Chapter 2
The General Noncompact Symmetric Space

"These things will become clear to you," said the old man gently, "at least," he added with a slight doubt in his voice, "clearer than they are at the moment."

From *The Hitchhiker's Guide to the Galaxy*, by Douglas Adams, Pocket Books, NY, 1981. Reprinted by permission of The Crown Publishing Group.

2.1 Geometry and Analysis on G/K

2.1.1 Symmetric Spaces, Lie Groups, and Lie Algebras

Volume I [612] and the first chapter of this tome considered various examples and applications of symmetric spaces X, along with harmonic analysis on X and X/Γ for discrete groups Γ of isometries of X. Here we consider some aspects of analysis on a general noncompact symmetric space $X = G/K$. Our discussion will be very sketchy. The main goal is to lay the groundwork for extension of the results of the preceding chapters to other symmetric spaces which are of interest for applications; in particular, the Siegel upper half space $\mathcal{H}_n$ [which can be identified with $Sp(n,\mathbb{R})/U(n)$] and hyperbolic three space $\mathcal{H}^c$ [which can be viewed as $SL(2,\mathbb{C})/SU(2)$]. We will also be interested in the fundamental domains $\mathcal{H}_n/Sp(n,\mathbb{Z})$ for the Siegel modular group as well as the fundamental domain $\mathcal{H}^c/SL(2,\mathbb{Z}[i])$ for the Picard modular group. It is possible to generalize just about everything we did in the earlier chapters for such examples; e.g., the Selberg trace formula. And our main motivations for doing so come from number theory. Because it is time consuming and sometimes not so enlightening to do each of these examples separately, we have decided to present some results on the general symmetric space. Those interested in number theoretic applications may find this equally tedious and attempt to jump to the next section. But I think it is useful to know what a general

A. Terras, *Harmonic Analysis on Symmetric Spaces—Higher Rank Spaces, Positive Definite Matrix Space and Generalizations*, DOI 10.1007/978-1-4939-3408-9_2

Iwasawa decomposition is, for example, in order to find the right coordinates to use in solving a given problem on the symmetric space. Of course, others will say that the discussion which follows is neither sufficiently general, detailed, nor rigorous. We refer those characters to the texts of other authors which are listed below.

Some topics in physics that lead one to study these other symmetric spaces are: quantum statistical mechanics and quantum field theory (see Hurt [312], Frenkel [187], and Ooguri's interview of Witten [480]), particle physics (see Wybourne [672, Ch. 21]), coherent states (see Hurt [312], Monastyrsky and Perelomov [457], and Perelomov [484]), boson fields (see Shale [551] and Cartier's article in Borel and Mostow [68, pp. 361–386]), solitons (see Dubrovin et al. [143], McKean and Trubowitz [440], Lonngren and Scott [407], and Novikov [474]), rotating tops (see Sofya Kovalevskaya [aka Sonya Kovalevsky] [368], Linda Keen [345], Pelageya Kochina [358], and Cooke [124]), and string theory (see Polyakov [491]).

Many branches of number theory steer one into these realms; e.g., the theory of quadratic forms (see Siegel [561–565]) and algebraic number theory (see Hecke [268] and Siegel [563]). The study of the ring $\mathbb{Z}[i]$ of Gaussian integers and similar rings for various algebraic number fields leads one to think that anything one can do for $\mathbb{Z}$ should be generalizable to $\mathbb{Z}[i]$. In particular, we will see that the theory of Maass wave forms for $SL(2,\mathbb{Z})$ has an analogue for $SL(2,\mathbb{Z}[i])$. This leads to some interesting formulas for the Dedekind zeta function of $\mathbb{Q}(i)$, among other things. See the Corollaries to Theorem 2.2.1 in Section 2.2 which follows. There is also an analogue of Selberg's trace formula (see the last section in this volume or Elstrodt et al. [168]).

Finally electrical engineering has many applications of these symmetric spaces as well (see Blankenship [51] and Helton [284–286]). We saw in Section 3.1 of Volume I that 2-port microwave circuits lead to quantities in $SL(2,\mathbb{R})$. Similarly, more complicated circuits lead to higher rank Lie groups.

References for this section include: Baily [32], Barut and Raçzka [39], Broecker and tom Dieck [81], Chevalley [104], Yvonne Choquet-Bruhat, Cécile DeWitt-Morette, & Margaret Bleick [106], Dieudonné [137], Gangolli [195–198], Harish-Chandra [262, 263], Helgason [273–282], Hermann [289, 290], Hua [308], Loos [408], Maass [426], Piatetski-Shapiro [485], Sagle and Walde [524], Séminaire Cartan [547], Siegel [561–565], Varadarajan [623–625], Wallach [651–653], Warner [655], and Wybourne [672].

We will assume that the reader has had a decent course in multivariable calculus. Our favorite books for this are Lang [385, 388]. The notions of differential, tangent space, matrix exponential, Taylor's formula are all covered there. You may also need to refer to a book like that of Sagle and Walde [524] for more details on various arguments. The true story of everything is found in Helgason's big green books. Varadarajan [623] is also useful, for example, as a source for all the details of the root space calculations.

Élie Cartan obtained the basic theory of symmetric spaces between 1914 and 1927. Then, beginning in the 1950s, Harish-Chandra, Helgason, and others developed harmonic analysis and representation theory on these spaces and their Lie groups of isometries.

A **symmetric space** M is a connected Riemannian manifold (as in the discussion at the beginning of Chapter 2, Volume I) such that at each point $P \in M$ there is a geodesic-reversing isometry

$$s_P : M \to M;$$

i.e., s_P preserves the Riemannian metric and flips geodesics about the fixed point P.

Our first goal is to produce a multitude of examples of symmetric spaces. We always start with a **Lie group** G; i.e., a real analytic manifold which is also a group such that the mapping

$$\begin{aligned} G \times G &\to G \\ (g,h) &\mapsto gh^{-1} \end{aligned} \tag{2.1}$$

is analytic. We will only consider Lie groups of real or complex matrices here. As we have said earlier, it is often useful to replace $\mathbb{R}$ or $\mathbb{C}$ with a finite field, or a local field such as $\mathbb{Q}_p$, the field of p-adic numbers. Mostly, we will avoid doing this.

