mathfrak{p} \subset \mathfrak{g}$.*

Proof. We have already observed that H is a compact Lie group that acts on the Lie algebra $\mathfrak{g}$ via the restriction of Ad to $H \subset G$. Therefore, $\mathfrak{g}$ can be equipped with an $\mathrm{Ad}(H)$-invariant inner product, i.e. an inner product $\langle \cdot, \cdot \rangle$ such that $\langle \mathrm{Ad}(h)X, \mathrm{Ad}(h)Y \rangle = \langle X, Y \rangle$ holds for all $h \in H$ and $X, Y \in \mathfrak{g}$.[5] The orthogonal complement of the Lie subalgebra $\mathfrak{h} \subset \mathfrak{g}$ with respect to this inner product is then the required space $\mathfrak{p}$.

[5] This is a consequence of the fact that on every Lie group G there exists a *right-invariant Haar measure* dg which makes it possible to integrate compactly supported smooth functions on G. It is unique up to a scalar multiple, so if G is compact, it is uniquely determined by requiring $\int_G 1\, dg = 1$. In this case, if $\rho : G \to GL(V)$ is a representation of G on a vector space V and $(\cdot, \cdot)$ is any inner product on V, then $\langle v, w \rangle := \int_G (\rho(g)v, \rho(g)w)\, dg$ defines an inner product on V satisfying $\langle \rho(g)v, \rho(g)w \rangle = \langle v, w \rangle$ for all $g \in G$ and $v, w \in V$. More details about this construction can be found in [Kna96, Chapter VIII, Section 2] and [Kun19, Theorem 23.7].

For $g \in G$ we denote by $L_g : h \mapsto gh$ and $R_g : h \mapsto hg$ the left and right multiplication by g, respectively. Since the isometries of H fix p, we have $\tau \circ R_h = \tau$ for every $h \in H$, while the left multiplication merely satisfies $\tau \circ L_h = h \circ \tau$. This implies that the following holds for every $X \in \mathfrak{p}$.

$$\begin{aligned} T_e\tau(\mathrm{Ad}(h)X) &= T_e\tau(T_e(L_h \circ R_{h^{-1}})(X)) = T_{h^{-1}}(\tau \circ L_h)(T_e R_{h^{-1}}(X)) \\ &= T_{h^{-1}}(h \circ \tau)(T_e R_{h^{-1}}(X)) = T_p h(T_e(\tau \circ R_{h^{-1}})(X)) \\ &= T_p h(T_e\tau(X)) = \lambda(h)(T_e\tau(X)) \end{aligned}$$

Using that the restriction of $T_e\tau$ to $\mathfrak{p}$ is an isomorphism, the claim follows. □

Owing to the previous result, we will always assume that $\mathfrak{p}$ is $\mathrm{Ad}(H)$-invariant. The inner product Q_p on T_pM gives rise to an inner product on $\mathfrak{p}$ which we denote by $\langle\cdot,\cdot\rangle_{\mathfrak{p}}$. The property that the isotropy representation of H on T_pM takes values in $O(T_pM)$ then translates into the fact that $\langle\cdot,\cdot\rangle_{\mathfrak{p}}$ is $\mathrm{Ad}(H)$-invariant. By choosing an arbitrary inner product $\langle\cdot,\cdot\rangle_{\mathfrak{h}}$ on $\mathfrak{h}$, we may define an inner product $\langle\cdot,\cdot\rangle$ on $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ by setting

$$\langle X_1 + Y_1, X_2 + Y_2\rangle := \langle X_1, X_2\rangle_{\mathfrak{h}} + \langle Y_1, Y_2\rangle_{\mathfrak{p}}, \qquad X_1, X_2 \in \mathfrak{h},\ Y_1, Y_2 \in \mathfrak{p}, \tag{1.9}$$

which turns $\mathfrak{h}$ and $\mathfrak{p}$ into orthogonal subspaces. Finally, by point-wise defining

$$Q^G_g(X, Y) := \langle T_g L_{g^{-1}}(X), T_g L_{g^{-1}}(Y)\rangle, \qquad X, Y \in T_gG, \tag{1.10}$$

we obtain a Riemannian metric Q^G on the Lie group G that is *left-invariant* in the sense that the left multiplication L_g is an isometry for every $g \in G$. With this metric, we can view the submersions $\pi : G \to G/H$ and $\tau : G \to M$ as maps between Riemannian manifolds. By definition, the restrictions of their tangent maps at e to $\mathfrak{p}$ are not only isomorphisms but in fact linear isometries from $\mathfrak{p}$ to $T_{eH}(G/H)$ and T_pM, respectively. These are examples of *Riemannian submersions*. Since we will make use of one result in this direction, let us introduce the relevant terminology here.

Let $\overline{M}$ and M be Riemannian manifolds and let $\pi : \overline{M} \to M$ be a submersion. Then the fiber $\pi^{-1}(q)$ is a Riemannian submanifold of $\overline{M}$ for every $q \in M$. For a point $p \in \pi^{-1}(q)$ the tangent map $T_p\pi : T_p\overline{M} \to T_qM$ is surjective with kernel $V_p := T_p(\pi^{-1}(q))$ which is called the *vertical tangent space* at p. Its orthogonal complement $H_p := V_p^{\perp}$ is then isomorphic to T_qM under $T_p\pi$ and is called the *horizontal tangent space* at p. The submersion is said to be *Riemannian* if $T_p\pi|_{H_p}$ is a linear isometry for every $p \in \overline{M}$. In particular, since $T_p\pi|_{H_p}$ is bijective, it follows that for each vector field $X \in \mathfrak{X}(M)$ there exists a unique vector field $\overline{X} \in \mathfrak{X}(\overline{M})$ such that $T\pi \circ \overline{X} = X \circ \pi$ and $\overline{X}(p) \in H_p$ for all $p \in \overline{M}$, which is called the *horizontal lift* of X. For an arbitrary vector field $Y \in \mathfrak{X}(\overline{M})$ we denote by Y_H and Y_V its horizontal and vertical projections. Let $\overline{\nabla}$ and $\overline{R}$ denote the Levi-Civita connection and curvature tensor of $\overline{M}$, then it can be shown that the following formulas holds for all vector fields $X, Y, Z, W \in \mathfrak{X}(M)$ (cf. [O'N66, Lemma 1 and Theorem 2] where the opposite sign convention for the curvature tensor is employed).

$$\overline{\nabla_X Y} = (\overline{\nabla}_{\overline{X}}\overline{Y})_H \tag{1.11}$$

$$\begin{aligned} \overline{R}(\overline{X}, \overline{Y}, \overline{Z}, \overline{W}) &= (R(X, Y, Z, W) \circ \pi) + 2\overline{Q}((\overline{\nabla}_{\overline{X}}\overline{Y})_V, (\overline{\nabla}_{\overline{Z}}\overline{W})_V) \\ &\quad - \overline{Q}((\overline{\nabla}_{\overline{Z}}\overline{X})_V, (\overline{\nabla}_{\overline{Y}}\overline{W})_V) - \overline{Q}((\overline{\nabla}_{\overline{Y}}\overline{Z})_V, (\overline{\nabla}_{\overline{X}}\overline{W})_V) \end{aligned} \tag{1.12}$$

Specializing to the previous case where G is the isometry group of a homogeneous Riemannian manifold M that is equipped with the left-invariant Riemannian metric defined in (1.10), we can immediately read off that the horizontal and vertical tangent spaces at a point $g \in G$ are given by $H_g = T_e L_g(\mathfrak{p})$ and $V_g = T_e L_g(\mathfrak{h})$. In this situation, there is even more structure available which makes it possible to lift curves in M to G in a specific way. In fact, the canonical projection $\pi : G \to G/H$ is an *H-principal bundle* on which $g \mapsto T_e L_g(\mathfrak{p})$ defines a *connection.* In this setting, it can be shown that if $\gamma : [0,1] \to G/H$ is a smooth curve and $g \in \pi^{-1}(\gamma(0))$ is arbitrary, then there exists a unique smooth curve $\overline{\gamma} : [0,1] \to G$, called the *horizontal lift* of γ, such that $\pi \circ \overline{\gamma} = \gamma$, $\overline{\gamma}(0) = g$ and $\overline{\gamma}'(t) \in H_{\overline{\gamma}(t)}$ for all $t \in [0,1]$ (cf. [Bau14, Satz 3.7, Beispiel 2.6 and 3.2]). Since we may identify π with τ under the diffeomorphism $\Phi : G/H \to M$, every smooth curve in M can be lifted over τ to a horizontal curve on G which is unique upon specifying its initial value.
In general, it is difficult to explicitly describe the isometry group of a Riemannian manifold. Therefore, it is a priori unclear how rich the class of homogeneous Riemannian manifolds actually is. As we shall see shortly, there is an intuitive source of such spaces. The crucial assumption for these constructions to work is that the given Riemannian manifold is equipped with a particularly simple family of isometries.

1.2 Symmetric spaces

We have seen that if U is a normal neighbourhood of a point $p \in M$, then every $q \in U$ can be uniquely written as $q = \mathrm{Exp}_p(v)$ for some $v \in V = \mathrm{Exp}_p^{-1}(U) \subset T_pM$. If we take U small enough such that V is symmetric with respect to the origin, we may define a map $q \mapsto q^* := \mathrm{Exp}_p(-v)$, which can be interpreted as the reflection of geodesics at the point p. This map can a priori only be defined on U and is usually not an isometry on this set. However, it turns out that there is an interesting class of Riemannian manifolds for which every geodesic reflection extends to a global isometry of M. For convenience, we restrict our attention to connected manifolds.

Definition 1.2.1. A connected Riemannian manifold M is called a *(Riemannian) symmetric space* if for every $p \in M$ there exists an isometry $s_p : M \to M$ such that $s_p(p) = p$ and $T_p s_p = -\mathrm{id}_{T_pM}$.

Since an isometry of a connected manifold is uniquely determined by its value and tangent map in a single point, the map s_p in the above definition is unique. It is called the *symmetry* of M at p and is *involutive*, i.e. it satisfies $s_p^2 = \mathrm{id}_M$ but is not the identity itself. If U is a normal neighbourhood of p, then every point in U can be uniquely expressed as $\mathrm{Exp}_p(v)$ for some $v \in \mathrm{Exp}_p^{-1}(U)$ and we have $s_p(\mathrm{Exp}_p(v)) = \mathrm{Exp}_p(T_p s_p(v)) = \mathrm{Exp}_p(-v)$. Therefore, s_p coincides on U with the geodesic reflection at p and its only fixed point in U is p. This observation also shows that the symmetries depend smoothly on the base point, i.e. that the map $p \mapsto s_p(q)$ is smooth for every $q \in M$. Moreover, if $\gamma : I \to M$ is a geodesic and $t_0 \in I$, then uniqueness of geodesics shows that

$$s_{\gamma(t_0)}(\gamma(t)) = \gamma(2t_0 - t) \tag{1.13}$$

whenever both sides are defined. In particular, if the left-hand side is defined, so is the right-hand side, which implies that γ can be extended to all of $\mathbb{R}$ by repeatedly applying suitable symmetries. Thus, M is complete and since it is assumed to be connected, any two points in M can be joined by a distance-minimizing geodesic.
Every symmetry s_p is by definition contained in the isometry group $I(M)$. If $f \in I(M)$ is any other isometry, then

$$s_{f(p)} = f \circ s_p \circ f^{-1}, \tag{1.14}$$

since both sides are isometries that have the same value and tangent map at $f(p)$. If p and q are points in M and γ is a geodesic with $\gamma(0) = p$ and $\gamma(1) = q$, the symmetry $s_{\gamma(\frac{1}{2})}$ maps p to q. Hence, we have proved the following result.

Proposition 1.2.2. *Every symmetric space is homogeneous and complete.*

Conversely, to show that a homogeneous Riemannian manifold is a symmetric space, it suffices by (1.14) to find the symmetry s_p for a single point $p \in M$.

Example 1.2.3. Let us construct first examples of symmetric spaces.

(i) (Space forms): Let $\langle \cdot, \cdot \rangle$ be an inner product on $\mathbb{R}^{n+1}$. This turns $\mathbb{R}^{n+1}$ into a Riemannian manifold such that the map $q \mapsto -q$ is an involutive isometry, which constitutes the symmetry s_0 at the origin. To find the symmetry at an arbitrary point $p \in \mathbb{R}^{n+1}$, it suffices to conjugate s_0 with the isometry $q \mapsto p + q$ mapping 0 to p, which yields $s_p(q) = 2p - q$.
Similarly, the sphere $S^n \subset \mathbb{R}^{n+1}$ inherits a Riemannian structure by restricting $\langle \cdot, \cdot \rangle$ to its tangent spaces. For every $p \in S^n$ the reflection at the line spanned by p, given by $s_p(q) = 2\langle p, q\rangle p - q$, is an isometry of S^n and these maps turn the sphere into a symmetric space.
Alternatively, we can equip $\mathbb{R}^{n+1}$ with the non-degenerate symmetric bilinear form $\eta(x, y) := \sum_{i=1}^{n} x_i y_i - x_{n+1} y_{n+1}$ and consider the one-sheet hyperboloid $H^n := \{x \in \mathbb{R}^{n+1} : \eta(x, x) = -1,\ x_{n+1} > 0\}$. The restriction of η to its tangent spaces is positive definite and thus defines a Riemannian metric on H^n whose isometry group is the subgroup of the Lorentz group $O(n, 1)$ that preserves H^n. This metric turns H^n into a symmetric space with the symmetries $s_p(q) = -2\eta(p, q)p - q$.

(ii) (Lie groups): Let G be a connected Lie group with inversion map $\nu : G \to G$. If G can be endowed with a bi-invariant (i.e. both left and right-invariant) Riemannian metric Q, then (G, Q) becomes a symmetric space under the symmetries $s_g := L_g \circ \nu \circ L_{g^{-1}}$. In fact, the differential of the inversion map at e is given by $T_e\nu = -\mathrm{id}_{\mathfrak{g}}$ such that we also have $T_g s_g = -\mathrm{id}_{T_g G}$ for every $g \in G$. Since L_g is an isometry by left-invariance, it suffices to show that ν is an isometry, which follows at once from bi-invariance. Let $X, Y \in T_g G$, then we have

$$\begin{aligned} Q_{\nu(g)}(T_g\nu(X), T_g\nu(Y)) &= Q_{g^{-1}}(T_e(\nu \circ R_g)(T_g R_{g^{-1}}(X)), T_e(\nu \circ R_g)(T_g R_{g^{-1}}(Y))) \\ &= Q_{g^{-1}}(T_e(L_{g^{-1}} \circ \nu)(T_g R_{g^{-1}}(X)), T_e(L_{g^{-1}} \circ \nu)(T_g R_{g^{-1}}(Y))) \\ &= Q_e(-T_g R_{g^{-1}}(X), -T_g R_{g^{-1}}(Y)) = Q_g(X, Y). \end{aligned}$$

Every bi-invariant metric Q on G gives rise to an $\mathrm{Ad}(G)$-invariant inner product Q_e on its Lie algebra $\mathfrak{g}$ and vice versa. Hence, if G admits such a metric, then $\mathrm{Ad}(G)$ is a subgroup of the orthogonal group $O(\mathfrak{g})$ with respect to this inner product, so $\mathrm{Ad}(G) \subset GL(\mathfrak{g})$ is relatively compact. Conversely, if $\mathrm{Ad}(G)$ is relatively compact, then there exists an $\mathrm{Ad}(G)$-invariant inner product on $\mathfrak{g}$ which gives rise to a bi-invariant metric on G. In particular, every compact Lie group admits a bi-invariant metric and every such metric turns it into a symmetric space.

We will see a much more efficient way of constructing and detecting symmetric spaces in Section 1.4. For the moment, we prove another characteristic feature of their geometry.

Proposition 1.2.4. *If M is a symmetric space, then its Riemann curvature tensor is parallel, i.e. $\nabla R = 0$.*[6]

Proof. Let $p \in M$ and $v_1, \dots, v_4, w \in T_pM$ be arbitrary. The Levi-Civita connection and the Riemann curvature tensor are invariant under isometries of M, which implies that $s_p^*(\nabla R) = \nabla R$ and hence

$$\begin{aligned}\nabla_w R(v_1, v_2, v_3, v_4) &= (\nabla R)_p(w, v_1, v_2, v_3, v_4) = (s_p^*(\nabla R))_p(w, v_1, v_2, v_3, v_4) \\ &= (\nabla R)_p(T_p s_p(w), T_p s_p(v_1), T_p s_p(v_2), T_p s_p(v_3), T_p s_p(v_4)) \\ &= (-1)^5 \nabla_w R(v_1, v_2, v_3, v_4).\end{aligned}$$

This shows that $(\nabla R)_p = 0$ for every $p \in M$, so we conclude that $\nabla R = 0$. ☐

This condition almost characterizes symmetric spaces. In fact, it can be shown that if M is a complete, simply connected Riemannian manifold with $\nabla R = 0$, then M is a symmetric space (cf. [Bau14, Satz 5.12]). In general, a Riemannian manifold with parallel curvature tensor is called a *locally symmetric space*. These spaces are an interesting topic in their own right which we will not consider in our discussion. More details about the relation between locally and "globally" symmetric spaces (in the sense of Definition 1.2.1) can be found in [Hel01, Chapter IV, Section 5].
Since we assume that every symmetric space M is connected, it follows that $I_0(M)$, the connected component of the identity in $I(M)$, acts transitively on M as well. In particular, this allows us to identify M with the coset space G/H, where $G = I_0(M)$ and $H = G_p$ is the stabilizer of an arbitrary point $p \in M$ under the action of G. However, there may already be smaller subgroups of $I_0(M)$ that have this property since we only made use of distinguished isometries to prove homogeneity. We will be particularly interested in those subgroups that are compatible with the family of symmetries of M. Therefore, we fix an arbitrary *base point* $o \in M$ and introduce the map

$$\begin{aligned}\sigma : I(M) &\to I(M) \\ g &\mapsto s_o \circ g \circ s_o \equiv s_o g s_o.\end{aligned}$$

[6] The Levi-Civita connection uniquely extends to a map $T \mapsto \nabla T$, where T is an arbitrary tensor field on M, such that it commutes with contractions and satisfies a product rule with respect to the tensor product. In this setting, ∇R is a well-defined $\binom{0}{5}$-tensor field. Details about this construction can be found in [O'N83, p. 43-46 and p. 59-65].

Since $I(M)$ is a Lie group, left and right multiplication with s_o define smooth maps on $I(M)$, so it follows that σ is smooth as well. Moreover, $s_o^2 = \mathrm{id}_M$ implies that σ is in fact an involutive automorphism of $I(M)$ and maps $I_0(M)$ to itself.

Definition 1.2.5. Let M be a symmetric space with base point $o \in M$. A pair (G, H) of Lie groups is called an *associated pair* of M if:

(i) G is a σ-invariant, connected Lie subgroup[7] of $I(M)$ that acts transitively on M.

(ii) $H = G_o$ is the stabilizer of o under the action of G.

If (G, H) is an associated pair of M, then the inclusion of G into $I(M)$ is smooth, which implies that G acts smoothly on M as well and the stabilizer H is a closed subgroup of G (but not necessarily of $I(M)$). Hence, it follows exactly as in Section 1.1 that we obtain a diffeomorphism $\Phi : G/H \to M$, which becomes an isometry when G/H is endowed with the pullback metric of M. From the remark preceding this definition, it is immediate that the pair (G, H), where $G = I_0(M)$ and $H = G_o$ is the stabilizer of o under the action of G, is an associated pair of M, but in general it is not the only one. In fact, we will explicitly construct the smallest associated pair of a symmetric space in the next section. Before, we will derive some of their common features.

Proposition 1.2.6. *Let M be a symmetric space and (G, H) an associated pair. The stabilizer $H = G_o$ contains no non-trivial normal subgroup of G and satisfies*

$$G_0^\sigma \subset H \subset G^\sigma, \tag{1.15}$$

where $G^\sigma := \{g \in G : \sigma(g) = g\}$ denotes the set of fixed points of σ and G_0^σ the connected component of the identity in G^σ. The induced Lie algebra isomorphism $\sigma_* := T_e\sigma : \mathfrak{g} \to \mathfrak{g}$ has eigenvalues ± 1 and its eigenspaces have the following properties.

(i) The $+1$-eigenspace of σ_* coincides with the Lie algebra $\mathfrak{h}$ of H.

(ii) The -1-eigenspace $\mathfrak{p}$ of σ_* is $Ad(H)$-invariant.

(iii) The eigenspace decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ of $\mathfrak{g}$ satisfies

$$[\mathfrak{h}, \mathfrak{h}] \subset \mathfrak{h}, \qquad [\mathfrak{h}, \mathfrak{p}] \subset \mathfrak{p}, \qquad [\mathfrak{p}, \mathfrak{p}] \subset \mathfrak{h}. \tag{1.16}$$

Proof. Assume that H' was a normal subgroup of G contained in H. For every point $p \in M$ there exists some $g \in G$ with $g(o) = p$. Since $g^{-1}H'g \subset H' \subset H$, we have for every $h \in H'$:

$$p = g(o) = g(g^{-1}hg(o)) = h(g(o)) = h(p)$$

Since p was arbitrary, it follows that $h = \mathrm{id}_M$, which proves the first claim. If $h \in H$, then both h and $\sigma(h) = s_o h s_o$ are isometries of M that map o to itself and whose tangent

[7] Following the convention in [Bau14] and [Hel01], this terminology is supposed to mean that G is a Lie group such that the inclusion $G \hookrightarrow I(M)$ is an immersion. We do not require G to be closed in $I(M)$, so it may carry a finer topology than $I(M)$.

map at o is $-\mathrm{id}_{T_oM}$. It follows that $h = \sigma(h)$, so H is contained in the fixed point-set G^σ. Clearly, G^σ is a closed subgroup of G and thus a Lie subgroup. Let $\mathfrak{g}^\sigma$ be its Lie algebra, then for $X \in \mathfrak{g}^\sigma$ and $t \in \mathbb{R}$ we have

$$s_o(\exp(tX)(o)) = \sigma(\exp(tX))(o) = \exp(tX)(o),$$

which implies that $\exp(tX)(o)$ is contained in the fixed point-set of s_o for every $t \in \mathbb{R}$. However, o is an isolated fixed point of s_o, so we must have $\exp(tX)(o) = o$ and hence $\exp(tX) \in H$ for all $t \in \mathbb{R}$. Elements of the form $\exp(X)$, $X \in \mathfrak{g}^\sigma$, generate the connected component of the identity G_0^σ of G^σ, so it follows that $G_0^\sigma \subset H$.
Furthermore, we have $\sigma^2 = \mathrm{id}_G$ and hence also $\sigma_*^2 = \mathrm{id}_{\mathfrak{g}}$. The eigenvalues of σ_* can therefore only be ± 1 and $\mathfrak{g}$ decomposes into a direct sum of the corresponding eigenspaces. The $+1$-eigenspace is the Lie algebra of the fixed point-set G^σ, which coincides with $\mathfrak{h}$ by (1.15), so (i) is established. If $h \in H \subset G^\sigma$, then we have $hs_o = s_o h$ and hence $\sigma(hgh^{-1}) = h\sigma(g)h^{-1}$ for all $g \in G$. By differentiation, it follows that σ_* commutes with $\mathrm{Ad}(h) = T_e\mathrm{conj}_h$, so $\mathrm{Ad}(H)$ preserves the eigenspaces of σ_*, which proves (ii). Finally, since σ_* is a Lie algebra isomorphism, we have $\sigma_*([X,Y]) = [\sigma_*(X), \sigma_*(Y)]$ for all $X, Y \in \mathfrak{g}$ and (iii) follows immediately. □

Remark 1.2.7. A decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ of a Lie algebra $\mathfrak{g}$ into linear subspaces with the properties (1.16) is called a *symmetric decomposition* of $\mathfrak{g}$. Hence, if (G,H) is an associated pair of M, then the Lie algebra $\mathfrak{g}$ can be symmetrically decomposed into the eigenspaces of the involution σ_*. Let $\mathfrak{g}' = \mathfrak{h}' \oplus \mathfrak{p}'$ be the symmetric decomposition corresponding to the associated pair (G', H'), where $G' = I_0(M)$ and $H' = G'_o$. Then $G \subset G'$ by connectedness, which implies $\mathfrak{h} \subset \mathfrak{h}'$ and $\mathfrak{p} \subset \mathfrak{p}'$. It will follow from the arguments given in the next paragraph that the dimensions of $\mathfrak{p}$ and $\mathfrak{p}'$ agree. Thus, we always have $\mathfrak{p} = \mathfrak{p}'$ and the symmetric decompositions of two associated pairs only differ in the $\mathfrak{h}$-component.
We have made an arbitrary choice of base point o in the definition of σ. If $p \in M$ is any other point, then there exists some $g \in G$ with $g(o) = p$. Let $\tilde{\sigma} : G \to G$ denote the involution $g' \mapsto s_p g' s_p$, then we deduce from (1.14) that

$$\tilde{\sigma}(gg'g^{-1}) = s_p(gg'g^{-1})s_p = (gs_og^{-1})(gg'g^{-1})(gs_og^{-1}) = g\sigma(g')g^{-1}, \tag{1.17}$$

i.e. $\tilde{\sigma} \circ \mathrm{conj}_g = \mathrm{conj}_g \circ \sigma$. If $\tilde{H} = G_p$ denotes the stabilizer of p under the action of G, we have $\tilde{H} = gHg^{-1}$ and (1.17) implies that the symmetric decompositions $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ and $\mathfrak{g} = \tilde{\mathfrak{h}} \oplus \tilde{\mathfrak{p}}$ induced by σ_* and $\tilde{\sigma}_*$, respectively, are related by $\tilde{\mathfrak{h}} = \mathrm{Ad}(g)\mathfrak{h}$ and $\tilde{\mathfrak{p}} = \mathrm{Ad}(g)\mathfrak{p}$. Since $\mathfrak{h}$ and $\mathfrak{p}$ are $\mathrm{Ad}(H)$-invariant, these relations are independent of the choice of g. To be precise about which point is used in the definition of σ, we will occasionally call the eigenspace decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ the symmetric decomposition of $\mathfrak{g}$ with respect to o.

If (G,H) is an associated pair of a symmetric space M, the identification of M with the quotient space G/H was only based on the property that G acts transitively on M. However, also the symmetric structure of M is mirrored in the algebraic structure of G/H. On the one hand, we can pull the symmetries of M over to the quotient space under the isometry $\Phi : G/H \to M$, which turns G/H into a symmetric space, where the symmetry at a point $gH \in G/H$ is given by $s_{gH} = \Phi^{-1} \circ s_{g(o)} \circ \Phi$. Since $s_{g(o)} = g s_o g^{-1} = g\sigma(g^{-1}) s_o$, we obtain the following explicit formula.

$$s_{gH}(g'H) = \Phi^{-1}(g\sigma(g^{-1})s_o(g'(o))) = \Phi^{-1}(g\sigma(g^{-1})\sigma(g')s_o(o)) = g\sigma(g^{-1}g')H \qquad (1.18)$$

In particular, we have $s_{eH}(g'H) = \sigma(g')H$. On the other hand, it follows exactly as in the previous section that the map $\tau : G \to M$, $g \mapsto g(o)$, is a surjective submersion and that the kernel of its differential at e is $\mathfrak{h}$. In the case where M is symmetric, the previous result implies that there is a natural choice of complementary subspace $\mathfrak{p}$, namely the -1-eigenspace of σ_*, which is in addition $\mathrm{Ad}(H)$-invariant. Hence, the more refined constructions from the first section also apply. First, we have a canonical isomorphism $j_o := T_e\tau|_{\mathfrak{p}} : \mathfrak{p} \to T_oM$ such that the following diagram commutes for every $h \in H$.

$$\begin{array}{ccc} \mathfrak{p} & \xrightarrow{\mathrm{Ad}(h)|_{\mathfrak{p}}} & \mathfrak{p} \\ \downarrow{\scriptstyle j_o} & & \downarrow{\scriptstyle j_o} \\ T_oM & \xrightarrow{\lambda(h)} & T_oM \end{array}$$

As discussed in the previous remark, this observation implies that $\mathfrak{p}$ is independent of the chosen pair. Second, the inner product Q_o on T_oM induces an $\mathrm{Ad}(H)$-invariant inner product on $\mathfrak{p}$, which can be used to equip G with a left-invariant Riemannian metric as in (1.10). With this metric, the projection $\pi : G \to G/H$ and thus also the map $\tau : G \to M$ become Riemannian submersions. The horizontal and vertical tangent spaces at a point $g \in G$ are given by $H_g = T_eL_g(\mathfrak{p})$ and $V_g = T_eL_g(\mathfrak{h})$.
Finally, the space $\mathfrak{p}$ in the symmetric decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ satisfies $[[\mathfrak{p}, \mathfrak{p}], \mathfrak{p}] \subset \mathfrak{p}$, which allows us to define

$$[[X, Y], Z] := j_o([[j_o^{-1}(X), j_o^{-1}(Y)], j_o^{-1}(Z)]) \qquad (1.19)$$

for $X, Y, Z \in T_oM$. This purely algebraic operation is tightly related to the geometry of the symmetric space as the next proposition shows.

Proposition 1.2.8. *Let M be a symmetric space, then the curvature tensor of M at the base point o is given by*

$$R_o(X, Y)Z = -[[X, Y], Z], \qquad \forall X, Y, Z \in T_oM. \qquad (1.20)$$

Proof. Let (G, H) be any associated pair[8] of M, then there exists a left-invariant Riemannian metric Q^G on G such that $\tau : G \to M$ is a Riemannian submersion. We denote by $\langle \cdot, \cdot \rangle := Q^G_e$ the corresponding inner product on the Lie algebra $\mathfrak{g}$ and we write ∇^G and R^G for the Levi-Civita connection and curvature tensor of (G, Q^G). In order to determine the curvature tensor of M at o, it suffices to compute R^G_e and insert into (1.12).

[8] The -1-eigenspace $\mathfrak{p}$ is independent of the chosen pair, so one may from now on just as well consider the case $G = I_0(M)$.

Every $X \in \mathfrak{g}$ is naturally associated to the left-invariant vector field $L_X : g \mapsto T_e L_g(X)$ on G. We denote L_X also by X and thus identify each $X \in \mathfrak{g}$ with its corresponding left-invariant vector field. Consequently, we also write $\nabla^G_X Y \equiv (\nabla^G_{L_X} L_Y)(e)$. Since the metric Q^G is left-invariant, the inner product of two left-invariant vector fields is constant, so for $X, Y, Z \in \mathfrak{g}$, the Koszul formula (1.4) in G evaluated at e simplifies to

$$2\langle \nabla^G_X Y, Z\rangle = \langle [Z,X], Y\rangle + \langle X, [Z,Y]\rangle + \langle [X,Y], Z\rangle. \tag{1.21}$$

We recall that the differential of the adjoint representation $\mathrm{ad} := T_e\mathrm{Ad} : \mathfrak{g} \to L(\mathfrak{g},\mathfrak{g})$ is given by $\mathrm{ad}(X)Y = [X,Y]$ for $X, Y \in \mathfrak{g}$. The $\mathrm{Ad}(H)$-invariance of the restriction of $\langle\cdot,\cdot\rangle$ to $\mathfrak{p}\times\mathfrak{p}$ implies by differentiation $\mathrm{ad}(\mathfrak{h})$-invariance in the following sense.

$$\langle \mathrm{ad}(Z)X, Y\rangle + \langle X, \mathrm{ad}(Z)Y\rangle = 0, \qquad X, Y \in \mathfrak{p},\ Z \in \mathfrak{h} \tag{1.22}$$

Let $X, Y \in \mathfrak{p}$ be horizontal. If $Z \in \mathfrak{p}$ is also horizontal, then all of the Lie brackets in (1.21) are vertical and every term vanishes. In particular, it follows that $\nabla^G_X Y \in \mathfrak{h}$. If $Z \in \mathfrak{h}$ is vertical, then the first two terms cancel by (1.22) and we are left with $2\langle \nabla^G_X Y, Z\rangle = \langle [X,Y], Z\rangle$, which implies $\nabla^G_X Y = \frac{1}{2}[X,Y]$. Using this, we can compute the curvature tensor of G for $X, Y, Z, W \in \mathfrak{p}$.

$$\begin{aligned}
R^G_e(X,Y,Z,W) &= \langle \nabla^G_X \nabla^G_Y Z - \nabla^G_Y \nabla^G_X Z - \nabla^G_{[X,Y]} Z, W\rangle \\
&= -\langle \nabla^G_Y Z, \nabla^G_X W\rangle + \langle \nabla^G_X Z, \nabla^G_Y W\rangle - \langle \nabla^G_{[X,Y]} Z, W\rangle \\
&= -\frac{1}{4}\langle [Y,Z], [X,W]\rangle + \frac{1}{4}\langle [X,Z], [Y,W]\rangle \\
&\quad - \frac{1}{2}\Big(\langle [W,[X,Y]], Z\rangle + \langle [X,Y], [W,Z]\rangle + \langle [[X,Y],Z], W\rangle\Big) \\
&= -\frac{1}{4}\langle [Y,Z], [X,W]\rangle + \frac{1}{4}\langle [X,Z], [Y,W]\rangle \\
&\quad - \frac{1}{2}\langle [X,Y], [W,Z]\rangle - \langle [[X,Y],Z], W\rangle
\end{aligned}$$

In the second line we have used that the Levi-Civita connection ∇^G is compatible with the metric Q^G as well as the fact that the inner product between left-invariant vector fields is constant. The remaining derivation uses (1.21) and (1.22) for the Lie bracket $[X,Y] \in [\mathfrak{p},\mathfrak{p}] \subset \mathfrak{h}$. Having established this, we deduce from (1.12) that (upon suppressing the isomorphism $j_o : \mathfrak{p} \to T_oM$) the curvature tensor of M at o is given by

$$\begin{aligned}
Q_o(R_o(X,Y)Z, W) = R_o(X,Y,Z,W) &= R^G_e(X,Y,Z,W) - \frac{1}{2}\langle [X,Y], [Z,W]\rangle \\
&\quad + \frac{1}{4}\langle [Z,X], [Y,W]\rangle + \frac{1}{4}\langle [Y,Z], [X,W]\rangle \\
&= -\langle [[X,Y],Z], W\rangle = Q_o(-[[X,Y],Z], W),
\end{aligned}$$

which proves the claim by non-degeneracy of Q_o. □

Theorem 1.2.9. *Let $X \in \mathfrak{p}$, then the unique geodesic γ of M with $\gamma(0) = o$ and $\dot\gamma(0) = j_o(X)$ is given by $\gamma(t) = \exp(tX)(o)$.*

Proof. Throughout this proof we write s_t for the symmetry $s_{\gamma(t)}$ and thus also s_0 instead of s_o. Since the symmetries depend smoothly on the base point, the map $t \mapsto r_t := s_{t/2}s_0$ defines a smooth curve in $I_0(M)$. We claim that $r_{t_1+t_2} = r_{t_1} \circ r_{t_2}$ for all $t_1, t_2 \in \mathbb{R}$. Both of these isometries map $\gamma(0) = o$ to $\gamma(t_1 + t_2)$, so it suffices to show that their tangent maps at o coincide.
Let X be a parallel vector field along γ and let $c \in \mathbb{R}$ be arbitrary. Using that isometries preserve geodesics and parallel vector fields, it follows that $Ts_c \circ X$ is parallel along $s_c \circ \gamma$. Since $s_c(\gamma(c + t)) = \gamma(c - t)$ by (1.13) and $T_{\gamma(c)}s_c(X(c)) = -X(c)$, we obtain $(Ts_c \circ X)(c+t) = -X(c-t)$ from uniqueness of parallel vector fields. Let $a, b \in \mathbb{R}$, then this implies

$$T_{\gamma(a)}r_b(X(a)) = T_{\gamma(-a)}s_{\frac{b}{2}}(T_{\gamma(a)}s_0(X(a))) = -T_{\gamma(-a)}s_{\frac{b}{2}}(X(-a)) = X(a + b), \quad (1.23)$$

i.e. that $T_{\gamma(a)}r_b = \mathcal{P}^{\gamma}_{a,a+b}$ is parallel transport along γ. The claim now follows since

$$T_o(r_{t_1} \circ r_{t_2}) = T_{\gamma(t_2)}r_{t_1} \circ T_o r_{t_2} = \mathcal{P}^{\gamma}_{t_2,t_1+t_2} \circ \mathcal{P}^{\gamma}_{0,t_2} = \mathcal{P}^{\gamma}_{0,t_1+t_2} = T_o r_{t_1+t_2}.$$

We conclude that $t \mapsto r_t$ is a one-parameter subgroup of $I_0(M)$ and can therefore be written as $r_t = \exp(tZ)$ for some Z in the Lie algebra $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ of $I_0(M)$. Moreover, we obtain $\sigma(r_t) = s_0 s_t = s_{-t}s_0 = r_{-t}$ from (1.14) and thus $\sigma_*(Z) = -Z$, i.e. $Z \in \mathfrak{p}$. Altogether, we have $\gamma(t) = r_t(o) = \exp(tZ)(o) = \tau(\exp(tZ))$ and hence $j_o(X) = \dot\gamma(0) = T_e\tau(Z) = j_o(Z)$, implying $X = Z$. □

This result allows us to relate the Riemannian exponential map $\mathrm{Exp}_o : T_oM \to M$ with the Lie exponential map $\exp : \mathfrak{g} \to G$. For $v \in T_oM$ the unique geodesic γ of M with $\gamma(0) = o$ and $\dot\gamma(0) = v$ is given by $\gamma(t) = \exp(tX)(o) = \tau(\exp(tX))$, where $X \in \mathfrak{p}$ is the unique element with $j_o(X) = v$. By definition, $\mathrm{Exp}_o(v) = \gamma(1)$ and therefore

$$\mathrm{Exp}_o = \tau \circ \exp \circ j_o^{-1}. \quad (1.24)$$

Depending on the context, we will sometimes suppress the isomorphism $j_o : \mathfrak{p} \to T_oM$ and write (1.24) simply as $\mathrm{Exp}_o = \tau \circ \exp$, i.e. we identify $X \in \mathfrak{p}$ with $j_o(X) \in T_oM$ and write $\mathrm{Exp}_o(X) \equiv \exp(X)(o)$. This identification will be in force in the following statement.

Corollary 1.2.10. *The derivative of the Riemannian exponential map at o is given by*

$$T_X Exp_o = T_o(\exp(X)) \circ T_e\tau \circ \sum_{n=0}^{\infty} \frac{ad(X)^{2n}}{(2n+1)!}, \qquad X \in \mathfrak{p}. \quad (1.25)$$

Proof. It is well-known from the theory of Lie groups that the derivative of the Lie exponential map is given by (cf. [Hel01, Chapter II, Theorem 1.7])

$$T_X \exp = T_e L_{\exp(X)} \circ \frac{1 - e^{-\mathrm{ad}(X)}}{\mathrm{ad}(X)} = T_e L_{\exp(X)} \circ \sum_{n=0}^{\infty} \frac{(-1)^n \mathrm{ad}(X)^n}{(n+1)!}. \quad (1.26)$$

The chain rule then implies

$$\begin{aligned} T_X \mathrm{Exp}_o &= T_{\exp(X)}\tau \circ T_e L_{\exp(X)} \circ \sum_{n=0}^{\infty} \frac{(-1)^n \mathrm{ad}(X)^n}{(n+1)!} \\ &= T_e(\tau \circ L_{\exp(X)}) \circ \sum_{n=0}^{\infty} \frac{(-1)^n \mathrm{ad}(X)^n}{(n+1)!} = T_e(\exp(X) \circ \tau) \circ \sum_{n=0}^{\infty} \frac{(-1)^n \mathrm{ad}(X)^n}{(n+1)!} \\ &= T_o(\exp(X)) \circ T_e \tau \circ \sum_{n=0}^{\infty} \frac{(-1)^n \mathrm{ad}(X)^n}{(n+1)!}. \end{aligned}$$

If $X, Y \in T_oM \cong \mathfrak{p}$, then repeatedly applying (1.16) shows that $\mathrm{ad}(X)^{2n-1}(Y) \in \mathfrak{h}$ and thus $T_e\tau(\mathrm{ad}(X)^{2n-1}(Y)) = 0$ for all $n \geq 1$, which proves the desired formula. □

The previous results show that a lot of geometric information of a symmetric space can be encoded algebraically in properties of a subspace of a certain Lie algebra. It will be a prominent theme throughout this exposition to use the structure theory of Lie algebras for geometric constructions in symmetric spaces.

1.3 The group of displacements

Our next aim will be to construct the smallest associated pair of a symmetric space. The proof of Theorem 1.2.9 shows that if (G, H) is an associated pair of M, then G contains all isometries of the form $s_p \circ s_o$, $p \in M$. A natural candidate for the smallest associated pair is therefore the following.

Definition 1.3.1. Let M be a symmetric space. The subgroup $G(M)$ of $I(M)$ generated by all isometries of the form $s_p \circ s_q$, $p, q \in M$, is called the *group of displacements* of M.

It follows from (1.13) that if p and q lie on a geodesic γ, then $s_p \circ s_q$ maps γ to itself and performs an affine shift in the parametrization, which explains the terminology.

