Essays on representations of real groups

Symmetric spaces of semi-simple groups

Bill Casselman
University of British Columbia
cass@math.ubc.ca

Suppose G to be the group of $\mathbb{R}$-rational points on a Zariski-connected, reductive, algebraic group defined over $\mathbb{R}$. It is determined through Galois descent by a Cartan involution θ whose fixed points make up a maximal compact subgroup K. This essay will be concerned with a classic theorem due originally to E. Cartan:

Theorem. *Every compact subgroup of G is contained in a conjugate of K.*

The proof follows very roughly the same lines that Cartan's did. The first half is a very general theorem about spaces which are, in some sense, of non-positive curvature. It says that any compact group acting on such a space possesses a fixed point. In the second, it is shown that the space G/K satisfies this curvature condition. The two together imply Cartan's theorem.

In Cartan's original proof, the fixed point theorem was about Riemannian spaces of negative curvature, but I'll use instead a simpler geometric notion due to [Bruhat-Tits:1972]. This allows an argument that is somewhat shorter and more direct than the standard one presented, for example, in [Helgason:1968]. The original application of the criterion of Bruhat-Tits was to buildings, as explained in [Bruhat-Tits:1972] and also the book [Brown:1989]. But Bruhat and Tits also observed that the criterion would apply to manifolds of negative curvature.

My treatment is somewhat similar to that in [Lang:1999], but differs substantially in the second part, where Lang follows closely [Mostow:1953]. The argument I'll present seems to be somewhat simpler than anything in the literature. The proof reduces to the easy case of $\mathrm{SL}_2(\mathbb{R})$. I'll show first how this works in terms of non-Euclidean geometry, then again in terms of root systems.

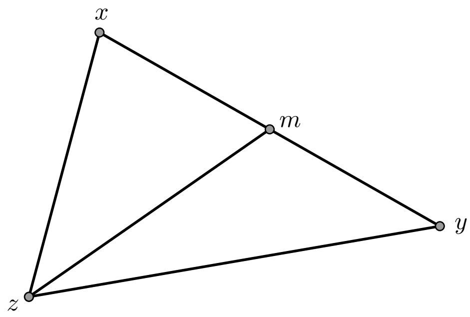
Contents

1. The fixed point theorem of Bruhat-Tits

Suppose X to be any complete metric space with distance function $\overline{xy}$. I'll call it **semi-hyperbolic** if it satisfies the criterion of §3.2 of [Bruhat-Tits:1972]:

> *Whenever one is given two points x and y there exists a third point m (for* **midpoint***) such that for all z in X*

$$2\,\overline{zm}^2 + \frac{\overline{xy}^2}{2} \le \overline{zx}^2 + \overline{zy}^2\,.$$



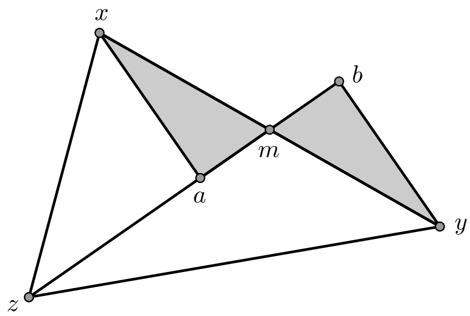
images/parallelogram.eps

This definition is motivated to some extent by an elementary theorem in Euclidean geometry, which asserts equality in that case:

[pappus] **1.1. Proposition.** *Given two points x and y in the Euclidean plane, let m be the midpoint of the segment xy. Then*

$$2\,\overline{zm}^2 + \frac{\overline{xy}^2}{2} = \overline{zx}^2 + \overline{zy}^2\,.$$