The **Lie algebra** $\mathfrak{g}$ of a Lie group G is the tangent space to G at the identity, once it has been provided with an additional operation called "the Lie bracket." It is traditional that the Lie algebra is written as the lowercase German letter (fraktur) corresponding to the uppercase Latin letter which is the group. The fraktur letters used here should be recognizable—except $\mathfrak{k}$, which is k.

The Lie bracket operation is defined by identifying the Lie algebra

$$\mathfrak{g} = T_e(G) = \text{ the tangent space to } G \text{ at the identity } e \in G,$$

with the space of left-invariant vector fields on G. These vector fields are first order differential operators on G (with real analytic coefficients) which commute with left translation. This identification is achieved by making use of the left translation $L_g(x) = gx$, for $x, g \in G$. If $\tilde{X}$ is a left-invariant vector field, then

$$\tilde{X}_g = dL_g(X), \qquad \text{for } g \in G,\ X \in \mathfrak{g}.$$

Here dL_g denotes the differential of left multiplication on G. Now we define the **Lie bracket** of two left invariant vector fields $\tilde{X}, \tilde{Y}$ by;

$$[\tilde{X}, \tilde{Y}] = \tilde{X}\tilde{Y} - \tilde{Y}\tilde{X},$$

which is also a left invariant vector field; i.e., the bracket of two first order differential operators is actually a first order and not a second order differential operator. Write $[X, Y]$ for the corresponding bracket of elements X, Y in the Lie algebra $\mathfrak{g}$.

What makes $\mathfrak{g}$ a **Lie algebra**? The answer is that the bracket can be shown to have the following defining properties of such an algebra:

(1) $[X, Y]$ is a bilinear map of $\mathfrak{g} \times \mathfrak{g}$ into $\mathfrak{g}$;
(2) $[X, Y] = -[Y, X]$;
(3) $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$ (**Jacobi's identity**).

Then one defines subalgebra, ideal, homomorphism, isomorphism, etc., for Lie algebras in the usual way (see the references). For example, an ideal $\mathfrak{a} \subset \mathfrak{g}$ is a vector subspace $\mathfrak{a}$ of $\mathfrak{g}$ such that $[\mathfrak{a}, \mathfrak{g}] \subset \mathfrak{a}$.

Exercise 2.1.1. Prove that if $G = GL(n, \mathbb{R})$, then the corresponding Lie algebra can be identified with the vector space $\mathbb{R}^{n\times n}$ of all $n \times n$ real matrices with bracket defined by

$$[A, B] = AB - BA, \qquad \text{for } A, B \in \mathbb{R}^{n\times n}.$$

Here AB denotes the usual matrix product. Thus $\mathfrak{gl}(n, \mathbb{R})$ is identified with $\mathbb{R}^{n\times n}$.
Hint. See Dieudonné [137, Vol. VI, pp. 145–146] or Sagle and Walde [524, pp. 117–118]. The vector space $\mathbb{R}^{n\times n}$ can certainly be identified with the tangent space to $GL(n, \mathbb{R})$, since $GL(n, \mathbb{R})$ is an open subset of $\mathbb{R}^{n\times n}$. In fact, using the matrix exponential, we can make the identification as follows. Suppose $A \in \mathbb{R}^{n\times n}$, $g \in GL(n, \mathbb{R})$, and $f : GL(n, \mathbb{R}) \to \mathbb{C}$. Then

$$(\tilde{A}f)(g) = \left.\frac{d}{dt} f(g \exp tA)\right|_{t=0}.$$

One has for $A, B \in \mathbb{R}^{n\times n}$ and $g \in GL(n, \mathbb{R})$:

$$(\tilde{A}\tilde{B}f)(g) = \left.\frac{\partial^2}{\partial t\, \partial s} f(g \exp tA \exp sB)\right|_{t=s=0}.$$

Use the chain rule to see that at $g = e$ this is $f_e''(A, B) + f_e'(AB)$. If you interchange A and B and then subtract, the second order terms cancel and you get $f_e'(AB - BA)$. This shows that the identification of $\mathfrak{g}$ with $\mathbb{R}^{n\times n}$ does preserve brackets.

There is a representation of any (real) Lie algebra $\mathfrak{g}$ in $\mathfrak{gl}(n, \mathbb{R})$, where $n = \dim_{\mathbb{R}} \mathfrak{g}$. This representation is called the **adjoint representation** defined as follows, thinking of $\mathfrak{gl}(n, \mathbb{R})$ as the space of linear transformations of $\mathfrak{g}$ into itself:

$$\begin{aligned} &\mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}(n, \mathbb{R}), \quad && n = \dim_{\mathbb{R}} \mathfrak{g}, \\ &(\mathrm{ad}\, X)Y = [X, Y], \quad && \text{for } X, Y \in \mathfrak{g}. \end{aligned} \tag{2.2}$$

Exercise 2.1.2. Show that the adjoint representation defined by (1.1) above does indeed preserve brackets; i.e., $[ad\, X, ad\, Y] = ad[X, Y]$.

The **Killing form** of a Lie algebra $\mathfrak{g}$ is defined to be the bilinear form:

$$B_{\mathfrak{g}} = B(X,Y) = \operatorname{Tr}(\operatorname{ad} X \operatorname{ad} Y), \quad \text{for } X, Y \in \mathfrak{g}. \tag{2.3}$$

A Lie algebra is called **semisimple** if the Killing form B is nondegenerate; i.e., $B(X,Y) = 0$ for all $Y \in \mathfrak{g}$ implies $X = 0$. A Lie algebra $\mathfrak{g}$ is **simple** if it is semisimple, and if, in addition, it has no ideals but $\{0\}$ and itself.

Example 2.1.1 ($GL(n,\mathbb{R})$).