Lemma 1.3.2. *The group of displacements is a connected Lie subgroup of $I(M)$.*

Proof. By [Bau14, Satz 1.23] it suffices to prove that every $g \in G(M)$ can be connected to the neutral element by a smooth curve in $I(M)$ with values in $G(M)$. Moreover, it is enough to show this for the generators $s_p \circ s_q$ of $G(M)$. Since M is complete and connected, there is a (smooth) geodesic $\gamma : [0,1] \to M$ with $\gamma(0) = p$ and $\gamma(1) = q$. If we define $r_t := s_{\gamma(t)} \circ s_q$, then $t \mapsto r_t$ is a curve in $I(M)$ with values in $G(M)$ and connects $s_p \circ s_q$ to the identity. We have already seen that the symmetries depend smoothly on the base point, so it follows that this curve is smooth. □

In particular, the result implies that every product of an even number of symmetries is contained in $I_0(M)$. However, in general $I_0(M)$ need not contain any individual symmetry. For example, the symmetries of the sphere $S^n \subset \mathbb{R}^{n+1}$ are induced by linear maps with determinant $(-1)^n$. Hence, they are not contained in $I_0(S^n) = SO(n+1)$ if n is odd.

Proposition 1.3.3. *The group of displacements is a σ-invariant subgroup of $I(M)$ that acts transitively on M and it is the smallest subgroup with these properties.*

Proof. Let $p, q \in M$, then (1.14) implies that

$$\sigma(s_p \circ s_q) = s_o(s_p s_q)s_o = (s_o s_p s_o)(s_o s_q s_o) = s_{s_o(p)} \circ s_{s_o(q)} \in G(M).$$

Since $G(M)$ is generated by elements of this form, it follows that $G(M)$ is σ-invariant. Moreover, since M is connected and complete, there is a geodesic $\gamma : \mathbb{R} \to M$ such that $\gamma(0) = p$ and $\gamma(1) = q$. Then the isometry $s_{\gamma(\frac{1}{2})} \circ s_p \in G(M)$ maps p to q, so the action of $G(M)$ is transitive.
Let G be any σ-invariant subgroup of $I(M)$ that acts transitively on M. Then for every $p \in M$ there exists some $g \in G$ with $g(o) = p$ and thus

$$s_p s_o = s_{g(o)} s_o = g s_o g^{-1} s_o = g\sigma(g^{-1}) \in G$$

as another consequence of (1.14). This implies

$$s_p s_q = s_p s_o s_o s_q = (s_p s_o)(s_q s_o)^{-1} \in G$$

for all $p, q \in M$ and shows that $G(M) \subset G$. □

If $H(M)$ denotes the stabilizer of the base point $o \in M$ under the action of $G(M)$, then the above results imply that $(G(M), H(M))$ is an associated pair of M. The minimality of $G(M)$ in this situation is also reflected on the Lie algebra level.

Lemma 1.3.4. *Let $\mathfrak{g}(M) = \mathfrak{h}(M) \oplus \mathfrak{p}$ be the symmetric decomposition corresponding to the associated pair $(G(M), H(M))$ of M. Then $G(M)$ is generated by $\exp(\mathfrak{p})$ and we have $[\mathfrak{p}, \mathfrak{p}] = \mathfrak{h}(M)$.*

Proof. Since we have $s_p s_q = (s_p s_o)(s_q s_o)^{-1}$, it suffices to prove that every isometry of the form $s_p s_o$ is contained in the subgroup of $G(M)$ generated by $\exp(\mathfrak{p})$. As above, we may choose a geodesic $\gamma : \mathbb{R} \to M$ with $\gamma(0) = o$ and $\gamma(1) = p$. Since every geodesic through o is of the form $t \mapsto \exp(tX)(o)$, we have $p = \exp(X)(o)$ for a suitable $X \in \mathfrak{p}$. This implies

$$\begin{aligned}\exp(2X) &= \exp(X)\exp(\sigma_*(-X)) = \exp(X)\sigma(\exp(-X)) \\ &= \exp(X)s_o\exp(-X)s_o = s_{\exp(X)(o)}s_o = s_p s_o,\end{aligned} \tag{1.27}$$

which proves the first claim. It remains to show that $[\mathfrak{p}, \mathfrak{p}] = \mathfrak{h}(M)$. The commutator relations (1.16) of a symmetric decomposition and the Jacobi identity directly imply that $\mathfrak{p} \oplus [\mathfrak{p}, \mathfrak{p}]$ is a Lie subalgebra of $\mathfrak{g}(M)$. Since $G(M)$ is generated by $\exp(\mathfrak{p})$ as a Lie group, it follows that $\mathfrak{g}(M)$ is generated by $\mathfrak{p}$ as a Lie algebra. This implies that $\mathfrak{p} \oplus [\mathfrak{p}, \mathfrak{p}] = \mathfrak{g}(M)$ and since $[\mathfrak{p}, \mathfrak{p}] \subset \mathfrak{h}(M)$, we must have equality. □

The group of displacements is not only significant from an algebraic point of view, but is also closely tied to the geometry of the symmetric space M as the next result shows.

Proposition 1.3.5. *Let M be a symmetric space and $G(M)$ its group of displacements.*

(i) For every smooth curve $\gamma : [0,1] \to M$ with $\gamma(0) = o$ and $\gamma(1) = p$ there exists an element $g \in G(M)$ with $g(o) = p$ such that $T_o g : T_o M \to T_p M$ is parallel transport along γ.

(ii) Conversely, for every $g \in G(M)$ there is a piecewise smooth curve $\gamma : [0,1] \to M$ with $\gamma(0) = o$ and $\gamma(1) = g(o)$ such that $T_o g$ is parallel transport along γ.

Proof. Throughout this proof we write G for $G(M)$ and we equip G with a left-invariant Riemannian metric such that $\tau : G \to M$ becomes a Riemannian submersion. Then every vector field on M can be uniquely lifted over τ to a horizontal vector field on G. In the proof of Proposition 1.2.8 we have shown that the Levi-Civita connection ∇^G of G satisfies $\nabla^G_X Y = \frac{1}{2}[X, Y]$ for $X, Y \in \mathfrak{p}$.

(i) It follows exactly as in the discussion at the end of Section 1.1 that there exists a unique horizontal lift $\overline{\gamma} : [0,1] \to G$ of γ over τ with $\overline{\gamma}(0) = e$, so we have $\overline{\gamma}(t)(o) = \tau(\overline{\gamma}(t)) = \gamma(t)$. If $X \in \mathfrak{p}$ and $v := j_o(X) \in T_o M$, then $Y(t) := T_o(\overline{\gamma}(t))(v)$ defines a smooth vector field Y along γ and we claim that it is parallel. Having shown this, it follows that $T_o(\overline{\gamma}(1))$ coincides with parallel transport along γ from o to p and that $\overline{\gamma}(1) \in G$ is the required element.
To prove the claim, we first observe that the unique horizontal lift of Y is given by $\overline{Y}(t) = T_e L_{\overline{\gamma}(t)}(X)$. In fact, it is horizontal because the horizontal tangent space at a point $g \in G$ is given by $H_g = T_e L_g(\mathfrak{p})$ and it is a lift since $\tau \circ L_g = g \circ \tau$. Next, we define the curve
$$c(t) := T_{\overline{\gamma}(t)} L_{\overline{\gamma}(t)^{-1}}(\overline{\gamma}'(t)) \in \mathfrak{p}.$$
We can now compute the horizontal lift of the derivative of Y along γ using (1.11).
$$\overline{\nabla_{\dot\gamma} Y}(t) = (\nabla^G_{\overline{\gamma}'} \overline{Y})_H(t) = \frac{1}{2}[\overline{\gamma}'(t), \overline{Y}(t)]_H = \frac{1}{2}[c(t), X]_{\mathfrak{p}} = 0$$
Therefore, Y has to be parallel along γ, which proves assertion (i).

(ii) Assume first that g can be written as $g = \exp(X)$ for some $X \in \mathfrak{p}$ and consider the smooth curve $\gamma(t) := \exp(tX)(o)$. Then $\gamma(0) = o$, $\gamma(1) = g(o)$ and we have $g = s_{\exp(X/2)(o)} s_o = s_{\gamma(1/2)} s_o$ by (1.27). Hence, it follows directly from (1.23) that $T_o g = \mathcal{P}^\gamma_{0,1}$ is parallel transport along γ. Suppose that we have already constructed curves γ_1, γ_2 satisfying (ii) for two elements $g_1, g_2 \in G$. Then the curve
$$\gamma(t) := \begin{cases} \gamma_1(2t) & 0 \le t \le \frac{1}{2} \\ g_1(\gamma_2(2t-1)) & \frac{1}{2} < t \le 1 \end{cases}$$
is piecewise smooth and satisfies $\gamma(0) = o$ and $\gamma(1) = (g_1 \circ g_2)(o)$. Moreover, parallel transport along γ is given by
$$\mathcal{P}^\gamma_{0,1} = \mathcal{P}^{g_1 \circ \gamma_2}_{0,1} \circ \mathcal{P}^{\gamma_1}_{0,1} = T_{g_1(o)}(g_1 \circ g_2 \circ g_1^{-1}) \circ T_o g_1 = T_o(g_1 \circ g_2),$$
so γ satisfies (ii) for $g_1 \circ g_2$. Since $G(M)$ is generated by $\exp(\mathfrak{p})$, the claim follows by induction.

□

1.4 The Lie-theoretic viewpoint

Every symmetric space M can be realized as a homogeneous space G/H for any associated pair (G, H) and this representation is minimal if $G = G(M)$ is the group of displacements of M. We will now examine the converse question of whether a given quotient of Lie groups can be equipped with a Riemannian metric in which it becomes a symmetric space. In order to do this, we are going to mimic the symmetric structure from (1.18).

Remark 1.4.1. Let G be a Lie group with Lie algebra $\mathfrak{g}$ and H a Lie subgroup of G. Since $\mathfrak{g}$ is a finite-dimensional vector space, the group $GL(\mathfrak{g})$ of all invertible linear maps on $\mathfrak{g}$ is a Lie group. The set $\mathrm{Ad}(H) = \{\mathrm{Ad}(h) : h \in H\}$ is a Lie subgroup of $GL(\mathfrak{g})$ and its Lie algebra is $\mathrm{ad}(\mathfrak{h}) = \{\mathrm{ad}(X) : X \in \mathfrak{h}\} \subset \mathfrak{gl}(\mathfrak{g})$. If G is connected, then the center Z of G coincides with the kernel of $\mathrm{Ad} : G \to GL(\mathfrak{g})$, which implies that $\mathrm{Ad}(H) \cong H/(H \cap Z)$. If H is compact, so is $\mathrm{Ad}(H)$ and the converse holds if $H \cap Z$ is finite.

If (G, H) is an associated pair of a symmetric space M and G is closed in $I(M)$, then H is already compact and contains no non-trivial normal subgroup of G, so in particular we have $H \cap Z = \{e\}$. In general, however, the following properties will suffice.

Definition 1.4.2. A pair (G, H) of Lie groups is called a *symmetric pair* if:

(i) G is a connected Lie group and H a closed subgroup of G such that there exists an involutive automorphism $\sigma : G \to G$ that satisfies $G_0^\sigma \subset H \subset G^\sigma$.

(ii) The subgroup $\mathrm{Ad}(H) \subset GL(\mathfrak{g})$ is compact.

The first part of the above definition implies that the proof of properties (i)-(iii) in Proposition 1.2.6 applies to any symmetric pair (G, H). In particular, we obtain a symmetric decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ of the Lie algebra of G, where $\mathfrak{p}$ is invariant under the adjoint representation of H on $\mathfrak{g}$. Since H is assumed to be closed in G, the coset space G/H can be equipped with a smooth structure such that the natural projection $\pi : G \to G/H$ becomes a submersion. The next proposition shows that the second property allows us to endow G/H with the structure of a symmetric space. As before, for $g \in G$ we denote by l_g the diffeomorphism $g'H \mapsto gg'H$ of G/H.

Theorem 1.4.3. *Let (G, H) be a symmetric pair, $\sigma : G \to G$ an involutive automorphism such that $G_0^\sigma \subset H \subset G^\sigma$ and set $o := \pi(e) = eH$. The quotient space G/H can be equipped with a G-invariant Riemannian metric and in every such metric, G/H becomes a symmetric space with base point o, where the symmetry s_o satisfies:*

$$s_o \circ \pi = \pi \circ \sigma \tag{1.28}$$

$$l_{\sigma(g)} \circ s_o = s_o \circ l_g \tag{1.29}$$

Proof. Let $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ be the symmetric decomposition induced by σ_*. The differential at e of the submersion $\pi : G \to G/H$ is surjective with kernel $\mathfrak{h}$, so it induces an isomorphism $\mathfrak{p} \cong T_o(G/H)$ and the eigenspace $\mathfrak{p}$ is $\mathrm{Ad}(H)$-invariant. As discussed in the proof of Lemma 1.1.2, compactness of $\mathrm{Ad}(H)$ implies that there exists an $\mathrm{Ad}(H)$-invariant inner product on $\mathfrak{p}$, which induces an inner product Q_o on $T_o(G/H)$. Let $p = gH$, then we define an inner product Q_p on $T_p(G/H)$ by setting

$$Q_p(X,Y) := Q_o(T_p l_{g^{-1}}(X), T_p l_{g^{-1}}(Y)), \qquad X, Y \in T_p(G/H).$$

The fundamental feature of $\mathrm{Ad}(H)$-invariance of Q_o is that this definition is independent of the representative of $p = gH$. Smoothness of the map $(p, X) \mapsto T_p l_{g^{-1}}(X)$ implies that $p \mapsto Q_p$ does indeed define a Riemannian metric on G/H which is G-invariant by construction, i.e. each l_g is an isometry.
We define the map s_o by the desired property $s_o(g'H) := \sigma(g')H$, which is motivated by (1.18). Clearly, this defines a diffeomorphism of G/H that satisfies $T_o s_o \circ T_e\pi = T_e\pi \circ \sigma_* = -T_e\pi$ since the $+1$-eigenspace $\mathfrak{h}$ of σ_* is the kernel of $T_e\pi$. It follows that $T_o s_o = -\mathrm{id}$ and by definition we also have $s_o \circ l_g = l_{\sigma(g)} \circ s_o$. Finally, s_o is an isometry because for $p = gH$ and $X, Y \in T_p(G/H)$ we have

$$\begin{aligned} Q_{s_o(p)}(T_p s_o(X), T_p s_o(Y)) &= Q_o(T_p(l_{\sigma(g)^{-1}} \circ s_o)(X), T_p(l_{\sigma(g)^{-1}} \circ s_o)(Y)) \\ &= Q_o(T_p(s_o \circ l_{g^{-1}})(X), T_p(s_o \circ l_{g^{-1}})(Y)) \\ &= Q_o(-T_p l_{g^{-1}}(X), -T_p l_{g^{-1}}(Y)) = Q_p(X,Y). \end{aligned}$$

Therefore, s_o is the required symmetry at o. Since each l_g is an isometry, the symmetry s_p at an arbitrary point $p = gH$ is given by $s_p = l_g \circ s_o \circ l_{g^{-1}}$, where $H \subset G^\sigma$ shows that this is independent of the representative of p. Explicitly, it reads $s_p(g'H) = g\sigma(g^{-1}g')H$, which coincides with (1.18) and turns G/H into a symmetric space. □

The formula $s_o \circ \pi = \pi \circ \sigma$ shows that the symmetries do not depend on the initial choice of the $\mathrm{Ad}(H)$-invariant inner product on $\mathfrak{p}$, but only on the automorphism σ of the symmetric pair (G, H) which need not be unique in general. This result suggests to define symmetric spaces more generally as homogeneous spaces G/H of symmetric pairs, which is a common approach in the literature (e.g. in [KN69]). As we have seen, every symmetric space in the sense of Definition 1.2.1 gives rise to such a space, where in addition H is compact.

Remark 1.4.4. If (G, H) is a symmetric pair and G/H is endowed with a G-invariant Riemannian metric, then G acts on G/H by isometries. More precisely, the subgroup $\tilde{G} := \{l_g : g \in G\} \subset I(G/H)$ acts transitively on G/H and is invariant under the involution $g \mapsto s_o \circ g \circ s_o$ of $I(G/H)$ by (1.29). The stabilizer of $o = eH$ under the action of G is then $\tilde{H} := \{l_h : h \in H\} \subset \tilde{G}$, so $(\tilde{G}, \tilde{H})$ is an associated pair of the symmetric space G/H. If we set $N := \{g \in G : l_g = \mathrm{id}_{G/H}\}$, then N is a closed subgroup of G such that $\tilde{G} \cong G/N$ and $\tilde{H} \cong H/N$. If Z denotes the center of G, then $H \cap Z \subset N$, so $\tilde{H}$ is contained in the compact group $H/(H \cap Z) \cong \mathrm{Ad}(H)$. Now H differs from $\tilde{H}$ by a possibly non-discrete quotient, which is precisely the reason for the more general requirement that $\mathrm{Ad}(H)$ is compact in the definition of a symmetric pair. If (G, H) is an associated pair of a symmetric space M, then $G \subset I(M)$ and the only isometry of $M \cong G/H$ that fixes every point is the identity. Thus, G acts effectively on G/H, so the map $g \mapsto l_g$ is injective and $N = \{e\}$.

Example 1.4.5. The purely Lie-theoretic characterization in Theorem 1.4.3 enables us to construct many more examples of symmetric spaces where it is otherwise difficult to guess a natural Riemannian metric directly. A particularly nice application of this strategy is the case for Grassmannian manifolds.
Let $\mathrm{Gr}(k, \mathbb{R}^n)$ denote the set of all k-dimensional linear subspaces of $\mathbb{R}^n$. The special orthogonal group $SO(n)$ acts smoothly on $\mathrm{Gr}(k, \mathbb{R}^n)$, where the action of $A \in SO(n)$ on a k-dimensional subspace $V \subset \mathbb{R}^n$ is given by $A \cdot V := A(V)$. Since every subspace of $\mathbb{R}^n$ possesses an orthonormal basis, the action is transitive. If A maps V to itself, then it also preserves the orthogonal complement $V^\perp$. Thus, the stabilizer of V under the action can be identified with $S(O(k) \times O(n-k))$, so we obtain the homogeneous space $\mathrm{Gr}(k, \mathbb{R}^n) \cong SO(n)/S(O(k) \times O(n-k))$.
Let $I_{k,n-k}$ be the diagonal matrix whose first k diagonal entries are equal to 1 and whose last $n-k$ entries are equal to -1. The map $\sigma : SO(n) \to SO(n)$, $A \mapsto I_{k,n-k} A I_{k,n-k}$ defines an involutive automorphism of $SO(n)$ whose fixed point-set $SO(n)^\sigma$ coincides with $S(O(k) \times O(n-k))$ which is already compact. Hence, we can equip $\mathrm{Gr}(k, \mathbb{R}^n)$ with an $SO(n)$-invariant Riemannian metric which turns it into a symmetric space. If n is odd, the action of $SO(n)$ is effective. If n is even, this is not the case since $-I_n \in SO(n)$ acts trivially, but this is the only non-identity element with this property.
In the case $k = 1$ we obtain a realization of real projective space $\mathbb{R}P^n$ as a symmetric space. Moreover, the construction can evidently be carried out over the complex numbers as well by replacing orthogonal by unitary groups, in which case $\mathrm{Gr}(k, \mathbb{C}^n)$ is identified with the symmetric space $SU(n)/S(U(k) \times U(n-k))$.

1.5 Hermitian symmetric spaces

In the final section of this chapter we look at a complex analogue of Riemannian symmetric spaces. As a preparation, let us outline some basics about smooth manifolds in the complex setting. A detailed account of these notions is given in [Hel01, Chapter VIII] and [KN69, Chapter IX].
To begin with, we recall that a *complex structure* on a finite-dimensional real vector space V is a linear map $J : V \to V$ such that $J^2 = -\mathrm{id}_V$. If such a structure exists, then V is necessarily even-dimensional and can be turned into a complex vector space by defining multiplication with complex scalars by $(a + bi)v := av + bJ(v)$ for $a, b \in \mathbb{R}$, $v \in V$. Conversely, for every complex vector space, multiplication by $i \in \mathbb{C}$ defines a complex structure on the underlying real vector space.
In order to generalize this to the manifold setting, it is natural to consider a smooth $2n$-dimensional real manifold M, where every tangent space is equipped with a complex structure $J_p : T_pM \to T_pM$ such that $p \mapsto J_p$ defines a smooth $\binom{1}{1}$-tensor field J on M. This can equivalently be viewed as an automorphism $J : TM \to TM$ of the tangent bundle satisfying $J^2 = -\mathrm{id}_{TM}$. In this case, J is called an *almost complex structure* on M and the pair (M, J) is called an *almost complex manifold.* A smooth map $f : M \to M'$ between almost complex manifolds (M, J) and (M', J') is said to be *almost complex* if $Tf \circ J = J' \circ Tf$.

A smooth manifold is called *complex* if it possesses a holomorphic atlas, i.e. a family $\{(\varphi_i, U_i) : i \in I\}$, where φ_i is a homeomorphism from an open set $U_i \subset M$ onto an open subset of $\mathbb{C}^n$ such that the transition functions $\varphi_i \circ \varphi_j^{-1}$ are holomorphic whenever $U_i \cap U_j \neq \emptyset$. A map between complex manifolds is said to be *holomorphic* if its chart expressions with respect to a holomorphic atlas are holomorphic.
Every complex manifold can also be viewed as a real manifold of double dimension, in which case it carries a canonical almost complex structure arising in the following way. Let $(z_1, \dots, z_n)$ be complex local coordinates in a chart domain $U \subset M$ and split them into real and imaginary part $z_j = x_j + iy_j$. Then $(x_1, y_1, \dots, x_n, y_n)$ are local coordinates for the underlying real manifold and the tangent vectors $\{\frac{\partial}{\partial x^i}\big|_p, \frac{\partial}{\partial y^j}\big|_p : i.j = 1, \dots, n\}$ form a basis of the tangent space T_pM for every $p \in U$. The linear map $J_p : T_pM \to T_pM$ defined by

$$J_p\left(\frac{\partial}{\partial x^i}\bigg|_p\right) = \frac{\partial}{\partial y^i}\bigg|_p, \qquad J_p\left(\frac{\partial}{\partial y^j}\bigg|_p\right) = -\frac{\partial}{\partial x^j}\bigg|_p, \qquad i, j = 1, \dots, n$$

is a complex structure on T_pM and $p \mapsto J_p$ is smooth in U. It is a straightforward consequence of the Cauchy-Riemann equations that this definition of J_p is independent of the chosen chart and hence, these maps patch together to define an almost complex structure on all of M. In general, an almost complex structure J on an almost complex manifold need not be inherited from a complex manifold in this way and if it is, then J is said to be *integrable*. Moreover, if (M, J) and (M', J') are complex manifolds with their canonical almost complex structures, then a map $f : M \to M'$ is almost complex if and only if it is holomorphic.
A Riemannian metric on an almost complex manifold (M, J) is said to be *Hermitian* if $Q(JX, JY) = Q(X, Y)$ holds for all vector fields $X, Y \in \mathfrak{X}(M)$. In this case, we can define a Hermitian inner product $\tilde{Q}_p$ on each tangent space by setting

$$\tilde{Q}_p(v, w) := Q_p(v, w) - iQ_p(J_p(v), w), \qquad v, w \in T_pM \tag{1.30}$$

and $p \mapsto \tilde{Q}_p$ is again smooth. Conversely, given a smooth $\binom{0}{2}$-tensor field $\tilde{Q}$ on M such that each $\tilde{Q}_p$ is a Hermitian inner product on T_pM, then we obtain a Hermitian metric Q on M by setting $Q_p(v, w) := \operatorname{Re} \tilde{Q}_p(v, w)$. After these preparations, we can now define the complex analogue of a Riemannian symmetric space.

Definition 1.5.1. A connected complex manifold M that is equipped with a Hermitian metric Q is called a *Hermitian symmetric space* if for every $p \in M$ there exists a holomorphic isometry $s_p : M \to M$ such that $s_p(p) = p$ and $T_ps_p = -\mathrm{id}_{T_pM}$.

The definition is a verbatim adaptation of Definition 1.2.1 to the complex setting, so by viewing M as a real manifold, it is evident that every Hermitian symmetric space is also a Riemannian symmetric space. Hence, all the results obtained in the previous sections are also applicable in this situation. In particular, a Hermitian symmetric space M can be realized as a homogeneous space of the identity component of its isometry group. Moreover, every symmetry of M is even holomorphic and thus contained in the subgroup $A(M) \subset I(M)$ consisting of all holomorphic isometries of M. This is a closed subgroup of $I(M)$ and thus a Lie group which also acts smoothly on M. The holomorphy property implies that the canonical almost complex structure on M satisfies

$$J_{f(p)} = T_p f \circ J_p \circ (T_p f)^{-1} \tag{1.31}$$

for all $f \in A(M)$ and $p \in M$. We denote by $A_0(M)$ the connected component of the identity in $A(M)$. By connectedness of M, the action of $A_0(M)$ is transitive and we can realize M as a homogeneous space of $A_0(M)$.
Let us choose a base point $o \in M$, set $G = A_0(M)$ and let $H = G_o$ be the stabilizer of o under the action of G. Then (G, H) is an associated pair of M and we have a symmetric decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ on the Lie algebra level, which is induced by the involutive automorphism $\sigma : g \mapsto s_o g s_o$ of G exactly as in the real case. The space $\mathfrak{p}$ is canonically identified with T_oM, which implies that the complex structure J_o of T_oM gives rise to a linear map $\tilde{J} : \mathfrak{p} \to \mathfrak{p}$ that satisfies $\tilde{J}^2 = -\mathrm{id}_{\mathfrak{p}}$. The diffeomorphism $M \cong G/H$ turns G/H into a Hermitian symmetric space with base point $o = eH$ such that the left-action l_g is a holomorphic isometry for every $g \in G$. Therefore, the almost complex structure on G/H is G-invariant in the sense that

$$J_{gH} \circ T_o l_g = T_o l_g \circ J_o, \qquad \forall g \in G. \tag{1.32}$$

Moreover, we can identify the restriction to $\mathfrak{p}$ of the adjoint representation of H on $\mathfrak{g}$ with the isotropy representation of H on $T_oM \cong T_o(G/H)$ and the holomorphy condition (1.32) then implies that

$$\tilde{J} \circ \mathrm{Ad}(h)|_{\mathfrak{p}} = \mathrm{Ad}(h)|_{\mathfrak{p}} \circ \tilde{J}, \qquad \forall h \in H. \tag{1.33}$$

Conversely, given a symmetric pair (G, H), these properties already suffice to endow G/H with the structure of a Hermitian symmetric space.[9]

Theorem 1.5.2. *Let (G, H) be a symmetric pair, $o := eH$ and let Q be any G-invariant Riemannian metric on $M := G/H$. If $A : T_oM \to T_oM$ is a linear map satisfying*

(i) $A^2 = -id_{T_oM}$

(ii) $Q_o(A(X), A(Y)) = Q_o(X, Y)$ *for all* $X, Y \in T_oM$

(iii) $A \circ T_o l_h = T_o l_h \circ A$ *for all* $h \in H$,

then there exists a unique G-invariant almost complex structure J on M such that Q is Hermitian, $J_o = A$ and every symmetry s_p, $p \in M$, is almost complex.

[9]The following result does not quite imply that G/H has the structure of a Hermitian symmetric space since it is merely an almost complex manifold at this point. However, the almost complex structure J can be shown to be integrable (cf. [Hel01, Chapter VIII, Proposition 4.2]).

Proof. Since G/H is a symmetric space in every G-invariant Riemannian metric, it remains to construct the almost complex structure J. Requiring G-invariance forces us to define it by $J_{gH} := T_o l_g \circ A \circ (T_o l_g)^{-1}$, which is independent of the representative of gH by (iii). Moreover, it is evidently smooth and satisfies $J^2 = -\mathrm{id}_{TM}$, so it indeed defines an almost complex structure on G/H. Since both Q and J are G-invariant, it follows directly from (ii) that Q is Hermitian with respect to J. To show that each symmetry is almost complex, it suffices to consider s_o by G-invariance. To this end, let $\sigma : G \to G$ be an involutive automorphism such that $G_0^\sigma \subset H \subset G^\sigma$ and let $\pi : G \to G/H$ be the natural projection. Then we have $s_o \circ \pi = \pi \circ \sigma$ and $s_o \circ l_g = l_{\sigma(g)} \circ s_o$ by Theorem 1.4.3. Let $p = gH \in M$ and $X \in T_pM$ be arbitrary, then these relations imply

$$\begin{aligned} T_p s_o(J_p(X)) &= T_p s_o(T_o l_g \circ A \circ (T_o l_g)^{-1}(X)) = T_o(s_o \circ l_g)(A \circ (T_o l_g)^{-1}(X)) \\ &= T_o(l_{\sigma(g)} \circ s_o)(A \circ (T_o l_g)^{-1}(X)) = (T_o l_{\sigma(g)} \circ T_o s_o \circ A \circ (T_o l_g)^{-1})(X) \\ &= (T_o l_{\sigma(g)} \circ A \circ T_o s_o \circ (T_o l_g)^{-1})(X) = (T_o l_{\sigma(g)} \circ A \circ T_p(s_o \circ l_{g^{-1}}))(X) \\ &= (T_o l_{\sigma(g)} \circ A \circ T_{s_o(p)} l_{\sigma(g)^{-1}} \circ T_p s_o)(X) = J_{s_o(p)}(T_p s_o(X)), \end{aligned}$$

where we have used that $T_o s_o = -\mathrm{id}_{T_oM}$ commutes with A in the third line and the definition of the almost complex structure in the last step. This shows that s_o is almost complex and concludes the proof. □

As before, we may identify A with a map $\tilde{A} : \mathfrak{p} \to \mathfrak{p}$, in which case condition (iii) of the preceding theorem is equivalent to requiring that $\tilde{A} \circ \mathrm{Ad}(h)|_{\mathfrak{p}} = \mathrm{Ad}(h)|_{\mathfrak{p}} \circ \tilde{A}$ holds for all $h \in H$. The fact that the almost complex structure induced by A is always integrable is a remarkable property and again emphasizes the claim that symmetric spaces form a very special class of Riemannian manifolds, even in the complex setting. For the moment, we restrict our discussion of Hermitian symmetric spaces to this observation. We will illustrate some more aspects of their theory in the following chapters, but for our purposes it will mostly suffice to view them as a distinguished class of Riemannian symmetric spaces. In fact, there is a simple criterion for detecting whether or not a given symmetric space is Hermitian, which we will prove in Section 2.3.1.

2 Symmetric spaces and Lie algebras

Having established the most important results about the global structure of a symmetric space, we now investigate more closely a sort of "infinitesimal data" associated to it. As we will see, many elements from the structure theory of Lie algebras have a corresponding geometric realization in the theory of symmetric spaces. In order not to confuse these algebraic notions with their geometric counterparts, we first derive some purely algebraic results about certain types of Lie algebras in Section 2.1 and 2.2. Afterwards, these concepts shall be applied in Section 2.3 to prove results about symmetric spaces. In particular, we explain how the structure theory of real semisimple Lie algebras gives rise to a complete classification of symmetric spaces.

2.1 Orthogonal symmetric Lie algebras

In the previous chapter we have seen that one of the most basic properties of a symmetric space M is that it can be expressed as a homogeneous space $M \cong G/H$, where (G, H) is an associated pair of M which is not necessarily unique. According to Proposition 1.2.6, the Lie algebra of G can be symmetrically decomposed as $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ into the ± 1-eigenspaces of an involutive automorphism of $\mathfrak{g}$. The Riemannian structure on M induces a G-invariant Riemannian metric on G/H and an $\mathrm{Ad}(H)$-invariant inner product on $\mathfrak{p}$. Conversely, given a symmetric pair (G, H), the homogeneous space G/H admits a G-invariant Riemannian metric and every such metric turns it into a symmetric space. Moreover, G-invariant metrics on G/H are in bijective correspondence with $\mathrm{Ad}(H)$-invariant inner products on $\mathfrak{p}$. Differentiating the invariance condition as in (1.22) shows that the restrictions to $\mathfrak{p} \subset \mathfrak{g}$ of the linear maps $\mathrm{ad}(\mathfrak{h}) = \{\mathrm{ad}(X) : X \in \mathfrak{h}\}$ are skew-symmetric with respect to such an inner product. Looking only at the Lie algebra-level, it is therefore natural to study Lie algebras with a symmetric decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$, where $\mathfrak{p}$ can be equipped with an inner product such that $\mathrm{ad}(\mathfrak{h})$ acts on $\mathfrak{p}$ by skew-symmetric linear maps.

In order to guarantee the existence of such an inner product, we need one auxiliary notion. Let $\mathfrak{g}$ be a real Lie algebra, then an element $X \in \mathfrak{g}$ is contained in the center $\mathfrak{z}$ of $\mathfrak{g}$ if and only if $\mathrm{ad}(X) \equiv 0$. If $\mathfrak{h} \subset \mathfrak{g}$ is a Lie subalgebra, then $\mathrm{ad}(\mathfrak{h})$ is a Lie subalgebra of the set $\mathfrak{gl}(\mathfrak{g})$ of linear maps on $\mathfrak{g}$ and $\mathrm{ad}(\mathfrak{h}) \cong \mathfrak{h}/(\mathfrak{h} \cap \mathfrak{z})$. We say that $\mathfrak{h}$ is *compactly embedded* in $\mathfrak{g}$ if the unique connected Lie subgroup of $\mathrm{GL}(\mathfrak{g})$ with Lie algebra $\mathrm{ad}(\mathfrak{h})$ is compact. Furthermore, the Lie algebra $\mathfrak{g}$ is called *compact* if it is compactly embedded in itself. Having this notion at hand, the preceding discussion suggests to consider the following objects.

Definition 2.1.1. Let $\mathfrak{g}$ be a real Lie algebra, $\rho : \mathfrak{g} \to \mathfrak{g}$ an involutive automorphism of $\mathfrak{g}$ and $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ the decomposition of $\mathfrak{g}$ into the ± 1-eigenspaces of ρ.

(i) The pair $(\mathfrak{g}, \rho)$ is called an *orthogonal symmetric Lie algebra* if $\mathfrak{h}$ is compactly embedded in $\mathfrak{g}$. It is called *effective* if $\mathfrak{h} \cap \mathfrak{z} = \{0\}$, where $\mathfrak{z}$ denotes the center of $\mathfrak{g}$, and *reduced* if $\mathfrak{h}$ contains no non-zero ideal of $\mathfrak{g}$.

(ii) A pair (G, H) of Lie groups is said to be *associated* to $(\mathfrak{g}, \rho)$ if G is a connected Lie group with Lie algebra $\mathfrak{g}$ and H a Lie subgroup of G with Lie algebra $\mathfrak{h}$.

(iii) Two orthogonal symmetric Lie algebras $(\mathfrak{g}_1, \rho_1)$ and $(\mathfrak{g}_2, \rho_2)$ are said to be *isomorphic* if there exists a Lie algebra isomorphism $\psi : \mathfrak{g}_1 \to \mathfrak{g}_2$ such that $\psi \circ \rho_1 = \rho_2 \circ \psi$.

Remark 2.1.2. The property of being reduced is slightly stronger than effectiveness. In fact, $(\mathfrak{g}, \rho)$ is reduced if and only if the restriction of the adjoint representation of $\mathfrak{h}$ to $\mathfrak{p} \subset \mathfrak{g}$ is injective, whereas it is effective if and only if the adjoint representation of $\mathfrak{h}$ on all of $\mathfrak{g}$ is injective. The above definition can be viewed as an "infinitesimal version" of Definition 1.4.2. If (G, H) is a symmetric pair and $\sigma : G \to G$ an involutive automorphism with $G_0^\sigma \subset H \subset G$, then $(\mathfrak{g}, \sigma_*)$ is an orthogonal symmetric Lie algebra since $\mathrm{Ad}(H) \subset GL(\mathfrak{g})$ is compact. It is reduced if and only if H contains no non-discrete normal subgroup of G. In particular, this is the case if (G, H) is associated to a symmetric space M.

On the one hand, if $(\mathfrak{g}, \rho)$ is an orthogonal symmetric Lie algebra, then the connected Lie subgroup H^* of $GL(\mathfrak{g})$ whose Lie algebra is $\mathrm{ad}(\mathfrak{h})$ is compact by assumption, so there exists an H^*-invariant inner product on $\mathfrak{g}$. The linear maps in $\mathrm{ad}(\mathfrak{h})$ are then skew-symmetric with respect to this inner product and we sloppily say that the inner product is $\mathrm{ad}(\mathfrak{h})$-*invariant* in this sense. On the other hand, every Lie algebra $\mathfrak{g}$ can also be equipped with the symmetric bilinear form

$$\begin{aligned} B : \mathfrak{g} \times \mathfrak{g} &\to \mathbb{C} \\ (X, Y) &\mapsto \mathrm{tr}(\mathrm{ad}(X) \circ \mathrm{ad}(Y)), \end{aligned}$$

which is called the *Killing form* of $\mathfrak{g}$ and is a natural object on $\mathfrak{g}$ for several reasons. First, if $\varphi : \mathfrak{g} \to \mathfrak{g}$ is a Lie algebra automorphism, then $\mathrm{ad}(\varphi(X)) = \varphi \circ \mathrm{ad}(X) \circ \varphi^{-1}$ implies that $B(\varphi(X), \varphi(Y)) = B(X, Y)$ for all $X, Y \in \mathfrak{g}$. Hence, the Killing form is invariant under every automorphism of $\mathfrak{g}$. If G is a Lie group with Lie algebra $\mathfrak{g}$, then we may apply this observation to $\varphi := \mathrm{Ad}(g)$ for every $g \in G$, in which case differentiating the invariance condition yields $B(\mathrm{ad}(Z)X, Y) = -B(X, \mathrm{ad}(Z)Y)$ for all $X, Y, Z \in \mathfrak{g}$. Second, the Killing form contains crucial information about the structure of the underlying Lie algebra. A particularly important result is *Cartan's criterion*, which states that a Lie algebra is semisimple if and only if its Killing form is non-degenerate (cf. [Kna96, Theorem 1.42]). Among all Lie algebras, the semisimple ones form a very special class and have many additional properties that simplify their structure theory. This is also reflected in the fact that semisimple Lie algebras can be completely classified, which is far from being possible for general Lie algebras.

If $(\mathfrak{g}, \rho)$ is an orthogonal symmetric Lie algebra with symmetric decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$, then for all $X \in \mathfrak{h}$ and $Y \in \mathfrak{p}$ we have that

$$B(X, Y) = B(\rho(X), \rho(Y)) = B(X, -Y) = -B(X, Y), \tag{2.1}$$

which implies that $\mathfrak{h}$ and $\mathfrak{p}$ are orthogonal with respect to the Killing form and explains the terminology. In general, the restriction of B to $\mathfrak{p}$ may well be degenerate, but its restriction to $\mathfrak{h}$ is often well-behaved.

Lemma 2.1.3. *Let $\mathfrak{g}$ be a real Lie algebra with center $\mathfrak{z}$ and let $\mathfrak{h}$ be a compactly embedded Lie subalgebra of $\mathfrak{g}$. If $\mathfrak{h} \cap \mathfrak{z} = \{0\}$, then the restriction to $\mathfrak{h}$ of the Killing form of $\mathfrak{g}$ is negative definite.*

Proof. Let B denote the Killing form of $\mathfrak{g}$ and let H^* be the connected Lie subgroup of $GL(\mathfrak{g})$ whose Lie algebra is $\mathrm{ad}(\mathfrak{h})$. As mentioned above, there exists an H^*-invariant inner product on $\mathfrak{g}$. With respect to an orthonormal basis of $\mathfrak{g}$, every linear map $\mathrm{ad}(X)$, $X \in \mathfrak{h}$, is then represented by a skew-symmetric matrix $(a_{ij}(X))$. Hence, we have

$$B(X, X) = \mathrm{tr}(\mathrm{ad}(X)^2) = -\sum_{i,j} a_{ij}(X)^2 \leq 0, \qquad \forall X \in \mathfrak{h}$$

and equality holds if and only if $\mathrm{ad}(X) \equiv 0$, i.e. $X \in \mathfrak{h} \cap \mathfrak{z} = \{0\}$. □

In particular, this result applies to every effective orthogonal symmetric Lie algebra. For these objects it therefore suffices to control the Killing form on $\mathfrak{p}$, which suggests the following distinction.

Definition 2.1.4. Let $(\mathfrak{g}, \rho)$ be an effective orthogonal symmetric Lie algebra with symmetric decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ and Killing form B. Then $(\mathfrak{g}, \rho)$ is said to be

(i) of the *compact type* if $B|_{\mathfrak{p} \times \mathfrak{p}}$ is negative definite.

(ii) of the *non-compact type* if $B|_{\mathfrak{p} \times \mathfrak{p}}$ is positive definite.

(iii) of the *Euclidean type* if $\mathfrak{p}$ is an abelian ideal in $\mathfrak{g}$.

A pair (G, H) of Lie groups that is associated to $(\mathfrak{g}, \rho)$ is said to be of the *compact*, *non-compact* or *Euclidean type* according to the type of $(\mathfrak{g}, \rho)$.