Proof. This is a straightforward consequence of the cosine formula for triangles or, equivalently, some simple vector dot-product calculations. Indeed, Bruhat and Tits make a point of referring in their article to Pappus of Alexandria (Book VII, Proposition 122 of *The Collection*), who apparently first proved the cosine formula. But there is a more direct proof, indicated by the figure below:



images/pappusx.eps

It tells us that

$$\begin{aligned}
\overline{zx}^2 &= \overline{za}^2 + \overline{ax}^2 \\
&= (\overline{zm} - \overline{am})^2 + \overline{ax}^2 \\
&= \overline{zm}^2 + \overline{xm}^2 - \overline{zm}\cdot\overline{am} \\
\overline{zy}^2 &= \overline{zb}^2 + \overline{by}^2 \\
&= (\overline{zm} + \overline{mb})^2 + \overline{by}^2 \\
&= \overline{zm}^2 + \overline{ym}^2 + \overline{zm}\cdot\overline{mb} \\
\overline{zx}^2 + \overline{zy}^2 &= 2\,\overline{zm}^2 + \overline{xm}^2 + \overline{my}^2 \\
&= 2\,\overline{zm}^2 + (\overline{xy}^2/2)\,. \qquad \square
\end{aligned}$$

One can rewrite the semi-hyperbolic inequality as

$$\frac{\overline{xy}^2}{2} \le (\overline{zx}^2 - \overline{zm}^2) + (\overline{zy}^2 - \overline{zm}^2)\,.$$

The important consequence is that if the differences between $\overline{zx}$ and $\overline{zm}$ and between $\overline{zy}$ and $\overline{zm}$ are small, then the distance $\overline{xy}$ is small as well. This is not true for geometry on the sphere, as one can see easily—points on the equator all have the same distance from the north pole, but may be quite far from each other.

What I call a *semi-hyperbolic space* is called in the current literature a $CAT(0)$ space, satisfying one of a hierarchy of curvature conditions $CAT(\kappa)$, with κ being a rough measure of curvature, so that $CAT(-1)$ means strictly negative curvature. This hierarchy was introduced in [Gromov:1987] to unite many different approaches to similar problems, and is an active subject of research.

[unique] **1.2. Proposition.** *In the criterion for a semi-hyperbolic space, the point m lies midway between x and y, and is unique.*

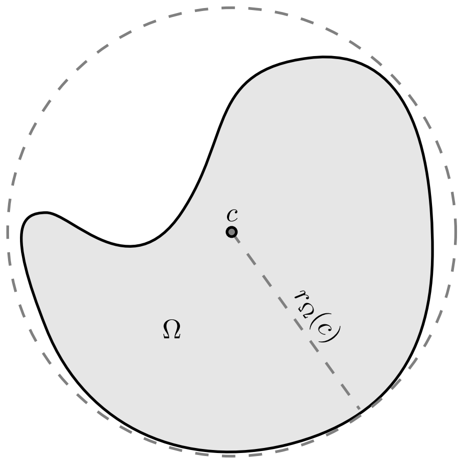
Proof. Letting $z = x$ leads to $\overline{xm}^2 \le \overline{xy}^2/4$, and letting $z = y$ to $\overline{ym}^2 \le \overline{xy}^2/4$. The triangle inequality then gives $\overline{xm} = \overline{ym} = \overline{xy}/2$. If m_* is a second point satisfying the criterion, then setting $z = m_*$ gives $\overline{mm_*} = 0$. ▯

For a given x, y let $m_{x,y}$ be this unique point, their midpoint.

In correspondence with Ken Brown, Serre proposed an elegant simplification of the argument of Bruhat-Tits, and it is an amplification of this that can be found in [Brown:1989]. I follow Brown's argument. If X is a metric space, the **disk** in X with centre c and radius r is the region

$$D_c(r) = \{x \,|\, \overline{cx} \le r\}\,.$$

Suppose Ω to be a bounded subset of a metric space X. For any c, define $r_\Omega(c)$ to be the least upper bound of the distances $\overline{cx}$ for x in Ω. Thus if $r \ge r_\Omega(c)$ then $\Omega \subseteq D_c(r)$, but if $r < r_\Omega(c)$ there exist points of Ω outside $D_c(r)$.



images/serre.eps

The **radius** r_Ω of Ω is the greatest lower bound of the radii of disks containing it, or equivalently the greatest lower bound of all $r_\Omega(c)$. This does not immediately imply that there exists a point c with $\Omega \subseteq D_c(r_\Omega)$, but if it does exists it will be called a **circumcentre** of Ω.