Since matrices of the form aI, $a \in \mathbb{R}$, commute with $n \times n$ matrices, it is clear that $\operatorname{ad}(aI) = 0$ and thus that $B(aI, Y) = 0$ for all $Y \in \mathbb{R}^{n\times n}$. Thus $\mathfrak{gl}(n,\mathbb{R}) \cong \mathbb{R}^{n\times n}$ is not semisimple.

It will be useful to compute the Killing form for $\mathfrak{gl}(n,\mathbb{R})$. One can do this as follows. Let E_{ij} be the $n \times n$ matrix with i,j entry equal to one and the rest zero. Let H be the diagonal matrix with ith diagonal entry h_i. Then $\operatorname{ad}(H)E_{ij} = (h_i - h_j)E_{ij}$. Therefore

$$B(H,H) = \operatorname{Tr}(\operatorname{ad} H \operatorname{ad} H) = \sum_{i,j=1}^{n} (h_i - h_j)^2 = 2n \operatorname{Tr}(H^2) - 2(\operatorname{Tr} H)^2.$$

Note that it suffices to compute the Killing form on diagonal matrices. For the map $X \mapsto gXg^{-1}$, with $g \in GL(n,\mathbb{R})$ and $X \in \mathfrak{gl}(n,\mathbb{R})$, leaves the Killing form invariant. Moreover matrices conjugate to a diagonal matrix are dense in $\mathfrak{gl}(n,\mathbb{R})$.

Exercise 2.1.3. (a) Show that if σ is a Lie algebra automorphism of $\mathfrak{g}$, then

$$B(X,Y) = B(\sigma X, \sigma Y), \text{ for all } X, Y \in \mathfrak{g}.$$

(b) Show that

$$B(X,[Y,Z]) = B(Y,[Z,X]) = B(Z,[X,Y]), \text{ for all } X,Y,Z \in \mathfrak{g}.$$

There is an analogue of the matrix exponential for any Lie group G. It is, appropriately enough, called the **exponential map** and it maps the Lie algebra into the Lie group such that if $X \in \mathfrak{g}$, $g \in G$, and $f : G \to \mathbb{C}$ is infinitely differentiable, then

$$\tilde{X}_g f = \frac{d}{dt} f(g \exp tX)\big|_{t=0}. \tag{2.4}$$

For matrix groups the matrix exponential is the Lie group exponential map. For general Lie groups, the existence of $\exp : \mathfrak{g} \to G$ comes from standard results in ordinary differential equations.

Let us list a few properties of exp. The curve

$$\begin{aligned} &\mathbb{R} \to G \\ &t \mapsto \exp(tX), \text{ for } X \in \mathfrak{g}, \end{aligned}$$

is a **one-parameter subgroup** of G; i.e., $\exp(0) = e$, the identity in G, and

$$\exp(tX)\exp(sX) = \exp(t+s)X, \quad \text{for all real numbers} \quad s,t. \tag{2.5}$$

Taylor's formula for G says that:

$$f(g\exp X) = \sum_{n\geq 0} \frac{1}{n!}(\tilde{X}^n f)(g), \quad \text{for } g \in G,\ X \in \mathfrak{g}, \tag{2.6}$$

where f is a real analytic function on G. **The exponential map allows one to relate multiplication on the Lie group with bracket on the Lie algebra via**:

$$\exp tX \exp tY = \exp\left\{t(X+Y) + \frac{1}{2}t^2[X,Y] + O(t^3)\right\}, \tag{2.7}$$

for $X, Y \in \mathfrak{g}$, and $t \in \mathbb{R}$. It is possible to continue the expansion inside the braces in (2.7) and the result is called the Campbell–Hausdorff formula.

Exercise 2.1.4. Prove formula (2.7).
Hint. First consider the case of $GL(n,\mathbb{R})$. The same sort of proof works in general using Taylor's formula (2.6).

It is possible to compute the **differential of exp** and obtain:

$$(d\exp)_X Y = (dL_{\exp X})_e \left(\frac{1-e^{-\operatorname{ad} X}}{\operatorname{ad} X}\right)(Y), \quad \text{for} \quad X, Y \in \mathfrak{g}. \tag{2.8}$$

Formula (2.8) implies in particular that the mapping from X to $\exp X$ is a diffeomorphism from an open neighborhood of 0 in $\mathfrak{g}$ onto an open neighborhood of the identity e in G.

Let us prove (2.8) in the case of matrix exp. First note that

$$\begin{aligned} \lim_{t\to 0}(e^{X+tY} - e^X)/t &= \lim_{t\to 0}\sum_{n\geq 0}\frac{1}{n!t}\left\{(X+tY)^n - X^n\right\} \\ &= \sum_{n\geq 0}\frac{1}{(n+1)!}\left\{X^nY + X^{n-1}YX + \cdots + YX^n\right\}. \end{aligned}$$

Beware that $XY \neq YX$, in general, so that you cannot blindly use the binomial theorem. However, it is possible to be clever (although that is unworthy of a Vulcan), since right and left multiplication by X do commute as operators. Define $R_X Y = YX$ and $L_X Y = XY$. Observe that ad$X = L_X - R_X$. The three operators $R_X,\ L_X$, and ad X will commute. Thus we can apply the binomial theorem to obtain:

$$R_X^m = (L_X - \ \mathrm{ad}\, X)^m = \sum_{k=0}^{m} \binom{m}{k} L_X^{m-k} (-\mathrm{ad}\, X)^k.$$

This allows us to write:

$$\begin{aligned} X^n Y + X^{n-1} YX + \cdots + YX^n &= \sum_{i=0}^{n} X^i \sum_{k=0}^{n-i} \binom{n-i}{k} X^{n-i-k} (-\mathrm{ad}\, X)^k Y \\ &= \sum_{k=0}^{n} \sum_{i=0}^{n-k} \binom{n-i}{k} X^{n-k} (-\mathrm{ad}\, X)^k Y, \end{aligned}$$

upon reversing sums. It is an **exercise** in the properties of binomial coefficients to show that

$$\sum_{i=0}^{n-k} \binom{n-i}{k} = \binom{n+1}{k+1}. \tag{2.9}$$

Therefore

$$\begin{aligned} (d\exp)_X Y &= \sum_{n\geq 0} \frac{1}{(n+1)!} \sum_{k=0}^{n} \binom{n+1}{k+1} X^{n-k} (-\mathrm{ad}\, X)^k Y \\ &= \sum_{k\geq 0} \sum_{n\geq k} \frac{1}{(k+1)!(n-k)!} X^{n-k} (-\mathrm{ad}\, X)^k Y \\ &= \sum_{k\geq 0} \sum_{r\geq 0} \frac{1}{(k+1)!r!} X^r (-\mathrm{ad}\, X)^k Y \\ &= e^X \sum_{k\geq 0} \frac{1}{(k+1)!} (-\mathrm{ad}\, X)^k Y. \end{aligned}$$

This completes the proof of (2.8) in the case of the matrix exponential. The general result is proved in a similar way (see Helgason [275]).