Since $\mathfrak{h}$ and $\mathfrak{p}$ are orthogonal with respect to the Killing form, it follows that if $(\mathfrak{g}, \rho)$ is of the compact or the non-compact type, then the Killing form is non-degenerate on all of $\mathfrak{g}$, which implies that $\mathfrak{g}$ is semisimple. In the compact case, the Killing form is negative definite on all of $\mathfrak{g}$ while in the non-compact case, the Lie subalgebra $\mathfrak{h}$ is a maximal subspace of $\mathfrak{g}$ on which the Killing form is negative definite.[1] The symmetric

[1] These observations justify the terminology since it can be shown that a Lie algebra is compact and semisimple if and only if its Killing form is negative definite (cf. [Kna96, Corollary 4.26 and Proposition 4.27]). Hence, if $(\mathfrak{g}, \rho)$ is of the compact type, then $\mathfrak{g}$ is compact and if it is of the non-compact type, then $\mathfrak{h}$ is a maximal compactly embedded subalgebra of $\mathfrak{g}$.

decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ is then a *Cartan decomposition* of $\mathfrak{g}$, a notion which we will study in more detail in the next section. The following theorem shows that these classes indeed form the building blocks of effective orthogonal symmetric Lie algebras. We refer to [Hel01, Chapter V, Theorem 1.1] for a proof.

Theorem 2.1.5. *Let $(\mathfrak{g}, \rho)$ be an effective orthogonal symmetric Lie algebra. There exist ideals $\mathfrak{g}_0$, $\mathfrak{g}_-$ and $\mathfrak{g}_+$ in $\mathfrak{g}$ such that:*

(i) $\mathfrak{g} = \mathfrak{g}_0 \oplus \mathfrak{g}_- \oplus \mathfrak{g}_+$

(ii) The ideals $\mathfrak{g}_0$, $\mathfrak{g}_-$ and $\mathfrak{g}_+$ are invariant under ρ and orthogonal with respect to the Killing form of $\mathfrak{g}$.

(iii) Let ρ_0, ρ_- and ρ_+ denote the restrictions of ρ to $\mathfrak{g}_0$, $\mathfrak{g}_-$ and $\mathfrak{g}_+$, respectively. Then $(\mathfrak{g}_0, \rho_0)$, $(\mathfrak{g}_-, \rho_-)$ and $(\mathfrak{g}_+, \rho_+)$ are effective orthogonal symmetric Lie algebras of the Euclidean, compact and non-compact type, respectively.

Combining the compact and non-compact summand in particular shows that $\mathfrak{g}$ can be written as a direct sum $\mathfrak{g} = \mathfrak{g}_0 \oplus \mathfrak{g}'$ where $\mathfrak{g}'$ is semisimple. Thus, for many purposes it suffices to consider semisimple Lie algebras, which is a tremendous simplification. In order to properly exploit this, we will add the requirement that $(\mathfrak{g}, \rho)$ is reduced, which is fulfilled in many situations. In this case, the symmetric decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ has additional properties.

Proposition 2.1.6. *Let $(\mathfrak{g}, \rho)$ be a reduced orthogonal symmetric Lie algebra such that $\mathfrak{g}$ is semisimple. Then the symmetric decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ satisfies $[\mathfrak{p}, \mathfrak{p}] = \mathfrak{h}$ and $\mathfrak{g} = [\mathfrak{p}, \mathfrak{p}] \oplus \mathfrak{p}$.*

Proof. Set $\mathfrak{g}' := [\mathfrak{p}, \mathfrak{p}] \oplus \mathfrak{p}$, then it follows from (1.16) and the Jacobi identity that $\mathfrak{g}'$ is an ideal in $\mathfrak{g}$. Hence, $\mathfrak{g}'$ is also semisimple and the restriction of the Killing form B of $\mathfrak{g}$ to $\mathfrak{g}'$ is non-degenerate. If $\mathfrak{m}$ denotes the orthogonal complement of $\mathfrak{g}'$ with respect to B, then B is also non-degenerate on $\mathfrak{m}$ and there is a direct sum decomposition $\mathfrak{g} = \mathfrak{g}' \oplus \mathfrak{m}$. Moreover, we must have $\mathfrak{m} \subset \mathfrak{h}$ since it is orthogonal to $\mathfrak{p}$ and we claim that $\mathfrak{m}$ is an ideal in $\mathfrak{g}$. To see this, let $X \in \mathfrak{g}$ and $Z \in \mathfrak{m}$ be arbitrary, then for every $Y \in \mathfrak{g}'$ we have $[X, Y] \in \mathfrak{g}'$ and thus $B([X, Z], Y) = -B(Z, [X, Y]) = 0$. This implies that $[X, Z] \in \mathfrak{m}$ and proves the claim. However, since $\mathfrak{m} \subset \mathfrak{h}$ and $(\mathfrak{g}, \rho)$ is reduced, it follows that $\mathfrak{m} = \{0\}$ and hence $\mathfrak{g} = [\mathfrak{p}, \mathfrak{p}] \oplus \mathfrak{p}$. Finally, the general property $[\mathfrak{p}, \mathfrak{p}] \subset \mathfrak{h}$ then also forces $[\mathfrak{p}, \mathfrak{p}] = \mathfrak{h}$. □

Another useful property of semisimple Lie algebras that will be important for our discussion is the following.

Lemma 2.1.7. *Let $\mathfrak{g}$ be a semisimple Lie algebra and D a derivation of $\mathfrak{g}$, i.e. a linear map $D : \mathfrak{g} \to \mathfrak{g}$ satisfying*

$$D[X, Y] = [D(X), Y] + [X, D(Y)], \qquad \forall X, Y \in \mathfrak{g}. \tag{2.2}$$

Then there exists an element $Z \in \mathfrak{g}$ such that $D = ad(Z)$.

Proof. First, we note that every linear map of the form $\mathrm{ad}(Z)$, $Z \in \mathfrak{g}$, is indeed a derivation of $\mathfrak{g}$ as a consequence of the Jacobi identity. The map $X \mapsto \mathrm{tr}(D \circ \mathrm{ad}(X))$ defines a linear functional $\mathfrak{g} \to \mathbb{R}$, so since the Killing form is non-degenerate, there exists an element $Z \in \mathfrak{g}$ such that $\mathrm{tr}(D \circ \mathrm{ad}(X)) = B(Z, X)$ for all $X \in \mathfrak{g}$. The derivation property (2.2) is equivalent to $\mathrm{ad}(D(X)) = D \circ \mathrm{ad}(X) - \mathrm{ad}(X) \circ D = [D, \mathrm{ad}(X)]$, which implies that

$$\begin{aligned} B(D(X), Y) &= \mathrm{tr}(\mathrm{ad}(D(X)) \circ \mathrm{ad}(Y)) = \mathrm{tr}([D, \mathrm{ad}(X)] \circ \mathrm{ad}(Y)) \\ &= \mathrm{tr}(D \circ \mathrm{ad}(X) \circ \mathrm{ad}(Y)) - \mathrm{tr}(\mathrm{ad}(X) \circ D \circ \mathrm{ad}(Y)) \\ &= \mathrm{tr}(D \circ \mathrm{ad}(X) \circ \mathrm{ad}(Y)) - \mathrm{tr}(D \circ \mathrm{ad}(Y) \circ \mathrm{ad}(X)) \\ &= \mathrm{tr}(D \circ [\mathrm{ad}(X), \mathrm{ad}(Y)]) = \mathrm{tr}(D \circ \mathrm{ad}[X, Y]) \\ &= B(Z, [X, Y]) = B([Z, X], Y) \end{aligned}$$

holds for all $X, Y \in \mathfrak{g}$. Hence, it follows that $D = \mathrm{ad}(Z)$ by non-degeneracy of B. □

As mentioned, the study of orthogonal symmetric Lie algebras can be reduced to the semisimple case, but there is even a much more powerful simplification available. In fact, there is a remarkable *duality* between orthogonal symmetric Lie algebras of the compact and non-compact type, which in many cases can be used to study both types simultaneously. Let $(\mathfrak{g}, \rho)$ be an orthogonal symmetric Lie algebra and $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ its symmetric decomposition. Then $\mathfrak{g}$ is a real Lie algebra and we denote by $\mathfrak{g}^{\mathbb{C}} := \mathfrak{g} \otimes_{\mathbb{R}} \mathbb{C}$ its complexification. The complex bilinear extension of the Lie bracket of $\mathfrak{g}$ turns $\mathfrak{g}^{\mathbb{C}}$ into a complex Lie algebra and the subspace $\mathfrak{g}^* := \mathfrak{h} \oplus i\mathfrak{p} \subset \mathfrak{g}^{\mathbb{C}}$ becomes a real Lie subalgebra under this operation. This decomposition induces a conjugation map

$$\begin{aligned} \rho^* : \mathfrak{g}^* = \mathfrak{h} \oplus i\mathfrak{p} &\to \mathfrak{g}^* \\ X + iY &\mapsto X - iY, \end{aligned} \tag{2.3}$$

which defines an involutive automorphism of $\mathfrak{g}^*$. Clearly, we have $\mathfrak{g}^{**} = \mathfrak{g}$ and $\rho^{**} = \rho$.

Lemma 2.1.8. *Let $(\mathfrak{g}, \rho)$ be an orthogonal symmetric Lie algebra and define $(\mathfrak{g}^*, \rho^*)$ as above, then $(\mathfrak{g}^*, \rho^*)$ is an orthogonal symmetric Lie algebra as well. If $(\mathfrak{g}, \rho)$ is of the compact type, then $(\mathfrak{g}^*, \rho^*)$ is of the non-compact type and vice versa. Two orthogonal symmetric Lie algebras $(\mathfrak{g}_1, \rho_1)$ and $(\mathfrak{g}_2, \rho_2)$ are isomorphic if and only if $(\mathfrak{g}_1^*, \rho_1^*)$ and $(\mathfrak{g}_2^*, \rho_2^*)$ are isomorphic.*

Proof. The Killing form $B^{\mathbb{C}}$ of the complexification $\mathfrak{g}^{\mathbb{C}}$ is the complex bilinear extension of the Killing form B of $\mathfrak{g}$. Since $(\mathfrak{g}^*)^{\mathbb{C}} = \mathfrak{g}^{\mathbb{C}}$, the same is true for $\mathfrak{g}^*$ and the Killing forms of $\mathfrak{g}$ and $\mathfrak{g}^*$ are simply the restrictions of $B^{\mathbb{C}}$ to the respective subspace. Evidently, $B^{\mathbb{C}}$ is negative definite on $\mathfrak{p} \times \mathfrak{p}$ if and only if it is positive definite on $i\mathfrak{p} \times i\mathfrak{p}$ and vice versa, so it remains to show that $\mathfrak{h}$ is compactly embedded in $\mathfrak{g}^*$. Let $(\mathfrak{g}^{\mathbb{C}})^{\mathbb{R}}$ denote the complexification of $\mathfrak{g}$ viewed as a Lie algebra over $\mathbb{R}$, then multiplication by $i \in \mathbb{C}$ defines a complex structure J on $(\mathfrak{g}^{\mathbb{C}})^{\mathbb{R}}$. Every $\mathbb{R}$-linear map on $\mathfrak{g}$ or $\mathfrak{g}^*$ uniquely extends to a $\mathbb{C}$-linear map on $\mathfrak{g}^{\mathbb{C}}$, which can be viewed as an $\mathbb{R}$-linear map on $(\mathfrak{g}^{\mathbb{C}})^{\mathbb{R}}$ that commutes with J. In this way, $\mathrm{GL}(\mathfrak{g})$ and $\mathrm{GL}(\mathfrak{g}^*)$ become closed subgroups of $\mathrm{GL}((\mathfrak{g}^{\mathbb{C}})^{\mathbb{R}})$. Let H^*

denote the connected Lie subgroup of $\mathrm{GL}(\mathfrak{g})$ with Lie algebra $\mathrm{ad}(\mathfrak{h})$. Then H^* becomes a compact Lie subgroup of $\mathrm{GL}((\mathfrak{g}^{\mathbb{C}})^{\mathbb{R}})$ that is contained in $\mathrm{GL}(\mathfrak{g}^*)$ since $\mathrm{ad}(\mathfrak{h}) \subset \mathfrak{gl}(\mathfrak{g}^*)$. Finally, if $\psi : \mathfrak{g}_1 \to \mathfrak{g}_2$ is an isomorphism of orthogonal symmetric Lie algebras, then the condition $\psi \circ \rho_1 = \rho_2 \circ \psi$ shows that $\psi(\mathfrak{h}_1) = \mathfrak{h}_2$ and $\psi(\mathfrak{p}_1) = \mathfrak{p}_2$. Hence, the map $\psi^* : \mathfrak{g}_1^* \to \mathfrak{g}_2^*$ defined by $\psi^*(X + iY) := \psi(X) + i\psi(Y)$, where $X \in \mathfrak{h}_1$ and $Y \in \mathfrak{p}_1$, is well-defined and an isomorphism. The converse follows in the same way. □

Definition 2.1.9. The orthogonal symmetric Lie algebra $(\mathfrak{g}^*, \rho^*)$ defined in (2.3) is called the *dual* of $(\mathfrak{g}, \rho)$.

The previous lemma shows that the duality induces an equivalence of categories between orthogonal symmetric Lie algebras of the compact and the non-compact type. Therefore, a classification of one of these types would immediately lead to a complete classification of the other as well. We will now describe that such a classification is indeed possible. To this end, it will be convenient to decompose orthogonal symmetric Lie algebras a little bit further.

Definition 2.1.10. An orthogonal symmetric Lie algebra $(\mathfrak{g}, \rho)$ with eigenspace decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ is called *irreducible* if:

(i) $\mathfrak{g}$ is semisimple and $(\mathfrak{g}, \rho)$ is reduced.

(ii) The restriction to $\mathfrak{p} \subset \mathfrak{g}$ of the adjoint representation of $\mathfrak{h}$ is irreducible.

Before turning to the main result, we observe that the irreducibility condition has a particular feature that will be very useful later. We have seen that every orthogonal symmetric Lie algebra $(\mathfrak{g}, \rho)$ admits an $\mathrm{ad}(\mathfrak{h})$-invariant inner product. If $(\mathfrak{g}, \rho)$ is irreducible, then - at least on the subspace $\mathfrak{p} \subset \mathfrak{g}$ - this inner product is essentially unique.

Proposition 2.1.11. *If $(\mathfrak{g}, \rho)$ is irreducible, then every ad$(\mathfrak{h})$-invariant inner product on $\mathfrak{p}$ is a scalar multiple of the restriction to $\mathfrak{p}$ of the Killing form of $\mathfrak{g}$.*

Proof. Let $\langle \cdot, \cdot \rangle$ be an $\mathrm{ad}(\mathfrak{h})$-invariant inner product on $\mathfrak{p}$, then there is a unique linear map $f : \mathfrak{p} \to \mathfrak{p}$ such that $\langle f(X), Y \rangle = B(X, Y)$ for all $X, Y \in \mathfrak{p}$. Then f is self-adjoint with respect to $\langle \cdot, \cdot \rangle$ and thus diagonalizable over $\mathbb{R}$. Moreover, non-degeneracy of the inner product shows that f commutes with $\mathrm{ad}(Z)$ for every $Z \in \mathfrak{h}$. Every eigenspace of f is then an invariant subspace for the adjoint representation of $\mathfrak{h}$ on $\mathfrak{p}$, so it follows from irreducibility that f can only have one eigenvalue, i.e. $f = c \cdot \mathrm{id}_{\mathfrak{p}}$ for some $c \in \mathbb{R}$ and $c\langle X, Y \rangle = B(X, Y)$ for all $X, Y \in \mathfrak{p}$. Finally, we note that $B|_{\mathfrak{p} \times \mathfrak{p}}$ is non-degenerate since $\mathfrak{g}$ is semisimple and $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ is an orthogonal decomposition. Hence, we have $c \neq 0$ and $\langle X, Y \rangle = \frac{1}{c} B(X, Y)$. □

In particular, we deduce from this result that an irreducible orthogonal symmetric Lie algebra is necessarily of the compact or non-compact type. As described in Theorem 2.1.5, every effective orthogonal symmetric Lie algebra can be decomposed into a Euclidean, a compact and a non-compact summand. The next proposition shows that in the reduced case, the latter can be further decomposed into irreducible factors. Again, we refer to [Hel01, Chapter VIII, Proposition 5.2] for a proof.

Theorem 2.1.12. *Let $(\mathfrak{g}, \rho)$ be a reduced orthogonal symmetric Lie algebra where $\mathfrak{g}$ is semisimple. There exist ideals $\mathfrak{g}_1, \ldots, \mathfrak{g}_n$ in $\mathfrak{g}$ such that:*

(i) $\mathfrak{g} = \oplus_{i=1}^n \mathfrak{g}_i$

(ii) The ideals $\mathfrak{g}_1, \ldots, \mathfrak{g}_n$ are invariant under ρ and mutually orthogonal with respect to the Killing form of $\mathfrak{g}$.

(iii) Let ρ_i denote the restriction of ρ to $\mathfrak{g}_i$. Then $(\mathfrak{g}_i, \rho_i)$ is an irreducible orthogonal symmetric Lie algebra for every $i = 1, \ldots, n$.

In order to classify reduced orthogonal symmetric Lie algebras, it therefore suffices to determine the irreducible ones. Moreover, by duality it is even enough to restrict to the compact or the non-compact type. We conclude this section by presenting the general classification result from [Hel01, Chapter VIII, Theorem 5.3 and 5.4].

Theorem 2.1.13 (Classification of orthogonal symmetric Lie algebras)**.** *Let $(\mathfrak{g}, \rho)$ be an irreducible orthogonal symmetric Lie algebra, then it is of one of the following four types:*

(i) $\mathfrak{g}$ is a compact, simple Lie algebra and ρ is any involutive automorphism of $\mathfrak{g}$.

(ii) $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2$ is compact and a direct sum of isomorphic simple ideals that are interchanged by ρ.

(iii) $\mathfrak{g}$ is a non-compact, simple Lie algebra over $\mathbb{R}$ whose complexification is simple and ρ is an involutive automorphism of $\mathfrak{g}$ whose $+1$-eigenspace $\mathfrak{h}$ is a compactly embedded subalgebra of $\mathfrak{g}$.

(iv) $\mathfrak{g} = \tilde{\mathfrak{g}}^{\mathbb{R}}$ is the underlying real Lie algebra of a complex simple Lie algebra $\tilde{\mathfrak{g}}$ and ρ is conjugation with respect to a maximal compactly embedded subalgebra of $\mathfrak{g}$.

If $(\mathfrak{g}, \rho)$ is of the compact type, then it is of type (i) or (ii) and if $(\mathfrak{g}, \rho)$ is of the non-compact type, then it is of type (iii) or (iv). Under the duality $(\mathfrak{g}, \rho) \leftrightarrow (\mathfrak{g}^*, \rho^*)$, the compact Lie algebras of type (i) correspond to the non-compact Lie algebras of type (iii) and type (ii) corresponds to type (iv).

The classification of irreducible orthogonal symmetric Lie algebras of the compact type is therefore essentially equivalent to the classification of involutive automorphisms of real compact, simple Lie algebras. This is indeed possible and was achieved by Élie Cartan in the 1920's. Let us briefly describe some of the underlying principles, more details can be found in [Hel01, Chapter IX]. Starting from a real Lie algebra, its complexification is naturally a complex Lie algebra, but it may well happen that non-isomorphic real Lie algebras have isomorphic complexifications. For example, the complexification of both $\mathfrak{su}(n)$ and $\mathfrak{sl}(n, \mathbb{R})$ is $\mathfrak{sl}(n, \mathbb{C})$. Given a complex Lie algebra $\mathfrak{g}$, every real Lie algebra $\mathfrak{g}_0$ whose complexification is isomorphic to $\mathfrak{g}$ is called a *real form* of $\mathfrak{g}$. In this case, $\mathfrak{g}_0$ is semisimple if and only if $\mathfrak{g}$ is semisimple and if $\mathfrak{g}_0$ is simple, then $\mathfrak{g}$ is either simple or a direct sum of two isomorphic simple ideals. It turns out that every complex semisimple

Lie algebra possesses a compact real form, which is unique up to an automorphism of $\mathfrak{g}$. If ρ is an involutive automorphism of a compact real form $\mathfrak{g}_0$ of $\mathfrak{g}$, then $\mathfrak{g}_0 = \mathfrak{h} \oplus \mathfrak{p}$ decomposes into the ± 1-eigenspaces of ρ and we can form the dual $\mathfrak{g}_0^* = \mathfrak{h} \oplus i\mathfrak{p} \subset \mathfrak{g}$, which is a non-compact real form of $\mathfrak{g}$. If ρ varies through all involutions of $\mathfrak{g}_0$, then $\mathfrak{g}_0^*$ varies through all non-compact real forms of $\mathfrak{g}$. To classify real compact, simple Lie algebras and their involutive automorphisms, it therefore suffices to classify complex semisimple Lie algebras and their real forms. This is a classical topic in the structure theory of Lie algebras, which is usually established using the theory of root systems and their corresponding Dynkin and Satake diagrams.
We will see in Section 2.3.1 how these results can be applied to decompose and classify symmetric spaces. Before, however, we make one more algebraic detour that will be very useful for studying the geometry of symmetric spaces.

2.2 Cartan decompositions

In the previous section we have introduced the Killing form and used it to decompose orthogonal symmetric Lie algebras into Euclidean, compact and non-compact factors. We now take a closer look at the non-compact type, but for transparency we work in a less restricted setting. Recall that if $(\mathfrak{g}, \rho)$ is an orthogonal symmetric Lie algebra of the non-compact type, then the eigenspace decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ induced by ρ has the property that the Killing form of $\mathfrak{g}$ is negative definite on $\mathfrak{h}$ and positive definite on $\mathfrak{p}$. This decomposition gives rise to several interesting features.
Let $\mathfrak{g}$ be a semisimple Lie algebra over $\mathbb{R}$ with Killing form B, then an involutive automorphism $\theta : \mathfrak{g} \to \mathfrak{g}$ is called a *Cartan involution* of $\mathfrak{g}$ if the symmetric bilinear form $B_\theta(X, Y) := -B(X, \theta Y)$ is positive definite and thus defines an inner product on $\mathfrak{g}$. The decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ into the ± 1-eigenspaces of θ then has the properties $[\mathfrak{h}, \mathfrak{h}] \subset \mathfrak{h}$, $[\mathfrak{h}, \mathfrak{p}] \subset \mathfrak{p}$ and $[\mathfrak{p}, \mathfrak{p}] \subset \mathfrak{h}$ and that $B|_{\mathfrak{h} \times \mathfrak{h}}$ is negative definite and $B|_{\mathfrak{p} \times \mathfrak{p}}$ is positive definite. We call any decomposition of $\mathfrak{g}$ into linear subspaces with these properties a *Cartan decomposition* of $\mathfrak{g}$. Note that the restriction of B_θ to $\mathfrak{p}$ coincides with the corresponding restriction of the Killing form. Conversely, given a Cartan decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$, the map $T + X \mapsto T - X$, where $T \in \mathfrak{h}$ and $X \in \mathfrak{p}$, defines a Cartan involution of $\mathfrak{g}$.
If G is a Lie group with Lie algebra $\mathfrak{g}$, then every Cartan involution of $\mathfrak{g}$ can be lifted to an involutive automorphism of G that automatically has simple structural properties. We mention here the general result whose proof can be found in [Kna96, Theorem 6.31]. However, we will encounter a special case of this result in the context of symmetric spaces in Section 2.3.2 and give a geometric proof there.

Theorem 2.2.1 (Global Cartan decomposition)**.** *Let $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ be a Cartan decomposition of a real semisimple Lie algebra $\mathfrak{g}$ with Cartan involution $\theta : \mathfrak{g} \to \mathfrak{g}$. If G is a Lie group with Lie algebra $\mathfrak{g}$, then there exists an involutive automorphism $\Theta : G \to G$ such that $\Theta_* = \theta$. The fixed point-set H of Θ is a connected and closed subgroup of G with Lie algebra $\mathfrak{h}$ and contains the center Z of G. The group H is compact if and only if Z is finite and in that case it is a maximal compact subgroup of G. Moreover, the map $(X, h) \mapsto \exp(X)h$ defines a diffeomorphism $\mathfrak{p} \times H \to G$.*

This theorem shows that all the information about the topology of a semisimple Lie group G with finite center is encoded in a maximal compact subgroup H. In particular, the quotient G/H is diffeomorphic to a vector space and hence contractible. Moreover, since H is connected, it is the unique connected Lie subgroup of G with Lie algebra $\mathfrak{h}$. However, the Cartan decomposition is not well-adapted to the group structure of G, so it is natural to look for a refinement.
If $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ is a Cartan decomposition and $\theta : \mathfrak{g} \to \mathfrak{g}$ the corresponding Cartan involution, then we have

$$B_\theta(\mathrm{ad}(\theta X)Y, Z) = -B([\theta X, Y], \theta Z) = B(Y, [\theta X, \theta Z]) = -B_\theta(Y, \mathrm{ad}(X)Z)$$

for all $X, Y, Z \in \mathfrak{g}$. Hence, the adjoint of $\mathrm{ad}(\theta X)$ with respect to the inner product B_θ is $-\mathrm{ad}(X)$. This implies that $\mathrm{ad}(X)$ is always diagonalizable over $\mathbb{R}$ for $X \in \mathfrak{p}$, but never diagonalizable for $X \in \mathfrak{h}$. If we choose a maximal abelian subspace $\mathfrak{a} \subset \mathfrak{p}$, which exists since $\mathfrak{p}$ is finite-dimensional, then the Jacobi identity shows that $\{\mathrm{ad}(A) : A \in \mathfrak{a}\}$ forms a family of pairwise commuting symmetric linear maps on $\mathfrak{g}$, which are therefore simultaneously orthogonally diagonalizable. If $X \in \mathfrak{g}$ is a joint eigenvector of all these maps, the corresponding eigenvalue $\alpha(A)$ of $\mathrm{ad}(A)$ (i.e. $\mathrm{ad}(A)X = \alpha(A)X$) depends linearly on A. For every linear map $\alpha : \mathfrak{a} \to \mathbb{R}$, this suggests to define

$$\mathfrak{g}_\alpha := \{X \in \mathfrak{g} : \mathrm{ad}(A)X = \alpha(A)X \ \ \forall A \in \mathfrak{a}\}.$$

The map α is called a *(restricted) root* of $\mathfrak{g}$ with respect to $\mathfrak{a}$ if $\alpha \neq 0$ and $\mathfrak{g}_\alpha \neq \{0\}$, in which case $\mathfrak{g}_\alpha$ is called the *(restricted) root space* of α.[2] In other words, the root spaces (together with $\mathfrak{g}_0$) are precisely the joint eigenspaces of the family of linear maps $\{\mathrm{ad}(A) : A \in \mathfrak{a}\}$. We denote by Δ the set of roots, then we obtain the decomposition

$$\mathfrak{g} = \mathfrak{g}_0 \oplus \bigoplus_{\alpha \in \Delta} \mathfrak{g}_\alpha, \tag{2.4}$$

which is orthogonal with respect to B_θ. The Jacobi identity implies that $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta}$, which is $\{0\}$ if $\alpha + \beta \notin \Delta$. If $\alpha \in \Delta$ is a root, then $\mathrm{ad}(A)$ acts on $\mathfrak{g}_\alpha$ by multiplication with $\alpha(A)$. In addition, if $X \in \mathfrak{g}_\alpha$, then

$$\mathrm{ad}(A)(\theta X) = [A, \theta X] = \theta[\theta A, X] = \theta[-A, X] = -\alpha(A)\theta X$$

for all $A \in \mathfrak{a}$. This shows that $-\alpha$ is a root as well and θ restricts to an isomorphism between $\mathfrak{g}_\alpha$ and $\mathfrak{g}_{-\alpha}$. In particular, $\theta(\mathfrak{g}_0) \subset \mathfrak{g}_0$ implies that $\mathfrak{g}_0$ can be decomposed as

$$\mathfrak{g}_0 = (\mathfrak{g}_0 \cap \mathfrak{h}) \oplus (\mathfrak{g}_0 \cap \mathfrak{p}).$$

Since $\mathfrak{a}$ is a maximal abelian subspace of $\mathfrak{p}$, we have $\mathfrak{g}_0 \cap \mathfrak{p} = \mathfrak{a}$. Moreover, $\mathfrak{g}_0 \cap \mathfrak{h}$ is by definition the space of all $X \in \mathfrak{h}$ such that $[X, A] = 0$ holds for all $A \in \mathfrak{a}$, which is precisely the centralizer $\mathfrak{m} := Z_{\mathfrak{h}}(\mathfrak{a})$ of $\mathfrak{a}$ in $\mathfrak{h}$.

[2] Usually, the notions of roots and root spaces are first defined in the context of complex semisimple Lie algebras. The restricted roots of a real semisimple Lie algebra can then be viewed as restrictions of the roots of its complexification, which explains the name. Since we only consider the real case in our discussion, we will often omit the term "restricted".

Definition 2.2.2. An element $X \in \mathfrak{p}$ is called *regular* if $Z_{\mathfrak{p}}(X) = \{Y \in \mathfrak{p} : [X, Y] = 0\}$, the centralizer of X in $\mathfrak{p}$, is an abelian subspace of $\mathfrak{p}$.

Since every abelian subspace of $\mathfrak{p}$ containing X is contained in $Z_{\mathfrak{p}}(X)$, it follows that $Z_{\mathfrak{p}}(X)$ is a maximal abelian subspace of $\mathfrak{p}$ for every regular element $X \in \mathfrak{p}$. In this case, it is the unique maximal abelian subspace of $\mathfrak{p}$ containing X. Conversely, every maximal abelian subspace of $\mathfrak{p}$ arises in this way.

Lemma 2.2.3. *Let $\mathfrak{a} \subset \mathfrak{p}$ be a maximal abelian subspace and let Δ be the set of roots of $\mathfrak{g}$ with respect to $\mathfrak{a}$. If $A \in \mathfrak{a}$ satisfies $\alpha(A) \neq 0$ for all $\alpha \in \Delta$, then the centralizer of A in $\mathfrak{g}$ is given by $Z_{\mathfrak{g}}(A) = \mathfrak{m} \oplus \mathfrak{a}$. In particular, its centralizer in $\mathfrak{p}$ is $Z_{\mathfrak{p}}(A) = \mathfrak{a}$, so A is regular.*

Proof. It follows from the root space decomposition (2.4) and $\mathfrak{g}_0 = \mathfrak{m} \oplus \mathfrak{a}$ that every $X \in \mathfrak{g}$ can be uniquely written as

$$X = X_0 + A_0 + \sum_{\alpha \in \Delta} X_\alpha \tag{2.5}$$

for some elements $X_0 \in \mathfrak{m}$, $A_0 \in \mathfrak{a}$ and $X_\alpha \in \mathfrak{g}_\alpha$, $\alpha \in \Delta$. If $X \in Z_{\mathfrak{g}}(A)$, then

$$0 = [A, X] = \sum_{\alpha \in \Delta} \alpha(A) X_\alpha,$$

which implies $X_\alpha = 0$ for all $\alpha \in \Delta$. Hence, $X \in \mathfrak{m} \oplus \mathfrak{a}$ and the converse is obvious. □

Proposition 2.2.4. *Let $\mathfrak{a}$ and $\mathfrak{a}'$ be two maximal abelian subspaces of $\mathfrak{p}$ and let H denote the Lie group from Theorem 2.2.1. There exists an element $h_0 \in H$ such that $Ad(h_0)\mathfrak{a} = \mathfrak{a}'$. Consequently, the space $\mathfrak{p}$ satisfies*

$$\mathfrak{p} = \bigcup_{h \in H} Ad(h)\mathfrak{a}. \tag{2.6}$$

Proof. Since there are only finitely many restricted roots, there exists an element $A \in \mathfrak{a}$ such that $\alpha(A) \neq 0$ for all $\alpha \in \Delta$ and thus $Z_{\mathfrak{p}}(A) = \mathfrak{a}$. By carrying out the same procedure for the root space decomposition with respect to $\mathfrak{a}'$, we find some $A' \in \mathfrak{a}'$ such that $Z_{\mathfrak{p}}(A') = \mathfrak{a}'$. Let B denote the Killing form of $\mathfrak{g}$ and consider the smooth map

$$h \mapsto B(\mathrm{Ad}(h)A', A), \qquad h \in H.$$

By Theorem 2.2.1, the adjoint group $\mathrm{Ad}(H) \cong H/Z$, where Z denotes the center of G, is compact, so there exists some $h_0 \in H$ such that $B(\mathrm{Ad}(h_0)A', A)$ is a local extremum. If $X \in \mathfrak{h}$, then the smooth function

$$t \mapsto B(\mathrm{Ad}(\exp(tX))\mathrm{Ad}(h_0)A', A), \qquad t \in \mathbb{R}$$

has a critical point at $t = 0$, so we obtain by differentiation

$$0 = B(\mathrm{ad}(X)\mathrm{Ad}(h_0)A', A) = B(X, [\mathrm{Ad}(h_0)A', A]).$$

Now $X \in \mathfrak{h}$ is arbitrary and $[\mathrm{Ad}(h_0)A', A] \in \mathfrak{h}$, so we conclude that $[\mathrm{Ad}(h_0)A', A] = 0$ and $\mathrm{Ad}(h_0)A' \in Z_{\mathfrak{p}}(A) = \mathfrak{a}$. Therefore, we have

$$\mathfrak{a} \subset Z_{\mathfrak{p}}(\mathrm{Ad}(h_0)A') = \mathrm{Ad}(h_0)Z_{\mathfrak{p}}(A') = \mathrm{Ad}(h_0)\mathfrak{a}'$$

and since both sides are maximal abelian subspaces, equality must hold.
Finally, if $Y \in \mathfrak{p}$ is arbitrary, then the one-dimensional abelian subspace $\mathrm{span}(Y) \subset \mathfrak{p}$ can be extended to a maximal abelian subspace $\mathfrak{a}'$ of $\mathfrak{p}$ and then $Y \in \mathfrak{a}' = \mathrm{Ad}(h_0)\mathfrak{a}$ for some $h_0 \in H$, which proves (2.6). □

Since roots are by definition non-zero, every root $\alpha \in \Delta$ determines a hyperplane

$$\mathfrak{a}_\alpha := \ker(\alpha) = \{A \in \mathfrak{a} : \alpha(A) = 0\} \subset \mathfrak{a}$$

and an element in $\mathfrak{a} \subset \mathfrak{p}$ is regular if and only if it is not contained in any such plane. Let $R(\mathfrak{a})$ denote the set of regular elements of $\mathfrak{a}$, then we define the following equivalence relation on $R(\mathfrak{a})$.

$$A_1 \sim A_2 :\Longleftrightarrow \forall \alpha \in \Delta : \alpha(A_1)\alpha(A_2) > 0 \tag{2.7}$$

The corresponding equivalence classes are called the *Weyl chambers* of $\mathfrak{a}$ and they are precisely the connected components of $R(\mathfrak{a})$. Moreover, if $A_1 \sim A_2$, then we clearly have $A_1 \sim \lambda A_2$ for every $\lambda > 0$, so Weyl chambers are *cones*. If $W \subset \mathfrak{a}$ is any Weyl chamber, then by definition of the equivalence relation, the restriction of any root $\alpha \in \Delta$ to W is either strictly positive or strictly negative.
Since the negative of every root $\alpha \in \Delta$ is a root as well and $\dim(\mathfrak{g}_\alpha) = \dim(\mathfrak{g}_{-\alpha})$, it is natural to distinguish them by introducing a notion of positivity on the set of roots, which can be done in the following way. Let $W \subset \mathfrak{a}$ be an arbitrary Weyl chamber, then a root $\alpha \in \Delta$ is called *positive* if $\alpha(A) > 0$ for all $A \in W$. We then define $\Delta^+ \subset \Delta$ to be the set of all positive roots and deduce that Δ can be written as the disjoint union of Δ^+ and $\Delta^- := \{-\alpha : \alpha \in \Delta^+\}$. With these notions at hand, we can formulate the following theorem.

Proposition 2.2.5 (Iwasawa decomposition)**.** *Let $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ be a Cartan decomposition of a real semisimple Lie algebra $\mathfrak{g}$. Let $\mathfrak{a} \subset \mathfrak{p}$ be a maximal abelian subspace, Δ the set of roots of $\mathfrak{g}$ with respect to $\mathfrak{a}$ and $\Delta^+ \subset \Delta$ a set of positive roots. Then*

$$\mathfrak{n} := \bigoplus_{\alpha \in \Delta^+} \mathfrak{g}_\alpha \tag{2.8}$$

is a nilpotent subalgebra and $\mathfrak{a} \oplus \mathfrak{n}$ is a solvable subalgebra of $\mathfrak{g}$. Moreover, we have the direct sum decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{a} \oplus \mathfrak{n}$.

Proof. It follows directly from $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta}$ that $\mathfrak{n}$ is a nilpotent and $\mathfrak{a} \oplus \mathfrak{n}$ is a solvable subalgebra of $\mathfrak{g}$. We further deduce from $\mathfrak{a} \subset \mathfrak{g}_0$ that $\mathfrak{a} \cap \mathfrak{n} = \{0\}$. It remains to show that $\mathfrak{h}$ and $\mathfrak{a} \oplus \mathfrak{n}$ are complementary subspaces. On the one hand, if $X \in \mathfrak{h} \cap (\mathfrak{a} \oplus \mathfrak{n})$, then $X = \theta X \in \mathfrak{a} \oplus \theta(\mathfrak{n})$ since $\mathfrak{a} \subset \mathfrak{p}$ is stable under θ. However, since θ maps positive to negative roots, we have $X = \theta X = (\mathfrak{a} \oplus \mathfrak{n}) \cap (\mathfrak{a} \oplus \theta(\mathfrak{n})) = \mathfrak{a}$ and hence $X \in \mathfrak{h} \cap \mathfrak{a} = \{0\}$.

On the other hand, every root is either positive or negative and θ is an isomorphism mapping the root spaces of positive roots to those of the corresponding negative roots, so we can write the root space decomposition (2.4) equivalently as $\mathfrak{g} = \mathfrak{g}_0 \oplus \mathfrak{n} \oplus \theta(\mathfrak{n})$. Using the notation in (2.5), we can express every $X \in \mathfrak{g}$ as

$$X = X_0 + A_0 + \sum_{\alpha \in \Delta} X_\alpha.$$

By the previous observation, this sum can be re-written as

$$X = \Big(X_0 + \sum_{\alpha \in \Delta^+} (X_{-\alpha} + \theta(X_{-\alpha}))\Big) + A_0 + \Big(\sum_{\alpha \in \Delta^+} (X_\alpha - \theta(X_{-\alpha}))\Big),$$

which is an element of $\mathfrak{h} \oplus \mathfrak{a} \oplus \mathfrak{n}$ because the first summand is an eigenvector for the eigenvalue $+1$ of θ and $\theta(X_{-\alpha}) \in \mathfrak{g}_\alpha$ for all $\alpha \in \Delta$. □

The decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{a} \oplus \mathfrak{n}$ is called an *Iwasawa decomposition* of $\mathfrak{g}$. Although this will not be relevant for our purposes, we mention here for completeness that - as in the case of a Cartan decomposition - the Iwasawa decomposition can be lifted to the group level. The corresponding proof can be found in [Kna96, Theorem 6.46].

Theorem 2.2.6 (Global Iwasawa decomposition)**.** *Let G be a semisimple Lie group and $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{a} \oplus \mathfrak{n}$ an Iwasawa decomposition of its Lie algebra. Let H be the subgroup from Theorem 2.2.1 and let A and N be the connected Lie subgroups of G with Lie algebras $\mathfrak{a}$ and $\mathfrak{n}$, respectively. Then the multiplication map*

$$\begin{aligned} H \times A \times N &\to G \\ (h, a, n) &\mapsto han \end{aligned}$$

is a diffeomorphism. Moreover, the subgroups A and N are simply connected.

2.3 Applications to symmetric spaces

Having collected all these purely algebraic results, we now return to our discussion of symmetric spaces. In the present section we demonstrate that this machinery can be applied in many ways to obtain detailed information about the algebraic and geometric structure of symmetric spaces.

2.3.1 An overview of the classification

As a first application, we illustrate how the classification of orthogonal symmetric Lie algebras gives rise to a classification of symmetric spaces. Our first step in this direction will be to clarify the relation between symmetric spaces and associated pairs. A priori, a symmetric space M can have several different associated pairs (G, H) and such a pair is minimal if $G = G(M)$ is the group of displacements of M, which need not coincide with the identity component $I_0(M)$ of its isometry group. However, we will show that these groups do in fact coincide if $I_0(M)$ is semisimple. To do this, we start with a preparatory lemma that also clarifies some of the ambiguity in constructing symmetric spaces from symmetric pairs.

Lemma 2.3.1. *Let (G, H) be a symmetric pair and let $\mathfrak{z}$ denote the center of the Lie algebra $\mathfrak{g}$ of G. If $\mathfrak{h} \cap \mathfrak{z} = \{0\}$, then there exists exactly one involutive automorphism $\sigma : G \to G$ such that $G_0^\sigma \subset H \subset G^\sigma$.*

Proof. The existence of such an automorphism is built into the definition of a symmetric pair, so it remains to show uniqueness. Suppose that σ_1 and σ_2 are two such automorphisms, then they induce two symmetric decompositions $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}_i$, $i = 1, 2$, where $\mathfrak{p}_i$ is the -1-eigenspace of the differential $(\sigma_i)_*$. As shown in (2.1), both decompositions are orthogonal with respect to the Killing form B of $\mathfrak{g}$. Every $X_1 \in \mathfrak{p}_1$ can then be uniquely written as $X_1 = Y + X_2$ for some $Y \in \mathfrak{h}$ and $X_2 \in \mathfrak{p}_2$. Since $\mathfrak{h} \cap \mathfrak{z} = \{0\}$, Lemma 2.1.3 shows that $B|_{\mathfrak{h}\times\mathfrak{h}}$ is negative definite. However, $Y = X_1 - X_2 \in \mathfrak{h}$ is orthogonal to $\mathfrak{h}$ and thus satisfies $B(Y, Y') = 0$ for all $Y' \in \mathfrak{h}$, which implies $Y = 0$ and $X_1 = X_2$. It now follows that $\mathfrak{p}_1 = \mathfrak{p}_2$ and $(\sigma_1)_* = (\sigma_2)_*$, which yields $\sigma_1 = \sigma_2$ since G is connected. □

The condition that $\mathfrak{h} \cap \mathfrak{z} = \{0\}$ is satisfied in particular if G acts effectively on G/H. In this case, it follows from Theorem 1.4.3 that all G-invariant Riemannian metrics on G/H possess the same family of symmetries. If G is semisimple, this is already sufficient to identify it with $I_0(G/H)$.