[circumcentre] **1.3. Proposition.** (Serre) *In a complete semi-hyperbolic space, every bounded subset possesses a unique circumcentre.*

Proof. Let Ω be a bounded subset of X. Suppose x and y to be any points in X, with $m = m_{x,y}$. Then

$$\Omega \subseteq D_x\big(r_\Omega(x)\big) \cap D_y\big(r_\Omega(y)\big)\,.$$

By the semi-hyperbolic inequality

$$\overline{mz}^2 \le \frac{\overline{xz}^2 + \overline{yz}^2}{2} - \frac{\overline{xy}^2}{4} \le \frac{r_\Omega^2(x) + r_\Omega^2(y)}{2} - \frac{\overline{xy}^2}{4}$$

for all z in Ω. But then

$$\begin{aligned} r_\Omega^2(m) &\le \frac{r_\Omega^2(x) + r_\Omega^2(y)}{2} - \frac{\overline{xy}^2}{4} \\ \overline{xy}^2 &\le 2\big[\big(r_\Omega^2(x) - r_\Omega^2(m)\big) + \big(r_\Omega^2(y) - r_\Omega^2(m)\big)\big] \\ &\le 2\big[\big(r_\Omega^2(x) - r_\Omega^2\big) + \big(r_\Omega^2(y) - r_\Omega^2\big)\big] \,. \end{aligned}$$

We can find a sequence of centres c_i of radii r_i such that (a) the limit of the r_i is $r(\Omega)$ and (b) $\Omega \subseteq D_{c_i}(r_i)$ for each i. The inequality above with $x = c_i$, $y = c_j$ implies this to be a Cauchy sequence, so there exists a limit c, and one can verify that $\Omega \subseteq D_c(r)$.

The same inequality immediately implies uniqueness. ∎

[fixed-serre] **1.4. Corollary.** *Any compact group of isometries of a complete semi-hyperbolic space possesses a fixed point.*

Proof. The circumcentre of any orbit will be fixed. ∎

2. Non-Euclidean geometry

In this section let $G = \mathrm{SL}_2(\mathbb{R})$, $K = \mathrm{SO}(2)$. I'll show here that G/K is semi-hyperbolic, thus proving Cartan's theorem in this case. I'll do this geometrically, interpreting the group G as a group of non-Euclidean isometries. This is best done in terms of the realization of G/K as the space $\mathfrak{X}$ of positive definite, symmetric, 2×2 matrices of determinant 1. The group $\mathrm{GL}_2(\mathbb{R})$ acts on the space of all 2×2 symmetric matrices according to the specification

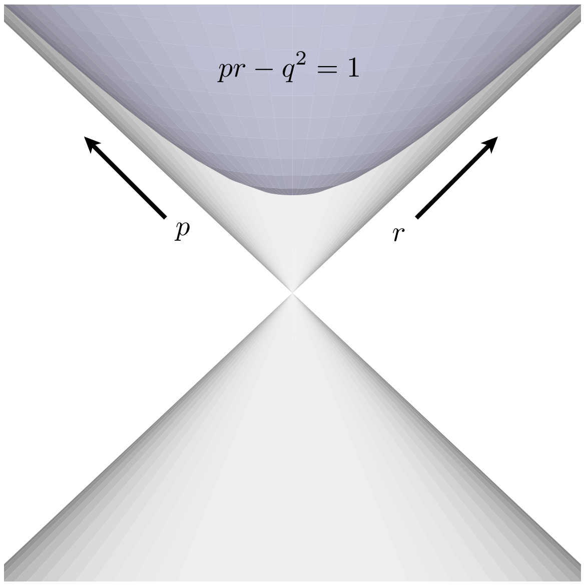
$$x \longmapsto g\, x\, {}^t g \,,$$

with $\mathrm{O}(2)$ the isotropy subgroup of the identity matrix I. The space $\mathfrak{X}$ is the $\mathrm{SL}_2(\mathbb{R})$-orbit of I.