Exercise 2.1.5. Prove formula (2.8) for a general Lie group.

One of the most important tools in Lie group theory is the **dictionary** that allows one to translate between Lie groups and Lie algebras. We list a few results from the dictionary. For the proofs, see references such as Helgason's big green books, Sagle and Walde [524], or Varadarajan [623].

For each Lie group G with Lie algebra $\mathfrak{g}$ and for each Lie subalgebra $\mathfrak{h}$ of $\mathfrak{g}$, there is a unique connected Lie subgroup H of G with Lie algebra $\mathfrak{h}$. However, H may not have the induced topology; e.g., consider the densely wound line in the torus:

$$\begin{aligned} &\mathbb{R} \to (\mathbb{R}/\mathbb{Z})^2 = \mathbb{T}^2 \\ &t \mapsto (e^{it}, e^{iat}), \quad \text{when} \quad a \in \mathbb{R} \text{ is irrational.} \end{aligned}$$

If $f : G_1 \to G_2$ is a Lie group homomorphism of connected Lie groups, then the differential $(df)_e : \mathfrak{g}_1 \to \mathfrak{g}_2$ is a Lie algebra homomorphism. Moreover, we have the following relations between images and kernels:

$$\begin{aligned} &\text{Lie Algebra } (f(G_1)) = (df)_e\, \mathfrak{g}_1, \\ &\text{Lie Algebra } (\ker f) = \ker(df)_e. \end{aligned}$$

If $\lambda : \mathfrak{g}_1 \to \mathfrak{g}_2$ is a Lie algebra homomorphism and G_1, G_2 are connected Lie groups with Lie algebras $\mathfrak{g}_1, \mathfrak{g}_2$, respectively, and if, in addition, G_1 is simply connected, then there exists a unique Lie group homomorphism $f : G_1 \to G_2$ such that $(df)_e = \lambda$.

The hypothesis that G_1 be simply connected cannot be removed in the preceding result. For example, $\mathbb{R}/\mathbb{Z}$ and $\mathbb{R}$ have the same Lie algebra. But the identity mapping of $\mathbb{R}$ onto itself cannot be the differential of a Lie group homomorphism from $\mathbb{R}/\mathbb{Z}$ to $\mathbb{R}$.

Exercise 2.1.6. (a) Show that the exponential map $\mathfrak{g} \to G$ need not be onto.
(b) Show that $\exp : \mathfrak{gl}(n, \mathbb{C}) \to GL(n, \mathbb{C})$ is onto.

Hints.

(a) Take $G = SL(2, \mathbb{R})$ and consider

$$A = \begin{pmatrix} r & 0 \\ 0 & 1/r \end{pmatrix} \quad \text{for} \quad r < -1.$$

If $A = \exp(X)$, consider the eigenvalues of X.
(b) Use the Jordan canonical form.

The final dictionary entry that we list here concerns a closed subgroup H of a Lie group G. Then H must have the induced topology and

$$\text{Lie Algebra } (H) = \mathfrak{h} = \{X \in \mathfrak{g} \mid \exp(tX) \in H, \text{ for all } t \in \mathbb{R}\}. \tag{2.10}$$

Formula (2.10) provides a quick way to compute Lie algebras. For example, since $\det(e^X) = e^{\operatorname{Tr} X}$, for matrices X, it follows that the Lie algebra $\mathfrak{sl}(n, \mathbb{R})$ consists of all $n \times n$ real matrices of trace zero. As we said, we use the notation that the Lie algebra of a group G is in lowercase German Fraktur letters so that $\mathfrak{sl}(n, \mathbb{R})$ is the Lie algebra of $SL(n, \mathbb{R})$. One can show that the Killing form of $\mathfrak{sl}(n, \mathbb{R})$ is:

$$B_{\mathfrak{sl}(n,\mathbb{R})}(X,Y) = 2n\mathrm{Tr}(XY), \quad \text{for } \ X,Y \in \mathfrak{sl}(n,\mathbb{R}).$$

Therefore $\mathfrak{sl}(n,\mathbb{R})$ is a semisimple Lie algebra. In fact, it is actually a simple Lie algebra.

Exercise 2.1.7. (a) Verify the comments made in the last paragraph.
(b) Find the Lie algebra of the **symplectic group**[1]:

$$Sp(n,\mathbb{R}) = \left\{ g \in \mathbb{R}^{2n\times 2n} \mid {}^t g J_n g = J_n \right\}, \quad \text{for} \quad J_n = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}.$$

(c) Find the Lie algebra of the Lorentz-type group

$$O(p,q) = \left\{ g \in \mathbb{R}^{n\times n} \mid {}^t g I_{p,q} g = I_{p,q} \right\},$$

where $n = p + q$, and

$$I_{p,q} = \begin{pmatrix} I_p & 0 \\ 0 & -I_q \end{pmatrix}.$$

Hint. The answer to part (b) is given in formula (2.12).