Theorem 2.3.2. *Let (G, H) be a symmetric pair and let Q be any G-invariant Riemannian metric on the quotient space G/H. If G is semisimple and acts effectively on G/H, then $G = I_0(G/H)$.*

Proof. Since G acts effectively on G/H, there exists a unique involutive automorphism $\sigma : G \to G$ such that $G_0^\sigma \subset H \subset G^\sigma$. In every G-invariant metric on G/H, the left-action l_g is by definition an isometry of G/H for every $g \in G$. The map $g \mapsto l_g$ is injective by effectiveness, so we may view G as a subgroup of $I(G/H)$, which is contained in $G' := I_0(G/H)$ since G is connected. It can be shown that G is closed in G' and thus a Lie subgroup (cf. [Hel01, p. 176, Remark 2]). Its Lie algebra $\mathfrak{g}$ can therefore be viewed as a subalgebra of $\mathfrak{g}'$. Moreover, G/H becomes a symmetric space in every G-invariant metric such that the symmetry at the base point $o := eH$ is given by $s_o(gH) = \sigma(g)H$. By identifying $g \in G$ with $l_g \in G'$ we may write (1.29) as $\sigma(g) = s_o g s_o$.
Let $\tilde{\sigma}$ denote the automorphism $g \mapsto s_o g s_o$ of G', then we have $\tilde{\sigma}|_G = \sigma$ and both σ and $\tilde{\sigma}$ induce symmetric decompositions

$$\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p} \qquad \text{and} \qquad \mathfrak{g}' = \mathfrak{h}' \oplus \mathfrak{p}',$$

where $\mathfrak{h}'$ is the Lie algebra of $H' := G'_o$. Moreover, we have $\mathfrak{h} \subset \mathfrak{h}'$ and $\mathfrak{p} = \mathfrak{p}'$ since $H \subset H'$ and both $\mathfrak{p}$ and $\mathfrak{p}'$ can be identified with the tangent space $T_o(G/H)$. Since H' contains no non-trivial normal subgroup of G', the pair $(\mathfrak{g}', \tilde{\sigma}_*)$ is a reduced orthogonal symmetric Lie algebra and can therefore be decomposed as $\mathfrak{g}' = \mathfrak{g}'_0 \oplus \mathfrak{g}'_- \oplus \mathfrak{g}'_+$ as in

Theorem 2.1.5. The Euclidean factor $\mathfrak{g}'_0 = \mathfrak{h}'_0 \oplus \mathfrak{p}'_0$ has the property that $\mathfrak{p}'_0$ is even an abelian ideal in $\mathfrak{g}'$ (cf. [Hel01, Chapter V, Lemma 1.3]) and $\mathfrak{p}'_0 \subset \mathfrak{p}' = \mathfrak{p} \subset \mathfrak{g}$ implies that it is an abelian ideal in $\mathfrak{g}$, so we have $\mathfrak{p}'_0 = \{0\}$ because $\mathfrak{g}$ is semisimple. Therefore, $\mathfrak{h}'_0$ is an ideal of $\mathfrak{g}'$ contained in $\mathfrak{h}'$, which must be trivial since $(\mathfrak{g}', \tilde{\sigma}_*)$ is reduced. Hence, $\mathfrak{g}'$ is the sum of two semisimple ideals and Proposition 2.1.6 implies that $[\mathfrak{p}', \mathfrak{p}'] = \mathfrak{h}'$. Finally, $[\mathfrak{p}', \mathfrak{p}'] = [\mathfrak{p}, \mathfrak{p}] \subset \mathfrak{h} \subset \mathfrak{h}'$ now shows that $\mathfrak{h} = \mathfrak{h}'$ and thus $\mathfrak{g} = \mathfrak{g}'$. Since G and G' are connected, they must coincide. □

Corollary 2.3.3. *If M is a symmetric space such that $I_0(M)$ is semisimple, then $G(M) = I_0(M)$. In particular, $I_0(M)$ is generated by the family of displacements $\{s_p \circ s_q : p, q \in M\}$.*

Proof. Let $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ be the usual symmetric decomposition of the Lie algebra of $I_0(M)$. We have seen in Lemma 1.3.4 that the Lie algebra of the group of displacements satisfies $\mathfrak{g}(M) = [\mathfrak{p}, \mathfrak{p}] \oplus \mathfrak{p}$. This shows that $\mathfrak{g}(M)$ is an ideal in $\mathfrak{g}$ and therefore semisimple as well. If $H(M)$ denotes the stabilizer of o under the action of $G(M)$, then we have an effective isometric action of the semisimple Lie group $G(M)$ on $M \cong G(M)/H(M)$, so Theorem 2.3.2 implies that $G(M) = I_0(G(M)/H(M)) = I_0(M)$. □

Every symmetric space with a semisimple group of isometries therefore has a unique associated pair and a unique orthogonal symmetric Lie algebra associated to it. This motivates the following more general definition.

Definition 2.3.4. Let M be a symmetric space, $G = I_0(M)$, $H = G_o$ and $\sigma : G \to G$ be the involution $g \mapsto s_o g s_o$. Then (G, H) is a symmetric pair and $(\mathfrak{g}, \sigma_*)$ is a reduced orthogonal symmetric Lie algebra. We say that M is of the *Euclidean*, *compact* or *non-compact type* according to the type of $(\mathfrak{g}, \sigma_*)$ and that M is *irreducible* if $(\mathfrak{g}, \sigma_*)$ is irreducible.

We will now establish a converse result which shows that it is possible to construct a symmetric space from a given orthogonal symmetric Lie algebra. In order to do this, we recall some aspects from the classification theory of Lie groups, which can be found in [Kun19, Chapter 22]. A fundamental result in this setting is *Lie's third theorem*, which states that every (abstract) Lie algebra $\mathfrak{g}$ can be realized as the Lie algebra of a connected Lie group G. Moreover, its universal cover $\tilde{G}$ possesses a unique Lie group structure such that its Lie algebra is $\mathfrak{g}$ as well and the covering $p : \tilde{G} \to G$ is a Lie group homomorphism. If H is another Lie group with Lie algebra $\mathfrak{h}$, then every Lie algebra homomorphism $\rho : \mathfrak{g} \to \mathfrak{h}$ can be uniquely lifted to a Lie group homomorphism $\sigma : \tilde{G} \to H$ such that $T_e\sigma = \rho$. In particular, these observations imply that there exists a simply connected Lie group with Lie algebra $\mathfrak{g}$ which is unique up to isomorphism. Using this, we obtain the following partial converse to the constructions we have met so far.

Theorem 2.3.5. *Let $(\mathfrak{g}, \rho)$ be an orthogonal symmetric Lie algebra and let (G, H) be a pair of Lie groups that associated to $(\mathfrak{g}, \rho)$. Suppose that G is simply connected and that H is connected. Then H is closed in G and the quotient space G/H is simply connected. Moreover, G/H can be equipped with a G-invariant Riemannian metric in which it becomes a symmetric space.*

Proof. It follows from the preceding discussion that a Lie group G with the stated properties does indeed exist. Since $\mathfrak{h} \subset \mathfrak{g}$ is a subalgebra, there also exists a connected Lie subgroup $H \subset G$ whose Lie algebra is $\mathfrak{h}$. Now G is simply connected and ρ is an automorphism of the Lie algebra $\mathfrak{g}$ of G, so there exists a unique Lie group automorphism $\sigma : G \to G$ such that $\sigma_* = \rho$. Then the Lie algebra $\mathfrak{h}$ of H coincides with the Lie algebra of the fixed point-group G^σ, which implies $H = G_0^\sigma$ since H is connected. In particular, H is closed in G and G/H can be naturally equipped with a smooth structure as a homogeneous space of G. Moreover, $\mathrm{Ad}(H) \subset GL(\mathfrak{g})$ is compact since it is connected and $\mathfrak{h}$ is compactly embedded in $\mathfrak{g}$. Hence, (G, H) is a symmetric pair and G/H can be turned into a symmetric space as in Theorem 1.4.3.
To prove that G/H is simply connected, let $\gamma : [0, 1] \to G/H$ be a continuous closed curve. We may without loss of generality assume that $\gamma(0) = \gamma(1) = eH$. Since the quotient map $\pi : G \to G/H$ is a submersion, it possesses local sections, which can be used to lift γ to a continuous curve $\overline{\gamma} : [0, 1] \to G$ such that $\pi \circ \overline{\gamma} = \gamma$. Then $\overline{\gamma}(0)$ and $\overline{\gamma}(1)$ are elements of H which can be connected by a continuous curve in H.[3] The concatenation of these curves is then closed in G and therefore homotopic to a constant path. Projecting the homotopy to G/H shows that γ is nullhomotopic as well. □

Therefore, every orthogonal symmetric Lie algebra $(\mathfrak{g}, \rho)$ gives rise to a simply connected symmetric space. In general, there is a lot of freedom in choosing the Riemannian metric on this space, but if $(\mathfrak{g}, \rho)$ is irreducible, then Proposition 2.1.11 implies that all G-invariant Riemannian metrics on G/H are scalar multiples of each other. This observation is a crucial step towards a classification of symmetric spaces. In fact, we claim that the construction from Theorem 2.3.5 induces an equivalence of categories between irreducible, simply connected symmetric spaces and irreducible orthogonal symmetric Lie algebras.
Let M be an irreducible, simply connected symmetric space, $G = I_0(M)$, $H = G_o$ and $\sigma : G \to G$ be the involution $g \mapsto s_o g s_o$. By definition, this means that the unique orthogonal symmetric Lie algebra $(\mathfrak{g}, \sigma_*)$ associated to M is irreducible and we can apply Theorem 2.3.5 to it. Let $\tilde{G}$ be a simply connected Lie group with Lie algebra $\mathfrak{g}$ and $\tilde{H}$ its connected Lie subgroup with Lie algebra $\mathfrak{h}$. The identity map on $\mathfrak{g}$ then lifts to a covering $p : \tilde{G} \to G$, which is the universal cover of G, such that $\tilde{H}$ is the identity component of $p^{-1}(H)$. Hence, p induces a covering[4]

$$\tilde{p} : \tilde{G}/\tilde{H} \to G/H, \qquad g\tilde{H} \mapsto p(g)H,$$

which is a diffeomorphism since $G/H \cong M$ is simply connected. The Riemannian structures on $\tilde{G}/\tilde{H}$ and G/H induce $\mathrm{ad}(\mathfrak{h})$-invariant inner products on $\mathfrak{p}$, which are scalar multiples of each other since $(\mathfrak{g}, \sigma_*)$ is irreducible. If we normalize the Riemannian metric on $\tilde{G}/\tilde{H}$ in such a way that these inner products coincide, then $\tilde{p}$ is an isometry.

[3] Since H is in particular a smooth manifold, connectedness of H is equivalent to path-connectedness.

[4] This follows from the fact that the discrete group $p^{-1}(H)/\tilde{H}$ acts discontinuously on the simply connected space $\tilde{G}/\tilde{H}$ and the corresponding orbit space can be identified with $\tilde{G}/p^{-1}(H) \cong G/H$ (cf. [Kun19, Proposition 21.4]).

Conversely, let $(\mathfrak{g}, \rho)$ be an irreducible orthogonal symmetric Lie algebra and G/H its corresponding simply connected symmetric space. Then $N := \{g \in G : l_g = \mathrm{id}_{G/H}\}$ is a normal subgroup of G contained in H, which must be discrete since $(\mathfrak{g}, \rho)$ is reduced by definition. Consequently, G/N and H/N are Lie groups with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$, respectively, and Theorem 2.3.2 implies that $G/N = I_0(G/H)$ since $\mathfrak{g}$ is semisimple. It follows from the proof of Theorem 2.3.5 that the differential of the involution $g \mapsto s_o g s_o$ of G/N coincides with ρ, so the orthogonal symmetric Lie algebra corresponding to the unique associated pair $(G/N, H/N)$ of G/H is $(\mathfrak{g}, \rho)$ as well.
Hence, these operations are indeed inverse to each other and we now show that they are even compatible with the respective notions of isomorphism. On the one hand, if $f : M_1 \to M_2$ is an isometry between irreducible symmetric spaces, then $g \mapsto f \circ g \circ f^{-1}$ defines a Lie group isomorphism $\Psi : I_0(M_1) \to I_0(M_2)$. Let σ_1 and σ_2 denote the involutions $g \mapsto s_p g s_p$ of $I_0(M_1)$ and $g \mapsto s_{f(p)} g s_{f(p)}$ of $I_0(M_2)$, respectively, where $p \in M$ is arbitrary. Then we have $\Psi \circ \sigma_1 = \sigma_2 \circ \Psi$ and the irreducible orthogonal symmetric Lie algebras $(\mathfrak{g}_1, (\sigma_1)_*)$ and $(\mathfrak{g}_2, (\sigma_2)_*)$ are isomorphic under the differential $T_e\Psi$. On the other hand, let $\psi : \mathfrak{g}_1 \to \mathfrak{g}_2$ be an isomorphism of irreducible orthogonal symmetric Lie algebras and let (G_1, H_1) and (G_2, H_2) be the corresponding associated pairs, where G_1, G_2 are simply connected and H_1, H_2 connected. Then ψ can be lifted to a Lie group isomorphism $\Psi : G_1 \to G_2$ which factors to a diffeomorphism $\bar{\Psi} : G_1/H_1 \to G_2/H_2$ between the corresponding simply connected symmetric spaces. Furthermore, for $i = 1, 2$, the G_i-invariant Riemannian metric on G_i/H_i arises from an $\mathrm{ad}(\mathfrak{h}_i)$-invariant inner product on $\mathfrak{p}_i$, which must be a scalar multiple of the restriction to $\mathfrak{p}_i$ of the Killing form B_i of $\mathfrak{g}_i$. Since we have $B_2(\psi(X), \psi(Y)) = B_1(X, Y)$ for all $X, Y \in \mathfrak{g}_1$, it follows that $\bar{\Psi}$ is an isometry if the Riemannian structures are normalized in the same way. This finishes the proof of our claim.
In particular, we conclude that if M is irreducible and $(\mathfrak{g}, \sigma_*)$ its orthogonal symmetric Lie algebra, then there exists a unique irreducible, simply connected symmetric space M^* whose orthogonal symmetric Lie algebra is the dual of $(\mathfrak{g}, \sigma_*)$, which is called the *dual* of M. Consequently, irreducible, simply connected symmetric spaces of the compact and the non-compact type are in bijective correspondence. We will see an example of this geometric duality in Section 3.5.
Summarizing our discussion shows that the classification of irreducible, simply connected symmetric spaces is equivalent to the classification of irreducible orthogonal symmetric Lie algebras, which we described in Theorem 2.1.13. Moreover, a general symmetric space M gives rise to a reduced orthogonal symmetric Lie algebra, which can be decomposed into a Euclidean, compact and non-compact factor as in Theorem 2.1.5. If M is simply connected, this decomposition can again be lifted to a geometric level.

Theorem 2.3.6. *Let M be a simply connected symmetric space. Then M is isometric to a product*

$$M \cong M_0 \times M_- \times M_+,$$

where M_0, M_- and M_+ are simply connected symmetric spaces of the Euclidean, compact and non-compact type, respectively.

Proof. Let $p : \tilde{G} \to G$ denote the universal covering of the identity component of the isometry group $G = I_0(M)$, set $H = G_o$ and let $\tilde{H}$ be the identity component of $p^{-1}(H)$. As in the discussion above, p factors to a covering $\tilde{p} : \tilde{G}/\tilde{H} \to G/H$ which is a diffeomorphism since $G/H \cong M$ is simply connected. Let $\sigma : G \to G$ denote the involution $g \mapsto s_o g s_o$. The reduced orthogonal symmetric Lie algebra $(\mathfrak{g}, \sigma_*)$ can be decomposed as $\mathfrak{g} \cong \mathfrak{g}_0 \oplus \mathfrak{g}_- \oplus \mathfrak{g}_+$ as in Theorem 2.1.5. Since $\mathfrak{g}_0, \mathfrak{g}_-, \mathfrak{g}_+$ are ideals in $\mathfrak{g}$, this decomposition is not only a direct sum of vector spaces but in fact a Lie algebra isomorphism. The Lie algebra of $\tilde{G}$ is also $\mathfrak{g}$, so there exist connected Lie subgroups G_0, G_-, G_+ of $\tilde{G}$ corresponding to these ideals. We claim that there is even a global Lie group isomorphism $\tilde{G} \cong G_0 \times G_- \times G_+$.
First, we show that elements from two of these subgroups always commute with each other. For example, let $a \in G_-$ and $b \in G_+$ and choose a continuous curve $c : [0, 1] \to G_-$ such that $c(0) = e$ and $c(1) = a$. Then we have $d(t) := c(t)bc(t)^{-1}b^{-1} \in G_- \cap G_+$ for all $t \in [0, 1]$ as both subgroups correspond to ideals in $\mathfrak{g}$ and are hence normal in $\tilde{G}$, but we also have $\mathfrak{g}_- \cap \mathfrak{g}_+ = \{0\}$, so $G_- \cap G_+$ must be discrete. It follows that $d(1) = d(0) = e$, which implies $aba^{-1}b^{-1} = e$ and $ab = ba$. The same argument works for a different combination of these subgroups. This commutativity shows that $(g^0, g^-, g^+) \mapsto g^0 g^- g^+$ defines a Lie group homomorphism $F : G_0 \times G_- \times G_+ \to \tilde{G}$. By construction, the differential of F at the neutral element is the Lie algebra isomorphism of $\mathfrak{g}_0 \oplus \mathfrak{g}_- \oplus \mathfrak{g}_+$ with $\mathfrak{g}$, which implies that F is a covering map. However, since $G_0 \times G_- \times G_+$ is connected and $\tilde{G}$ is simply connected, it follows that F must be a diffeomorphism, which proves the claim. In particular, these three subgroups are all simply connected.
The Lie group isomorphism $\tilde{G} \cong G_0 \times G_- \times G_+$ induces a corresponding decomposition $\tilde{H} \cong H_0 \times H_- \times H_+$ where H_0, H_-, H_+ are all connected. Moreover, we can lift the differential σ_* to an involutive automorphism $\tilde{\sigma} : \tilde{G} \to \tilde{G}$. The ideals $\mathfrak{g}_0, \mathfrak{g}_-, \mathfrak{g}_+$ are invariant under σ_*, which implies that G_0, G_-, G_+ are invariant under $\tilde{\sigma}$. The proof of Theorem 2.3.5 shows that (G_0, H_0), (G_-, H_-) and (G_+, H_+) are symmetric pairs and the quotient spaces $M_0 := G_0/H_0$, $M_- := G_-/H_-$ and $M_+ := G_+/H_+$ can be turned into simply connected symmetric spaces with the required properties. □

As a complement to the above result, the compact and non-compact factors can be further decomposed into irreducible ones. This can be achieved by an almost identical proof based on the Lie algebra splitting in Theorem 2.1.12 and using the fact that the isometry group of a symmetric space of the compact or the non-compact type is semisimple by definition (cf. [Hel01, Chapter VIII, Proposition 5.5]). Altogether, every simply connected symmetric space can be decomposed into a Euclidean part and a finite number of irreducible factors, each of which is either of the compact or the non-compact type. As the next result shows, the Euclidean factor is a simply connected Riemannian manifold whose sectional curvature is identically zero and hence isometric to $\mathbb{R}^n$ for some $n \geq 0$. The classification of simply connected symmetric spaces therefore reduces to the irreducible case, which is equivalent to the classification of irreducible orthogonal symmetric Lie algebras. As discussed in Section 2.1, this problem was solved by Élie Cartan and gives rise to a complete list of irreducible symmetric spaces. Both the result and a detailed discussion of Cartan's classification can be found in [Hel01, Chapter IX].

There is also a geometric aspect to the notions of Euclidean, compact and non-compact type. We have shown in Proposition 1.2.8 that the Riemann curvature tensor of a symmetric space can be described algebraically via the Lie bracket of its associated orthogonal symmetric Lie algebra. The type of a symmetric space is then reflected in its sectional curvature.

Theorem 2.3.7. *Let M be a symmetric space.*

(i) If M is of the compact type, then its sectional curvature is non-negative.

(ii) If M is of the non-compact type, then its sectional curvature is non-positive.

(iii) If M is of the Euclidean type, then its sectional curvature is identically zero.

Proof. Let S be a two-dimensional plane in T_oM and let $X, Y \in T_oM$ be an orthonormal basis of S. We directly consider X and Y as elements of $\mathfrak{p} \cong T_oM$ and we also view the inner product Q_o induced by the Riemannian metric as an $\mathrm{ad}(\mathfrak{h})$-invariant inner product on $\mathfrak{p}$. By (1.6) and Proposition 1.2.8, the sectional curvature at o along S is given by

$$K(S) = Q_o(R_o(X,Y)Y, X) = -Q_o([[X,Y],Y],X) = Q_o([[X,Y],X],Y). \tag{2.9}$$

If $\mathfrak{p}$ is abelian, we immediately see that $K(S) = 0$, so (iii) follows. Hence, we may assume that the Lie algebra $\mathfrak{g}$ of $I_0(M)$ is semisimple. There is a unique linear map $f : \mathfrak{p} \to \mathfrak{p}$ that satisfies $Q_o(f(X), Y) = B(X, Y)$ for all $X, Y \in \mathfrak{p}$, where B denotes the Killing form of $\mathfrak{g}$. Symmetry of B shows that $Q_o(f(X), Y) = Q_o(X, f(Y))$, which means that f is self-adjoint with respect to Q_o and thus diagonalizable with real eigenvalues $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$. Moreover, all eigenvalues are non-zero and have the same sign since $B|_{\mathfrak{p}\times\mathfrak{p}}$ is positive or negative definite. Let $E_1, \ldots, E_n \subset \mathfrak{p}$ denote the corresponding eigenspaces which are orthogonal with respect to Q_o and B.

We claim that $[E_i, E_j] = \{0\}$ whenever $i \neq j$. First, if $T \in \mathfrak{h}$ and $X_i \in E_i$, then we have

$$\begin{aligned} Q_o(f([T,X_i]),Y) &= B([T,X_i],Y) = -B(X_i,[T,Y]) = -Q_o(f(X_i),[T,Y]) \\ &= -Q_o(\lambda_i X_i,[T,Y]) = Q_o(\lambda_i[T,X_i],Y) \end{aligned}$$

for all $Y \in \mathfrak{p}$ and therefore $[\mathfrak{h}, E_i] \subset E_i$. Second, if $X_j \in E_j$ for some $i \neq j$, then $[X_i, X_j] \in [\mathfrak{p}, \mathfrak{p}] \subset \mathfrak{h}$ and

$$B(T,[X_i,X_j]) = B([T,X_i],X_j) = 0$$

since the eigenspaces are mutually orthogonal. However, $T \in \mathfrak{h}$ is arbitrary and $B|_{\mathfrak{h}\times\mathfrak{h}}$ is non-degenerate, so it follows that $[E_i, E_j] = \{0\}$.

We expand X and Y into eigenspaces of f in order to compute the sectional curvature. For $i = 1, \ldots, n$ let X_i and Y_i denote the projections of X and Y onto the eigenspace E_i. The previous claim now implies $[X,Y] = \sum_{i=1}^n [X_i, Y_i]$ and $[[X_i,Y_i],X] = [[X_i,Y_i],X_i]$. Substituting into (2.9) then yields

$$K(S) = \sum_{i=1}^n Q_o([[X_i,Y_i],X_i],Y_i) = \sum_{i=1}^n \frac{1}{\lambda_i} B([[X_i,Y_i],X_i],Y_i),$$

which is equivalent to

$$K(S) = \sum_{i=1}^{n} \frac{1}{\lambda_i} B([X_i, Y_i], [X_i, Y_i]). \tag{2.10}$$

By the respective definiteness of the Killing form on $\mathfrak{p}$ we have $\lambda_i < 0$ in case (i) and $\lambda_i > 0$ in case (ii) for all $i = 1, \ldots, n$. Furthermore, the Killing form $B|_{\mathfrak{h}\times\mathfrak{h}}$ is negative definite. Hence, the sectional curvature at o is non-negative or non-positive, respectively, and since the choice of o was arbitrary, the assertions follow. □

To conclude this section, we indicate how the classification of irreducible Hermitian symmetric spaces can be deduced from the general Riemannian case. To this end, we derive a simple criterion to detect whether a given symmetric space has a Hermitian structure. As a preparation, let us clarify the relation between holomorphic and non-holomorphic isometries.

Proposition 2.3.8. *If M is a Hermitian symmetric space, then $I_0(M)$ is semisimple if and only if $A_0(M)$ is semisimple and in this case we have $I_0(M) = A_0(M)$.*

Proof. The group $A(M)$ of holomorphic isometries is a Lie subgroup of $I(M)$ and contains all the symmetries of M. In particular, $A_0(M)$ contains the group of displacements $G(M)$ which coincides with $I_0(M)$ if the latter is semisimple. Conversely, if $A_0(M)$ is semisimple, then $A_0(M) = I_0(M)$ follows from Theorem 2.3.2. □

This result allows us to treat Hermitian symmetric spaces with a semisimple isometry group in the same framework as general symmetric spaces. In particular, there is no ambiguity between $I_0(M)$ and $A_0(M)$ when defining the notions of compact type, non-compact type and irreducibility for these spaces. The next result shows that in order to decide whether an irreducible symmetric space is Hermitian, it suffices to determine the center of the stabilizer of a point under the action of $I_0(M)$.

Theorem 2.3.9. *Let M be an irreducible symmetric space, $G = I_0(M)$, $H = G_o$ and $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ the symmetric decomposition of the Lie algebra of G induced by the involution $\sigma : g \mapsto s_o g s_o$ of G.*

(i) If M is a Hermitian symmetric space and J its canonical almost complex structure, then there exists an element $Z_0 \in \mathfrak{g}$, which is contained in the center of $\mathfrak{h}$, such that the complex structure $\tilde{J} : \mathfrak{p} \to \mathfrak{p}$ induced by J is given by $\tilde{J} = ad(Z_0)|_{\mathfrak{p}}$. In particular, the center of H is not discrete.

(ii) Conversely, if the center of H is not discrete, then M can be endowed with an almost complex structure which turns it into a Hermitian symmetric space.

Proof. (i) The Riemannian metric Q of M induces an $\mathrm{ad}(\mathfrak{h})$-invariant inner product on $\mathfrak{p}$, which is a scalar multiple of the Killing form $B|_{\mathfrak{p}\times\mathfrak{p}}$ by irreducibility. Since Q is Hermitian, we have $B(\tilde{J}(X), \tilde{J}(Y)) = B(X, Y)$ for all $X, Y \in \mathfrak{p}$. We extend $\tilde{J}$ linearly to all of $\mathfrak{g}$ by setting $\tilde{J}|_{\mathfrak{h}} \equiv 0$ and claim that this turns $\tilde{J}$ into a derivation of $\mathfrak{g}$.

We have seen in (1.33) that $\tilde{J}$ commutes with the adjoint representation of H on $\mathfrak{p}$, so differentiating this identity shows that $\tilde{J} \circ \operatorname{ad}(X)|_{\mathfrak{p}} = \operatorname{ad}(X) \circ \tilde{J}|_{\mathfrak{p}}$ holds for all $X \in \mathfrak{h}$, which is equivalent to

$$\tilde{J}[X, Y] = [X, \tilde{J}(Y)] \qquad \forall X \in \mathfrak{h},\, Y \in \mathfrak{p}. \tag{2.11}$$

Moreover, let $X, Y \in \mathfrak{p}$ and $Z \in \mathfrak{h}$, then $\operatorname{ad}(\mathfrak{g})$-invariance of the Killing form implies

$$\begin{aligned} B([X,Y] - [\tilde{J}(X), \tilde{J}(Y)], Z) &= B(Y, [Z, X]) - B(\tilde{J}(Y), [Z, \tilde{J}(X)]) \\ &= B(\tilde{J}(Y), \tilde{J}[Z, X] - [Z, \tilde{J}(X)]) = 0. \end{aligned}$$

Since $[\mathfrak{p}, \mathfrak{p}] \subset \mathfrak{h}$ and the restriction of B to $\mathfrak{h}$ is non-degenerate, it follows that

$$[\tilde{J}(X), \tilde{J}(Y)] = [X, Y] \qquad \forall X, Y \in \mathfrak{p}. \tag{2.12}$$

The derivation property now follows immediately from the relations (2.11) and (2.12) and noting that $\tilde{J}|_{\mathfrak{h}} \equiv 0$. The Lie algebra $\mathfrak{g}$ is semisimple by definition of irreducibility, so Lemma 2.1.7 shows that we can find $Z_0 \in \mathfrak{g}$ such that $\tilde{J} = \operatorname{ad}(Z_0)$. Since $\tilde{J}$ commutes with the involution σ_*, we must have $Z_0 \in \mathfrak{h}$. Finally, $\tilde{J}$ vanishes on $\mathfrak{h}$, which implies that Z_0 is in fact contained in the center of $\mathfrak{h}$.

(ii) Let H_0 denote the identity component of H. Since the inner product on $\mathfrak{p}$ induced by the Riemannian metric is $\operatorname{Ad}(H)$-invariant, we can view $\operatorname{Ad}(H_0)$ as a connected Lie subgroup of the orthogonal group $O(\mathfrak{p})$ acting irreducibly on $\mathfrak{p}$. In this case, the proof of [KN63, Appendix 5, Theorem 2] shows that the center of $\operatorname{Ad}(H_0)$ - and thus the center of $\operatorname{Ad}(H)$ - can at most be one-dimensional. However, since $h \mapsto \operatorname{Ad}(h)|_{\mathfrak{p}}$ is injective on H which has non-discrete center, it follows that the center of H is one-dimensional. Therefore, the identity component of the center of H is a compact, connected, one-dimensional abelian Lie group and thus isomorphic to the circle group $U(1) = \{z \in \mathbb{C} : |z| = 1\}$ (cf. [Bau14, Satz 1.10]). Hence, there exists $h_0 \in H$ such that $\operatorname{Ad}(h_0)|_{\mathfrak{p}}^2 = -\operatorname{id}_{\mathfrak{p}}$. Altogether, $\tilde{J} := \operatorname{Ad}(H_0)|_{\mathfrak{p}}$ defines an $\operatorname{Ad}(H)$-invariant complex structure on $\mathfrak{p}$ which turns $M \cong G/H$ into a Hermitian symmetric space as in Theorem 1.5.2.

□

The proof shows in particular that the center of H is either discrete or one-dimensional and it is one-dimensional if and only if M is a Hermitian symmetric space. The classification of irreducible symmetric spaces therefore immediately provides a classification of the Hermitian case as well, which can be found in [Hel01, p. 354].

2.3.2 The non-compact type

Having established the classification result for general symmetric spaces, we take a closer look at the non-compact type in this section. We recall that if M is a symmetric space of the non-compact type, then $G = I_0(M)$ is semisimple by definition and the symmetric decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ induced by the involution $\sigma : g \mapsto s_o g s_o$ of G is a Cartan decomposition of its Lie algebra. This has remarkable consequences for the topology and the algebraic structure of the symmetric space.

Theorem 2.3.10. *Let M be a symmetric space of the non-compact type, $G = I_0(M)$, $H = G_o$ and let $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ be the symmetric decomposition induced by the involution $\sigma : g \mapsto s_o g s_o$ of G. The map*

$$\begin{aligned} F : \mathfrak{p} \times H &\to G \\ (X, h) &\mapsto \exp(X)h \end{aligned}$$

is a diffeomorphism.

Proof. To be consistent with our notation from Section 2.2, we denote the Cartan involution σ_* corresponding to the Cartan decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ by θ. As usual, we denote by B the Killing form of $\mathfrak{g}$.
First, we prove that every $g \in G$ can be uniquely expressed as $g = \exp(X)h$ for some $X \in \mathfrak{p}$ and $h \in H$. Since M is complete, there is a geodesic γ in M such that $\gamma(0) = o$ and $\gamma(1) = g(o)$ and it is of the form $\gamma(t) = \mathrm{Exp}_o(tv)$ for some $v \in T_oM$. Let $X \in \mathfrak{p}$ be the unique element with $j_o(X) = v$ and set $g' := \exp(X)$. Then $g'(o) = \exp(X)(o) = \mathrm{Exp}_o(v) = \gamma(1) = g(o)$, so $h := (g')^{-1}g \in H$ and $g = g'h = \exp(X)h$. To show uniqueness, assume that $g = \exp(X_1)h_1 = \exp(X_2)h_2$, then

$$\mathrm{Ad}(g) = \exp(\mathrm{ad}(X_1))\mathrm{Ad}(h_1) = \exp(\mathrm{ad}(X_2))\mathrm{Ad}(h_2).$$

As explained in Section 2.2, $B_\theta(X, Y) := -B(X, \theta Y)$ defines an inner product on $\mathfrak{g}$ and the matrices of the linear maps $\mathrm{ad}(X_1)$ and $\mathrm{ad}(X_2)$ are symmetric with respect to a B_θ-orthonormal basis of $\mathfrak{g}$. Hence, their exponential images are symmetric and positive definite. Moreover, the matrices of $\mathrm{Ad}(h_1)$ and $\mathrm{Ad}(h_2)$ are orthogonal since the elements of $\mathrm{Ad}(H)$ commute with θ and the Killing form is invariant under automorphisms of $\mathfrak{g}$. The decomposition of an invertible matrix into a positive definite and an orthogonal matrix is unique (cf. [Lan02, Chapter XV, Theorem 6.9]), so we deduce that $\exp(\mathrm{ad}(X_1)) = \exp(\mathrm{ad}(X_2))$. Every symmetric matrix can be orthogonally diagonalized over $\mathbb{R}$, which easily implies that the matrix exponential map is injective on the set of symmetric matrices. It follows that $\mathrm{ad}(X_1) = \mathrm{ad}(X_2)$ and that $X_1 - X_2$ is contained in the center of $\mathfrak{g}$ which is trivial since $\mathfrak{g}$ is semisimple. Hence, we obtain $X_1 = X_2$ and then $h_1 = h_2$ follows as well.
It remains to show that F is a diffeomorphism. It is clearly smooth and since it is bijective, it suffices to prove that its tangent map is invertible at every point $(X, h) \in \mathfrak{p} \times H$. Every element in T_hH is of the form $T_eL_h(Z)$ for some $Z \in T_eH = \mathfrak{h}$ and the tangent space $T_X\mathfrak{p}$ can be canonically identified with $\mathfrak{p}$. Now we have for $Y \in \mathfrak{p}$ and $t \in \mathbb{R}$:

$$\begin{aligned} F(X + tY, h) &= \exp(X + tY)h = h\exp(\mathrm{Ad}(h^{-1})(X + tY)) \\ &= (L_h \circ \exp)(\mathrm{Ad}(h^{-1})(X + tY)) \\ F(X, h\exp(tZ)) &= \exp(X)h\exp(tZ) = L_{\exp(X)h}(\exp(tZ)) \end{aligned}$$

We set $X' := \mathrm{Ad}(h^{-1})X$ and $Y' := \mathrm{Ad}(h^{-1})Y$, then differentiating these identities at $t = 0$ using the chain rule and (1.26) implies

$$\begin{aligned} T_{(X,h)}F(Y, T_eL_h(Z)) &= T_{(X,h)}F(Y,0) + T_{(X,h)}F(0, T_eL_h(Z)) \\ &= T_eL_{\exp(X)h} \circ \left(\frac{1 - e^{-\mathrm{ad}(X')}}{\mathrm{ad}(X')}(Y') + Z \right). \end{aligned} \tag{2.13}$$

Since $L_{\exp(X)h}$ is a diffeomorphism of G, this expression vanishes if and only if the term in brackets is zero, which is an element of $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$. We have $Z \in \mathfrak{h}$ and from the proof of Corollary 1.2.10 we know that the $\mathfrak{p}$-component of the first summand is given by

$$\sum_{n=0}^{\infty} \frac{\mathrm{ad}(X')^{2n}}{(2n+1)!}(Y'). \tag{2.14}$$

However, the linear map $\mathrm{ad}(X')^2|_{\mathfrak{p}}$ is symmetric and positive semi-definite with respect to the inner product $B|_{\mathfrak{p}\times\mathfrak{p}}$ because

$$\begin{aligned} &B(\mathrm{ad}(X)^2Y, Z) = B([X,[X,Y]], Z) = B(Y, [X,[X,Z]]) = B(Y, \mathrm{ad}(X)^2 Z) \\ &B(\mathrm{ad}(X)^2Y, Y) = -B([X,Y],[X,Y]) \geq 0 \end{aligned}$$

holds for all $X, Y, Z \in \mathfrak{p}$. Hence, the expression (2.14) can only be zero if $Y' = 0$, which is equivalent to $Y = 0$. But then (2.13) only vanishes if $Y = Z = 0$. This shows that $T_{(X,h)}F$ is injective and that F is a local diffeomorphism at every point $(X, h) \in \mathfrak{p} \times H$. □

It is clear from the proof that the map F defined above is smooth and surjective for every symmetric space, but injectivity is only true in the non-compact case. In this situation, connectedness of G implies that H is connected as well. Moreover, $\mathfrak{p}$ is a vector space and hence contractible, so the entire topological information about $G = I_0(M)$ is contained in the stabilizer $H = G_o$. In particular, F induces a diffeomorphism $\mathfrak{p} \to G/H \cong M$, $X \mapsto \exp(X)(o)$. Together with the isomorphism $\mathfrak{p} \cong T_oM$ this proves the following fundamental result.

Corollary 2.3.11. *If M is a symmetric space of the non-compact type, then the Riemannian exponential map $Exp_p : T_pM \to M$ is a diffeomorphism for every $p \in M$. In particular, M is contractible and hence simply connected.*

Combining this result with Theorem 2.3.7 shows that a symmetric space of the non-compact type is a *Hadamard manifold*, i.e. a complete, simply connected Riemannian manifold of non-positive sectional curvature. These spaces are classical objects of study in Riemannian geometry and are in a sense characterized by the fact that their geodesics behave in a particularly simple way. We will make this more precise in Section 3.1 where we use the characteristic geometric properties of Hadamard manifolds to construct a method of compactification which can then be applied to compactify symmetric spaces of the non-compact type. To conclude this section, we derive some interesting algebraic properties that arise naturally from the previous observations. Just like Theorem 2.3.10, the following result is a special case of the global Cartan decomposition in Theorem 2.2.1.

Proposition 2.3.12. *Let M be a symmetric space of the non-compact type, $G = I_0(M)$ and $H = G_o$, then H is a connected, maximal compact subgroup of G that coincides with the fixed point set G^σ of the involution $\sigma : g \mapsto s_o g s_o$ of G. Conversely, every maximal compact subgroup H' of G is of the form $H' = G_p$ for some $p \in M$. In particular, every pair of maximal compact subgroups is conjugate by an element of G.*

Proof. Using the diffeomorphism $F : \mathfrak{p} \times H \to G$ we have already observed that H is connected. Next, we show that $H = G^\sigma$. Assume that there exists an element $g \in G^\sigma \backslash H$, then g can be uniquely expressed as $g = \exp(X)h$ for some $h \in H$ and $0 \neq X \in \mathfrak{p}$. However, we know that $H \subset G^\sigma$, which implies

$$g = \sigma(g) = \sigma(\exp(X))\sigma(h) = \exp(\sigma_*(X))h = \exp(-X)h.$$

This is a contradiction since the representation of g is unique and $X \neq 0$.
Furthermore, G is closed in $I(M)$ and the stabilizer $I(M)_o$ is compact, which implies that H is compact as well. Assume that K was another compact subgroup of G such that $H \subset K$. By Cartan's fixed point theorem (cf. [Ebe96, Theorem 1.4.6]), the action of K on M has a common fixed point $p \in M$, i.e. $k(p) = p$ for all $k \in K$. Since G acts transitively on M, we can find some $g \in G$ with $g(p) = o$. Then we have $gHg^{-1} \subset gKg^{-1} \subset gG_pg^{-1} = G_o = H$, so g normalizes H. Since conjugation by g is an automorphism of G, it follows that $gHg^{-1} = H$ and thus also $H = K$.
Conversely, if H' is a maximal compact subgroup of G, then its action on M fixes some point $p \in M$, which means $H' \subset G_p$. Compactness of G_p and maximality of H' then imply $H' = G_p$. Finally, for every pair of points $p, q \in M$ we can find $g \in G$ with $g(q) = p$, which implies $G_p = gG_qg^{-1}$, so G_p and G_q are conjugate in G. □

The results of this section show that symmetric spaces of the non-compact type are rather simple objects, both from a topological and from an algebraic point of view. The situation is very different in the compact case where none of the analogous statements remain true in general (cf. [Hel01, Chapter VII, Proposition 1.2] for examples). Nevertheless, compact spaces are often more convenient in applications, e.g. in analysis where integration plays an important role. It is therefore natural to ask how one might be able to compactify a symmetric space of the non-compact type. We will return to this question in Chapter 3 where we illustrate different approaches and examples to this problem.