This set can be pictured in three dimensions, since the space of symmetric matrices has dimension three. Choose coordinates

$$\begin{bmatrix} p & q \\ q & r \end{bmatrix} .$$

The matrices here with determinant 0 make up the homogeneous cone $pr - q^2 = 0$; the positive definite ones are those where $pr - q^2 > 0$ and both $p > 0$, $r > 0$; the negative definite ones are where $pr - q^2 > 0$, $p < 0$, $r < 0$; the indefinite ones are where $pr - q^2 < 0$. The positive definite ones of determinant 1 make one sheet of the two-sheeted hyperboloid $\mathfrak{X}$ where $pr - q^2 = 1$, that where $p > 0$.



images/forms.eps

Since K is the isotropy subgroup of I, a G-invariant Riemannian metric on $\mathfrak{X}$ is determined uniquely by a K-invariant metric on the tangent plane at I. The full three-dimensional tangent space at I may be identified with the vector space itself. We choose on the space of symmetric 2×2 matrices the metric

$$\left(\frac{1}{2}\right)\mathrm{trace}(X^2) = \frac{p^2 + r^2}{2} + q^2 \,.$$

The basis

$$\sqrt{2}\,\frac{\partial}{\partial p}, \quad \sqrt{2}\,\frac{\partial}{\partial r}, \quad \frac{\partial}{\partial q}$$

is orthonormal at I, and

$$\frac{\partial}{\partial p} - \frac{\partial}{\partial r}, \quad \frac{\partial}{\partial q}$$

is an orthonormal basis for the tangent space of $\mathfrak{X}$ at I.

The group K acts by rotating the hyperboloid around the line $p = r$, $q = 0$. The geodesics passing through I are the intersections of $\mathfrak{X}$ with the planes passing through this line. Because the action of G is linear, the other geodesics are the intersections of $\mathfrak{X}$ with the planes passing through the origin that intersect the interior of the cone $pr - q^2 > 0$.

Slice the cone by a plane giving a circular section. Projection of the cone's interior onto the inside of this disk is a bijection. This gives the **Klein model** of non-Euclidean geometry, in which the geodesics are straight line segments inside the unit disk.

In particular, the orbit of I with respect to the group of diagonal matrices

$$a_t = \begin{bmatrix} e^{t/2} & \circ \\ \circ & e^{-t/2} \end{bmatrix}$$

is a geodesic. The non-Euclidean distance from I to $A_t = a_t \cdot I = (e^t, 0, e^{-t})$ is t. The differential of a_t takes the vector $\partial/\partial b$ to itself. Therefore the length of the circumference of the non-Euclidean circle at A_t is just its Euclidean circumference, which is 2π times the radial distance from the centre line $p = r$, $q = 0$. This radial distance is the same as half the distance from A_t to A_{-t}, or $(1/2)|e^t - e^{-t}| = \big|\sinh(t)\big|$. The non-Euclidean circumference of the circle of radius $t > 0$ is therefore $2\pi\sinh(t)$. As $t \to 0$ this is asymptotically $2\pi t$, as it should be.

This turns out to be the crucial point in proving that $\mathfrak{X}$ is semi-hyperbolic. If v is a tangent vector at a point x on $\mathfrak{X}$, there exists a unique geodesic starting out from x in the direction of v. The exponential map $\exp_x$ at x takes v to the point at distance $\|v\|$ along that geodesic. The exponential map $\exp_x$ is a bijection of T_x with $\mathfrak{X}$ for each x, since G acts transitively on $\mathfrak{X}$, and at the point $x = (1, 0, 1)$ this is clear. The differential-geometric exponential map defined here agrees with the matrix exponential map

$$I + X + \frac{X^2}{2!} + \frac{X^3}{3!} + \cdots$$

on the subspace of symmetric matrices of trace 0 at I. Let $\log_x$ be the map inverse to $\exp_x$.