There is an analogue of the adjoint representation on the group level, denoted ***Ad***. To obtain it, proceed as follows. If $g \in G$, define

$$\mathrm{Int}(g)x = gxg^{-1}, \text{ for all } x \in G \text{ and } \mathrm{Ad}(g) = (d\,\mathrm{Int}(g))_e, \text{ where } e \text{ is the identity of } G. \tag{2.11}$$

Then we have a commutative diagram:

$$\begin{array}{ccc} \mathfrak{g} & \xrightarrow{\mathrm{Ad}(g)} & \mathfrak{g} \\ \exp\downarrow & & \downarrow\exp \\ G & \xrightarrow{\mathrm{Int}(g)} & G. \end{array}$$

It can be proved that $(d\ \mathrm{Ad})_e X = \mathrm{ad}X$, for all $X \in \mathfrak{g}$. Thus we have another commutative diagram:

$$\begin{array}{ccc} \mathfrak{g} & \xrightarrow{\mathrm{ad}} & \mathfrak{gl}(\mathfrak{g}) \\ \exp\downarrow & & \downarrow \ \text{matrix}\exp \\ G & \xrightarrow{\mathrm{Ad}} & G. \end{array}$$

[1]Beware! Some authors write $Sp(2n,F)$ instead of $Sp(n,F)$.

If G is a matrix group already, then the matrix $\mathrm{Ad}(g)$ is the matrix of $\mathrm{Int}(g)$, since $\mathrm{Int}(g)x$ is a linear function of x. If $\mathfrak{g}$ is semisimple, then the kernel of ad is $\{0\}$ and the kernel of Ad is the center of G, which must then be discrete.

Exercise 2.1.8. Prove that $(d\ \mathrm{Ad})_e X = \mathrm{ad}X$, for all $X \in \mathfrak{g}$.

It is possible to classify all the simple Lie algebras over the complex numbers. Except for a finite number of exceptional Lie algebras, the **simple Lie algebras over** $\mathbb{C}$ are in the following list (with J_n as in Exercise 2.1.7 above):

$$\left.\begin{aligned}
&\mathfrak{a}_n = \mathfrak{sl}(n+1,\mathbb{C}), \quad n \geq 1;\\
&\mathfrak{b}_n = \mathfrak{so}(2n+1,\mathbb{C}) = \left\{X \in \mathbb{C}^{(2n+1)\times(2n+1)} \;\middle|\; {}^tX = -X\right\}, \quad n \geq 2;\\
&\mathfrak{c}_n = \mathfrak{sp}(n,\mathbb{C}) = \left\{X \in \mathbb{C}^{(2n)\times(2n)} \;\middle|\; {}^tXJ_n + J_nX = 0\right\}\\
&\quad = \left\{\begin{pmatrix} A & B \\ C & -{}^tA \end{pmatrix} \;\middle|\; A,B,C \in \mathbb{R}^{n\times n}, B, C \text{ symmetric}\right\}, \quad n \geq 3;\\
&\mathfrak{d}_n = \mathfrak{so}(2n,\mathbb{C}) = \left\{X \in \mathbb{C}^{(2n)\times(2n)} \;\middle|\; {}^tX = -X\right\}, \quad n \geq 4.
\end{aligned}\right\} \tag{2.12}$$

The indices n are restricted because in low dimensions some strange things happen; e.g., $\mathfrak{d}_1$ and $\mathfrak{d}_2$ are not simple, since $\mathfrak{d}_1$ is abelian and $\mathfrak{d}_2 \cong \mathfrak{a}_1 \oplus \mathfrak{a}_1$. Also $\mathfrak{a}_1 \cong \mathfrak{b}_1 \cong \mathfrak{c}_1$, $\mathfrak{b}_2 \cong \mathfrak{c}_2$, $\mathfrak{d}_3 \cong \mathfrak{a}_3$. These things can be proved using Dynkin diagrams. You can find the details in Varadarajan [623]. The Lie groups corresponding to the Lie algebras in this list are $SL(n,\mathbb{C})$, the special linear group of $n\times n$ complex matrices of determinant one, $SO(n,\mathbb{C})$, the special orthogonal group of $n \times n$ complex matrices g of determinant one such that ${}^tgg = I$, and $Sp(n,\mathbb{C})$, the complex symplectic group of $(2n)\times(2n)$ matrices g with the property that ${}^tgJ_ng = J_n$, for J_n as in Exercise 2.1.7 above. We should perhaps note again that some authors write $Sp(2n,\mathbb{C})$ instead of $Sp(n,\mathbb{C})$. This is rather confusing.

Cartan's classification of symmetric spaces makes use of the preceding classification of complex simple Lie algebras. It also uses the surprising, but simple, observation that **the group $I(M)$ of isometries of a symmetric space M acts transitively on M**. To see this fact, it helps to recall the **Hopf–Rinow theorem** in differential geometry (see Helgason [273, p. 56]) which says that if M is a Riemannian manifold, then the following are equivalent:

(a) M is a complete metric space;
(b) each maximal geodesic $\gamma(t)$ in M can be extended to all $t \in \mathbb{R}$;
(c) each bounded closed subset of M is compact.

If M is a complete Riemannian manifold, then any two points P, Q in M can be joined by a geodesic whose length is the metric space distance between P and Q. To see that a symmetric space M must be complete, note that if a point P lies on the geodesic γ of M and s_P denotes the geodesic-reversing isometry at P, then $s_P\gamma$ is an extension of γ. Thus each maximal geodesic of a symmetric space must have domain the set of all real numbers. Then to see that the group $I(M)$ of isometries of M acts transitively on M, note that if P and Q are in M, then the geodesic-reversing isometry at the midpoint of the geodesic connecting P to Q will exchange P and Q.

It is possible to prove that $I(M)$ is a Lie group such that the connected component of the identity in $I(M)$ still acts transitively on M (see Helgason [273, Ch. 4]). **Let G be the connected component of the identity in $I(M)$.** Now fix a point o to be called the **origin** of the symmetric space M. And let K denote the subgroup of G consisting of elements which fix o. Then K is compact and we can identify M with G/K. Suppose next that s_o denotes the **geodesic-reversing isometry at the origin**. Now consider the map:

$$\begin{aligned} \sigma : G &\to G \\ g &\mapsto s_o g s_o. \end{aligned}$$

Note that σ is an involutive automorphism of G (i.e., σ is an automorphism in the sense of Lie groups such that σ^2 is the identity). Moreover, setting

$$K_\sigma = \{g \in G \mid \sigma g = g\}$$

and

$$(K_\sigma)_o = \text{the connected component of the identity in } K_\sigma,$$

we have

$$(K_\sigma)_o \subset K \subset K_\sigma.$$

This means that K and K_σ have the same Lie algebra.