2.3.3 Totally geodesic submanifolds

Before passing to this question, however, we need to study the intrinsic geometry of symmetric spaces in more detail. Our goal is to show that symmetric spaces contain a distinguished class of submanifolds which are rather rare in general Riemannian manifolds. Although this appears to be entirely a question of geometry, the algebraic machinery that we developed will prove to be very powerful.

Definition 2.3.13. Let M be a Riemannian manifold and $S \subset M$ a submanifold, then S is called *totally geodesic* if every geodesic of M that is tangent to S at some point is already entirely contained in S.

If M is a symmetric space, we are going to show that this geometric notion has the following algebraic counterpart.

Definition 2.3.14. Let $\mathfrak{g}$ be a Lie algebra. A linear subspace $\mathfrak{s} \subset \mathfrak{g}$ is called a *Lie triple system* if $[[\mathfrak{s}, \mathfrak{s}], \mathfrak{s}] \subset \mathfrak{s}$.

If $\mathfrak{g}$ is a Lie algebra and ρ an involutive automorphism of $\mathfrak{g}$, then the -1-eigenspace of ρ is obviously a Lie triple system. In particular, this shows that if (G, H) is an associated pair of a symmetric space M and $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ is the symmetric decomposition induced by the involution $\sigma : g \mapsto s_o g s_o$ of G, then $\mathfrak{p}$ is a Lie triple system which is isomorphic to the tangent space T_oM. Since the triple bracket in $\mathfrak{p}$ corresponds to the Riemann curvature tensor at o, a linear subspace $V \subset T_oM$ is a Lie triple system[5] if and only if it is *curvature invariant* in the sense that $R_o(V, V)V \subset V$. This already suggests that Lie triple systems have a geometric significance.

Theorem 2.3.15. *Let M be a symmetric space with base point $o \in M$.*

(i) If V is a Lie triple system in T_oM, then $S := Exp_o(V)$ has a natural smooth structure such that it becomes a totally geodesic submanifold of M containing o with tangent space $T_oS = V$.

(ii) Conversely, if S is a totally geodesic submanifold of M containing o, then its tangent space $T_oS \subset T_oM$ is a Lie triple system.

Proof. (i) As usual, we set $G = I_0(M)$, $H = G_o$ and let $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ be the symmetric decomposition induced by the involution $g \mapsto s_o g s_o$ of G. By definition, the space $\mathfrak{s} := j_o^{-1}(V) \subset \mathfrak{p}$ is a Lie triple system in $\mathfrak{g}$ and the Jacobi identity shows that $\mathfrak{g}' := [\mathfrak{s}, \mathfrak{s}] \oplus \mathfrak{s}$ is a Lie subalgebra of $\mathfrak{g}$. Let G' be the connected Lie subgroup of G with Lie algebra $\mathfrak{g}'$, set $S' := G' \cdot o$ and denote by H' the stabilizer of o under the action of G'. Then H' is a closed subgroup of G', so there is a natural smooth structure on the quotient space G'/H' and we define a smooth structure on S' by declaring the bijection $G'/H' \to S'$, $gH' \mapsto g(o)$, to be a diffeomorphism. This turns S' into an immersed submanifold of M such that $T_oS' = j_o(\mathfrak{s}) = V$. By Theorem 1.2.9, the geodesics of M through o are all of the form $t \mapsto \exp(tX)(o)$, $X \in \mathfrak{p}$. Such a geodesic is tangent to S' at o if and only if $X \in \mathfrak{s}$ and these geodesics are contained in S' by definition. Since G' is a group of isometries of M that maps S' to itself and acts transitively on it, every geodesic of M that is tangent to S' at an arbitrary point $p \in S'$ can be mapped to a geodesic of M that is tangent to S' at o by an element of G'. Since these geodesics are entirely contained in S' and G' maps S' to itself, it follows that S' is totally geodesic. Finally, we have $S' = G' \cdot o = \exp(\mathfrak{s})(o) = \mathrm{Exp}_o(V) = S$.

(ii) If $v, w \in T_oS$, then $t \mapsto \mathrm{Exp}_o(tv)$ and $t \mapsto \mathrm{Exp}_o(tw)$ are geodesics of M that are tangent to S at o and thus contained in S. The exponential map at o therefore

[5] We will also call a subspace $V \subset T_oM$ a Lie triple system if the corresponding isomorphic image $\mathfrak{s} := j_o^{-1}(V) \subset \mathfrak{p}$ is a Lie triple system in $\mathfrak{g}$.

restricts to a map $T_oS \to S$, which implies that $T_{tw}\mathrm{Exp}_o(v) \in T_{\mathrm{Exp}_o(tw)}S$. If we set $X := j_o^{-1}(v)$ and $Y := j_o^{-1}(w)$, then it follows from Corollary 1.2.10 that[6]

$$T_{tw}\mathrm{Exp}_o(v) = T_o(\exp(tY)) \circ T_e\tau \circ \sum_{n=0}^{\infty} \frac{\mathrm{ad}(tY)^{2n}(X)}{(2n+1)!}.$$

The proof of part (ii) of Proposition 1.3.5 shows that $T_o(\exp(tY))$ is parallel transport along the geodesic $s \mapsto \exp(stY)(o)$, so it maps T_oS to $T_{\mathrm{Exp}_o(tw)}S$. Setting $\mathfrak{s} := j_o^{-1}(T_oS) \subset \mathfrak{p}$, the fact that $T_{tw}\mathrm{Exp}_o(v) \in T_{\mathrm{Exp}_o(tw)}S$ now implies

$$\sum_{n=0}^{\infty} \frac{\mathrm{ad}(Y)^{2n}(X)}{(2n+1)!} t^{2n} \in \mathfrak{s} \qquad \forall t \in \mathbb{R}.$$

Taking the second derivative at $t = 0$ shows that $\mathrm{ad}(Y)^2(X) = [Y, [Y, X]] \in \mathfrak{s}$. In particular, we have $\mathrm{ad}(Y+Z)^2(X) \in \mathfrak{s}$ for all $X, Y, Z \in \mathfrak{s}$, which explicitly reads

$$\mathrm{ad}(Y+Z)^2 = \mathrm{ad}(Y)^2 + \mathrm{ad}(Y)\mathrm{ad}(Z) + \mathrm{ad}(Z)\mathrm{ad}(Y) + \mathrm{ad}(Z)^2$$

and the Jacobi identity implies that

$$2[Y, [Z, X]] + [X, [Y, Z]] = [Y, [Z, X]] + [Z, [Y, X]] \in \mathfrak{s}.$$

Interchanging X and Y shows that $4[X, [Z, Y]] + 2[Y, [X, Z]] \in \mathfrak{s}$ and adding this to the left-hand side of the expression above now finally yields $-3[X, [Y, Z]] \in \mathfrak{s}$. Consequently, $\mathfrak{s}$ and therefore also $T_oS = j_o(\mathfrak{s})$ are Lie triple systems.

□

Hence, the maximal connected, totally geodesic submanifolds of M containing o are precisely the sets of the form $\exp(\mathfrak{s})(o)$, where $\mathfrak{s} \subset \mathfrak{p}$ is a Lie triple system, and the proof of (i) shows that every such submanifold is homogeneous and complete. Moreover, the connected Lie subgroup G' of $G = I_0(M)$ with Lie algebra $\mathfrak{g}' = [\mathfrak{s}, \mathfrak{s}] \oplus \mathfrak{s}$ is invariant under the involution $\sigma : g \mapsto s_o g s_o$ of G since $\mathfrak{g}'$ is invariant under the differential $\sigma_* : \mathfrak{g} \to \mathfrak{g}$. This shows that (G', H') is a symmetric pair and $S \cong G'/H'$ is a symmetric space in its own right. Alternatively, one may note that for every $p \in S$ the geodesic reflection of M at p maps S to itself and thus restricts to an involutive isometry of S which constitutes the symmetry of S at p. Conversely, if G' is a σ-invariant, connected Lie subgroup of G, then its Lie algebra is invariant under σ_* and therefore decomposes as

$$\mathfrak{g}' = (\mathfrak{g}' \cap \mathfrak{h}) \oplus (\mathfrak{g}' \cap \mathfrak{p}),$$

where it is easily verified that $\mathfrak{s} := \mathfrak{g}' \cap \mathfrak{p}$ is a Lie triple system. Hence, it gives rise to a totally geodesic submanifold of M that can be expressed as $\exp(\mathfrak{s})(o) = G'' \cdot o$ for some σ-invariant, connected Lie subgroup G'' of G. The Lie algebra of G'' satisfies

$$\mathfrak{g}'' = [\mathfrak{s}, \mathfrak{s}] \oplus \mathfrak{s} = [\mathfrak{g}' \cap \mathfrak{p}, \mathfrak{g}' \cap \mathfrak{p}] \oplus (\mathfrak{g}' \cap \mathfrak{p}) \subset (\mathfrak{g}' \cap \mathfrak{h}) \oplus (\mathfrak{g}' \cap \mathfrak{p}) = \mathfrak{g}',$$

[6]Note that there we have suppressed the isomorphism $j_o : \mathfrak{p} \to T_oM$.

which shows that $G'' \cdot o \subset G' \cdot o$. However, both orbits are connected and have the same dimension, so completeness of $G'' \cdot o$ implies that they coincide. Collecting these observations and suppressing the isomorphism $j_o : \mathfrak{p} \to T_oM$, we have proved the following characterization.

Corollary 2.3.16. *Let M be a symmetric space, $G = I_0(M)$ and let $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ be the symmetric decomposition induced by the involution $\sigma : G \to G$, $g \mapsto s_o g s_o$.*

(i) If $V \subset T_oM$ is a Lie triple system, then $\mathfrak{g}' := [V, V] \oplus V$ is a Lie subalgebra of $\mathfrak{g}$ and the connected Lie subgroup G' of G with Lie algebra $\mathfrak{g}'$ is σ-invariant. The orbit $S := G' \cdot o = Exp_o(V)$ is the connected, complete, totally geodesic submanifold of M with $T_oS = V$ and is intrinsically a symmetric space.

(ii) Conversely, if G' is a connected, σ-invariant Lie subgroup of G, then $S := G' \cdot o$ is a connected, complete, totally geodesic submanifold of M and $T_oS \subset T_oM$ is a Lie triple system.

Hence, we have arrived at a completely Lie-theoretic description of totally geodesic submanifolds of symmetric spaces in terms of Lie triple systems or, equivalently, σ-invariant, connected Lie subgroups of the isometry group. We will be particularly interested in submanifolds with vanishing curvature.

Definition 2.3.17. (i) A Riemannian manifold M is called *flat* if its Riemann curvature tensor vanishes identically. If M is arbitrary, a submanifold $S \subset M$ is said to be *flat* if S has this property when viewed as a Riemannian manifold on its own.[7]

(ii) Let M be a symmetric space. A flat, totally geodesic submanifold of dimension r will be called an *r-flat* in M. The *rank* of M is the maximal dimension of a flat, totally geodesic submanifold of M and will be denoted by $\mathrm{rk}(M)$.

Corollary 2.3.18. *Let M be a symmetric space of the compact or the non-compact type. Let $\mathfrak{s} \subset \mathfrak{p}$ be a Lie triple system and $S := \exp(\mathfrak{s})(o)$ the corresponding totally geodesic submanifold. Then S is flat if and only if $\mathfrak{s}$ is abelian.*

Proof. The totally geodesic submanifold S is itself a symmetric space with tangent space $T_oS \cong \mathfrak{s}$, so its curvature tensor is given by $R_o(X,Y)Z = -[[X,Y],Z]$ for $X, Y, Z \in \mathfrak{s}$. If $\mathfrak{s}$ is abelian, this directly implies that R_o vanishes. Since the choice of base point is arbitrary, it follows that S is flat. Conversely, if S is flat, then the sectional curvature of S along any two-dimensional plane vanishes. Since S is totally geodesic, the sectional curvature of M along any plane in $T_oS \subset T_oM$ is zero as well. Now M is of the compact or the non-compact type, so we may apply (2.10) to deduce that $[\mathfrak{s}, \mathfrak{s}] = 0$ because the Killing form of $\mathfrak{g}$ is negative definite on $\mathfrak{h}$. □

[7] Every submanifold $S \subset M$ inherits a Riemannian structure by restricting the metric of M to the tangent spaces of S. However, the Riemann curvature tensor of S is in general not just the restriction of the curvature tensor of M to S, but depends on an additional quantity called the *second fundamental form*. The precise relation between the curvature tensors is expressed by the *Gauss equation* (cf. [Lee97, Theorem 8.4]). If S is totally geodesic, then the second fundamental form vanishes identically.

In particular, the rank of a symmetric space of the compact or the non-compact type can equivalently be defined as the dimension of a maximal abelian subspace, which is automatically a Lie triple system, in the -1-eigenspace $\mathfrak{p}$ of its associated orthogonal symmetric Lie algebra. Moreover, the first part of the previous proof clearly applies to every symmetric space. A one-dimensional subspace of $\mathfrak{p}$ is trivially abelian and the corresponding totally geodesic submanifold is a geodesic through o, which is trivially flat. Hence, the rank of a symmetric space M is at least 1 and if $\mathrm{rk}(M) = 1$, then we have $[\mathfrak{s}, \mathfrak{s}] \neq \{0\}$ for every subspace $\mathfrak{s} \subset \mathfrak{p}$ that is at least two-dimensional. In this case, if M is of the compact or the non-compact type, then (2.10) implies that the sectional curvature along any two-dimensional plane in T_oM is non-zero. Therefore, a symmetric space M of the compact or the non-compact type satisfies $\mathrm{rk}(M) = 1$ if and only if its sectional curvature is strictly positive or negative, respectively.
If $S \subset M$ is a totally geodesic submanifold and $f : M \to M$ an isometry, then $f(M)$ is totally geodesic as well. We will now prove that if M is of the compact or the non-compact type and $\mathrm{rk}(M) = k$, then $I_0(M)$ acts transitively on the set of k-flats in M. To do this, we recall from Proposition 2.2.4 that if $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ is a Cartan decomposition of a real semisimple Lie algebra, then any two maximal abelian subspaces of $\mathfrak{p}$ are related by the adjoint action of an element of a suitable Lie group with Lie algebra $\mathfrak{h}$. The geometric analogue of that result is the following.

Proposition 2.3.19. *Let M be a symmetric space of the compact or the non-compact type, $rk(M) = k$ and let S and S' be two k-flats in M.*

(i) For all $p \in S$ and $p' \in S'$ there exists an element $g \in I_0(M)$ such that $g(S) = S'$ and $g(p) = p'$.

(ii) For all $p \in S$ and $v \in T_pM$ there exists an element $h \in I_0(M)$ such that $h(p) = p$ and $T_ph(v) \in T_pS$.

Proof. Let us first assume that $p = p'$, that M is of the non-compact type and set $G = I_0(M)$ and $H = G_p$. Then T_pS and T_pS' correspond to maximal abelian subspaces $\mathfrak{a}, \mathfrak{a}' \subset \mathfrak{p}$ in the Cartan decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ induced by the involution $\sigma : g \mapsto s_p g s_p$ of G. We have shown in Proposition 2.3.12 that $H = G^\sigma$, so we may apply Proposition 2.2.4 to deduce that there exists an element $h \in H$ with $\mathrm{Ad}(h)\mathfrak{a} = \mathfrak{a}'$. Under the isomorphism $j_p : \mathfrak{p} \to T_pM$ the adjoint representation of H on $\mathfrak{p}$ corresponds to the isotropy representation of H on T_pM, so this relation is equivalent to $T_ph(T_pS) = T_pS'$. The isometry h then maps $S = \exp(\mathfrak{a})(p) = \mathrm{Exp}_p(T_pS)$ to $S' = \exp(\mathfrak{a}')(p) = \mathrm{Exp}_p(T_pS')$. If $p \neq p'$, we can first find $g \in G$ such that $g(p) = p'$ and then the problem is reduced to the previous situation, which proves (i) in the non-compact case. The claim in (ii) follows similarly from (2.6).
If M is of the compact type, then $\mathfrak{g}^* = \mathfrak{h} \oplus i\mathfrak{p}$ is a Cartan decomposition of its dual orthogonal symmetric Lie algebra. A subspace $\mathfrak{a} \subset \mathfrak{p}$ is abelian if and only if $i\mathfrak{a} \subset i\mathfrak{p}$ is abelian. Therefore, the results from Proposition 2.2.4 can also be applied in this context and the claim is proved in the same way as for the non-compact type. □

Corollary 2.3.20. *If M is a symmetric space of the compact or the non-compact type and rk$(M) = k$, then every geodesic γ of M is contained in a k-flat of M.*

Proof. Let $p = \gamma(0)$ and $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ be the symmetric decomposition induced by the involution $g \mapsto s_p g s_p$ of $I_0(M)$, then there is a unique $X \in \mathfrak{p}$ such that $\gamma(t) = \exp(tX)(p)$ for all $t \in \mathbb{R}$. There exists a maximal abelian subspace $\mathfrak{a} \subset \mathfrak{p}$ containing X and γ is contained in the flat, totally geodesic submanifold $\exp(\mathfrak{a})(p)$, which has dimension $\dim(\mathfrak{a}) = \mathrm{rk}(M) = k$. □

These results conclude our discussion of the classical theory of symmetric spaces. As we will see in several instances in Chapter 3, the rank of a symmetric space is a fundamental invariant which in a sense measures the "complexity" of the space in question. In particular, symmetric spaces of rank 1 are in many situations much simpler objects than those of higher rank. This is especially true in the context of compactifications of symmetric spaces of the non-compact type.

3 Compactifications of symmetric spaces

In the previous chapter we have seen that there are two distinguished classes of symmetric spaces, the compact and the non-compact type. One of the major results was that every symmetric space of the non-compact type is a Hadamard manifold, i.e. a complete, simply connected Riemannian manifold of non-positive sectional curvature. In particular, it is diffeomorphic to a vector space and hence not compact as the name suggests. It is therefore a natural question how one might compactify such a space. Initially, there are two evident approaches to this problem, which clearly apply to every non-compact manifold M.

- *The intrinsic method:* Define a set $M(\infty)$ of "boundary points" and a topology on $\overline{M} := M \cup M(\infty)$ that turns $\overline{M}$ into a compact topological space containing M as a dense, open subset. The explicit construction then often makes it possible to understand the relation between M and its boundary $M(\infty)$ well. However, the downside of this approach is that it is a priori not clear if $\overline{M}$ can still be viewed as a manifold (possibly with boundary) and whether extra structures available on M (e.g. a Riemannian metric) extend to the boundary.
- *The extrinsic method:* Find a compact manifold N such that there exists an open embedding $M \hookrightarrow N$. The closure $\overline{M} \subset N$ is then automatically compact and contains M as a dense, open subset. On the one hand, additional structures on N can then often be restricted both to M and to $\overline{M}$, so such an embedding may even give rise to new information about M itself. On the other hand, it is more difficult to investigate the structure of the boundary $\overline{M} \setminus M$ in this setting since it results from taking the topological closure which may be very complicated.

Another level of complication stems from the fact that both methods can usually be carried out in several different ways, which poses the problem of relating compactifications obtained by different constructions. In this chapter we give some examples of compactification methods of symmetric spaces and discuss their similarities and differences. We do not intend to develop a general theory, but instead focus on concrete cases that illustrate various phenomena that occur in this setting. For a systematic treatment of this subject, the reader may consult [AO05] or the extensive book [BJ06]. It should be pointed out that much of the general framework is not restricted to the class of symmetric spaces and the symmetric structure will not play a major role in most constructions. Nevertheless, symmetric spaces often arise naturally in this context and provide interesting examples that can be treated with a wide range of tools from different areas of mathematics.

3.1 The geodesic compactification

In the first section of this chapter we describe a simple example of the "intrinsic" compactification method. This construction can be applied to any Hadamard manifold, which is why we leave aside symmetric spaces for the moment and formulate the procedure in this more general setting. We return to symmetric spaces at the end of this section and describe a concrete example thereafter. Intuitively, a natural notion of "boundary" of a non-compact Riemannian manifold may be considered to be the "set of points that geodesics converge to". Our goal is to make this intuitive idea into a rigorous construction that actually leads to a compactification. We follow the discussion in [EO73].

Remark 3.1.1. The construction we have in mind is based on the property that geodesics in a Hadamard manifold behave in a particularly simple way, which is expressed in several classical theorems of Riemannian geometry. More precisely, we will make use of the following general results. Let M be a complete, connected Riemannian manifold of non-positive sectional curvature. A *geodesic triangle* in M is a set $\Delta \subset M$ formed by three distinct points, called its *vertices*, and three geodesic segments connecting them, called the *sides* of Δ.

(i) (Cartan-Hadamard Theorem, [Lee97, Theorem 11.5]): The Riemannian exponential map $\mathrm{Exp}_p : T_pM \to M$ is a covering map for every $p \in M$. If M is simply connected, it is a diffeomorphism and every $q \in M$ can be connected to p by a unique distance-minimizing geodesic.

(ii) (Law of cosines, [Hel01, Chapter I, Corollary 13.2]): Let a, b, c be the side lengths of a geodesic triangle in M and let α, β, γ be the angles opposite to a, b, c, respectively. Then we have $a^2 + b^2 - 2ab\cos(\gamma) \leq c^2$ and $\alpha + \beta + \gamma \leq \pi$.

(iii) (CAT(0)-property, [BH99, Chapter I, Lemma 2.14 and Chapter II, Theorem 1A.6]): Let $\Delta \subset M$ be a geodesic triangle with vertices $p, q, r \in M$. There exist points $p', q', r' \in \mathbb{R}^2$ such that $|p' - q'| = d(p,q)$, $|p' - r'| = d(p,r)$ and $|q' - r'| = d(q,r)$. The triangle $\Delta' \subset \mathbb{R}^2$ formed by p', q', r' is unique up to isometry and we have $d(x,y) \leq |x' - y'|$ for all $x, y \in \Delta$, where $x', y' \in \Delta'$ are the points corresponding to x and y.[1]

Since M is complete, every triple of points indeed determines a geodesic triangle. In the simply connected case, the uniqueness result in (i) implies that this triangle is unique and its side lengths are precisely the Riemannian distances between its vertices.

Throughout this section we assume that M is a Hadamard manifold with Riemannian metric Q, distance function $d : M \times M \to \mathbb{R}$ and that all geodesics of M are parametrized by arc length. For all $p, q \in M$ we denote by $\gamma_{p,q} : \mathbb{R} \to M$ the unique unit-speed geodesic that satisfies $\gamma_{p,q}(0) = p$ and $\gamma_{p,q}(l) = q$ where $l = d(p,q)$. The fundamental result for the following construction is the property that geodesics in a Hadamard manifold either approach each other or spread apart with increasing speed.

[1] A point in Δ is uniquely determined by its distance to the vertices. Therefore, every point in Δ can be uniquely identified with a point in Δ'.

Lemma 3.1.2. *For every pair of geodesics* γ_1, γ_2 *of* M*, the function*

$$t \mapsto d(\gamma_1(t), \gamma_2(t))$$

is continuous and convex.

Proof. The continuity of the Riemannian distance function follows from the triangle inequality, so it remains to show convexity. Let us first assume that $\gamma_1(0) = \gamma_2(0)$. Consider the triangle $\Delta \subset M$ with vertices $p := \gamma_1(0)$, $q := \gamma_1(1)$ and $r := \gamma_2(1)$. In a comparison triangle $\Delta' \subset \mathbb{R}^2$ with the same side lengths as Δ we have[2] $|\gamma_1(t) - \gamma_2(t)| = t|\gamma_1(1) - \gamma_2(1)|$ for every $t \in [0, 1]$, so the CAT(0)-property implies

$$d(\gamma_1(t), \gamma_2(t)) \leq |\gamma_1(t) - \gamma_2(t)| = t|\gamma_1(1) - \gamma_2(1)| = td(\gamma_1(1), \gamma_2(1)).$$

Assume now that $\gamma_1(0) \neq \gamma_2(0)$, then we define p, q, r as above and set $\sigma(t) := \gamma_{p,r}(lt)$ where $l = d(p, r)$. The argument above is valid even if the geodesics are only parametrized proportional to arc length. Hence, we may apply this result first to γ_1 and σ and then to γ_2^- and σ^-, where we set $\gamma_2^-(t) := \gamma_2(1 - t)$ and $\sigma^-(t) := \sigma(1 - t)$. This yields

$$\begin{aligned} d(\gamma_1(t), \gamma_2(t)) &\leq d(\gamma_1(t), \sigma(t)) + d(\sigma(t), \gamma_2(t)) \\ &= d(\gamma_1(t), \sigma(t)) + d(\sigma^-(1 - t), \gamma_2^-(1 - t)) \\ &\leq td(\gamma_1(1), \sigma(1)) + (1 - t)d(\sigma(0), \gamma_2(0)) \\ &= (1 - t)d(\gamma_1(0), \gamma_2(0)) + td(\gamma_1(1), \gamma_2(1)) \end{aligned}$$

for every $t \in [0, 1]$. The general case

$$d(\gamma_1((1 - t)a + tb), \gamma_2((1 - t)a + tb)) \leq (1 - t)d(\gamma_1(a), \gamma_2(a)) + td(\gamma_1(b), \gamma_2(b)),$$

where $t \in [0, 1]$ and $a, b \in \mathbb{R}$, follows from this by an affine shift $t \mapsto a + t(b - a)$ in the parametrization of the geodesics. □

Definition 3.1.3. Let M be a Hadamard manifold. Two geodesics $\gamma_1, \gamma_2 : \mathbb{R} \to M$ are said to be *asymptotic* if

$$\sup_{t \geq 0} d(\gamma_1(t), \gamma_2(t)) < \infty. \tag{3.1}$$

This defines an equivalence relation on the set of all maximally extended, unit-speed geodesics of M and the equivalence class of a geodesic γ will be denoted by $\gamma(\infty)$. If γ^- is the "opposite" geodesic defined by $\gamma^-(t) := \gamma(-t)$, then we set $\gamma(-\infty) := \gamma^-(\infty)$. We have $d(\gamma(t), \gamma^-(t)) = 2t$ for all $t \geq 0$, which implies that γ and γ^- are never asymptotic and hence $\gamma(\infty) \neq \gamma(-\infty)$. We define $M(\infty)$ to be the set of all such equivalence classes and $\overline{M} := M \cup M(\infty)$. It follows from Lemma 3.1.2 that if γ_1 and γ_2 are asymptotic, then $t \mapsto d(\gamma_1(t), \gamma_2(t))$ is non-increasing. In particular, if two asymptotic geodesics have a point in common, then they are already the same up to a unit-speed re-parametrization by the law of cosines. Given a geodesic γ and a point $p \in M$, there is therefore at most one geodesic through p that is asymptotic to γ. The following statement shows that such a geodesic does indeed exist.

[2] Here, we denote points in Δ and Δ' by the same letter. The context is determined by whether the Riemannian distance d or the Euclidean norm $|\cdot|$ is employed.

Proposition 3.1.4. *Let $p \in M$ be any point, then $M(\infty)$ can be identified with the unit sphere in T_pM.*

Proof. If γ is any geodesic that contains p, then we may assume that $p = \gamma(0)$, which implies that γ can be written as $\gamma(t) = \mathrm{Exp}_p(tv) =: \gamma_v(t)$ for some $v \in T_pM$ with $||v|| = 1$. If $v \neq w$, then γ_v and γ_w are not asymptotic since they both contain p but do not coincide. Hence, the equivalence classes of geodesics through p are in one-to-one correspondence with the unit sphere in T_pM and it remains to show that every geodesic of M is asymptotic to one that passes through p.
Let γ be any geodesic and consider the sequence of points $(x_n)_{n=1}^{\infty}$ where $x_n := \gamma(n)$. If we set $t_n = d(p, x_n)$, then for every $n \in \mathbb{N}$ there exists a unique geodesic γ_n with $\gamma_n(0) = p$ and $\gamma_n(t_n) = x_n$. The triangle inequality yields

$$d(\gamma(n), \gamma(t_n)) = |n - t_n| = |d(\gamma(0), x_n) - d(p, x_n)| \leq d(p, \gamma(0)),$$

which implies that $t_n \to \infty$. Fix $t \in \mathbb{R}$ and choose $N \in \mathbb{N}$ so large that $t \leq t_n$ for all $n \geq N$. Then convexity shows that

$$\begin{aligned} d(\gamma_n(t), \gamma(t)) &\leq \max\{d(\gamma_n(0), \gamma(0)), d(\gamma_n(t_n), \gamma(t_n))\} \\ &= \max\{d(p, \gamma(0)), d(\gamma(n), \gamma(t_n))\} \leq d(p, \gamma(0)). \end{aligned} \tag{3.2}$$

We can write γ_n as $\gamma_n(t) = \mathrm{Exp}_p(tv_n) = \gamma_{v_n}(t)$ for some sequence $(v_n)_{n=1}^{\infty}$ in the unit sphere of T_pM. By compactness, there exists a sub-sequence of $(v_n)_{n=1}^{\infty}$ that converges to some $v \in T_pM$ with $||v|| = 1$. We set $\gamma_\infty := \gamma_v$, then γ_∞ passes through p and it follows from continuity of the exponential map and (3.2) that γ_∞ is asymptotic to γ. By the first part of the proof, this asymptote is unique, so every convergent sub-sequence of $(v_n)_{n=1}^{\infty}$ has the same limit. $\square$

The previous result shows that for every $p \in M$ and $\xi \in M(\infty)$ there is a unique unit-speed geodesic $\gamma_{p,\xi}$ with $\gamma_{p,\xi}(0) = p$ and $\gamma_{p,\xi}(\infty) = \xi$. Our next goal is to construct an appropriate topology on $\overline{M}$. A useful notion of distance between points in $M(\infty)$ is given by the following.

Definition 3.1.5. Let $p \in M$ and $x, y \in \overline{M} \setminus \{p\}$, then the *angle subtended by x and y at p*, written as $\angle_p(x, y)$, is defined to be the angle between the vectors $\dot{\gamma}_{p,x}(0)$ and $\dot{\gamma}_{p,y}(0)$ in T_pM measured with respect to the inner product Q_p induced by the Riemannian metric.

If γ is a geodesic with $\gamma(0) = p$, then $\gamma = \gamma_{p,\gamma(t)}$ implies that $\angle_p(\gamma(t), y) = \angle_p(\gamma(\infty), y)$ for all $t > 0$ and $y \in \overline{M}$. For an arbitrary point $p \in M$ we now define for each $v \in T_pM$ with $||v|| = 1$ and every $\varepsilon \in (0, \pi)$ a set

$$C(v, \varepsilon) := \{y \in \overline{M} \setminus \{p\} : \angle_p(\gamma_v(\infty), y) < \varepsilon\}, \tag{3.3}$$

where $\gamma_v(t) := \mathrm{Exp}_p(tv)$, which is called a *cone with vertex p, axis v and angle ε*. Note that the vertex is not explicitly included in the notation since it is determined by the axis $v \in T_pM$. We are going to show that these cones can be used to define a sensible topology on $\overline{M}$. The following properties will be useful.

Lemma 3.1.6. *(i) Let γ be a geodesic in M. If $s \leq t$ and $\delta \leq \varepsilon$, then we have $C(\dot\gamma(t), \delta) \subset C(\dot\gamma(s), \varepsilon)$.*

(ii) Let V be a cone with vertex p. For every $\xi \in V \cap M(\infty)$ and every $q \in M$ there exist $T > 0$ and $\delta > 0$ such that $C(\dot\gamma_{q,\xi}(t), \delta) \subset V$ for all $t \geq T$.

Proof. (i) The claim is obvious if $s = t$, so we may assume that $s < t$ such that the vertices $\gamma(s)$ and $\gamma(t)$ of the cones are distinct. If $y \in C(\dot\gamma(t), \delta)$, then we have $\angle_{\gamma(t)}(\gamma(\infty), y) < \delta$, so the complementary angle satisfies $\angle_{\gamma(t)}(\gamma(s), y) = \angle_{\gamma(t)}(\gamma(-\infty), y) > \pi - \delta$. Applying the angle sum property from Remark 3.1.1 yields $\angle_{\gamma(s)}(\gamma(\infty), y) = \angle_{\gamma(s)}(\gamma(t), y) < \delta \leq \varepsilon$, so $y \in C(\dot\gamma(s), \varepsilon)$.

(ii) Let us set $\alpha := \gamma_{p,\xi}$ and $\beta := \gamma_{q,\xi}$. Since $\xi \in V$ and the vertex of V is p, we can find $\varepsilon > 0$ such that $C(\dot\alpha(0), \varepsilon) \subset V$. By definition, α and β are asymptotic, which implies $\angle_p(\alpha(t), \beta(t)) \to 0$ as $t \to \infty$ by the law of cosines. Similarly, we also have $\angle_{\beta(t)}(p, q) \to 0$ as $t \to \infty$. Hence, for $\delta := \frac{\varepsilon}{3}$ there exists $T > 0$ such that

$$\angle_p(\alpha(t), \beta(t)) < \delta \qquad \text{and} \qquad \angle_{\beta(t)}(p, q) < \delta \qquad \forall t \geq T. \tag{3.4}$$

The second property also implies $\angle_{\beta(t)}(p, \xi) > \pi - \delta$. We claim that this choice of T and δ satisfies the assertion. Let $t \geq T$ and $y \in C(\dot\beta(t), \delta)$, then we have $\angle_{\beta(t)}(\xi, y) < \delta$. Combining the last two inequalities gives $\angle_{\beta(t)}(p, y) > \pi - 2\delta$ and the angle sum property yields $\angle_p(\beta(t), y) < 2\delta$. From this, we finally obtain that $\angle_p(\xi, y) \leq \angle_p(\xi, \beta(t)) + \angle_p(\beta(t), y) < 3\delta = \varepsilon$. This shows that $y \in C(\dot\alpha(0), \varepsilon) \subset V$, which proves the claim.

□

Geometrically, the first assertion states that if one slides a cone along a geodesic and reduces its angle, then the translated cone is always contained in the initial one. The second property shows that if $\xi \in M(\infty)$ is contained in a cone V, then V contains cones "along" every geodesic γ with $\gamma(\infty) = \xi$. This information is already sufficient to define a topology on $\overline{M}$.

Proposition 3.1.7. *Let $\xi \in M(\infty)$ and denote by $\mathcal{B}_\xi$ the family of cones containing ξ. There is a unique topology $\mathcal{T}$ on $\overline{M}$ such that:*

(i) M is a dense, open subset of $\overline{M}$ and the trace topology of $\mathcal{T}$ on M coincides with the natural manifold topology of M.

(ii) $\mathcal{B}_\xi$ is a neighbourhood basis at ξ.

Proof. Let us denote by $\mathcal{U}_\xi$ the family of subsets of $\overline{M}$ that contain a set from $\mathcal{B}_\xi$. If $U \in \mathcal{U}_\xi$, then we obviously have $\xi \in U$ and $U' \in \mathcal{U}_\xi$ for every $U' \subset \overline{M}$ with $U \subset U'$. Next, we claim that for all $V, W \in \mathcal{B}_\xi$ there exists a cone $U \in \mathcal{B}_\xi$ such that $U \subset V \cap W$. Let p and q be the vertices of V and W, respectively. On the one hand, $\xi \in V \cap W$ implies that we can find $\varepsilon_1, \varepsilon_2 > 0$ such that $C(\dot\gamma_{p,\xi}(0), \varepsilon_1) \subset V$ and $C(\dot\gamma_{q,\xi}(0), \varepsilon_2) \subset W$. On the other hand, part (ii) of Lemma 3.1.6 shows that there exist $T > 0$ and $\delta > 0$ such

that $C(\dot{\gamma}_{q,\xi}(t), \delta) \subset C(\dot{\gamma}_{p,\xi}(0), \varepsilon_1) \subset V$ holds for all $t \geq T$. We set $\lambda := \min\{\delta, \varepsilon_2\}$ and $U := C(\dot{\gamma}_{q,\xi}(T), \lambda)$. Part (i) of that lemma yields $C(\dot{\gamma}_{q,\xi}(t), \lambda) \subset C(\dot{\gamma}_{q,\xi}(0), \varepsilon_2) \subset W$ for $t \geq 0$, which shows that $U \subset V \cap W$. It is clear that $\xi \in U$, so the claim is proved. Thus, we have $U_1 \cap U_2 \in \mathcal{U}_\xi$ for all $U_1, U_2 \in \mathcal{U}_\xi$. By induction, an analogous result holds for the intersection of any finite number of sets from $\mathcal{U}_\xi$. Finally, let $U \in \mathcal{U}_\xi$ and choose a cone $V \in \mathcal{B}_\xi$ with $V \subset U$. If $\eta \in V \cap M(\infty)$, then we have $V \in \mathcal{U}_\eta$ by definition. Continuity of the exponential map at the vertex of V implies that $M \cap V$ is an open neighbourhood of all of its points in the natural manifold topology of M.
For every $p \in M$ we now define $\mathcal{U}_p$ to be family of subsets of $\overline{M}$ that contain a set which is an open neighbourhood of p in the natural manifold topology of M. By the properties of $\mathcal{U}_\xi$ shown above, there is a unique topology $\mathcal{T}$ on $\overline{M}$ such that the families $\mathcal{U}_p$ and $\mathcal{U}_\xi$ constitute the system of neighbourhoods of finite points $p \in M$ and boundary points $\xi \in M(\infty)$. By definition, the trace topology of $\mathcal{T}$ on M coincides with the natural manifold topology of M and the family $\mathcal{B}_\xi$ is a neighbourhood basis at ξ. Since the intersection of every cone with $M \subset \overline{M}$ is non-empty and open in the natural manifold topology of M, it follows that M is a dense, open subset of $\overline{M}$. □

We will from now on assume that $\overline{M}$ is equipped with the topology $\mathcal{T}$ constructed above, which is called the *cone topology*. It should be noted that it has the slightly uncomfortable property that different cones will usually have different vertices, which makes it difficult to compare them. However, it is possible to refine the construction in order to remedy this issue. Fix a point $p \in M$, then every cone (with an arbitrary vertex) contains a *truncated cone* with vertex p, i.e. a set of the form

$$T(v, \varepsilon, r) := C(v, \varepsilon) \setminus \{q \in M : d(p, q) \leq r\},$$

where $v \in T_pM$, $||v|| = 1$, $\varepsilon \in (0, \pi)$ and $r > 0$. The proof of this assertion is similar to Lemma 3.1.6 and can be found in [EO73, Proposition 2.6]. Clearly, every truncated cone is open in the cone topology. Given a point $\xi \in M(\infty)$, the set of truncated cones with fixed vertex p containing ξ therefore also forms a neighbourhood basis at ξ in the cone topology. Since p is arbitrary, we see in this picture that a sequence $(p_n)_{n=1}^\infty$ in $M \subset \overline{M}$ converges to $\xi \in M(\infty)$ if and only if $d(p, p_n) \to \infty$ and $\dot{\gamma}_{p,p_n}(0) \to \dot{\gamma}_{p,\xi}(0)$ holds for all $p \in M$. In particular, every geodesic $\gamma : \mathbb{R} \to M$ extends continuously to $[-\infty, \infty]$ in the obvious way. In Proposition 3.1.4 we have identified $M(\infty)$ with the unit sphere in T_pM. This observation can now be extended to see that $\overline{M}$ is in fact a compact topological space. To formulate the result, we consider the interval $[0, \infty]$ to be equipped with its standard topology, which turns it into a compact topological space that is homeomorphic to $[0, 1]$, e.g. via the map $x \mapsto 1 - \frac{1}{1+x}$ for $x \geq 0$ and mapping ∞ to 1.

Corollary 3.1.8. *Let $D(p)$ denote the closed unit ball and $S(p)$ the unit sphere in T_pM and let $f : [0, 1] \to [0, \infty]$ be a homeomorphism with $f(0) = 0$ and $f(1) = \infty$. The map*

$$\begin{aligned} \varphi : D(p) &\to \overline{M} \\ v &\mapsto \gamma_v(f(||v||)), \end{aligned}$$

where $\gamma_v(t) := Exp_p(tv)$, is a homeomorphism that maps $S(p)$ to $M(\infty)$.

Proof. We have already seen in the proof of Proposition 3.1.4 that the map φ restricts to a bijection $S(p) \to M(\infty)$. Moreover, it is injective on $D(p) \setminus S(p)$ since f is necessarily strictly increasing and different geodesics do not intersect. To show that φ is surjective, we note that every finite point $q \in M$ can be uniquely written as $q = \gamma_v(l) = \mathrm{Exp}_p(lv)$ for some $v \in S(p)$ where $l = d(p, q)$. For every $\lambda \in (0, 1)$ we then have $q = \gamma_{\lambda v}(l/\lambda)$ and we can choose λ in such a way that $f(\lambda) = l/\lambda$, which implies that $\varphi(\lambda v) = q$. Therefore, φ is bijective as a map $D(p) \to \overline{M}$ and by definition it can be written as $\varphi(v) = \mathrm{Exp}_p(f(||v||) \cdot v)$ on $D(p) \setminus S(p)$, which is continuous. Let $v \in S(p)$ and T be a truncated cone with vertex p containing $\varphi(v) \in M(\infty)$. Since the exponential map at p maps straight lines through 0 and balls centered at the origin to geodesics through p and metric balls centered at p, it follows that $\varphi^{-1}(T)$ is a truncated cone in the Euclidean sense in $D(p)$ containing v, so φ is continuous at v and hence continuous on all of $D(p)$. Finally, $D(p)$ is compact and $\overline{M}$ is Hausdorff since distinct points in $M(\infty)$ can be separated by cones with the same vertex by choosing the angles small enough, which implies that φ is a homeomorphism. □

Definition 3.1.9. The compact topological space $\overline{M}$ equipped with the cone topology is called the *geodesic compactification* of M.