A linear map from one Euclidean vector space to another is expanding if lengths are not decreased. A map φ from one Riemannian manifold to another is **locally expanding** if its derivative at any point is expanding. It follows immediately from the definition that a locally expanding map is expanding in the non-local sense that if $\gamma(t)$ is a path in the first manifold than the image path $\varphi\big(\gamma(t)\big)$ has length at least that of $\gamma(t)$.

[expanding-exp] **2.1. Proposition.** *Suppose M to be a Riemannian manifold such that (a) the exponential map at any x on M is a diffeomorphism of T_x with M; (b) the exponential map at any point is locally expanding. Then M is semi-hyperbolic.*

Proof. Given x and y in M, there exists a unique point m half-way on the geodesic from x to y. There exists a unique X in T_m such that $x = \exp(X)$; let $Y = -X$, so that $y = \exp(Y)$. Given z on M, let Z in T_m be such that $\exp(Z) = z$.

We have

$$\begin{aligned} \overline{xm} &= \|X\| \\ \overline{ym} &= \|Y\| \\ \overline{zm} &= \|Z\| \end{aligned}$$

since the appropriate paths are geodesics through m. Let γ be the geodesic segment from z to x. Since the exponential map is expanding, its inverse shrinks, and so

$$\overline{zx} = \overline{\gamma} \ge \overline{\log(\gamma)} \ge \overline{ZX}\,.$$

The last is because the shortest path in T_m is a straight line segment. Similarly for z, y. Therefore, because of the parallelogram equality in Euclidean space

$$\overline{zm}^2 = \|Z\|^2 = \frac{\overline{ZX}^2 + \overline{ZY}^2}{2} - \|X\|^2 \le \frac{\overline{zx}^2 + \overline{zy}^2}{2} - \frac{\overline{xy}^2}{4}\,. \qquad \blacksquare$$

It now remains to show that the exponential map at $(1, 0, 1)$ is expanding. But it follows from the calculation of the non-Euclidean circumference that at a point on the tangent plane at radius r the derivative of the exponential map has matrix

$$\begin{bmatrix} 1 & \circ \\ \circ & \sinh(r)/r \end{bmatrix},$$

and $|\sinh(r)/r| \ge 1$. $\blacksquare$

3. The Riemannian symmetric space G/K

We now take up the second part of the proof of the initial Theorem for an arbitrary semi-simple group G. Like all other proofs I am aware of, it shows that the exponential maps on G/K are locally expanding, and thus that G/K is semi-hyperbolic.

First we describe a particular model for G/K. Recall that θ is the Cartan involution defining G, with K equal to the group of its fixed points. Let $\mathfrak{g}$ be the Lie algebra of G. Then

$$\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{s}$$

where $\mathfrak{k}$ is the $+1$-eigenspace of θ (also the Lie algebra of K) and $\mathfrak{s}$ is the -1-eigenspace. Restricted to $\mathfrak{s}$, the exponential map is an isomorphism with its image $\mathfrak{S}$, which is precisely the set of g with $g^\theta = g^{-1}$. The group G acts on $\mathfrak{S}$ according to the formula

$$s \longmapsto g\, s\, g^{-\theta} \quad \big(g^{-\theta} = \theta(g^{-1})\big)\,.$$

For example, if $G = \mathrm{SL}_n(\mathbb{R})$ then we may choose θ to be the map $g \mapsto {}^t g^{-1}$ and $\mathfrak{s}$ to be the space of symmetric matrices of trace 0, so that $\mathfrak{S}$ is the set of all positive definite $n \times n$ matrices.