Now consider the consequences of the preceding remarks about symmetric spaces and Lie groups of isometries on the Lie algebras of these groups, using the dictionary relating Lie groups and Lie algebras. One sees that:

$$(d\sigma)_e : \mathfrak{g} \to \mathfrak{g}$$

is an involutive Lie algebra automorphism which fixes $\mathfrak{k}$, the Lie algebra of K. Moreover, the eigenspace decomposition of $(d\sigma)_e$ on $\mathfrak{g}$ is:

$$\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p},$$

where $\mathfrak{k} = \{X \in \mathfrak{g} \mid (d\sigma)_e X = X\}$ and $\mathfrak{p} = \{X \in \mathfrak{g} \mid (d\sigma)_e X = -X\}$; that is, $\mathfrak{k}$ is the space of eigenvectors corresponding to the eigenvalue $+1$ while $\mathfrak{p}$ consists of eigenvectors corresponding to the eigenvalue -1.

If $\pi : G/K \to M$ is the natural identification, then $(d\pi)_o$ maps $\mathfrak{k}$ to $\{0\}$ and identifies $\mathfrak{p}$ with the tangent space $T_o(M)$.

To proceed further with the classification of symmetric spaces, one must reduce to semisimple Lie algebras, using the following result of E. Cartan (see Helgason [273, Ch. 5]).

Symmetric Space Decomposition Suppose that M is a simply connected symmetric space. Then M is a product:

$$M = M_e \times M_c \times M_n,$$

where M_e is of **Euclidean** type, M_c is **compact with semisimple Lie group** of isometries, and M_n is **noncompact with semisimple Lie group** of isometries having a Lie algebra with a Cartan decomposition described below.

We say that a semisimple Lie algebra is of **compact type** if its Killing form is negative definite (see Helgason [273, p. 122]).

A **Cartan decomposition** of a noncompact semisimple Lie algebra $\mathfrak{g}$ is a vector space direct sum decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ such that the Killing form of $\mathfrak{g}$ is negative definite on $\mathfrak{k}$ and positive definite on $\mathfrak{p}$. Also the mapping $\theta : \mathfrak{g} \to \mathfrak{g}$ with $\theta(X+Y) = X - Y$, for $X \in \mathfrak{k}$ and $Y \in \mathfrak{p}$, must be an automorphism of $\mathfrak{g}$. We call θ the **Cartan involution**.

Example 2.1.2. Consider the simple Lie algebra $\mathfrak{sl}(n, \mathbb{R})$ of all trace zero $n \times n$ real matrices. Set

$$\mathfrak{so}(n) = \{\text{skew-symmetric } n \times n \text{ real matrices}\}$$

and

$$\mathfrak{p}_n = \{\text{symmetric } n \times n \text{ real matrices of trace } 0\}.$$

Clearly we have the direct sum decomposition:

$$\mathfrak{sl}(n, \mathbb{R}) = \mathfrak{so}(n) \oplus \mathfrak{p}_n,$$

with Cartan involution $\theta(X) = -{}^tX$. The Killing form on $\mathfrak{sl}(n, \mathbb{R})$ is $B(X, Y) = 2n\mathrm{Tr}(XY)$, and it is easy to see that this is negative definite on $\mathfrak{so}(n)$ and positive definite on $\mathfrak{p}_n$. It is also easy to check that the Cartan involution preserves the Lie bracket in $\mathfrak{sl}(n, \mathbb{R})$, which is $[X, Y] = XY - YX$.

Exercise 2.1.9. Prove all the claims made in the preceding example.

There is a mirror image of the **Cartan decomposition on the Lie group level**:

$$G = KP,$$

where $P = \exp \mathfrak{p}$. For the example above, we have

$$SL(n, \mathbb{R}) = SO(n)\, \mathcal{SP}_n, \tag{2.13}$$

where, as usual, $\mathcal{SP}_n$ denotes the positive $n \times n$ real matrices of determinant one. The proof of (2.13) is easy (see Exercise 1.1.5 of Section 1.1.2).

We have not given more than a rough sketch of the preceding arguments on classification of symmetric spaces because we are more interested in studying the examples. Thus we will give more attention to the question: **How does one obtain symmetric spaces out of Cartan decompositions of semisimple Lie algebras?** Suppose that $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ is a Cartan decomposition of the semisimple Lie algebra $\mathfrak{g}$ with Cartan involution θ. Let G be a connected real semisimple Lie group with Lie algebra $\mathfrak{g}$, and let K be a Lie subgroup of G having Lie algebra $\mathfrak{k}$. Then G/K has a unique analytic manifold structure such that the mapping of $\mathfrak{p}$ into G/K defined by sending X to $(\exp X)K$ is a diffeomorphism. If $\mathfrak{g}$ is of noncompact type, it can be proved that K is closed, connected, and equal to the fixed point set of an involutive automorphism $t : G \to G$ such that $(dt)_e$ is the Cartan involution θ. Such a map t clearly exists if G is simply connected (making use of the dictionary between group and algebra). But one does not really have to assume that G is simply connected in the noncompact case. Moreover K is compact if and only if the center of G is finite and then K is a **maximal compact subgroup** of G. For proofs of these results, see Helgason [273, Ch. 6].

To make G/K a symmetric space, we use the Killing form B of $\mathfrak{g}$. Let $\pi : G \to G/K$ be defined by $\pi(g) = gK$. Define the Riemannian metric Q on G/K by translating the Killing form on the space $\mathfrak{p}$:

$$Q_{gK}((d\pi)_g \tilde{X}_g, (d\pi)_g \tilde{Y}_g) = B(X, Y), \quad \text{for all } X, Y \in \mathfrak{p}. \tag{2.14}$$

Here $\tilde{X}$ denotes the left invariant vector field corresponding to $X \in \mathfrak{p}$. The metric Q is well defined because the Killing form is invariant under $\mathrm{Ad}\,(k)$, for $k \in K$. It is clear that the metric is positive from the definition of the Cartan decomposition. And it is easily seen that the metric is G-invariant.