Lemma 3.1.10. *Every isometry of M extends to a homeomorphism of $\overline{M}$ and the smooth action of $I(M)$ on M extends continuously to $\overline{M}$.*

Proof. If $f \in I(M)$ is an isometry and γ a unit-speed geodesic of M, then $f \circ \gamma$ is a unit-speed geodesic as well and we define $f \cdot \gamma(\infty) := (f \circ \gamma)(\infty)$. That f is an isometry directly implies that this definition only depends on the equivalence class of γ, meaning that the action is well-defined. Let $\xi \in M(\infty)$ and $C(v, \varepsilon)$ be a cone with vertex $p \in M$ containing ξ. Then we have $f(C(v, \varepsilon)) = C(T_p f(v), \varepsilon)$, which shows that f extends to a homeomorphism of $\overline{M}$. In particular, $I(M)$ becomes a subset of the space $C(\overline{M}, \overline{M})$ of continuous maps on $\overline{M}$. When this set is endowed with the compact-open topology, the evaluation map $(f, x) \mapsto f(x)$ is continuous since $\overline{M} \cong D(p)$ is locally compact. Since the topology on $I(M)$ is the compact-open topology as well (cf. [Hel01, Chapter IV, Section 2]), it follows that $I(M)$ acts continuously on $\overline{M}$. □

By construction, the extended action of $I(M)$ preserves the subsets M and $M(\infty)$ and conjugation with φ gives rise to an action of $I(M)$ on $D(p)$ which maps $S(p)$ to itself. If $H = I(M)_p$ denotes the stabilizer of p, then the action of H on $M(\infty)$ is conjugate to the action of H on $S(p) \subset T_pM$ under the isotropy representation. Therefore, H acts transitively on $M(\infty)$ if and only if it acts transitively on $S(p)$ under the isotropy representation on T_pM.

Remark 3.1.11. The geodesic compactification can in particular be applied to every symmetric space M of the non-compact type. In this situation, the symmetry s_p at any point $p \in M$ extends to a homeomorphism of $\overline{M}$ that satisfies $s_p(\gamma(\infty)) = \gamma(-\infty)$ for all geodesics γ. Whereas in this case $G := I_0(M)$ always acts transitively on M, it is not clear whether the extended action on $\overline{M}$ is still transitive. However, we can obtain a weaker result about the action of the stabilizer $H := G_p \subset G$.

As usual, let $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ denote the Cartan decomposition of the Lie algebra of G induced by the involution $\sigma : g \mapsto s_p g s_p$ of G. The isotropy representation of H on T_pM can then be identified with the adjoint representation of $H \subset G$ on $\mathfrak{p} \subset \mathfrak{g}$. If $\mathrm{rk}(M) = 1$, then it follows from Proposition 2.2.4 that $H = G^\sigma$ acts transitively on the unit sphere in $\mathfrak{p} \cong T_pM$ and therefore also on $M(\infty)$. Trivially, the action of G on this set is then also transitive. If $\mathrm{rk}(M) \geq 2$, however, then H cannot act transitively on $M(\infty)$. In fact, an element $X \in \mathfrak{p}$ is regular if and only if $\mathrm{Ad}(h)X$ is regular for every $h \in H$, so $\mathrm{Ad}(H)$ cannot map a regular to a non-regular element. If $\mathfrak{a} \subset \mathfrak{p}$ is a maximal abelian subspace with $\dim(\mathfrak{a}) \geq 2$, then the roots of $\mathfrak{g}$ with respect to $\mathfrak{a}$ are not injective, so there exists an element $0 \neq X \in \mathfrak{a}$ that satisfies $\alpha(X) = 0$ for some root $\alpha \in \Delta$. Then X is not regular and hence cannot be mapped to a regular element of $\mathfrak{p}$. By rescaling we may assume that X is contained in the unit sphere of $\mathfrak{p}$, which shows that H does not act transitively on this set.
If the rank of M is at least 2, then the action of the entire group G is much more difficult to describe and we only briefly indicate some constructions in this direction. If $\mathfrak{a} \subset \mathfrak{p}$ is a maximal abelian subspace with $\dim(\mathfrak{a}) \geq 2$, then the Weyl chambers of $\mathfrak{g}$ with respect to $\mathfrak{a}$ are non-trivial. Hence, one can intersect these Weyl chambers with the unit sphere in $\mathfrak{p} \cong T_pM$ to define Weyl chambers in $M(\infty)$ and this process is repeated for all maximal abelian subspaces $\mathfrak{a} \subset \mathfrak{p}$. Moreover, since the choice of p in the definition of σ and in Corollary 3.1.8 is arbitrary, this procedure can be carried out for all Cartan decompositions $\mathfrak{g} = \tilde{\mathfrak{h}} \oplus \tilde{\mathfrak{p}}$ with respect to all possible points $p \in M$. Taking the union of all these Weyl chambers defines an elaborate structure on $M(\infty)$, which is called the *spherical Tits building* associated to M. The identification of maximal abelian subspaces with maximal flat, totally geodesic submanifolds of M then associates a geometric interpretation to this structure, which gives rise to the *Tits geometry* of the symmetric space and whose "complexity" is directly proportional to the rank of M. Furthermore, the extended action of G on $\overline{M}$ allows one to consider the stabilizers $G_\xi \subset G$ of boundary points $\xi \in M(\infty)$. These turn out to be precisely the *parabolic subgroups* of G and have very different properties compared to the stabilizers of finite points. The analysis of the action of G on $M(\infty)$ is characterized by the interplay between the theory of parabolic subgroups and the Tits geometry of the symmetric space. We will not explore this subject in more detail, a nice introduction can be found in [Ebe96, Chapter 2 and 3]. Nevertheless, this emphasizes that the rank of a symmetric space is in fact an important invariant and that symmetric spaces of rank 1 often have a much simpler behaviour than those of higher rank.

Altogether, we have turned an arbitrary Hadamard manifold M into a compact topological space by adding a certain "boundary at infinity". The construction was entirely based on the intrinsic geometry of the manifold and from a topological perspective it is compatible with the action of the isometry group. Moreover, declaring the map φ from Corollary 3.1.8 to be a diffeomorphism turns $\overline{M}$ into a smooth manifold with boundary, so it is natural to ask whether the continuous action of $I(M)$ on $\overline{M}$ is even smooth. This turns out to be a delicate question, especially when M is a symmetric space of the non-compact type.

3.2 Compactifications of hyperbolic space

After the general discussion of the previous section, we now turn to a concrete problem of compactifying a symmetric space of the non-compact type. The simplest example in this setting is hyperbolic space which we already briefly introduced in Example 1.2.3. Here, it is possible to construct compactifications directly from a geometric realization of the space. We first follow this approach and afterwards compare the resulting space to the abstract geodesic compactification.

Assume that $n \geq 2$ and consider $\mathbb{R}^{n+1}$ with the standard Lorentzian bilinear form $\eta(x,y) = \sum_{i=1}^{n} x_i y_i - x_{n+1}y_{n+1}$. As mentioned, the one-sheet hyperboloid

$$H^n = \{x \in \mathbb{R}^{n+1} : \eta(x,x) = -1,\, x_{n+1} > 0\}$$

is a symmetric space whose isometry group is a subgroup of the Lorentz group $O(n,1)$, which acts on H^n by restrictions of linear maps of $\mathbb{R}^{n+1}$. Its identity component is the special orthochronous Lorentz group $G := SO_0(n,1)$. If we choose the base point $o = (0,\dots,0,1) \in H^n$, the stabilizer of o under the action of $SO_0(n,1)$ can be identified with $SO(n)$, so we can realize H^n as the homogeneous space

$$H^n \cong SO_0(n,1)/SO(n).$$

The symmetry at o has the form $s_o(p) = 2p_{n+1}o - p$, which is represented by the matrix $-I_{n,1}$ in the standard basis of $\mathbb{R}^{n+1}$. The involution σ of $SO_0(n,1)$ induced by the symmetric structure is therefore given by $\sigma(A) = I_{n,1}AI_{n,1}$ and its symmetric decomposition on the Lie algebra level reads

$$\mathfrak{so}(n,1) = \mathfrak{so}(n) \oplus \mathfrak{p}, \qquad \mathfrak{p} = \left\{ \begin{pmatrix} 0 & v \\ v^T & 0 \end{pmatrix} : v \in \mathbb{R}^n \right\}. \tag{3.5}$$

The Lie algebra $\mathfrak{so}(n,1)$ is simple with Killing form $B(X,Y) = (n-1)\mathrm{tr}(XY)$,[3] which is positive definite on $\mathfrak{p} \times \mathfrak{p}$. Moreover, the Lie bracket on $\mathfrak{p} \times \mathfrak{p}$ is given by

$$\left[\begin{pmatrix} 0 & v \\ v^T & 0 \end{pmatrix}, \begin{pmatrix} 0 & w \\ w^T & 0 \end{pmatrix} \right] = \begin{pmatrix} vw^T - wv^T & 0 \\ 0 & v^Tw - w^Tv \end{pmatrix}.$$

Now we always have $v^Tw - w^Tv = 0$, but the requirement $vw^T - wv^T = 0$ easily implies that every abelian subspace of $\mathfrak{p}$ is one-dimensional. Hence, H^n is a symmetric space of the non-compact type and $\mathrm{rk}(H^n) = 1$. The geodesics through the point o are given by the matrix exponentials of elements of $\mathfrak{p}$

$$\gamma_v(t) = \exp\left(t \begin{pmatrix} 0 & v \\ v^T & 0 \end{pmatrix} \right) o = \cosh(|v|t)o + \sinh(|v|t)\frac{v}{|v|}, \tag{3.6}$$

[3] This is a consequence of the fact that the complexification of $\mathfrak{so}(n,1)$ is $\mathfrak{so}(n+1,\mathbb{C})$ which is one of the classical complex simple Lie algebras (cf. [FH91, Chapter 21 and 26]). The explicit computation of the Killing form can be found in [Bau14, Lösung zu Aufgabe 1.8].

where $|v|$ denotes the Euclidean norm of $v \in \mathbb{R}^n$ which we identify with $(v, 0) \in \mathbb{R}^{n+1}$ in the second term on the right-hand side. Geometrically, we can read off from this expression that such a curve is given by the intersection of H^n with a two-dimensional plane containing o and the origin. The geodesics through arbitrary points are then obtained from the geodesics through o using the G-action, which is induced by linear maps of $\mathbb{R}^{n+1}$. Thus, every geodesic of the hyperboloid is given by the intersection of H^n with a two-dimensional plane containing the origin. It is well-known that H^n with this set of geodesics constitutes a model for hyperbolic geometry which will be called the *hyperboloid model*. There are two more famous models of hyperbolic space arising from suitable projections of the hyperboloid that will be more useful for the purpose of compactification.

The Klein model

A central projection from the origin onto the hyperplane $\mathbb{R}^n \times \{1\} \subset \mathbb{R}^{n+1}$ yields a diffeomorphism between H^n and the open unit ball in that plane, which we may identify with the open unit ball $B^n \subset \mathbb{R}^n$ by leaving out the last coordinate. Since this map only rescales the last coordinate of a point in H^n, it is explicitly given by

$$\begin{aligned} \varphi_K : H^n &\to B^n \\ (x_1, \dots, x_{n+1}) &\mapsto \frac{1}{x_{n+1}}(x_1, \dots, x_n) \\ \frac{1}{\sqrt{1-r^2}}(y_1, \dots, y_n, 1) &\leftarrow\!\shortmid (y_1, \dots, y_n), \end{aligned} \tag{3.7}$$

where we set $r := \sqrt{y_1^2 + \ldots + y_n^2}$ and we always take the positive square root. By endowing B^n with the pullback of the Riemannian metric of H^n under φ_K^{-1}, we obtain a Riemannian manifold (B^n, Q_K) which is called the *Klein model* of hyperbolic geometry. It is isometric to H^n by definition, so its geodesics correspond precisely to the images of geodesics in H^n under φ_K. Since the central projection preserves planes through the origin, the geodesics of the Klein model are also given by the intersections of $B^n \times \{1\}$ with two-dimensional planes containing the origin. Under the identification with $B^n \subset \mathbb{R}^n$ they are represented by line segments contained in B^n. From the explicit expression in (3.6) we even obtain the parametrized form of geodesics passing through the origin

$$\varphi_K(\gamma_v(t)) = \frac{\sinh(|v|t)}{\cosh(|v|t)} \frac{v}{|v|}, \tag{3.8}$$

which corresponds to a diameter of the unit ball. In other words, the geodesics of the Klein model are represented by straight lines.

Definition 3.2.1. Two Riemannian metrics on a Riemannian manifold M are said to be *projectively equivalent* if the geodesics of their respective Levi-Civita connections coincide up to re-parametrization.

The previous observation now simply states that the Riemannian metric Q_K in the Klein model is projectively equivalent to the Euclidean metric on B^n. However, clearly these metrics cannot coincide altogether since B^n is complete with respect to the hyperbolic metric, which is not the case for the Euclidean metric. Moreover, the parametrization in (3.8) is obviously different from that of a Euclidean geodesic.
By conjugation with φ_K the action of $G = SO_0(n,1)$ on the hyperboloid gives rise to an isometric action on the Klein model which turns φ_K into a G-equivariant diffeomorphism. The construction directly implies that this action on B^n maps line segments to line segments, which also follows from the fact that isometries preserve geodesics. In particular, we can translate the symmetric structure of H^n into the Klein model. The symmetry of B^n at the origin is simply the antipodal map $y \mapsto -y$. In general, the symmetry at an arbitrary point $p \in B^n$ has to reflect line segments through p while scaling them in such a way at either side of p that they remain contained in B^n.
By passing from the hyperboloid to the Klein model we have realized hyperbolic space as an open, bounded set in $\mathbb{R}^n$, so a natural approach to compactify it is to consider its closure, which is the closed unit ball $D^n := \overline{B^n} = B^n \cup S^{n-1}$ carrying the trace topology of $\mathbb{R}^n$. Alternatively, we can also apply the geodesic compactification to B^n and endow the resulting space $B^n \cup B^n(\infty)$ with the cone topology. In fact, the two resulting spaces coincide.

Proposition 3.2.2. *The identity map on B^n extends to a homeomorphism between D^n and $B^n \cup B^n(\infty)$.*

Proof. Two unit-speed geodesics γ_1 and γ_2 in the Klein model are asymptotic if and only if $\gamma_1(t)$ and $\gamma_2(t)$ converge to the same boundary point $\xi \in S^{n-1}$ as $t \to \infty$ (see the explicit expression of the distance function in [Lou20, Proposition 6.30 and 6.31]). Clearly, every such point is realized in this way, so S^{n-1} is in bijective correspondence with $B^n(\infty)$. Moreover, every equivalence class in $B^n(\infty)$ can be represented by a geodesic passing through the origin. Choosing such representatives, the identification of S^{n-1} with $B^n(\infty)$ associates to each $\xi \in S^{n-1}$ the equivalence class $\gamma_\xi(\infty)$, where $\gamma_\xi(t) = \frac{\sinh(t)}{\cosh(t)}\xi$ is the unit-speed geodesic through the origin with initial velocity ξ. A sequence $(y_k)_{k=1}^\infty$ in B^n converges to a point $\xi \in S^{n-1}$ if and only if $d(0, y_k) \to \infty$ and the sequence of unit-speed geodesics $(\gamma_k)_{k=1}^\infty$ connecting the origin to y_k (i.e. the line segment from the origin in the direction of y_k) converges to γ_ξ, which means that $(y_k)_{k=1}^\infty$ converges to $\gamma_\xi(\infty)$ in the cone topology. A similar argument applies to sequences that are entirely contained in S^{n-1}. Thus, convergence in the Euclidean topology of D^n is equivalent to convergence in the cone topology in $B^n \cup B^n(\infty)$. □

As a result of our general theory, the isometric action of $SO_0(n,1)$ on B^n extends to a continuous action on $B^n \cup B^n(\infty) \cong D^n$. The question now arises whether this action is even smooth on D^n, which is a smooth manifold with boundary. To answer this, it will be useful to view the compactification from a different perspective.
The construction of the Klein model was based on a central projection from the hyperboloid onto the plane $\mathbb{R}^n \times \{1\}$. This projection can equivalently be regarded as

identifying a point in H^n with the unique line connecting it to the origin. In this picture, it is more natural to view the resulting space B^n as an open subset of real projective space $\mathbb{R}P^n$ by identifying each $y \in B^n$ with its homogeneous coordinates $[y_1 : \ldots : y_n : 1]$. A line $\ell \subset \mathbb{R}^{n+1}$ containing the origin is said to be *timelike* if $\eta|_{\ell\times\ell}$ is negative definite, *spacelike* if $\eta|_{\ell\times\ell}$ is positive definite and *lightlike* if $\eta|_{\ell\times\ell} \equiv 0$. If a point $x \in \mathbb{R}^{n+1}$ is contained in H^n, then the unique line connecting it to the origin is timelike by definition and conversely, every timelike line intersects H^n in a unique point. Hence, we can view the open unit ball as the set of all timelike lines in $\mathbb{R}^{n+1}$. Now $\mathbb{R}P^n$ is naturally a smooth, n-dimensional, compact manifold which is endowed with a *global* smooth action of $SO_0(n,1)$ induced by matrix multiplication.

$$\begin{aligned} SO_0(n,1) \times \mathbb{R}P^n &\to \mathbb{R}P^n \\ (A,[y]) &\mapsto [Ay] \end{aligned}$$

The sets of timelike, lightlike and spacelike lines in $\mathbb{R}^{n+1}$ are clearly invariant subsets for this action. In fact, it will follow from the more general results in Proposition 3.3.4 and Remark 3.3.5 that these sets are precisely the orbits for the action of $SO_0(n,1)$ on $\mathbb{R}P^n$. In particular, it restricts to a transitive action on the set of timelike lines, which by construction coincides with the G-action in the Klein model when the latter is expressed in homogeneous coordinates. The boundary sphere S^{n-1} corresponds to the G-invariant set of lightlike lines, so from this viewpoint we immediately see that the G-action on B^n can be smoothly extended to the boundary. Thus, we obtain a smooth compactification of hyperbolic space inside $\mathbb{R}P^n$ which will be called the *projective compactification*.

The Poincaré ball model

A similar construction yields another useful model of hyperbolic space. Instead of a central projection, we now consider a stereographic projection of H^n from the point $(0,\ldots,0,-1)$ onto the plane $\mathbb{R}^n \times \{0\}$. As in the previous case, this defines a diffeomorphism between H^n and the open unit ball in that plane, which we identify with the unit ball $B^n \subset \mathbb{R}^n$ as before. A straightforward computation shows that this map is explicitly given by

$$\begin{aligned} \varphi_P : H^n &\to B^n \\ (x_1,\ldots,x_{n+1}) &\mapsto \frac{1}{1+x_{n+1}}(x_1,\ldots,x_n) \\ \frac{1}{1-r^2}(2y_1,\ldots,2y_n,1+r^2) &\leftarrow\!\shortmid (y_1,\ldots,y_n), \end{aligned} \tag{3.9}$$

where r is defined as above. By pulling back the metric from H^n, we obtain another Riemannian metric Q_P on B^n such that (B^n, Q_P) is isometric to the hyperboloid, which is called the *Poincaré ball model* of hyperbolic geometry. The stereographic projection does not preserve planes through the origin, so the geodesics of the Poincaré model look rather different compared to the Klein ball. Nevertheless, they still admit a simple geometric description.

Lemma 3.2.3. *The geodesics of the Poincaré ball model are represented by diameters of B^n and circular arcs in B^n that intersect S^{n-1} orthogonally.*

Proof. The Poincaré model is by definition isometric to the Klein ball where we already know that geodesics are given by the intersections of $B^n \times \{1\}$ with two-dimensional planes containing the origin. To pass from one model to the other, we can on the one hand consider the map $\varphi_P \circ \varphi_K^{-1}$ that relates them both to the hyperboloid. We deduce from (3.7) and (3.9) that this is given by

$$(\varphi_P \circ \varphi_K^{-1})(y) = \frac{1}{1+\sqrt{1-r^2}} y = \frac{1-\sqrt{1-r^2}}{r^2} y, \qquad y \in B^n. \tag{3.10}$$

On the other hand, we can also directly project $B^n \times \{1\}$ vertically onto S^n and then project S^n stereographically from the point $(0, \dots, 0, -1)$ onto the plane $\mathbb{R}^n \times \{0\}$. It is straightforward to verify that these maps are given by

$$\begin{aligned} B^n \times \{1\} \ni (y_1, \dots, y_n, 1) &\mapsto (y_1, \dots, y_n, \sqrt{1-r^2}) \in S^n \\ S^n \setminus \{(0, \dots, 0, -1)\} \ni (z_1, \dots, z_{n+1}) &\mapsto \frac{1}{1+z_{n+1}} (z_1, \dots, z_n, 0) \in \mathbb{R}^n \times \{0\}, \end{aligned}$$

which shows that their composition coincides with $\varphi_P \circ \varphi_K^{-1}$ when the last coordinate is omitted. Therefore, we obtain the geodesics of the Poincaré ball from those of the Klein model via these two projections. Under the vertical projection, a geodesic of the Klein ball (i.e. a line segment in B^n) is mapped onto a circular arc on the upper hemisphere that intersects the equator $S^{n-1} \times \{0\} \subset S^n$ orthogonally. Since stereographic projection is a conformal map (cf. [Lou20, Corollary 7.23]), this arc is then projected to a circular arc in $B^n \times \{0\}$ that intersects the boundary sphere orthogonally. If the original line segment contains the base point $o = (0, \dots, 0, 1)$ which corresponds to the origin in the Klein model, then its vertical projection is the upper half of a great circle on S^n through the north pole, which is mapped to a diameter of B^n under stereographic projection. □

Although the description of geodesics in the Poincaré model is more complicated than in the Klein model, we will soon see that it has another distinctive feature.

Definition 3.2.4. Two Riemannian metrics Q_1 and Q_2 on a Riemannian manifold M are said to be *conformally equivalent* if there exists a smooth function $f : M \to (0, \infty)$ such that $Q_1 = fQ_2$.

Conformal equivalence of metrics has a nice geometric interpretation. If $v, w \in T_pM$ are tangent vectors at $p \in M$, then their inner products with respect to $(Q_1)_p$ and $(Q_2)_p$ differ by a factor of $f(p)^2$, which is cancelled after dividing by the product of their respective norms. Hence, the angle between v and w is the same in both metrics.

Proposition 3.2.5. *The Riemannian metric Q_P in the Poincaré ball model is conformally equivalent to the Euclidean metric on B^n.*

Proof. To prove the assertion, we will explicitly compute the Riemannian metric of the Poincaré model. Let $(x_1, \ldots, x_{n+1})$ denote the components of φ_P^{-1}, then differentiating the second expression in (3.9) shows that:

$$dx_k = \sum_{i=1}^{n} \frac{2(1-r^2)\delta_{ik} + 4y_i y_k}{(1-r^2)^2} dy_i, \qquad k = 1, \ldots, n$$

$$dx_{n+1} = \sum_{i=1}^{n} \frac{2y_i(1-r^2) + 2y_i(1+r^2)}{(1-r^2)^2} dy_i = \sum_{i=1}^{n} \frac{4y_i}{(1-r^2)^2} dy_i$$

We now substitute these expressions into the Lorentzian metric $dx_1^2 + \ldots + dx_n^2 - dx_{n+1}^2$ of $\mathbb{R}^{n+1}$ which induces the Riemannian structure of H^n. In order not to overload the notation with summation signs, we employ Einstein summation convention in the following calculation. In particular, we may write $r^2 = y^l y_l$ in this case.

$$\begin{aligned}(1-r^2)^4 Q_P &= (1-r^2)^4(dx_1^2 + \ldots + dx_n^2 - dx_{n+1}^2) \\ &= [4(1-y^l y_l)^2 \delta^{ik}\delta^j_k + 16y^k y_k y^i y^j + 16(1-y^l y_l)\delta^{ik} y_k y^j] dy_i dy_j - 16y^i y^j dy_i dy_j \\ &= 4(1-y^l y_l)^2 dy^k dy_k - 16(1-y^k y_k) y^i y^j dy_i dy_j + 16(1-y^l y_l) y^i y^j dy_i dy_j\end{aligned}$$

The last two terms cancel after re-labelling the indices. Thus, we finally obtain that the Poincaré metric is given by

$$Q_P = \frac{4}{(1-r^2)^2}(dy_1^2 + \ldots + dy_n^2), \tag{3.11}$$

which also proves conformal equivalence to the Euclidean metric. □

In complete analogy with the Klein model, conjugation with φ_P induces an isometric action of $SO_0(n,1)$ on the Poincaré ball. The previous result now shows that this action has to preserve the Euclidean angles between intersecting curves. An isometry of the Poincaré model can thus be regarded as a conformal diffeomorphism of B^n, viewed as a subset of Euclidean space. In order to characterize these maps, let us first pass to a larger space whose conformal diffeomorphisms are easier to describe. We have already mentioned in the proof of Lemma 3.2.3 that stereographic projection from the south pole defines a conformal map $\mathbb{R}^n \to S^n$ which maps B^n to the open upper hemisphere. Hence, every isometry of the Poincaré ball can be viewed as a conformal diffeomorphism of the open upper hemisphere of S^n. From this viewpoint, it is natural to first investigate conformal maps of the sphere, which we briefly discuss in the following. A thorough and very explicit treatment of these notions, including proofs of most of the rather elaborate results, can be found in [HJ03, Chapter I, Section 3 and 5]. The connection to hyperbolic geometry and complex analysis is made explicit in [Lou20, Chapter 7].
As a first step, it will be convenient to view the sphere from a different perspective. Exactly as in the situation discussed so far, we now consider $\mathbb{R}^{n+2}$ and endow it with the standard Lorentzian bilinear form η. Then $O(n+1,1)$ acts on $\mathbb{R}^{n+2}$ via linear maps, so it induces a smooth action on real projective space $\mathbb{R}P^{n+1}$. Let $[x] = [x_1 : \ldots : x_{n+2}]$ denote

homogeneous coordinates for a point in $\mathbb{R}P^{n+1}$ and set $U := \{[x] \in \mathbb{R}P^{n+1} : x_{n+2} \neq 0\}$. Then U is diffeomorphic to $\mathbb{R}^{n+1}$ under the chart

$$\psi : U \to \mathbb{R}^{n+1} \qquad [x_1 : \ldots : x_{n+2}] \mapsto \frac{1}{x_{n+2}}(x_1, \ldots, x_{n+1}).$$

Let $\mathcal{C} := \{x \in \mathbb{R}^{n+2} : \eta(x,x) = 0\}$ denote the set of all null vectors, then the projective equivalence classes of non-zero points in $\mathcal{C}$ define a quadric $P\mathcal{C} := \{[x] \in \mathbb{R}P^{n+1} : \sum_{i=1}^{n+1} x_i^2 - x_{n+2}^2 = 0\}$, called the *projectivized light cone.* Note that no point in $P\mathcal{C}$ satisfies $x_{n+2} = 0$, so that $P\mathcal{C} \subset U$ and ψ restricts to a diffeomorphism $P\mathcal{C} \to S^n$. Since the action of $O(n+1,1)$ on $\mathbb{R}P^{n+1}$ preserves $P\mathcal{C}$, it induces a smooth action on S^n as well. In this setting, it can be shown that $O(n+1,1)$ acts on S^n by conformal maps and conversely, every conformal diffeomorphism of S^n arises in this way from an element of $O(n+1,1)$ which is unique up to sign. It follows that the group $\mathrm{Conf}(S^n)$ of conformal diffeomorphisms of S^n is isomorphic to $PO(n+1,1) = O(n+1,1)/\{\pm I_{n+2}\}$. Moreover, if $n \geq 3$, then even locally defined conformal maps can be characterized in this way. In this case, every conformal diffeomorphism $f : V_1 \to V_2$ between connected, open subsets $V_1, V_2 \subset S^n$ uniquely extends to a conformal diffeomorphism of the entire sphere. This is a strong result which is sometimes known as *Liouville's theorem* and has some remarkable consequences. First, conjugation with the stereographic projection shows that every conformal diffeomorphism of $\mathbb{R}^n$ arises from a conformal diffeomorphism of S^n that fixes the south pole, so $\mathrm{Conf}(\mathbb{R}^n)$ can be identified with a closed subgroup of $PO(n+1,1)$. Second, every conformal diffeomorphism $B^n \to B^n$ can be viewed as a conformal diffeomorphism of the open upper hemisphere of S^n. Since such a map extends to all of S^n, it has to preserve the equator $S^{n-1} \times \{0\} \subset S^n$ and induces a conformal diffeomorphism of S^{n-1}. Conversely, every $f \in \mathrm{Conf}(S^{n-1})$ uniquely extends to a map $\tilde{f} \in \mathrm{Conf}(S^n)$ preserving the upper hemisphere, which gives an alternative proof that the isometry group of hyperbolic space is isomorphic to $PO(n,1)$. The case $n = 2$ is slightly different and will be discussed below.

It is even possible to obtain a more explicit description of $\mathrm{Conf}(S^n)$ in this setting. In fact, if $g \in O(n+1,1)$ is a reflection at a hyperplane of the form $\mathrm{span}\{x\}^{\perp}$ with $\eta(x,x) > 0$, then it can be shown that the conformal diffeomorphism of S^n induced by g is an inversion at a hypersphere in S^n. Since $O(n+1,1)$ is generated by such reflections and $-I_{n+2}$, it follows that every conformal diffeomorphism of S^n is a *Möbius transformation*, i.e. a finite composition of inversions at hyperspheres. Such an inversion preserves the open upper hemisphere if and only if the sphere in question intersects the equator $S^{n-1} \subset S^n$ orthogonally, so also the conformal diffeomorphisms of B^n can be nicely characterized in this framework.

Returning to our initial considerations, let $g \in SO_0(n,1)$ be an isometry of the Poincaré ball, then g naturally extends to a conformal diffeomorphism of S^n which preserves S^{n-1}. In this picture, it is now obvious that the isometric action of $SO_0(n,1)$ on B^n extends smoothly to the boundary. To compactify the Poincaré ball, it is therefore natural to consider the closed unit ball D^n and view it as the closed upper hemisphere of S^n. We obtain a global smooth action of $SO_0(n,1)$ on the closed ball and thus another smooth compactification of hyperbolic space which will be called the *conformal compactification*. Moreover, the exact analogue of Proposition 3.2.2 is also true for the Poincaré model with an identical proof. Hence, the conformal compactification is also homeomorphic to the geodesic compactification of H^n. Finally, this construction is also possible if $n = 2$ as a consequence of the following remark.

Remark 3.2.6. At first glance there are many more conformal maps in the case $n = 2$ where we may identify $\mathbb{R}^2$ with the complex plane $\mathbb{C}$. In fact, it is a standard result of complex analysis that a function $f : U \to \mathbb{C}$ defined on an open subset $U \subset \mathbb{C}$ is conformal if and only if it is holomorphic or anti-holomorphic with non-vanishing derivative. Despite this abundance of locally defined functions, the holomorphic conformal maps preserving the unit disk B^2 take a much simpler form. Every such map is a *fractional linear transformation* $z \mapsto \frac{az+b}{cz+d}$, where $a, b, c, d \in \mathbb{C}$ satisfy $ad - bc \neq 0$, and it preserves B^2 if and only if $c = \bar{b}$, $d = \bar{a}$ and $|a|^2 - |b|^2 = 1$. Thus, the action of every element of $SO_0(2, 1)$ on B^2 is realized by a fractional linear transformation of an element of

$$SU(1,1) = \left\{ \begin{pmatrix} a & b \\ \bar{b} & \bar{a} \end{pmatrix} : a, b \in \mathbb{C},\ |a|^2 - |b|^2 = 1 \right\}.$$

Note that $-I_2 \in SU(1, 1)$ acts trivially on B^2, but this is the only non-trivial element with that property, so the identity component of the isometry group of the Poincaré disk can also be identified with $PSU(1, 1) = SU(1, 1)/\{\pm I_2\}$. In particular, there is a double covering $SU(1, 1) \to SO_0(2, 1)$. Fractional linear transformations naturally extend to the Riemann sphere S^2 and are again Möbius transformations in the sense that they are finite compositions of circle inversions. Exactly as above, every such map that preserves the upper hemisphere also preserves the equator $S^1 \subset S^2$, so we obtain a smooth action of $SO_0(2, 1)$ on the closed unit ball D^2 in this situation as well.
Finally, we also note that there is a subtle anomaly hidden in this example that is a consequence of the low number of dimensions. Instead of embedding B^2 into S^2, it is also possible to embed it into the complex projective line $\mathbb{C}P^1$, which is homeomorphic to S^2, by mapping $z \in B^2 \subset \mathbb{C}$ to $[z : 1] \in \mathbb{C}P^1$. A fractional linear transformation then maps $[z : 1]$ to $[\frac{az+b}{cz+d} : 1] = [az + b : cz + d]$. Hence, the isometric action of $SU(1, 1)$ on B^2 corresponds again to matrix multiplication when B^2 is viewed as a subset of $\mathbb{C}P^1$. This should not be confused with the action of $SO_0(2, 1)$ on $\mathbb{R}P^2$ that was used in the Klein model. The difference between these actions is captured in the non-trivial task of associating a fractional linear transformation to an element of $SO_0(2, 1)$.

Comparison between the models

Both the Klein and the Poincaré model provide a natural way to compactify hyperbolic space and on a topological level both methods coincide with the geodesic compactification. Furthermore, we have also seen that in both cases it was possible to extend the isometric action of $G = SO_0(n, 1)$ smoothly to the boundary, so one might ask whether the two compactifications are distinguishable on a differentiable level. To answer this question, we start with a purely formal consideration. In order to pass from one model to the other, we may simply compose their identifications with the hyperboloid, which map $y \in B^n$ to

$$\begin{aligned} (\varphi_P \circ \varphi_K^{-1})(y) &= \frac{1-\sqrt{1-r^2}}{r^2} y, \\ (\varphi_K \circ \varphi_P^{-1})(y) &= \frac{2}{1+r^2} y. \end{aligned} \tag{3.12}$$

These maps are by construction isometries between the models which are equivariant with respect to the two different G-actions. However, we see from the formulas that only one of the maps is compatible with the compactifications. The second map can be smoothly extended to the closed ball D^n whereas the first map only admits a continuous extension. In particular, we do not obtain a diffeomorphism between the compactified spaces in this way and in fact, there is no diffeomorphism that is compatible with the group actions.

Corollary 3.2.7. *There is no diffeomorphism of the closed unit ball that is equivariant with respect to the actions of $SO_0(n,1)$ in the Klein and Poincaré model.*

Proof. Since $G = SO_0(n,1)$ acts transitively in both models, any G-equivariant diffeomorphism $F : D^n \to D^n$ is uniquely determined by its value at a single point, e.g. the origin. The stabilizer of the origin in both models is the subgroup $SO(n) \subset SO_0(n,1)$ which has no other fixed point since it contains arbitrary rotations. The action of $SO(n)$ also stabilizes $F(0)$ by G-equivariance so that we must have $F(0) = 0$. The restriction of F to the open ball has to map B^n to itself and thus induces a G-equivariant diffeomorphism of B^n with $F(0) = 0$. Now either $\varphi_P \circ \varphi_K^{-1}$ or its inverse is another map with these properties, so F coincides with one of these maps on B^n and hence also on D^n by continuity. However, we have already observed that $\varphi_P \circ \varphi_K^{-1}$ does not extend to a diffeomorphism of D^n. □

Therefore, the two compactifications obtained from the Klein and the Poincaré model are topologically conjugate, but distinct on the smooth level. This difference can also be seen from a geometric point of view. Every geodesic in the Poincaré model intersects the boundary sphere orthogonally, so if two geodesics are asymptotic, then their tangent vectors become collinear at the point of intersection. In the Klein model, however, two distinct asymptotic geodesics never have collinear tangent vectors since geodesics are represented by straight lines.

We have seen that both compactifications are homeomorphic to the geodesic compactification of H^n. This example illustrates that the geodesic compactification is in general not well-adapted to the smooth structure of the underlying manifold since the G-action admits two smooth extensions to $H^n(\infty)$ that are not smoothly conjugate.

Remark 3.2.8. In this section we have only considered two specific examples that arise naturally in the study of hyperbolic geometry, but the observations of the preceding paragraph can be generalized. More precisely, it is shown in [Klo06] that there are infinitely many compactifications of H^n into a closed ball endowed with an action of $G = SO_0(n,1)$ with the following properties.

- There is a G-equivariant diffeomorphism from the open ball to H^n.
- The G-action on the interior of the ball can be smoothly extended to the boundary.
- There is a G-equivariant homeomorphism between the closed ball and the projective compactification obtained from the Klein model, but no such map is a diffeomorphism.

Hence, the different extensions of the G-action in the Poincaré and Klein model are only two examples from an infinite family of non-equivalent smooth compactifications of H^n.

The compactifications that we discussed in this section made strong use of a visual understanding of hyperbolic space and hinged on a concrete model from which they could be constructed. For more general symmetric spaces, such intuitive approaches are usually not available and it is necessary to resort to different methods.

3.3 Embeddings in Grassmannian manifolds

As explained in the previous section, the projective compactification in the Klein model could equivalently be regarded as an embedding of B^n into $\mathbb{R}P^n$ that is endowed with a particularly simple action of $SO_0(n,1)$. We now want to describe a method of compactification that applies to a wider class of symmetric spaces and in a sense generalizes the hyperbolic case. This will be an example of the "extrinsic method" that we mentioned at the beginning of this chapter. To motivate the construction, we recall that a general symmetric space M can always be expressed as a homogeneous space G/H of an associated pair (G,H) of M. In many cases - for the so-called "classical" groups - the group G arising in this setting can be realized as a group of matrices that can be understood by methods of linear algebra. In order to find suitable embeddings of M, the idea is to consider other manifolds that carry a smooth G-action. A particularly simple example would be a Lie group U such that G is a closed subgroup of U. If K is a closed subgroup of U, then U/K is a smooth manifold on which G acts as a Lie transformation group. The orbit of a point $p = uK \in U/K$ under the action of G can then naturally be identified with G/G_p. In particular, for the base point $o = eK$ we have $G_o = G \cap K$. If this coincides with H, the inclusion $G \hookrightarrow U$ descends to an embedding $G/H \hookrightarrow U/K$ and if U/K is compact, we obtain a compactification of G/H by taking its closure in U/K. We summarize this approach in the following definition.

Definition 3.3.1. Let G, U be Lie groups and $H \subset G$, $K \subset U$ closed subgroups such that U/K is compact. Assume that $G \subset U$ is a closed subgroup satisfying $G \cap K = H$. Then the inclusion $G \hookrightarrow U$ descends to an embedding $G/H \hookrightarrow U/K$ into a compact space and the closure $\overline{G/H} \subset U/K$ is called a *homogeneous compactification* of G/H.

Given a homogeneous compactification, the space G/H is realized as the orbit of a point in U/K under the action of G, so $G/H \subset U/K$ is an invariant subset for the G-action on U/K. This implies that also the closure $\overline{G/H}$ and the boundary $\partial(G/H)$ are G-invariant. However, it is a priori unclear whether homogeneous compactifications even exist for a given homogeneous space G/H and what its properties would be. Our first goal is show that an important family of symmetric spaces admits a homogeneous compactification.