The group G acts transitively—in fact, if $s = \exp(X)$ then $\exp(X/2)$ takes I to $\exp(X/2)\cdot I\cdot\exp(X/2) = s$. The isotropy subgroup of I is K, so that the space $\mathfrak{S}$ may be identified with G/K. In fact, the group G is the direct product $K \times \mathfrak{S}$. The Killing form induces a K-invariant positive definite Euclidean metric on $\mathfrak{s}$, which in turn induces a G-invariant Riemannian metric on $\mathfrak{S}$. The space $\mathfrak{s}$ contains the Lie algebra $\mathfrak{a}$ of a maximal split Cartan subgroup A. The connected component $|A|$ is $\exp(\mathfrak{a})$. The space $\mathfrak{S}$ is the K-transform of $|A|$, as $\mathfrak{s}$ is of $\mathfrak{a}$. The subspace $\mathfrak{s}$ is spanned by $\mathfrak{a}$ and the elements $e_\lambda + e_\lambda^\theta$, and $\mathfrak{k}$ is spanned by the $e_\lambda - e_\lambda^\theta$, as λ varies over the roots of $\mathfrak{g}$ with respect to A and e_λ varies over the root space $\mathfrak{g}_\lambda$.

I shall prove:

[exp-exp] **3.1. Lemma.** *The exponential map from $\mathfrak{s}$ to $\mathfrak{S}$ is locally expanding.*

Proof. Given X and Y in $\mathfrak{s}$, we let $\gamma(t)$ be the path $X + tY$, and want to compute $\|d\exp\big(\gamma(t)\big)/dt\|$ at $t = 0$. We identify the tangent space at $\exp(X)$ with that at I by applying $\exp(-X/2)$ on left and right. In other words, we define

$$\Phi_X(Y) = (d/dt)\exp(-X/2)\exp\big(\gamma(t)\big)\exp(-X/2)$$

at $t = 0$, and want to show that $\|\Phi_X(Y)\| \ge \|Y\|$.

Since $\mathfrak{s}$ is the union of the $k\,\mathfrak{a}\,k^{-1}$ for k in K, it suffices to show this for X in $\mathfrak{a}$. If Y lies in the centralizer of $\mathfrak{a}$, it is immediate that $\Phi_X(Y) = Y$. So we may assume that $Y = e_\lambda + e_\lambda^\theta$. If $\langle\lambda, X\rangle = 0$ then again $\Phi_X(Y) = Y$. Otherwise, we may assume that X, e_λ, and e_λ^θ span a Lie algebra isomorphic to $\mathfrak{sl}_2$. The path $\exp\big(\gamma(t)\big)$ is in a copy of $\mathrm{SL}_2(\mathbb{R})/\mathrm{SO}(2)$ embedded in $\mathfrak{S}$. It suffices now to prove the assertion in the case $G = \mathrm{SL}_2$. We have already dealt with that case, but I'll give here a different argument.

Let

$$X = \begin{bmatrix} \tau/2 & \circ \\ \circ & -\tau/2 \end{bmatrix}$$

$$Y = \begin{bmatrix} \circ & 1 \\ 1 & \circ \end{bmatrix}$$

We must choose a path $\gamma(t)$ in $\mathfrak{s}$ with $\gamma(0) = X$, $\gamma'(0) = Y$ and then calculate

$$\Phi_X(Y) = \left[\frac{d}{dt}\,\exp(-X/2)\,\exp\big(\gamma(t)\big)\exp(-X/2)\right]_{t=0}\,.$$

To define $\gamma(t)$, set

$$\begin{aligned} k(t) &= \begin{bmatrix} \cos(t/\tau) & -\sin(t/\tau) \\ \sin(t/\tau) & \cos(t/\tau) \end{bmatrix} \\ \gamma(t) &= k(t)X\,k(t)^{-1} \\ &= k(t)Xk(-t) \end{aligned}$$

since $k(t)^{-1} = k(-t)$. To check our hypotheses:

$$\begin{aligned} k'(0) &= \begin{bmatrix} \circ & -1/\tau \\ 1/\tau & \circ \end{bmatrix} \\ \gamma'(0) &= k'(0)X - Xk'(0) \\ &= \begin{bmatrix} \circ & 1 \\ 1 & \circ \end{bmatrix} . \end{aligned}$$