The **geodesic-reversing isometry** s_o at the origin o, which is the coset K in G/K, is obtained from the involutive automorphism $t : G \to G$ as follows:

$$\begin{aligned} s_o : G/K &\to G/K \\ gK &\mapsto t(g)K. \end{aligned}$$

Translate by elements of G to obtain the geodesic-reversing isometries at other points of G/K.

Example 2.1.3. The Riemannian structure on $SL(n, \mathbb{R})/SO(n)$ obtained from (2.14) above is just the same as that defined in Chapter 1 of this volume. To see this, first note that one has an identification:

$$\begin{aligned} SL(n, \mathbb{R})/SO(n) &\to \mathcal{SP}_n \\ gSO(n) &\mapsto g\,{}^t g. \end{aligned}$$

The action of $g \in SL(n, \mathbb{R})$ on $Y \in \mathcal{SP}_n$ is given by $a_g(Y) = Y[{}^t g]$. The differential is $(da_g)_I = a_g$ since $a_g(Y)$ is a linear function of Y. So we find that if $Y = g\,{}^t g$, for

$g \in SL(n,\mathbb{R})$, and if u, v are in $T_Y(\mathcal{SP}_n)$, the tangent space to $\mathcal{SP}_n$ at the point Y, then:

$$Q_Y(u,v) = 2n\,\mathrm{Tr}\left((da_g)_I^{-1}u \cdot (da_g)_I^{-1}v\right) = 2n\,\mathrm{Tr}\left(g^{-1}u\,{}^t g^{-1} \cdot g^{-1}v\,{}^t g^{-1}\right)$$
$$= 2n\,\mathrm{Tr}(Y^{-1}uY^{-1}v).$$

This is exactly the Riemannian structure of Chapter 1 of this volume.

Before considering more examples, let us record a few more general facts. Suppose again that we have a Cartan decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ of a semisimple noncompact Lie algebra over the real numbers. Assume that the Lie group $\tilde{G}$ is the universal covering group of G. Then there is a unique involutive automorphism $\tilde{t} : \tilde{G} \to \tilde{G}$ such that the differential $(d\tilde{t})_e$ is θ, the Cartan involution. It can be proved that the center $\tilde{Z}$ of $\tilde{G}$ is contained in $\tilde{K}$, where $\tilde{K}$ is the analytic subgroup of $\tilde{G}$ with Lie algebra $\mathfrak{k}$ (see Helgason [273, p. 216]). Now G is a quotient $\tilde{G}/N$ for some $N \subset \tilde{Z}$. Thus $\tilde{t}$ induces an involution automorphism of G. Setting $K = \tilde{K}/N$, we have:

$$G/K \cong (\tilde{G}/N)/(\tilde{K}/N) \cong \tilde{G}/\tilde{K}.$$

So the symmetric space G/K is independent of the choice of Lie group G with Lie algebra $\mathfrak{g}$. So we may assume that G is simply connected whenever we need this. Furthermore, it can be proved that all the K's are conjugate (see Helgason [273, p. 256]). Note, however, that the K's need not be semisimple.

The preceding arguments fail for symmetric spaces of compact type. For example, the center of G need not lie in K; e.g., consider $G = SU(n),\ K = SO(n)$. Also K need not be connected; e.g., $SO(3)/K = \mathbb{P}^2$, the real projective plane, with K the subgroup of $SO(3)$ leaving a line through the origin invariant. Finally, the Cartan involution need not correspond to an automorphism of G in the compact case.

Another difference between compact and noncompact symmetric spaces is that the noncompact ones are topologically (though not geometrically) identifiable with the Euclidean space $\mathfrak{p}$. However, the compact symmetric spaces are not topologically trivial (see Greub et al. [244]). This fact makes the compact and noncompact symmetric spaces very different. However, there is a duality between the two types, as we shall see.

2.1.2 Examples of Symmetric Spaces

Now we intend to manufacture many examples of symmetric spaces by exploring the connection between real forms of complex simple Lie algebras and Cartan decompositions of real Lie algebras.

A **real form** $\mathfrak{g}$ of a complex simple Lie algebra $\mathfrak{g}^c$ is defined by the equality of the complexification of $\mathfrak{g}$ with $\mathfrak{g}^c$; i.e.,

$$\mathfrak{g}^c = \mathfrak{g} \otimes_{\mathbb{R}} \mathbb{C} = \mathfrak{g} \oplus i\mathfrak{g}.$$

It is possible to list the real forms $\mathfrak{g}$ of $\mathfrak{g}^c$ by listing the conjugations of $\mathfrak{g}^c$. By a **conjugation** of $\mathfrak{g}^c$, we mean a mapping $C : \mathfrak{g}^c \to \mathfrak{g}^c$ which is conjugate linear, bracket preserving, and such that C^2 is the identity. See Helgason [273, Chs. 3, 10] or Loos [408, Vol. II, Ch. VII].

It can also be shown that any complex semisimple Lie algebra $\mathfrak{g}^c$ has a compact real form $\mathfrak{u}$; i.e., the Killing form of $\mathfrak{u}$ is negative definite. Then to make a *list of symmetric spaces of noncompact type coming from complex simple Lie algebras* $\mathfrak{g}^c$, *one must follow through the following plan of action.*

Plan for Construction of Noncompact Symmetric Spaces of Type III

I. List the conjugations ; i.e., involutive automorphisms of $\mathfrak{g}^c$. The fixed points will be real forms of $\mathfrak{g}^c$. One of these real forms $\mathfrak{u}$ will be compact.