Remark 3.3.2. As a preparation for the following construction, it will be useful to review some notions about bilinear forms that are not necessarily positive definite. For convenience, we restrict our attention to symmetric bilinear forms on real vector spaces. A detailed and more general discussion can be found in [Lan02, Chapter XV].
Let V be an n-dimensional real vector space and $b : V \times V \to \mathbb{R}$ a symmetric bilinear form, then two vectors $v, w \in V$ are said to be *orthogonal* if $b(v, w) = 0$. Given a fixed linear subspace $W \subset V$, its *orthogonal space* $W^{\perp} := \{v \in V : b(v, w) = 0 \ \forall w \in W\}$ is a linear subspace of V. We say that b is *non-degenerate* if $V^{\perp} = \{0\}$. In this case, the map $v \mapsto b(v, \cdot)$ induces an isomorphism between V and its dual space V^*. Restricting these maps to linear functionals on W defines a surjective linear map $V \to W^*$ with kernel $W^{\perp}$, which implies the dimension formula

$$\dim(W) + \dim(W^{\perp}) = \dim(V). \tag{3.13}$$

Moreover, we have $W = W^{\perp\perp}$ since the inclusion $W \subset W^{\perp\perp}$ holds trivially and the dimension formula applied to W and $W^{\perp}$ shows that $\dim(W) = \dim(W^{\perp\perp})$. The subspace W is said to be *non-degenerate* if $b|_{W\times W}$ is non-degenerate, which is equivalent to $W \cap W^{\perp} = \{0\}$. In this case, the orthogonal space $W^{\perp}$ is non-degenerate as well and the dimension formula shows that $V = W \oplus W^{\perp}$. Otherwise, there is an element $w \in W$ such that $b(w, w') = 0$ holds for all $w' \in W$, in which case $W \cap W^{\perp}$ is a non-trivial linear subspace of V and the sum $W + W^{\perp}$ is not direct. Its orthogonal space is $(W + W^{\perp})^{\perp} = W^{\perp\perp} \cap W^{\perp} = W \cap W^{\perp}$, so the dimension formula implies that

$$\dim(W + W^{\perp}) + \dim(W \cap W^{\perp}) = \dim(V). \tag{3.14}$$

A vector $v \in V$ is said to be *isotropic* if $b(v, v) = 0$ and a subspace $N \subset V$ is called *totally isotropic* if $b|_{N\times N} = 0$, i.e. $N \subset N^{\perp}$. If $v \in V$ is not isotropic, then $L := \operatorname{span}\{v\}$ is non-degenerate and we have $V = L \oplus L^{\perp}$ where $L^{\perp}$ is non-degenerate as well. Hence, it follows immediately by induction that if b is non-degenerate, then V possesses a basis $\{v_1, \dots, v_n\}$ that is *b-orthonormal* in the sense that $b(v_i, v_j) = \pm\delta_{ij}$. We may assume that $b(v_i, v_i) = 1$ for $1 \leq i \leq p$ and $b(v_i, v_i) = -1$ for $p + 1 \leq i \leq n$, in which case $\{v_1, \dots, v_p\}$ span a p-dimensional subspace of V on which b is positive definite. Similarly, $\{v_{p+1}, \dots, v_n\}$ span a subspace of dimension $q = n - p$ on which b is negative definite. The pair (p, q) is independent of the chosen basis and is called the *signature* of b. In particular, the dimension of a totally isotropic subspace is at most $\min(p, q)$.
If b is non-degenerate on V with signature (p, q) and $W \subset V$ a non-degenerate subspace, then two sets of b-orthonormal bases of W and $W^{\perp}$ form a b-orthonormal basis of V. More generally, the observation above can be used for an arbitrary subspace $W \subset V$ to obtain a simple basis of V that is compatible with the increasing chain of subspaces $W \cap W^{\perp} \subset W \subset W + W^{\perp} \subset V$, which we will construct in Proposition 3.3.3. To prepare for the proof, we need two preliminary observations. First, we note that by factoring out

all the elements orthogonal to W, b induces a non-degenerate symmetric bilinear form

$$\begin{aligned} \underline{b} : W/(W \cap W^\perp) \times W/(W \cap W^\perp) &\to \mathbb{R} \\ \underline{b}([w_1],[w_2]) &= b(w_1, w_2). \end{aligned}$$

The dimension $k := \dim(W/(W \cap W^\perp))$ is called the *rank* of W and the *signature* of W is defined to be the signature of the induced form $\underline{b}$. If $\nu := \dim(W \cap W^\perp)$ and W has dimension d, rank k and signature (r, s), then $r + s = k = d - \nu$. Second, the isomorphism $V \cong V^*$ arising from the assignment $v \mapsto b(v, \cdot)$ induces a duality $W \cap W^\perp \cong (V/(W + W^\perp))^*$. In fact, given a vector $w \in W \cap W^\perp$, the map $b(w, \cdot)$ vanishes identically on $W + W^\perp$ and thus induces a linear map $V/(W + W^\perp) \to \mathbb{R}$. Conversely, pre-composing such a map with the projection $V \to V/(W + W^\perp)$ defines an element $\varphi \in V^*$, which is of the form $\varphi = b(v, \cdot)$ where we must have $v \in W \cap W^\perp$ since φ vanishes on $W + W^\perp$. Having these notions at hand, we can construct the desired basis.

Proposition 3.3.3. *Let $W \subset V$ be a linear subspace of dimension d, signature (r, s) and set $\nu = dim(W \cap W^\perp)$, $\tilde{r} = p - r - \nu$ and $\tilde{s} = q - s - \nu$. There exists a basis $\{w_1, \dots, w_n\}$ of V that contains a basis for each linear subspace in the increasing chain $W \cap W^\perp \subset W \subset W + W^\perp \subset V$ and satisfies*

$$(b(w_i, w_j))_{i,j=1}^n = \begin{pmatrix} 0 & 0 & 0 & I_\nu \\ 0 & I_{r,s} & 0 & 0 \\ 0 & 0 & I_{\tilde{r},\tilde{s}} & 0 \\ I_\nu & 0 & 0 & 0 \end{pmatrix}. \tag{3.15}$$

Proof. To begin with, we choose an arbitrary basis $B_1 = \{w_1, \dots, w_\nu\}$ of $W \cap W^\perp$. Since this space is totally isotropic, we indeed have $b(w_i, w_j) = 0$ for $1 \le i, j \le \nu$ as required. Moreover, B_1 gives rise to a basis of $(V/(W + W^\perp))^*$ by the duality discussed above and we consider the corresponding dual basis $\{[v_1], \dots, [v_\nu]\}$ of $V/(W + W^\perp)$. Every set of representatives $\{v_1, \dots, v_\nu\}$ from these equivalence classes then satisfies $b(w_i, v_j) = \delta_{ij}$ for $1 \le i, j \le \nu$. Hence, $\{w_1, \dots, w_\nu, v_1, \dots, v_\nu\}$ span a subspace of V of dimension 2ν and signature (ν, ν).
Next, b induces a non-degenerate symmetric bilinear form $\underline{b}$ on $(W + W^\perp)/(W \cap W^\perp)$. We have

$$(W + W^\perp)/(W \cap W^\perp) = W/(W \cap W^\perp) \oplus W^\perp/(W \cap W^\perp),$$

so W and $W^\perp$ descend to complementary, non-degenerate subspaces in that quotient which are orthogonal for the induced form $\underline{b}$. The dimension of this space is $n - 2\nu$ by (3.14) and the signature of $\underline{b}$ is $(p - \nu, q - \nu)$ by the observation in the first paragraph. Since the signature of W is (r, s), it follows that $W^\perp$ has signature $(\tilde{r}, \tilde{s})$. We may choose a $\underline{b}$-orthonormal basis $\{[w_{\nu+1}], \dots, [w_d], [w_{d+1}], \dots, [w_{n-\nu}]\}$ of $(W + W^\perp)/(W \cap W^\perp)$ consisting of two sets of $\underline{b}$-orthonormal bases for $W/(W \cap W^\perp)$ and $W^\perp/(W \cap W^\perp)$. Let $\{w_{\nu+1}, \dots, w_d, w_{d+1}, \dots, w_{n-\nu}\}$ be any set of representatives from these equivalence classes, then these vectors span a subspace of $W + W^\perp$ that is complementary to $W \cap W^\perp$.

Hence, they extend B_1 to a basis $B_2 = \{w_1, \dots, w_{n-\nu}\}$ of $W + W^\perp$, where in addition $\{w_1, \dots, w_d\}$ is a basis of W. By construction, we also have $b(w_i, w_j) = 0$ whenever $1 \le i \le \nu < j \le n-\nu$ and $b(w_i, w_j) = \underline{b}([w_i],[w_j]) = \pm\delta_{ij}$ for $\nu+1 \le i, j \le n-\nu$, so the matrix $(b(w_i, w_j))_{i,j=1}^{n-\nu}$ has the desired shape.
Although the representatives $\{v_1, \dots, v_\nu\}$ chosen in the beginning would extend B_2 to a basis of V, it is not automatic that they have the right orthogonality properties, but this problem can be solved by adding suitable linear combinations. Every representative of $[v_i]$ is of the form $v_i + \sum_{j=1}^{n-\nu} \lambda_{ij} w_j$ for some $\lambda_{ij} \in \mathbb{R}$. For $k = \nu+1, \dots, n-\nu$ we have

$$b\left(w_k, v_i + \sum_{j=1}^{n-\nu} \lambda_{ij} w_j\right) = b(w_k, v_i) + \lambda_{ik} b(w_k, w_k),$$

so that we are forced to set $\lambda_{ik} = -\frac{b(w_k, v_i)}{b(w_k, w_k)}$ for $i = 1, \dots, \nu$ and $k = \nu+1, \dots, n-\nu$ to make this expression vanish. To determine the remaining coefficients, we compute

$$\begin{aligned} b\left(v_i + \sum_{k=1}^{n-\nu} \lambda_{ik} w_k, v_j + \sum_{l=1}^{n-\nu} \lambda_{jl} w_l\right) &= b\left(v_i + \sum_{k=1}^{n-\nu} \lambda_{ik} w_k, v_j + \sum_{l=1}^{\nu} \lambda_{jl} w_l\right) \\ &= b(v_i, v_j) + \sum_{k=1}^{\nu} \lambda_{ik} b(w_k, v_j) + \sum_{k=\nu+1}^{n-\nu} \lambda_{ik} b(w_k, v_j) + \sum_{l=1}^{\nu} \lambda_{jl} b(v_i, w_l) \\ &= b(v_i, v_j) + \lambda_{ij} + \lambda_{ji} - \sum_{k=\nu+1}^{n-\nu} \frac{b(w_k, v_i) b(w_k, v_j)}{b(w_k, w_k)}. \end{aligned}$$

Requiring this expression to be zero for $i, j = 1, \dots, \nu$ amounts to a system of ν^2 equations for the coefficients λ_{ij} which clearly has a solution, e.g. via the ansatz $\lambda_{ij} = \lambda_{ji}$. Using these coefficients, we may set $w_{n-\nu+i} = v_i + \sum_{j=1}^{n-\nu} \lambda_{ij} w_j$ for $i = 1, \dots, \nu$ to extend B_2 to a basis $B_3 = \{w_1, \dots, w_n\}$ of V with the required properties. □

We will say that a basis of V that satisfies the assumptions of Proposition 3.3.3 is *adapted to* W. Finally, we note that the above discussion also applies to complex vector spaces that are equipped with a non-degenerate Hermitian bilinear form, which follows directly by inspecting the arguments. In particular, we can also find adapted bases in this setting. However, requiring the form to be Hermitian instead of symmetric is really necessary, since if b is a symmetric bilinear form on a complex vector space V and $v \in V$ satisfies $b(v, v) > 0$, then $b(iv, iv) = -b(v, v) < 0$. Therefore, the concept of signature is not well-defined in this situation.

We can use these results from linear algebra to construct a very important class of symmetric spaces that generalizes hyperbolic space. Let us consider a non-degenerate symmetric bilinear form $\langle \cdot, \cdot \rangle$ of signature (p, q) on $\mathbb{R}^n$, then there exists a symmetric matrix $M \in GL(n, \mathbb{R})$ such that $\langle v, w \rangle = v^T M w$ for all $v, w \in \mathbb{R}^n$. The group of linear isometries of this bilinear form is

$$\begin{aligned} O(p, q) &= \{g \in GL(n, \mathbb{R}) : \langle gv, gw \rangle = \langle v, w \rangle \ \forall v, w \in \mathbb{R}^n\} \\ &= \{g \in GL(n, \mathbb{R}) : g^T M g = M\}. \end{aligned}$$

This group is never connected, having four connected components if $p, q > 0$ and two if $p = 0$ or $q = 0$, and we denote by $SO_0(p,q)$ its identity component. For the rest of this section we assume that $n = p + q \geq 3$ and $p, q > 0$. The latter condition ensures that $SO_0(p,q)$ is not compact. By choosing an orthonormal basis, we may identify $\langle \cdot, \cdot \rangle$ with the standard non-degenerate symmetric bilinear form of signature (p,q) where $M = I_{p,q}$. Elements of $O(p,q)$ can then be written as block-matrices and the condition $g^T I_{p,q} g = I_{p,q}$ is equivalent to the following.

$$g = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in O(p,q) \iff \begin{cases} A^T A - C^T C = I_p \\ B^T B - D^T D = -I_q \\ A^T B - C^T D = B^T A - D^T C = 0 \end{cases} \tag{3.16}$$

Moreover, it can be shown that such a matrix belongs to $SO_0(p,q)$ if and only if it further satisfies $\det(A) > 0$ and $\det(D) > 0$ (cf. [O'N83, p. 238]). Given a matrix in $O(p,q)$, its columns form an orthonormal basis of $\mathbb{R}^n$, so by reversing the signs of one or two of its column vectors - one of positive and one of negative length - this matrix can be mapped to any connected component.
Conjugation with $I_{p,q}$ maps $SO_0(p,q)$ to itself, so as in the case of Grassmannians (cf. Example 1.4.5), the map $g \mapsto I_{p,q} g I_{p,q}$ defines an involutive automorphism of $SO_0(p,q)$ whose fixed point-set can be identified with the compact subgroup $SO(p) \times SO(q)$ consisting of those block-matrices where $B = C = 0$. Hence, these groups form a symmetric pair and the quotient space

$$SO_0(p,q)/SO(p) \times SO(q)$$

can be equipped with an $SO_0(p,q)$-invariant Riemannian metric which turns it into a symmetric space of dimension pq. Similarly to the hyperbolic case where $p = n-1$ and $q = 1$, the induced symmetric decomposition on the Lie algebra level is given by

$$\mathfrak{so}(p,q) = (\mathfrak{so}(p) \times \mathfrak{so}(q)) \oplus \mathfrak{p}, \qquad \mathfrak{p} = \left\{ \begin{pmatrix} 0 & Z \\ Z^T & 0 \end{pmatrix} : Z \in M_{p,q}(\mathbb{R}) \right\}. \tag{3.17}$$

Again, $\mathfrak{so}(p,q)$ is simple with Killing form $B(X,Y) = (n-2)\operatorname{tr}(XY)$ which is positive definite on $\mathfrak{p} \times \mathfrak{p}$. The commutator of matrices in $\mathfrak{p}$ is easily seen to be

$$\left[\begin{pmatrix} 0 & Z_1 \\ Z_1^T & 0 \end{pmatrix}, \begin{pmatrix} 0 & Z_2 \\ Z_2^T & 0 \end{pmatrix} \right] = \begin{pmatrix} Z_1 Z_2^T - Z_2 Z_1^T & 0 \\ 0 & Z_1^T Z_2 - Z_2^T Z_1 \end{pmatrix},$$

which is zero if and only if $Z_1 Z_2^T$ and $Z_1^T Z_2$ are symmetric matrices. Hence, it follows that $SO_0(p,q)/SO(p) \times SO(q)$ is a symmetric space of the non-compact type and rank $\min(p,q)$. Using Theorem 2.3.2, the fact that $SO_0(p,q)$ is simple also implies that the identity component of the isometry group of this space has the form $SO_0(p,q)/N$, where N is a discrete, normal subgroup.
A matrix $g \in SO_0(p,q)$ necessarily satisfies $\det(g) = 1$, so we can view $SO_0(p,q)$ as a closed subgroup of $SL(n,\mathbb{R})$. Therefore, every manifold that is endowed with a smooth

action of $SL(n,\mathbb{R})$ also carries an action of $SO_0(p,q)$. An obvious example of such a space is the Grassmannian $\mathrm{Gr}(q,\mathbb{R}^n)$ of q-dimensional subspaces of $\mathbb{R}^n$, where $SL(n,\mathbb{R})$ acts transitively via the standard action $A \cdot V := A(V)$. Let $\mathbb{R}^q \subset \mathbb{R}^n$ be the subspace spanned by the last q vectors $\{e_{p+1},\dots,e_n\}$ of the standard basis, then the stabilizer of $\mathbb{R}^q$ under the action of $SL(n,\mathbb{R})$ is the group of all block-lower-triangular matrices

$$P := \left\{ \begin{pmatrix} A & 0 \\ C & D \end{pmatrix} \in SL(n,\mathbb{R}) : A \in GL(p,\mathbb{R}),\, D \in GL(q,\mathbb{R}),\, C \in M_{q,p}(\mathbb{R}) \right\}, \qquad (3.18)$$

which implies that there is a diffeomorphism $\mathrm{Gr}(q,\mathbb{R}^n) \cong SL(n,\mathbb{R})/P$. We could now restrict this action to $SO_0(p,q)$, but it will be more convenient to first consider the slightly larger group $SO(p,q) = \{g \in O(p,q) : \det(g) = 1\} \subset SL(n,\mathbb{R})$. To avoid one slightly anomalous case, we also assume that $p \neq q$ for the rest of this section.

Proposition 3.3.4. *Let $SO(p,q)$ act on $Gr(q,\mathbb{R}^n)$ by the standard action $(g,V) \mapsto g(V)$ and let V be any q-dimensional subspace of $\mathbb{R}^n$. Let k be the rank and (r,s) be the signature of the restriction of $\langle\cdot,\cdot\rangle$ to V. Then the orbit of V under the action of $SO(p,q)$ is the set of all $W \in Gr(q,\mathbb{R}^n)$ such that the restriction of $\langle\cdot,\cdot\rangle$ to W also has rank k and signature (r,s). Such an orbit is open in $Gr(q,\mathbb{R}^n)$ if and only if $k = q$.*

Proof. Every $g \in SO(p,q)$ preserves the bilinear form, so the restrictions of $\langle\cdot,\cdot\rangle$ to V and to $g(V)$ necessarily have the same rank and signature. To show the converse, let V and W be q-dimensional subspaces of rank $k \leq q$ and signature (r,s). Then we can choose bases $\{v_1,\dots,v_q,v_{q+1},\dots,v_n\}$ and $\{w_1,\dots,w_q,w_{q+1},\dots,w_n\}$ of $\mathbb{R}^n$ that are adapted to V and W as in Proposition 3.3.3. We define a linear map $g : \mathbb{R}^n \to \mathbb{R}^n$ by $g(v_i) = w_i$ for $i = 1,\dots,n$ which is a linear isometry by construction and satisfies $g(V) = W$. This is equivalent to $g \in O(p,q)$ and we may even assume that $g \in SO(p,q)$. In fact, if $(r,s) \neq (0,0)$ or $(\tilde{r},\tilde{s}) \neq (0,0)$, it suffices to possibly replace a vector of non-zero length in the basis adapted to V by its negative. The remaining case is only possible if $p = q = \nu$ which we explicitly exclude. Hence, V and W lie in the same orbit.
To prove the final claim, let us first recall the definition of the topology on the Grassmannian. Let $\mathcal{V}(q,\mathbb{R}^n) \subset (\mathbb{R}^n)^q \cong M_{n,q}(\mathbb{R})$ denote the open set of linearly independent q-tuples of vectors in $\mathbb{R}^n$, then the topology on $\mathrm{Gr}(q,\mathbb{R}^n)$ is the quotient topology induced by the map $\zeta : \mathcal{V}(q,\mathbb{R}^n) \to \mathrm{Gr}(q,\mathbb{R}^n)$ sending $(v_1,\dots,v_q)$ to their span. Given an open set $\mathcal{U} \subset \mathcal{V}(q,\mathbb{R}^n)$, it is easy to see that $\zeta^{-1}(\zeta(\mathcal{U})) = \bigcup_{A \in GL(q,\mathbb{R})} \mathcal{U} \cdot A$ is a union of open sets, showing that ζ is an open map. If $V \in \mathrm{Gr}(q,\mathbb{R}^n)$ has rank k and signature (r,s), then we again choose a basis $\{v_1,\dots,v_n\}$ of $\mathbb{R}^n$ that is adapted to V. The top left $(q \times q)$-block of the corresponding matrix (3.15) has an invertible $(k \times k)$-submatrix which is then invertible in an open neighbourhood $\mathcal{N} \subset \mathcal{V}(q,\mathbb{R}^n)$ of $(v_1,\dots,v_q)$. In addition, its eigenvalues depend continuously on the matrix entries, so this block has the same number (r,s) of positive and negative eigenvalues throughout $\mathcal{N}$. This implies that for $(w_1,\dots,w_q) \in \mathcal{N}$ the subspace $W = \zeta(w_1,\dots,w_q)$ has rank $k' \geq k$ and signature (r',s') with $r' \geq r$ and $s' \geq s$. If $k = q$ is maximal, then rank and signature cannot increase any further, so that $\zeta(\mathcal{N}) \subset \mathrm{Gr}(q,\mathbb{R}^n)$ is an open neighbourhood of V on which rank and signature are constant, which implies that the orbit of V is open in this situation.

If $k < q$, then $\nu = q - k > 0$ and $v_1 \in V \cap V^\perp$, but for every open neighbourhood $\mathcal{N}'$ of V there exists some $\varepsilon > 0$ such that $(v_1 + \varepsilon v_{n-\nu+1}, v_2, \dots, v_q) \in \zeta^{-1}(\mathcal{N}')$. Now we have $\langle v_1 + \varepsilon v_{n-\nu+1}, v_1 + \varepsilon v_{n-\nu+1} \rangle = 2\varepsilon > 0$ while the rest of the matrix representing the bilinear form on $V_\varepsilon = \zeta(v_1 + \varepsilon v_{n-\nu+1}, v_2, \dots, v_q)$ remains unchanged. Therefore, the rank of $V_\varepsilon \in \mathcal{N}'$ is greater than k, which shows that the orbit of V is not open. □

Remark 3.3.5. With slightly improved arguments it is possible to show that even the group $SO_0(p,q)$ acts transitively on the set of subspaces of a given rank and signature. To this end, it suffices to prove that for every subspace $W \in \mathrm{Gr}(q, \mathbb{R}^n)$ of rank k and signature (r,s), every connected component of $O(p,q)$ contains an element mapping W to itself. This is particularly easy in the case $k = q$ and $(r,s) = (0,q)$. A basis of $\mathbb{R}^n$ that is adapted to W is then (up to a permutation) the same as an orthonormal basis, so it suffices to reverse the signs of one or two basis vectors - one of positive and one of negative length - to adjust for the right determinants in the blocks of (3.16). A similar argument applies more generally in the situation where $\max(r,\tilde{r}) > 0$ and $\max(s,\tilde{s}) > 0$ after first extending a pair of vectors of positive and negative length (that are each either contained in W or $W^\perp$) to an orthonormal basis of $\mathbb{R}^n$. If $r = \tilde{r} = 0$ and $s > 0$ or $\tilde{s} > 0$, then on the one hand it is possible to reverse the sign of a basis vector of negative length in W or $W^\perp$. On the other hand, this case also forces $\nu = p > 0$, so one can consider the basis vector $v_1 \in W \cap W^\perp$ and its corresponding "dual" vector $v_{\mathrm{n}-\nu+1} \in V \setminus (W + W^\perp)$. These elements span a subspace of signature $(1,1)$, so changing their signs amounts to reversing a positive and a negative vector in an orthonormal basis. The same argument works if $s = \tilde{s} = 0$ and $r > 0$ or $\tilde{r} > 0$ and the remaining case $(r,s) = (\tilde{r},\tilde{s}) = (0,0)$ is excluded from our considerations. It is therefore possible to replace $SO(p,q)$ by $SO_0(p,q)$ in the discussion below, but since we only need this observation in the case $r = 0$ and $s = q$ and the notation will become easier when considering $SO(p,q)$, we mostly restrict our attention to this group.
For completeness, we also remark that $SO(p,q)$ clearly acts smoothly on $\mathrm{Gr}(d, \mathbb{R}^n)$ for any $d = 1, \dots, n$. The proof of Proposition 3.3.4 carries over word-for-word to this more general situation, so the orbits of $SO(p,q)$ and $SO_0(p,q)$ on $\mathrm{Gr}(d, \mathbb{R}^n)$ are always distinguished by rank and signature and open orbits correspond to non-degenerate subspaces. However, only the case $d = q$ will be relevant for our purposes.

The previous result shows that we may label the orbits of the $SO(p,q)$-action on $\mathrm{Gr}(q, \mathbb{R}^n)$ as $\mathcal{O}_{(r,s)}$, where (r,s) is the signature and $r + s = k \leq q$ the rank of the spaces in that orbit. Note that we must have $\nu = q - (r+s) \leq \min(p,q)$, which means that some of the orbits are empty if $p < q$. If $V \in \mathrm{Gr}(q, \mathbb{R}^n)$ is a non-degenerate subspace of signature (r,s), then the orbit $\mathcal{O}_{(r,s)} = SO(p,q) \cdot V$ is an open subset of the compact space $\mathrm{Gr}(q, \mathbb{R}^n)$, so the closure of this orbit is compact. Moreover, the orthogonal space $V^\perp$ is non-degenerate of signature $(\tilde{r},\tilde{s}) = (p-r, q-s)$. If $g \in SO(p,q)$ stabilizes V, then it also fixes $V^\perp$, which implies that $g \in S(O(r,s) \times O(p-r,q-s))$. Hence, we obtain a diffeomorphism

$$\mathcal{O}_{(r,s)} \cong SO(p,q)/S(O(r,s) \times O(p-r,q-s)). \tag{3.19}$$

In general, these spaces are not symmetric in the Riemannian sense, but they can be viewed as *pseudo-Riemannian symmetric spaces* which are defined by an immediate generalization of Definition 1.2.1 to that setting. However, we can consider the case $r = 0$ and $s = q$ where $\mathcal{O}_{(0,q)}$ consists of all maximal negative definite subspaces. As discussed in the previous remark, $SO_0(p, q)$ acts transitively on that orbit, so that $\mathcal{O}_{(0,q)}$ is diffeomorphic to the Riemannian symmetric space $SO_0(p, q)/SO(p) \times SO(q)$ and we obtain an open embedding

$$SO_0(p, q)/SO(p) \times SO(q) \hookrightarrow SL(n, \mathbb{R})/P \tag{3.20}$$

whose closure gives a homogeneous compactification of that symmetric space. In this situation, we can also easily characterize the boundary of this compactification as the set of all negative semi-definite subspaces, which follows immediately from the following more general result.

Corollary 3.3.6. *Let $\mathcal{O}_{(r,s)}$ be an orbit for the action of $SO(p, q)$ on $Gr(q, \mathbb{R}^n)$, then its closure is a finite disjoint union of orbits that is explicitly given by*

$$\overline{\mathcal{O}_{(r,s)}} = \bigcup_{\substack{r' \leq r \\ s' \leq s}} \mathcal{O}_{(r',s')}. \tag{3.21}$$

Proof. We have already seen in the proof of Proposition 3.3.4 that for every $V \in \mathrm{Gr}(q, \mathbb{R}^n)$ there exists an open neighbourhood of V on which rank and signature do not decrease. Therefore, every element in $\overline{\mathcal{O}_{(r,s)}}$ has to have signature (r', s') with $r' \leq r$ and $s' \leq s$. To show the reverse inclusion, let $V \in \mathcal{O}_{(r',s')}$ with $r' \leq r$ and $s' \leq s$ and choose a basis $\{v_1, \dots, v_n\}$ of $\mathbb{R}^n$ that is adapted to V as in (3.15). We set $\nu = q - (r + s)$, $\nu' = q - (r' + s')$ and define for every $m \geq 1$

$$\tilde{v}_i^m := \begin{cases} v_i + \frac{1}{2}(1 - \frac{1}{m})v_{n-\nu'+i} & 1 \leq i \leq r - r' \\ v_i - \frac{1}{2}(1 - \frac{1}{m})v_{n-\nu'+i} & r - r' + 1 \leq i \leq \nu' - \nu \\ v_i & \nu' - \nu + 1 \leq i \leq q \end{cases}$$

and $V_m := \zeta(\tilde{v}_1^m, \dots, \tilde{v}_q^m) \in \mathrm{Gr}(q, \mathbb{R}^n)$. Then we have

$$(\langle \tilde{v}_i^m, \tilde{v}_j^m \rangle)_{i,j=1}^q = \begin{pmatrix} (1 - \frac{1}{m})I_{r-r'} & 0 & 0 & 0 \\ 0 & -(1 - \frac{1}{m})I_{s-s'} & 0 & 0 \\ 0 & 0 & 0_\nu & 0 \\ 0 & 0 & 0 & I_{r',s'} \end{pmatrix},$$

which shows that $V_m \in \mathcal{O}_{(r,s)}$ for every $m \geq 1$. Since $(\tilde{v}_1^m, \dots, \tilde{v}_q^m)$ converges to $(v_1, \dots, v_q)$ as $m \to \infty$, it follows that $\lim_{m\to\infty} V_m = V$ and $V \in \overline{\mathcal{O}_{(r,s)}}$. □

In particular, the previous result shows that there is at most one closed (and hence compact) orbit, namely the set $\mathcal{O}_{(0,0)}$ of totally isotropic q-dimensional subspaces and this is contained in the closure of every other orbit. Furthermore, we have realized the

symmetric space $SO_0(p,q)/SO(p) \times SO(q)$ as the open orbit $\mathcal{O}_{(0,q)} \subset \mathrm{Gr}(q,\mathbb{R}^n)$ and its boundary as the union $\bigcup_{q'<q} \mathcal{O}_{(0,q')}$. Every orbit is invariant under the action of $SO_0(p,q)$, so it is obvious that the action of $SO_0(p,q)$ on $\mathcal{O}_{(0,q)}$ extends smoothly to the boundary. To conclude this section, we take a closer look at the individual boundary components of this compactification. In order to do this, we examine more generally the non-open orbits $\mathcal{O}_{(r,s)}$ where $r+s<q$.
Thus, we let $\mathcal{O} := \mathcal{O}_{(r,s)}$ be one such orbit and set $\nu = q-(r+s)$ as before. Hence, if $V \in \mathcal{O}$, then $V \cap V^\perp$ is a ν-dimensional subspace of $\mathbb{R}^n$ that is totally isotropic for $\langle\cdot,\cdot\rangle$. If we denote by $\mathrm{IGr}(\nu,\mathbb{R}^n)$ the set of all ν-dimensional totally isotropic subspaces of $\mathbb{R}^n$, then we can define a map

$$\begin{aligned} \kappa : \mathcal{O} &\to \mathrm{IGr}(\nu,\mathbb{R}^n) \\ V &\mapsto V \cap V^\perp. \end{aligned}$$

It follows as in Proposition 3.3.4 that $G := SO(p,q)$ acts transitively on this space, so the isotropic Grassmannian has a natural manifold structure as a homogeneous space of G, which clearly admits a smooth inclusion into the full Grassmannian $\mathrm{Gr}(\nu,\mathbb{R}^n)$. If $\{v_1,\dots,v_n\}$ is a basis of $\mathbb{R}^n$ that is adapted to V, then $\kappa(V) = \mathrm{span}\{v_1,\dots,v_\nu\}$, which shows that κ is smooth. Note that for every $g \in G$ we have $g(V)^\perp = g(V^\perp)$ and $g(V)\cap g(V)^\perp = g(V\cap V^\perp)$. In terms of actions, this can be stated as $\kappa(g\cdot V) = g\cdot\kappa(V)$, which proves that κ is G-equivariant and surjective. In fact, for every $V \in \mathcal{O}$ and every $N \in \mathrm{IGr}(\nu,\mathbb{R}^n)$ there exists some $g \in G$ with $g\cdot\kappa(V) = N$ by transitivity and therefore $\kappa(g\cdot V) = N$. Let H_1 and H_2 be the stabilizers of $V \in \mathcal{O}$ and $V\cap V^\perp \in \mathrm{IGr}(\nu,\mathbb{R}^n)$ under the respective G-actions, then we have diffeomorphisms $\mathcal{O} \cong G/H_1$, $\mathrm{IGr}(\nu,\mathbb{R}^n) \cong G/H_2$ and an inclusion $H_1 \subset H_2$ as a closed subgroup. This induces a canonical G-equivariant projection map $\tilde{\kappa} : G/H_1 \to G/H_2$ via the following commutative diagram.

$$\begin{array}{ccc} G & \xrightarrow{\quad\pi_2\quad} & G/H_2 \\ & \searrow^{\pi_1} \qquad \nearrow_{\tilde{\kappa}} & \\ & G/H_1 & \end{array}$$

Here, π_1 and π_2 are the usual projections and thus surjective submersions, so the same is true for the induced map $\tilde{\kappa}$.[4] Using the above diffeomorphisms, we can identify κ with a G-equivariant map between the quotient spaces which by construction maps eH_1 to eH_2. Since G acts transitively on G/H_1, every G-equivariant map on this space is uniquely determined by its value at a single point, which implies that κ - as a map between homogeneous spaces of G - coincides with $\tilde{\kappa}$. In particular, κ is a locally trivial fiber bundle and our next goal is to investigate its fibers.

[4] More precisely, the projection $\pi_2 : G \to G/H_2$ is an H_2-principal bundle and there is a natural left-action of H_2 on H_2/H_1. Therefore, we can form the associated fiber bundle $E := G \times_{H_2} (H_2/H_1)$ whose total space can be identified with G/H_1 via the map $[g, hH_1] \mapsto ghH_1$ (cf. [KN63, Chapter I, Proposition 5.5]). In this picture, the natural bundle projection $E \to G/H_2$ corresponds to the map $\tilde{\kappa} : G/H_1 \to G/H_2$ which is therefore a fiber bundle as well

To this end, let us take $V \in \mathcal{O}$ and consider the sum $V + V^\perp$, then $\langle \cdot, \cdot \rangle$ induces a non-degenerate symmetric bilinear form on the quotient $\underline{V} := (V + V^\perp)/(V \cap V^\perp)$. As discussed in the proof of Proposition 3.3.3, the dimension of $\underline{V}$ is $p + q - 2\nu$ and the signature of the induced form is $(p - \nu, q - \nu)$. The images of V and $V^\perp$ under the projection $V + V^\perp \to \underline{V}$ are complementary subspaces of $\underline{V}$ on which the induced form has signature (r, s) and $(p - r - \nu, q - s - \nu)$, respectively. Similarly, if $N \in \mathrm{IGr}(\nu, \mathbb{R}^n)$ is totally isotropic, then $N \subset N^\perp$, $\dim(N^\perp) = p + q - \nu$ and $\langle \cdot, \cdot \rangle$ induces a non-degenerate bilinear form on $\underline{N} := N^\perp/N$ with signature $(p - \nu, q - \nu)$. Thus, we can choose a subspace of dimension $r + s$ in $\underline{N}$ on which this form has signature (r, s). Its pre-image under the projection $N^\perp \to \underline{N}$ is then a space of dimension $r + s + \nu = q$ and signature (r, s), so it defines an element $W \in \mathcal{O}$ that satisfies $W \cap W^\perp = N$.
Summarizing our observations shows that the elements $V \in \mathcal{O}$ with $\kappa(V) = N$ are in bijective correspondence with non-degenerate subspaces of signature (r, s) in $N^\perp/N$, which is a vector space of dimension $p + q - 2\nu$ that is endowed with a non-degenerate symmetric bilinear form of signature $(p - \nu, q - \nu)$. Thus, there is a natural group action of $SO(p-\nu, q-\nu)$ on the fiber $\kappa^{-1}(N)$ which is transitive by an immediate adaptation of Proposition 3.3.4 to this setting. Being the fiber of a submersion, $\kappa^{-1}(N)$ has a natural manifold structure as a submanifold of $\mathcal{O}$. To see that the action of $SO(p - \nu, q - \nu)$ is smooth with respect to this structure, let us relate it to the smooth G-action on $\mathcal{O}$. Every element from the stabilizer H_2 of N under the G-action on $\mathrm{IGr}(\nu, \mathbb{R}^n)$ also stabilizes $N^\perp$. Thus, the action of $H_2 \subset G$ on $\mathcal{O} \cong G/H_1$ restricts to an action on those subspaces that are contained in $N^\perp$. This in turn descends to a smooth action on the non-degenerate subspaces of signature (r, s) contained in $N^\perp/N$, which is precisely the fiber $\kappa^{-1}(N)$. Under this identification, H_2 acts on $\kappa^{-1}(N)$ via elements of $SO(p-\nu, q-\nu)$. Explicitly, let us choose a basis $\{v_1, \ldots, v_n\}$ of $\mathbb{R}^n$ that is adapted to N and let $h \in H_2$ be arbitrary. Since h preserves the chain of subspaces $N \subset N^\perp \subset \mathbb{R}^n$, it follows that the matrix representation of h with respect to this basis is block-upper-triangular

$$h = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ 0 & A_{22} & A_{23} \\ 0 & 0 & A_{33} \end{pmatrix},$$

where the diagonal blocks are square-matrices of size $\nu, n - 2\nu$ and ν, respectively. The element of $SO(p-\nu, q-\nu)$ that is associated to h in the construction above is the central $(n-2\nu) \times (n-2\nu)$-block of that matrix. Conversely, given a matrix $B \in SO(p-\nu, q-\nu)$, its column vectors form an orthonormal basis of $N^\perp/N$ which can be extended to a basis of $\mathbb{R}^n$ that is adapted to N as in the proof of Proposition 3.3.3. Setting $A_{22} = B$, $A_{11} = A_{33} = I_\nu$ and the remaining matrices equal to zero then defines an element of $SO(p, q)$ that stabilizes N. Hence, this element is contained in H_2 and induces the action of B on $\kappa^{-1}(N)$. Altogether, this correspondence yields a surjective homomorphism $H_2 \to SO(p-\nu, q-\nu)$, so the action of $SO(p-\nu, q-\nu)$ on $\kappa^{-1}(N)$ is actually induced by the stabilizer $H_2 \subset G$ and therefore smooth. Consequently, the fibers are diffeomorphic to the homogeneous spaces

$$\kappa^{-1}(N) \cong SO(p - \nu, q - \nu)/S(O(r, s) \times O(p - r - \nu, q - s - \nu)). \tag{3.22}$$

This description gives rise to a nice structural hierarchy between the orbits. In the extreme case $r = s = 0$, the quotient (3.22) is trivial, which is obvious from the fact that κ is bijective in this case. For increasing r and s, the fiber bundle structure of κ implies that the orbit $\mathcal{O}$ is locally diffeomorphic to the product of an open set in the isotropic Grassmannian and a pseudo-Riemannian symmetric space. If $r + s = q$ is maximal, then we have already observed in the discussion after Proposition 3.3.4 that $\mathcal{O}$ is globally diffeomorphic to a symmetric space. In the case $r = 0$ and $s = q$ where the orbits $\mathcal{O}_{(0,q')}$, $q' < q$, constitute the boundary of the homogeneous compactification of the Riemannian symmetric space $SO_0(p, q)/SO(p) \times SO(q)$, we have $\nu = q - q'$, so it follows that every boundary component of this compactification is locally diffeomorphic to the product of an open set in an isotropic Grassmannian manifold and a Riemannian symmetric space which is of the form $SO_0(p - q + q', q')/SO(p - q + q') \times SO(q')$.

Remark 3.3.7. (i) The results of this section can in particular be applied to a non-degenerate symmetric bilinear form of signature $(n, 1)$ on $\mathbb{R}^{n+1}$. In this case, there is an action of $SO(n, 1)$ on $\mathrm{Gr}(1, \mathbb{R}^{n+1}) = \mathbb{R}P^n$ with two open orbits $\mathcal{O}_{(1,0)}, \mathcal{O}_{(0,1)}$ corresponding to the sets of spacelike and timelike lines and one closed orbit $\mathcal{O}_{(0,0)}$ consisting of all lightlike lines in $\mathbb{R}^{n+1}$. The set of timelike lines is diffeomorphic to the symmetric space $SO_0(n, 1)/SO(n) \cong H^n$, so we see that the homogeneous compactification of that space coincides with the projective compactification of hyperbolic space arising from the Klein model.

(ii) In this simple case, the homogeneous compactification of $SO_0(n, 1)/SO(n)$ is also homeomorphic to the geodesic compactification which is no longer true in general. Although this can be seen directly by looking at the boundary components, it is also a consequence of a deeper property. If $p, q > 1$, then $SO_0(p, q)/SO(p) \times SO(q)$ is a symmetric space whose rank is at least 2. It is shown in [Klo10] that if M is a symmetric space of the non-compact type with $\mathrm{rk}(M) > 1$, then the smooth action of $I_0(M)$ cannot be smoothly extended to the boundary of the geodesic compactification $\overline{M} = M \cup M(\infty)$. This is in sharp contrast to the case $\mathrm{rk}(M) = 1$ where we have seen that the isometric action in the geodesic compactification of H^n admits infinitely many different smooth extensions. As we have indicated in Remark 3.1.11, this is a consequence of the fact that the Tits building of a symmetric space of rank 1 is rather trivial compared to spaces of higher rank.

(iii) Throughout this section we focused on symmetric bilinear forms in real vector spaces, but a similar construction can be carried out over the complex numbers. In this setting, one considers a non-degenerate Hermitian form of signature (p, q) on $\mathbb{C}^n$ and the corresponding indefinite unitary group $U(p, q)$. Its subgroup $SU(p, q) = \{g \in U(p, q) : \det(g) = 1\}$ is a connected subgroup of the complex special linear group $SL(n, \mathbb{C})$. The complex Grassmannian $\mathrm{Gr}(q, \mathbb{C}^n)$ can be expressed as a homogeneous space $SL(n, \mathbb{C})/P$, where P is the complex analogue of the group in (3.18). The orbits of the action of $SU(p, q)$ on $\mathrm{Gr}(q, \mathbb{C}^n)$ are again distinguished by rank and signature, which follows as in the proof of Proposition 3.3.4 using the fact that adapted bases also exist for Hermitian forms. In this case, it is even easier to reduce transitivity of the action of $U(p,q)$ to $SU(p,q)$ since it suffices to multiply by diagonal matrices of the form $e^{i\theta} I_n \in U(p,q)$, $\theta \in \mathbb{R}$, which act trivially on $\mathrm{Gr}(q,\mathbb{R}^n)$ and requiring that $p \neq q$ is not necessary. In particular, the open subset of maximal negative definite subspaces of $\mathbb{C}^n$ can be identified with $SU(p,q)/S(U(p)\times U(q))$ and this yields a homogeneous compactification of that space whose boundary can be described exactly as in Corollary 3.3.6. Finally, $SU(p,q)/S(U(p)\times U(q))$ is a symmetric space of rank $\min(p,q)$ by the same reasoning as in the real case. Its dimension is $2pq$, so the case $n = 2$ and $p = q = 1$ is also admissible in this setting. We will consider the complex case in more detail in Section 3.5.