Now

$$\begin{aligned} \exp\big(\gamma(t)\big) &= k(t) \begin{bmatrix} e^{\tau/2} & \circ \\ \circ & e^{-\tau/2} \end{bmatrix} k(t) \\ [\exp\big(\gamma(t)\big)]' &= k'(t) \begin{bmatrix} e^{\tau/2} & \circ \\ \circ & e^{-\tau/2} \end{bmatrix} k(t) - k(t) \begin{bmatrix} e^{\tau/2} & \circ \\ \circ & e^{-\tau/2} \end{bmatrix} k'(t) \\ [\exp\big(\gamma(t)\big)]'_{t=0} &= \begin{bmatrix} \circ & \sinh(\tau)/\tau \\ \sinh(\tau)/\tau & \circ \end{bmatrix} \\ &= \frac{\sinh(\tau)}{\tau} \cdot Y . \end{aligned}$$

Since

$$\exp(-X/2) \begin{bmatrix} \circ & 1 \\ 1 & \circ \end{bmatrix} \exp(-X/2) = \begin{bmatrix} \circ & 1 \\ 1 & \circ \end{bmatrix} ,$$

we can summarize this calculation:

[summary] **3.2. Lemma.** *If X lies in $\mathfrak{a}$ and Y in $\mathfrak{g}_\lambda$ then*

$$\Phi_X(Y) = \frac{\sinh(\tau)}{\tau} \cdot Y$$

where $\tau = \langle \lambda, X \rangle$.

♣ [exp-exp] Since $\sinh(x)/x = 1 + x^2/3! + \cdots \geq 1$ for all x, this concludes the proof of Lemma 3.1. ▯

We can summarize things even more succinctly:

[mostow] **3.3. Theorem.** *If X is in $\mathfrak{s}$ and $D = \mathrm{ad}_{X/2}$ then D^2 is a self-adjoint operator on $\mathfrak{s}$ and*

$$\Phi_X = I + \frac{D^2}{3!} + \frac{D^4}{5!} + \cdots = \frac{\sinh(D)}{D} .$$

4. References

1. Kenneth Brown, **Buildings**, Springer-Verlag, 1989.

2. François Bruhat and Jacques Tits, 'Groupes réductifs sur un corps local I. Données radicielles valuées', *Publications Mathématiques* **41** (1972), 5–252. Their fixed point lemma is in §3.2.

3. Eli Cartan, **Leçons sur la géométrie des espace de Riemann**, second edition, Gauthiers-Villars, 1946.

4. ——, 'Sur certains formes riemanniennes remarquables des géométries à groupe fondamental simple', *Ann. Sci. École Norm. Sup.* **44** (1947), 345–467.

5. Mischa Gromov, 'Hyperbolic groups', 75–264 in **Essays in group theory**, edited by S. M. Gersten, Springer-Verlag, 1988.

6. Sigurdur Helgason, **Differential geometry and symmetric spaces**, Academic Press, 1968.

7. Serge Lang, **Math talks for undergraduates**, Springer-Verlag, 1999. The fifth essay is about 'Bruhat-Tits spaces', and is concerned with Cartan's fixed point theorem for $\mathrm{GL}_n(\mathbb{R})$. It follows Brown's book closely. His proof that $\mathfrak{S}$ is semi-hyperbolic follows almost exactly that in the beginning of the paper of Mostow mentioned below.

8. Dan Mostow, 'Some new decomposition theorems for semi-simple groups', pages 31–54 in **Lie algebras and Lie groups**, *Memoirs of the American Mathematical Society* **14**, 1953.

9. Pappus of Alexandria, **La collection mathématique**, Albert Blanchard, 1933. Book VII is contained in volume 2. There is an English translation of Book VII by Alexander Jones, published by Springer-Verlag, 1985. But Proposition 122 has been rendered incorrectly there.