The advantage of this compactification method is that one has a clear understanding of the orbit structure and the boundary components of the compactified space. Moreover, it arises naturally from the inclusion $SO_0(p,q) \hookrightarrow SL(n,\mathbb{R})$ and generalizes the projective compactification of hyperbolic space. Finally, the number and structure of the non-empty boundary components increases with the rank of the symmetric space in question. This is a common phenomenon in many different methods of compactification, which again illustrates the importance of the rank as a characteristic invariant of a symmetric space.

3.4 The Baily-Borel compactification

Having studied concrete examples in the preceding sections, we now return to the abstract theory and describe an "extrinsic" compactification method that can be applied to the class of Hermitian symmetric spaces of the non-compact type. In contrast to the geodesic compactification, the construction will not build upon understanding the geometric properties of the space in question but on its algebraic structure. However, our main goal is to compare this method to the homogeneous compactification of $SU(p,q)/S(U(p)\times U(q))$ that we mentioned in the previous section. The construction for an arbitrary symmetric space requires a lot of machinery from the theory of Lie groups, which is why we only outline the general case and refer the reader to [Hel01, Chapter VIII, Section 7] and [Sat80, Chapter II, Section 4] for details. As we shall explain, Hermitian symmetric spaces of the non-compact type can be realized as the following objects.

Definition 3.4.1. A bounded domain $\Omega \subset \mathbb{C}^n$ in a complex vector space is called *symmetric* if for every point $p \in \Omega$ there exists an involutive holomorphic diffeomorphism $s_p : \Omega \to \Omega$ such that p is an isolated fixed point of s_p.

Throughout this section we assume that M is a Hermitian symmetric space of the non-compact type with base point $o \in M$ and we further suppose that M is irreducible. As usual, we let $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ be the symmetric decomposition of the Lie algebra $\mathfrak{g}$ of $I_0(M)$ induced by the involution $\sigma : g \mapsto s_o g s_o$ and we note that $I_0(M) = A_0(M)$ by Proposition 2.3.8. Then $\mathfrak{p} \cong T_oM$ and the Riemannian structure Q induces an $\mathrm{ad}(\mathfrak{h})$-invariant inner product $\langle \cdot,\cdot \rangle_{\mathfrak{p}}$ on $\mathfrak{p}$ which must be a scalar multiple of the restriction to $\mathfrak{p}$ of the Killing form of $\mathfrak{g}$. Moreover, the canonical almost complex structure J of M induces a linear map $\tilde{J} : \mathfrak{p} \to \mathfrak{p}$ that satisfies $\tilde{J}^2 = -\mathrm{id}_{\mathfrak{p}}$. In the proof of Theorem 2.3.9 we extended this map by zero to a derivation $\tilde{J} : \mathfrak{g} \to \mathfrak{g}$ and deduced that there exists an element $Z_0 \in \mathfrak{g}$, which is contained in the one-dimensional center of $\mathfrak{h}$, such that $\tilde{J} = \mathrm{ad}(Z_0)$. Let $\mathfrak{g}^{\mathbb{C}}$ denote the complexification of $\mathfrak{g}$ and set $\mathfrak{h}^{\mathbb{C}} := \mathfrak{h} \oplus i\mathfrak{h}$, $\mathfrak{p}^{\mathbb{C}} := \mathfrak{p} \oplus i\mathfrak{p}$. The $\mathbb{R}$-linear map $\mathrm{ad}(Z_0)$ on $\mathfrak{g}$ uniquely extends to a $\mathbb{C}$-linear map on $\mathfrak{g}^{\mathbb{C}}$ which maps $\mathfrak{p}^{\mathbb{C}}$ to itself and satisfies $(\mathrm{ad}(Z_0)|_{\mathfrak{p}^{\mathbb{C}}})^2 = -\mathrm{id}_{\mathfrak{p}^{\mathbb{C}}}$. Thus, we may split $\mathfrak{p}^{\mathbb{C}}$ into the eigenspaces $\mathfrak{p}_+$ and $\mathfrak{p}_-$ for the eigenvalues i and $-i$ of $\mathrm{ad}(Z_0)|_{\mathfrak{p}^{\mathbb{C}}}$ and obtain the decomposition $\mathfrak{g}^{\mathbb{C}} = \mathfrak{p}_+ \oplus \mathfrak{h}^{\mathbb{C}} \oplus \mathfrak{p}_-$, which satisfies

$$[\mathfrak{h}^{\mathbb{C}}, \mathfrak{p}_\pm] \subset \mathfrak{p}_\pm, \qquad [\mathfrak{p}_+, \mathfrak{p}_-] \subset \mathfrak{h}^{\mathbb{C}}, \qquad [\mathfrak{p}_+, \mathfrak{p}_+] = [\mathfrak{p}_-, \mathfrak{p}_-] = 0 \tag{3.23}$$

by (2.11) and (2.12). In particular, $\mathfrak{p}_+$ and $\mathfrak{p}_-$ are abelian subalgebras of $\mathfrak{g}^{\mathbb{C}}$. Let $(\mathfrak{g}^{\mathbb{C}})^{\mathbb{R}}$ denote the Lie algebra $\mathfrak{g}^{\mathbb{C}}$ viewed as a Lie algebra over $\mathbb{R}$. As discussed in Section 2.3.1, there exists a simply connected Lie group $G^{\mathbb{C}}$ whose Lie algebra is $(\mathfrak{g}^{\mathbb{C}})^{\mathbb{R}}$. Since the spaces $\mathfrak{g}, \mathfrak{h}, \mathfrak{h}^{\mathbb{C}}, \mathfrak{p}_+$ and $\mathfrak{p}_-$ are real Lie subalgebras of $(\mathfrak{g}^{\mathbb{C}})^{\mathbb{R}}$, there exist connected Lie subgroups $G, H, H^{\mathbb{C}}, P_+$ and P_- of $G^{\mathbb{C}}$ corresponding to these subalgebras. Note that G and $A_0(M)$ have the same Lie algebra $\mathfrak{g}$, but they do not necessarily coincide since $A_0(M)$ need not be contained in $G^{\mathbb{C}}$. Moreover, H has the same Lie algebra as the stabilizer of $o \in M$ under the action of $A_0(M)$. Nevertheless, it follows from the global Cartan decomposition in Theorem 2.2.1 that H is closed in G and that the manifold G/H is diffeomorphic to $\mathfrak{p}$. In fact, this result is an extension of Theorem 2.3.10 from which we deduced that M is diffeomorphic to $\mathfrak{p}$ as well, so we obtain a diffeomorphism $M \cong G/H$. The relations between these Lie groups form the heart of the construction, which we summarize in the following theorem. Its proof is given in [Hel01, Chapter VIII, Lemma 7.8, 7.9 and 7.10].[5]

Theorem 3.4.2. *The Lie groups $G, G^{\mathbb{C}}, H, H^{\mathbb{C}}, P_+$ and P_- from above have the following properties.*

(i) The exponential map of $G^{\mathbb{C}}$ restricts to a diffeomorphism $\mathfrak{p}_- \to P_-$ and $\mathfrak{p}_+ \to P_+$. Consequently, P_+ and P_- are simply connected.

(ii) The multiplication map

$$P_+ \times H^{\mathbb{C}} \times P_- \to G^{\mathbb{C}}, \qquad (p^+, h, p^-) \mapsto p^+ h p^-$$

is a diffeomorphism onto an open subset of $G^{\mathbb{C}}$ containing G.

(iii) The product $GH^{\mathbb{C}}P_-$ is open in $P_+H^{\mathbb{C}}P_-$ and we have $G \cap H^{\mathbb{C}}P_- = H$. Moreover, the group $H^{\mathbb{C}}P_-$ is closed in $G^{\mathbb{C}}$.

Note that part (ii) of this result in particular states that $P_+ \cap H^{\mathbb{C}}P_- = \{e\}$. We denote the inverses of the exponential maps of $P_\pm$ by $\log : P_\pm \to \mathfrak{p}_\pm$. The relations from the theorem can be collected in the following commutative diagram.

[5] The spaces $\mathfrak{p}_+$ and $\mathfrak{p}_-$ are constructed in a slightly different way in [Hel01, p. 312-313], but this is easily seen to coincide with our definition by comparing the root spaces used there with the eigenspaces of $\mathrm{ad}(Z_0) : \mathfrak{g}^{\mathbb{C}} \to \mathfrak{g}^{\mathbb{C}}$.

$$\begin{array}{ccccc} GH^{\mathbb{C}}P_-/H^{\mathbb{C}}P_- & \xrightarrow{\iota_1} & P_+H^{\mathbb{C}}P_-/H^{\mathbb{C}}P_- & \xrightarrow{\iota_2} & G^{\mathbb{C}}/H^{\mathbb{C}}P_- \\ \downarrow \psi_1 & & \downarrow \psi_2 & & \\ M \cong G/H & \xrightarrow{\psi} & P_+ & \xrightarrow{\log} & \mathfrak{p}_+ \end{array}$$

Here, the maps ι_1 and ι_2 are open embeddings, whereas ψ_1 and ψ_2 are diffeomorphisms. Therefore, $\psi := \psi_2 \circ \iota_1 \circ \psi_1^{-1}$ is an open embedding as well and $\Psi := \log \circ \psi$ is a diffeomorphism from G/H onto a simply connected, open subset of $\mathfrak{p}_+$ which we may identify with $\mathbb{C}^n$. Furthermore, Ψ is holomorphic or anti-holomorphic and by choosing a norm[6] it can be shown that the domain $\Psi(G/H) \subset \mathbb{C}^n$ is bounded (cf. [Hel01, Chapter VIII, Lemma 7.12]).
By construction, the map Ψ is given as follows: Consider a coset $gH \in G/H$ that is represented by an element $g \in G$, then there exist unique elements $p^{\pm} \in P_{\pm}$, $h \in H^{\mathbb{C}}$ such that $g = p^+hp^-$ and we have $\psi_1^{-1}(gH) = gH^{\mathbb{C}}P_- = p^+hp^-H^{\mathbb{C}}P_-$. The embedding ι_1 is induced by the inclusion $G \subset P_+H^{\mathbb{C}}P_-$ and $\psi_2(p^+hp^-H^{\mathbb{C}}P_-) = p^+$, so it follows that $\psi(gH) = p^+$ and $\Psi(gH) = \log(p_+)$. Equivalently, $\Psi(gH)$ is the unique element $X \in \mathfrak{p}_+$ such that $\exp(X)^{-1}g \in H^{\mathbb{C}}P_-$.

Definition 3.4.3. The domain $\Omega := \Psi(G/H) \subset \mathfrak{p}_+$ is called the *Harish-Chandra embedding* of the Hermitian symmetric space $M \cong G/H$. Its closure in $\mathfrak{p}_+$ is called the *Baily-Borel compactification* of M.

Remark 3.4.4. The almost complex structure J on M could also be replaced by $-J$ in which case the element $Z_0 \in \mathfrak{h}$ that realizes $\tilde{J}$ is replaced by $-Z_0$, which in turn amounts to interchanging $\mathfrak{p}_+$ and $\mathfrak{p}_-$. Every element in the center $\mathfrak{z}$ of $\mathfrak{h}$ is a real scalar multiple of Z_0, so $X = \pm Z_0$ are the only elements in $\mathfrak{z}$ that satisfy $(\mathrm{ad}(X)|_{\mathfrak{p}})^2 = -\mathrm{id}_{\mathfrak{p}}$. Conversely, starting from an irreducible symmetric space in the form $M = G/H$ where the center of H is one-dimensional, there are precisely two elements in $\mathfrak{z}$ with that property. Both of them give rise to almost complex structures J_1 and J_2 on G/H such that $J_1 = -J_2$.

The Harish-Chandra embedding realizes a Hermitian symmetric space M as a bounded domain Ω in a complex vector space which inherits all the properties of the symmetric space. In particular, there is a smooth action of $A_0(M)$ on Ω by biholomorphisms and the symmetries of M can be pushed forward to Ω which turns it into a bounded symmetric domain. Conversely, it can be shown that every bounded symmetric domain admits a Riemannian metric, called the *Bergman metric*, in which it becomes a Hermitian symmetric space of the non-compact type (cf. [Hel01, Chapter VIII, Section 3]).
The ambient space in the Baily-Borel compactification is a finite-dimensional vector space. On the one hand, this makes it possible to explicitly describe the resulting bounded symmetric domain as a set of matrices and we will construct one such example in the next section. On the other hand, the embedding is not satisfactory from a geometric point of view since the symmetric structure is not at all related to the linear structure of the surrounding vector space. Thus, it is usually not clear whether the action of $A_0(M)$ on its bounded symmetric domain can be extended to the boundary.

[6]Since $\mathbb{C}^n$ is finite-dimensional, all norms are equivalent.

We can go one step further in the above construction and obtain another compactification of M that remedies this issue (cf. [Hel01, Chapter VIII, Theorem 7.13]). The starting point is the observation that the dual $\mathfrak{g}^* = \mathfrak{h} \oplus i\mathfrak{p}$ of $\mathfrak{g}$ is a real subalgebra of the complexification $\mathfrak{g}^{\mathbb{C}}$ as well, so there exists a unique connected Lie subgroup $G^* \subset G^{\mathbb{C}}$ whose Lie algebra is $\mathfrak{g}^*$. The inclusion $\mathfrak{h} \subset \mathfrak{g}^*$ implies that H is a subgroup of G^*. It can be shown that G^* is simply connected, so that (G^*, H) is a symmetric pair by Theorem 2.3.5. Hence, the quotient space $M^* := G^*/H$ can be turned into a simply connected Hermitian symmetric space of the compact type which is the compact dual of $M \cong G/H$. In this setting, it turns out that there is a diffeomorphism

$$G^*/H \to G^{\mathbb{C}}/H^{\mathbb{C}}P_-, \qquad g^*H \mapsto g^*H^{\mathbb{C}}P_-, \tag{3.24}$$

which can be used in the commutative diagram above to realize M as an open subset of its compact dual M^*.

Definition 3.4.5. The open embedding of a Hermitian symmetric space M into its compact dual M^* constructed above is called the *Borel embedding* of M.

From a purely algebraic point of view, this is surprising since there is a priori no inclusion of G into G^* that would suggest such an embedding. The Borel embedding has the advantage that the ambient space is now a homogeneous space of a Lie group that is closely related to the isometry group of M. Moreover, G^*/H admits a global smooth G-action induced by the diffeomorphism $G^*/H \cong G^{\mathbb{C}}/H^{\mathbb{C}}P_-$ and the inclusion $G \hookrightarrow G^{\mathbb{C}}$. In particular, if G coincides with $A_0(M)$, then the isometry group of M also acts on the compact dual M^*.

3.5 The Borel embedding as a homogeneous compactification

In this final section we construct the Baily-Borel compactification of the symmetric space $SU(p,q)/S(U(p)\times U(q))$ explicitly and show that its Borel embedding coincides with the homogeneous compactification that we discussed in Section 3.3. Throughout this section we assume that $n = p+q \geq 2$ and $p, q > 0$. We write $G := SU(p,q)$, $H := S(U(p)\times U(q))$ and denote the Hermitian adjoint of a complex matrix $Z \in M_{p,q}(\mathbb{C})$ by $Z^* := \bar{Z}^T$.
To begin with, we let $\langle \cdot, \cdot \rangle$ be a non-degenerate Hermitian form of signature (p,q) on $\mathbb{C}^n$ and we may again assume that it is given by the standard form

$$\langle v, w \rangle = \sum_{i=1}^{p} v_i \bar{w}_i - \sum_{i=p+1}^{n} v_i \bar{w}_i, \qquad v, w \in \mathbb{C}^n.$$

As in the real case, we can then write matrices in $U(p,q)$ as block-matrices satisfying the following conditions.

$$g = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in U(p,q) \Longleftrightarrow \begin{cases} A^*A - C^*C = I_p \\ B^*B - D^*D = -I_q \\ A^*B - C^*D = B^*A - D^*C = 0 \end{cases} \tag{3.25}$$

By definition, such a matrix is contained in $SU(p,q)$ if and only if $\det(g) = 1$. As indicated in Remark 3.3.7, the space $SU(p,q)/S(U(p)\times U(q))$ can be described as a symmetric space of the non-compact type of dimension $2pq$ and rank $\min(p,q)$ exactly as in the real case. Moreover, we now see that it can be turned into a Hermitian symmetric space since the center of $S(U(p)\times U(q))$ is one-dimensional and consists of diagonal matrices of the form

$$\begin{pmatrix} e^{i\theta} I_p & 0 \\ 0 & e^{-i\frac{p}{q}\theta} I_q \end{pmatrix}, \qquad \theta \in \mathbb{R}.$$

The Lie algebra $\mathfrak{h}$ of $H = S(U(p)\times U(q))$ is given by

$$\mathfrak{h} = \left\{ \begin{pmatrix} X & 0 \\ 0 & Y \end{pmatrix} : X \in \mathfrak{u}(p),\, Y \in \mathfrak{u}(q),\, \operatorname{tr}(X) + \operatorname{tr}(Y) = 0 \right\},$$

where $\mathfrak{u}(p)$ denotes the set of all complex skew-Hermitian $(p\times p)$-matrices. Therefore, we obtain the symmetric decomposition $\mathfrak{g} = \mathfrak{su}(p,q) = \mathfrak{h} \oplus \mathfrak{p}$, where

$$\mathfrak{p} = \left\{ \begin{pmatrix} 0 & Z \\ Z^* & 0 \end{pmatrix} : Z \in M_{p,q}(\mathbb{C}) \right\}.$$

The element

$$Z_0 := i \begin{pmatrix} \frac{q}{p+q} I_p & 0 \\ 0 & -\frac{p}{p+q} I_q \end{pmatrix}$$

is contained in the center of $\mathfrak{h}$ and satisfies $\operatorname{ad}(Z_0)|_{\mathfrak{p}}^2 = -\operatorname{id}_{\mathfrak{p}}$ by a straightforward computation. Hence, it induces an almost complex structure on G/H and we may use Z_0 to construct the Baily-Borel compactification. To do this, we have to determine the complexifications of the spaces involved.

Every complex matrix can be written as the sum of a Hermitian and a skew-Hermitian matrix, which implies that the complexification of $\mathfrak{g} = \mathfrak{su}(p,q)$ is the Lie algebra $\mathfrak{sl}(n,\mathbb{C})$ of all complex trace-free matrices. The subspace $\mathfrak{p}^{\mathbb{C}} = \mathfrak{p} \oplus i\mathfrak{p}$ is then given by

$$\begin{aligned} \mathfrak{p}^{\mathbb{C}} &= \left\{ \begin{pmatrix} 0 & Z_1 + iZ_2 \\ Z_1^* + iZ_2^* & 0 \end{pmatrix} : Z_1, Z_2 \in M_{p,q}(\mathbb{C}) \right\} \\ &= \left\{ \begin{pmatrix} 0 & W_1 \\ W_2 & 0 \end{pmatrix} : W_1 \in M_{p,q}(\mathbb{C}),\, W_2 \in M_{q,p}(\mathbb{C}) \right\} \end{aligned}$$

and the eigenspaces $\mathfrak{p}_\pm \subset \mathfrak{p}^{\mathbb{C}}$ of $\operatorname{ad}(Z_0)$ for the eigenvalues $\pm i$ are

$$\mathfrak{p}_+ = \left\{ \begin{pmatrix} 0 & W \\ 0 & 0 \end{pmatrix} : W \in M_{p,q}(\mathbb{C}) \right\}, \qquad \mathfrak{p}_- = \left\{ \begin{pmatrix} 0 & 0 \\ W & 0 \end{pmatrix} : W \in M_{q,p}(\mathbb{C}) \right\}.$$

Consequently, the exponential images of these spaces have the form

$$P_+ = \left\{ \begin{pmatrix} I_p & W \\ 0 & I_q \end{pmatrix} : W \in M_{p,q}(\mathbb{C}) \right\}, \qquad P_- = \left\{ \begin{pmatrix} I_p & 0 \\ W & I_q \end{pmatrix} : W \in M_{q,p}(\mathbb{C}) \right\}$$

and we clearly see that $\exp : \mathfrak{p}_\pm \to P_\pm$ is a diffeomorphism. Finally, we also have

$$\mathfrak{h}^{\mathbb{C}} = \left\{ \begin{pmatrix} X & 0 \\ 0 & Y \end{pmatrix} : X \in M_{p,p}(\mathbb{C}),\, Y \in M_{q,q}(\mathbb{C}),\, \operatorname{tr}(X) + \operatorname{tr}(Y) = 0 \right\},$$
$$H^{\mathbb{C}} = \left\{ \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} : A \in GL(p,\mathbb{C}),\, B \in GL(q,\mathbb{C}),\, \det(A)\det(B) = 1 \right\}$$

and every matrix in $SU(p,q)$ uniquely decomposes into a product of three matrices[7]

$$SU(p,q) \ni g = \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} I_p & BD^{-1} \\ 0 & I_q \end{pmatrix} \begin{pmatrix} A - BD^{-1}C & 0 \\ 0 & D \end{pmatrix} \begin{pmatrix} I_p & 0 \\ D^{-1}C & I_q \end{pmatrix}, \quad (3.26)$$

where the factors belong to P_+, $H^{\mathbb{C}}$ and P_-, respectively. Let $o := eH$ be the base point of the Hermitian symmetric space G/H. Under the Harish-Chandra embedding $\Psi : G/H \to \mathfrak{p}_+$, a coset $gH \in G/H$ represented by a block-matrix g as above is mapped to the unique matrix $\Psi(gH) \in \mathfrak{p}_+$ such that $\exp(\Psi(gH))$ is the P_+-factor in the decomposition of g. Hence, we immediately obtain

$$\Psi(gH) = \begin{pmatrix} 0 & BD^{-1} \\ 0 & 0 \end{pmatrix} \in \mathfrak{p}_+,$$

which is easily seen to be smooth and independent of the representative of gH as it should be. We will identify $\mathfrak{p}_+$ with $M_{p,q}(\mathbb{C})$ and simply express the Harish-Chandra embedding as $\Psi(gH) = BD^{-1}$. Our general theory suggests that Ψ is a diffeomorphism onto a bounded, open set in $M_{p,q}(\mathbb{C})$ which we can explicitly describe in this situation.

Proposition 3.5.1. *The Harish-Chandra embedding of $SU(p,q)/S(U(p) \times U(q))$ is injective and its image is given by*

$$\mathcal{D}_{I_{p,q}} := \{Z \in M_{p,q}(\mathbb{C}) : I_q - Z^*Z \text{ is positive definite}\} \subset M_{p,q}(\mathbb{C}). \quad (3.27)$$

Proof. If $\Psi(g_1H) = \Psi(g_2H)$, then using the block-matrix notation from above we have $B_1D_1^{-1} = B_2D_2^{-1}$ and thus $B_1D_1^{-1}D_2 = B_2$. Applying the second relation of (3.25) twice shows that

$$\begin{aligned} -I_q = B_2^*B_2 - D_2^*D_2 &= D_2^*[(D_1^{-1})^*B_1^*B_1D_1^{-1} - I_q]D_2 \\ &= D_2^*[(D_1^{-1})^*(D_1^*D_1 - I_q)D_1^{-1} - I_q]D_2 \\ &= D_2^*[I_q - (D_1^{-1})^*D_1^{-1} - I_q]D_2 = -(D_1^{-1}D_2)^*(D_1^{-1}D_2), \end{aligned}$$

which is equivalent to $N := D_1^{-1}D_2 \in U(q)$. We then have $D_2 = D_1N$ and $B_2 = B_1N$, so unitarity of N implies that the unique matrix $h \in SU(p,q)$ satisfying $g_2 = g_1h$ has the block-form

$$h = \begin{pmatrix} M & 0 \\ 0 & N \end{pmatrix}, \qquad M \in U(p),\, N \in U(q),\, \det(M)\det(N) = 1.$$

[7] Note that $D^*D = I_q + B^*B$ by (3.25), so that D is positive definite and hence invertible.

Therefore, it is contained in $H = S(U(p) \times U(q))$, which shows that $g_1 H = g_2 H$ and that Ψ is injective. To characterize the image $\Psi(G/H)$, we make use of the fact that a Hermitian matrix $M \in GL(q, \mathbb{C})$ is positive definite if and only if there exists a matrix $N \in GL(q, \mathbb{C})$ such that $M = N^* N$. This matrix is unique up to unitary transformations and there is a unique choice of N that is positive definite as well, in which case we have $M = N^2$. Both assertions are simple consequences of the spectral theorem for Hermitian matrices (cf. [Lan02, Chapter XV, Section 6].
Using again the defining conditions in (3.25), we deduce that the matrix $\Psi(gH) = BD^{-1}$ satisfies

$$\begin{aligned} I_q - (BD^{-1})^*(BD^{-1}) &= I_q - (D^{-1})^* B^* B D^{-1} = I_q - (D^{-1})^*(D^* D - I_q) D^{-1} \\ &= I_q - (I_q - (D^{-1})^* D^{-1}) = (DD^*)^{-1}, \end{aligned}$$

which is positive definite since D is invertible. Conversely, let $Z \in M_{p,q}(\mathbb{C})$ be such that $I_q - Z^* Z$ is positive definite. Then the same is true[8] for $I_p - ZZ^*$ and their inverses, so there exist matrices $X \in GL(p, \mathbb{C})$, $Y \in GL(q, \mathbb{C})$ such that $X^* = X$, $X^2 = (I_p - ZZ^*)^{-1}$ and $Y^* = Y$, $Y^2 = (I_q - Z^* Z)^{-1}$. In this case, the block-matrix

$$g := \begin{pmatrix} X & ZY \\ Z^* X & Y \end{pmatrix}$$

is an element of $U(p, q)$, which follows directly by checking the conditions in (3.25). Moreover, we even have $g \in SU(p, q)$ since we can calculate the determinant using a splitting of g into a product as in (3.26).

$$\begin{aligned} \det(g) &= \det(X - (ZY)Y^{-1}(Z^* X)) \det(Y) \\ &= \det(X - ZZ^* X) \det(Y) = \det(I_p - ZZ^*) \det(X) \det(Y) \\ &= \frac{\det(I_p - ZZ^*)}{\sqrt{\det(I_p - ZZ^*) \det(I_q - Z^* Z)}} = 1 \end{aligned}$$

Here, the last step is justified as we have already noted that the matrices in the denominator have the same eigenvalues, only the multiplicities for the eigenvalue 1 may be different which is irrelevant for the determinant. Thus, g defines an element $gH \in G/H$ that satisfies $\Psi(gH) = Z$, which shows that $\Psi(G/H) = \mathcal{D}_{I_{p,q}}$. □

The condition that $I_q - Z^* Z$ is positive definite implies that all eigenvalues of $Z^* Z$ are smaller than 1. However, every such eigenvalue also has to be non-negative since $Z^* Z$ is positive semi-definite. Therefore, we have indeed realized G/H as a bounded domain in

[8] In fact, a Hermitian matrix is positive definite if and only if all its eigenvalues are positive. Now $Z^* Z$ and ZZ^* have the same non-zero eigenvalues because if $\lambda \neq 0$ is an eigenvalue for $Z^* Z$ with eigenvector x, then $Z^* Z x = \lambda x$, which implies $Zx \neq 0$ and $ZZ^*(Zx) = \lambda(Zx)$ and conversely. Hence, the eigenvalues of $Z^* Z$ are smaller than 1 if and only if the eigenvalues of ZZ^* have this property. Moreover, this observation also shows that the geometric multiplicities of ZZ^* and $Z^* Z$ for non-zero eigenvalues coincide and since these are Hermitian matrices, their algebraic multiplicities agree too.

a complex vector space. Moreover, it is evident that the closure of this domain, i.e. the Baily-Borel compactification of G/H, is

$$\overline{\mathcal{D}_{I_{p,q}}} = \{Z \in M_{p,q}(\mathbb{C}) : I_q - Z^*Z \text{ is positive semi-definite}\} \subset M_{p,q}(\mathbb{C}). \tag{3.28}$$

The left-action of G on G/H induces an action of G on the domain $\mathcal{D}_{I_{p,q}}$ which by definition of the diffeomorphism $\Psi : G/H \to \mathcal{D}_{I_{p,q}}$ arises in the following way: We have seen in the proof of Proposition 3.5.1 that every matrix $Z \in \mathcal{D}_{I_{p,q}} \subset M_{p,q}(\mathbb{C})$ can be identified with the representative

$$\begin{pmatrix} X & ZY \\ Z^*X & Y \end{pmatrix} \in SU(p,q)$$

from its pre-image in G/H on which a matrix $g \in G$ acts by left-multiplication. In terms of block-matrices, this takes the form

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} X & ZY \\ Z^*X & Y \end{pmatrix} = \begin{pmatrix} (A+BZ^*)X & (AZ+B)Y \\ (C+DZ^*)X & (CZ+D)Y \end{pmatrix}.$$

Applying again the diffeomorphism Ψ, it follows altogether that the conjugated action of G on $\mathcal{D}_{I_{p,q}}$ is given by *generalized Möbius transformations*

$$\begin{aligned} SU(p,q) \times \mathcal{D}_{I_{p,q}} &\to \mathcal{D}_{I_{p,q}} \\ \begin{pmatrix} A & B \\ C & D \end{pmatrix} \cdot Z &= (AZ+B)(CZ+D)^{-1}, \end{aligned} \tag{3.29}$$

which is isometric when $\mathcal{D}_{I_{p,q}}$ is endowed with the Riemannian metric induced by Ψ.

Remark 3.5.2. In the case $p = q = 1$ we obtain that $\mathcal{D}_{I_{1,1}} = \{z \in \mathbb{C} : |z|^2 < 1\} = B^2$ is the open unit disk on which $SU(1,1)$ acts by (usual) Möbius transformations as in the Poincaré ball model of hyperbolic geometry. Thus, the Riemannian metric on B^2 induced by the diffeomorphism with $SU(1,1)/S(U(1) \times U(1))$ is the Poincaré metric.

Having determined the Harish-Chandra embedding of G/H, let us also construct its Borel embedding into its compact dual. As we have seen, the complexification of the Lie algebra $\mathfrak{g} = \mathfrak{su}(p,q)$ is $\mathfrak{g}^{\mathbb{C}} = \mathfrak{sl}(n,\mathbb{C})$. Since the complex special linear group $SL(n,\mathbb{C})$ is simply connected (cf. [FH91, Proposition 23.1]) and indeed contains the connected groups $G, H, H^{\mathbb{C}}, P_+$ and P_- that we considered so far, we may use $G^{\mathbb{C}} := SL(n,\mathbb{C})$ in the construction of the Borel embedding. Moreover, the closed subgroup

$$H^{\mathbb{C}}P_- = \left\{ \begin{pmatrix} A & 0 \\ C & D \end{pmatrix} \in SL(n,\mathbb{C}) : A \in GL(p,\mathbb{C}),\ D \in GL(q,\mathbb{C}),\ C \in M_{q,p}(\mathbb{C}) \right\}$$

is the complex analogue of the group P in (3.18). Therefore, the space $G^{\mathbb{C}}/H^{\mathbb{C}}P_-$ coincides with the complex Grassmannian $\mathrm{Gr}(q,\mathbb{C}^n)$, viewed as the homogeneous space $SL(n,\mathbb{C})/P$, where P is the stabilizer of the maximal negative definite subspace $\mathbb{C}^q \subset \mathbb{C}^n$

that is spanned by the last q vectors of the standard basis. A coset $gP \in SL(n,\mathbb{C})/P$ is then identified with the subspace $g(\mathbb{C}^q) \subset \mathbb{C}^n$. Furthermore, the dual of $\mathfrak{g}$ is given by

$$\mathfrak{g}^* = \mathfrak{h} \oplus i\mathfrak{p} = \left\{ \begin{pmatrix} X & Z \\ -Z^* & Y \end{pmatrix} : X \in \mathfrak{u}(p),\ Y \in \mathfrak{u}(q),\ Z \in M_{p,q}(\mathbb{C}),\ \mathrm{tr}(X) + \mathrm{tr}(Y) = 0 \right\},$$

which obviously consists of all trace-free, skew-Hermitian $(n \times n)$-matrices. Thus, we have $\mathfrak{g}^* = \mathfrak{su}(n)$ and the corresponding connected Lie subgroup of $SL(n,\mathbb{C})$ is $G^* := SU(n)$. Hence, the compact dual of $G/H = SU(p,q)/S(U(p) \times U(q))$ is the space $G^*/H = SU(n)/S(U(p) \times U(q))$. As discussed in Example 1.4.5, $SU(n)$ also acts transitively on $\mathrm{Gr}(q,\mathbb{C}^n)$ and the stabilizer of the maximal negative definite subspace $\mathbb{C}^q \subset \mathbb{C}^n$ coincides with $S(U(p) \times U(q))$. In this example, the diffeomorphism

$$G^*/H = SU(n)/S(U(p) \times U(q)) \cong SL(n,\mathbb{C})/P = G^{\mathbb{C}}/H^{\mathbb{C}}P_- \tag{3.30}$$

that is used in the construction of the Borel embedding simply expresses the fact that every q-dimensional subspace of $\mathbb{C}^n$ possesses an orthonormal basis. Moreover, both P and $S(U(p) \times U(q))$ correspond to the stabilizer of the same space under the respective actions of $SL(n,\mathbb{C})$ and $SU(n)$ on $\mathrm{Gr}(q,\mathbb{C}^n)$, which implies that the diffeomorphism above reduces to the identity when viewed as a map $\mathrm{Gr}(q,\mathbb{C}^n) \to \mathrm{Gr}(q,\mathbb{C}^n)$. On the algebraic level, it identifies a coset $gP \in SL(n,\mathbb{C})/P$ with $uH \in G^*/H$, where $u \in SU(n)$ satisfies $u(\mathbb{C}^q) = g(\mathbb{C}^q)$.

Comparing this construction with the results from Section 3.3 now provides us with two embeddings of $SU(p,q)/S(U(p) \times U(q))$ into $\mathrm{Gr}(q,\mathbb{C}^n)$. On the one hand, there is a natural inclusion $SU(p,q) \hookrightarrow SL(n,\mathbb{C})$ and a corresponding left-action of $SU(p,q)$ on $SL(n,\mathbb{C})/P$. In the Grassmannian picture, this corresponds to the standard action and the coset eP is identified with the negative definite subspace $\mathbb{C}^q \subset \mathbb{C}^n$. The induced embedding $gH \mapsto g(\mathbb{C}^q)$ realizes G/H as the open set of maximal negative definite q-dimensional subspaces of $\mathbb{C}^n$ and the closure of this set is the homogeneous compactification that we discussed in Section 3.3. On the other hand, we have the Borel embedding where we view the Grassmannian as the compact dual $SU(n)/S(U(p) \times U(q))$. To determine the image and the action in this embedding, we first identify G/H with its Harish-Chandra embedding $\Psi(G/H) \subset \mathfrak{p}_+$. The entire space $\mathfrak{p}_+ \cong M_{p,q}(\mathbb{C})$ maps into $SL(n,\mathbb{C})/P \cong \mathrm{Gr}(q,\mathbb{C}^n)$ and by construction, this embedding is given by

$$M_{p,q}(\mathbb{C}) \ni Z \mapsto \begin{pmatrix} I_p & Z \\ 0 & I_q \end{pmatrix} P \mapsto \begin{pmatrix} I_p & Z \\ 0 & I_q \end{pmatrix} (\mathbb{C}^q) = \mathrm{span} \begin{pmatrix} Z \\ I_q \end{pmatrix} \in \mathrm{Gr}(q,\mathbb{C}^n).$$

The image of the restriction of this map to $\mathcal{D}_{I_{p,q}}$ clearly consists precisely of the q-dimensional negative definite subspaces as expected. Moreover, the action of $SU(p,q)$ on $\mathcal{D}_{I_{p,q}}$ by generalized Möbius transformations is equivalent to the usual left-action on $\mathrm{Gr}(q,\mathbb{C}^n)$ under the embedding since the corresponding subspaces coincide.

$$\mathrm{span} \begin{pmatrix} (AZ+B)(CZ+D)^{-1} \\ I_q \end{pmatrix} = \mathrm{span} \begin{pmatrix} AZ+B \\ CZ+D \end{pmatrix}$$

As mentioned before, passing from $SL(n,\mathbb{C})/P$ to $SU(n)/S(U(p) \times U(q))$ is the identity in the picture of Grassmannians, only the algebraic description of the G-action on $SU(n)/S(U(p) \times U(q))$ is cumbersome. Therefore, the image of $SU(p,q)/S(U(p) \times U(q))$ in its Borel embedding also consists of all maximal negative definite subspaces of $\mathbb{C}^n$ and the action of $SU(p,q)$ coincides with the usual left-action when $SU(n)/S(U(p) \times U(q))$ is identified with $\mathrm{Gr}(q,\mathbb{C}^n)$. Summarizing the discussion, we have proved the following.

Theorem 3.5.3. *The Borel embedding of $SU(p,q)/S(U(p) \times U(q))$ in its compact dual $SU(n)/S(U(p) \times U(q))$ coincides with its homogeneous compactification in $SL(n,\mathbb{C})/P$ in the sense that both embeddings agree when viewed as subsets of $Gr(q,\mathbb{C}^n)$ and carry the same action of $SU(p,q)$.*

This result concludes our survey about different compactification methods for symmetric spaces of the non-compact type. There are many other constructions from various areas of mathematics that could be discussed in this setting, but often require more powerful tools - especially from Lie group theory - that go beyond the goals of this thesis. Nevertheless, the elementary examples that we discussed in this chapter already display many key phenomena of the more general theory. For example, it often occurs that the number of boundary components increases with the rank of the symmetric space and that the components itself can be described as “smaller” symmetric spaces of a similar kind. Finally, our discussion also illustrates that tools from many different parts of mathematics have interesting applications in this framework.

Bibliography

[AO05] Jean-Philippe Anker and Bent Orsted, editors. *Lie theory: unitary representations and compactifications of symmetric spaces.* Birkhäuser, 2005.

[Bau14] Helga Baum. *Eichfeldtheorie: Eine Einführung in die Differentialgeometrie auf Faserbündeln.* Springer Berlin Heidelberg, 2014. 2., vollst. überarb. Aufl.

[BCO16] Jürgen Berndt, Sergio Console, and Carlos Enrique Olmos. *Submanifolds and holonomy.* Chapman and Hall/CRC, second edition, 2016.

[Ber02] Jürgen Berndt. Lie group actions on manifolds. Lecture Notes, 2002. `https://nms.kcl.ac.uk/juergen.berndt/sophia.pdf`, accessed: 07.05.2021.

[BH99] Martin R. Bridson and André Haefliger. *Metric spaces of non-positive curvature.* Springer, 1999.

[BJ06] Armand Borel and Lizhen Ji. *Compactifications of symmetric and locally symmetric spaces.* Birkhäuser, 2006.

[ČGH19] Andreas Čap, A. Rod Gover, and Matthias Hammerl. Parabolic compactification of homogeneous spaces. *Journal of the Institute of Mathematics of Jussieu*, pages 1–38, Oct 2019.

[Ebe96] Patrick B. Eberlein. *Geometry of nonpositively curved manifolds.* Univ. of Chicago Press, 1996.

[EO73] Patrick B. Eberlein and Barrett O'Neill. Visibility manifolds. *Pacific Journal of Mathematics*, 46(1):45–109, 1973.

[FH91] William Fulton and Joe Harris. *Representation theory: a first course.* Springer, 1991.

[Hel01] Sigurður Helgason. *Differential geometry and symmetric spaces.* American Math. Soc., 2001. Reprint. with corr.

[HJ03] Udo Hertrich-Jeromin. *Introduction to Möbius differential geometry.* Cambridge University Press, 2003.

[Klo06] Benoit Kloeckner. On differentiable compactifications of the hyperbolic space. *Transformation groups*, 11(2):185–194, 2006.

[Klo10] Benoît Kloeckner. Symmetric spaces of higher rank do not admit differentiable compactifications. *Mathematische Annalen*, 347(4):951–961, 2010.

[KN63] Shōshichi Kobayashi and Katsumi Nomizu. *Foundations of differential geometry: 1.* Wiley, 1963.

[KN69] Shōshichi Kobayashi and Katsumi Nomizu. *Foundations of differential geometry: 2.* Wiley, 1969.

[Kna96] Anthony W. Knapp. *Lie groups beyond an introduction.* Birkhäuser, 1996.

[Kun19] Michael Kunzinger. Lie groups. Lecture Notes, 2019. `https://www.mat.univie.ac.at/~mike/teaching/ws1920/lg.pdf`, accessed: 10.04.2021.

[Lan02] Serge Lang. *Algebra.* Springer, revised third edition, 2002.

[Lee97] John M. Lee. *Riemannian manifolds: an introduction to curvature.* Springer, 1997.

[Lou20] Brice Loustau. Hyperbolic geometry, 2020. `https://arxiv.org/pdf/2003.11180.pdf`, accessed: 10.04.2021.

[O'N66] Barrett O'Neill. The fundamental equations of a submersion. *The Michigan mathematical journal*, 13(4):459–469, 1966.

[O'N83] Barrett O'Neill. *Semi-Riemannian geometry: with applications to relativity.* Acad. Press, 1983.

[Sat80] Ichirō Satake. *Algebraic structures of symmetric domains.* Iwanami Shoten, Princeton University Press, 1980.