II. For the noncompact real forms $\mathfrak{g}$ of $\mathfrak{g}^c$, the Cartan decomposition is:

$$\mathfrak{g} = (\mathfrak{g} \cap \mathfrak{u}) \oplus (\mathfrak{g} \cap i\,\mathfrak{u}).$$

Note that the Killing form of $\mathfrak{g}$ has the correct behavior on the decomposition since $\mathfrak{u}$ is compact; i.e., the Killing form is negative definite on $\mathfrak{u}$. Furthermore, if τ is the conjugation of $\mathfrak{g}^c$ corresponding to the compact real form $\mathfrak{u}$, then the restriction of τ to $\mathfrak{g}$ is θ, the Cartan involution corresponding to this Cartan decomposition.

III. Form the symmetric space G/K by taking Lie groups $G \supset K$ with Lie algebras $\mathfrak{g}, \mathfrak{k}$, respectively. Here $\mathfrak{k} = \mathfrak{g} \cap \mathfrak{u}$.

Type $\mathfrak{a}$ Examples

I. Real Forms of $\mathfrak{sl}(n,\mathbb{C})$.

1. $\mathfrak{sl}(n,\mathbb{R})$=normal real form = fixed points of the conjugation $\tau(X) = \overline{X}$.
2. $\mathfrak{su}(n,\mathbb{R})$=compact real form = fixed points of the conjugation $\tau(X) = -\,{}^t\overline{X}$.
3. $\mathfrak{su}(p,q)$=fixed points of the conjugation $\tau(X) = -I_{p,q}\,{}^t\overline{X}\,I_{p,q}$, where

$$I_{p,q} = \begin{pmatrix} I_p & 0 \\ 0 & -I_q \end{pmatrix}, \quad n = p + q.$$

4. $\mathfrak{su}^*(2m)$=fixed points of the conjugation $\tau(X) = J_m \overline{X} J_m^{-1}$, where

$$J_m = \begin{pmatrix} 0 & I_m \\ -I_m & 0 \end{pmatrix}, \quad n = 2m \quad (\text{for even } \ n).$$

II. Cartan Decompositions of Noncompact Real Forms of $\mathfrak{sl}(n,\mathbb{C})$.

1. $\mathfrak{sl}(n,\mathbb{R}) = \mathfrak{so}(n) \oplus \mathfrak{p}_n$, where

$$\mathfrak{so}(n) = \{X \in \mathfrak{sl}(n,\mathbb{R}) \mid {}^tX = -X\},$$
$$\mathfrak{p}_n = \{X \in \mathfrak{sl}(n,\mathbb{R}) \mid {}^tX = X\}.$$

The Cartan involution is $\theta(X) = -\,{}^tX,\ X \in \mathfrak{sl}(n,\mathbb{R})$.

2. $\mathfrak{su}(p,q) = \mathfrak{k} \oplus \mathfrak{p}$, where

$$\mathfrak{k} = \left\{ \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \middle| \ A \in \mathfrak{u}(p),\ B \in \mathfrak{u}(q),\ \mathrm{Tr}(A+B) = 0 \right\},$$

$$\mathfrak{p} = \left\{ \begin{pmatrix} 0 & Z \\ {}^t\overline{Z} & 0 \end{pmatrix} \middle| \ Z \in \mathbb{C}^{p\times q} \right\}.$$

The Cartan involution is $\theta(X) = I_{p,q} X I_{p,q},\ X \in \mathfrak{su}(p,q)$.

3. $\mathfrak{su}^*(2m) = \mathfrak{k} \oplus \mathfrak{p}$, where

$$\mathfrak{k} = \mathfrak{sp}(m,\mathbb{C}) \cap \mathfrak{u}(2m) \doteqdot \mathfrak{sp}(m) \quad (\text{by definition}),$$
$$\mathfrak{p} = \mathfrak{su}^*(2m) \cap (i\,\mathfrak{u}(2m)).$$

The Cartan involution is $\theta(X) = -J_m\, {}^tX J_m^{-1}$.

III. The Noncompact Symmetric Spaces Corresponding to the Noncompact Real Forms.

1. $SL(n,\mathbb{R})/SO(n)$.
2. $SU(p,q)/S(U_p \times U_q)$, where $n = p+q$ and

$$SU(p,q) = \left\{ g \in SL(n,\mathbb{C}) \ \middle|\ {}^t\overline{g} I_{p,q} g = I_{p,q} \right\},$$
$$U(p) = \left\{ g \in \mathbb{C}^{p\times p} \ \middle|\ {}^t\overline{g} g = I_p \right\} = \text{ the unitary group},$$
$$S(U_p \times U_q) = \left\{ g \in SL(n,\mathbb{C}) \ \middle|\ g = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix},\ A \in U(p), B \in U(q) \right\}.$$

3. $SU^*(2n)/Sp(n)$, where

$$SU^*(2n) = \left\{ g \in SL(2n,\mathbb{C}) \ \middle|\ g = J_n \overline{g} J_n^{-1} \right\},$$
$$Sp(n) = Sp(n,\mathbb{C}) \cap U(2n),$$
$$Sp(n,\mathbb{C}) = \left\{ g \in \mathbb{C}^{2n\times 2n} \ \middle|\ {}^tg J_n g = J_n \right\} = \text{the complex symplectic group}.$$

IV. The Corresponding Compact Symmetric Spaces.

4. $SU(n)/SO(n)$.
5. $SU(p+q)/S(U_p \times U_q)$.
6. $SU(2n)/Sp(n)$.

Type $\mathfrak{c}$ Examples

I. Real Forms of $sp(n,\mathbb{C})$.

1. $\mathfrak{sp}(n,\mathbb{R})$ =normal real form=fixed points of the conjugation $\tau(X) = \overline{X}$.
2. $\mathfrak{sp}(n)$ =compact real form=fixed points of the conjugation $\tau(X) = -\,{}^t\overline{X}$. Note that $\mathfrak{sp}(n) = \mathfrak{sp}(n,\mathbb{C}) \cap \mathfrak{u}(2n)$.
3. $\mathfrak{sp}(p,q) =$ fixed points of the conjugation $\tau(X) = -K_{p,q}\,{}^t\overline{X}K_{p,q}$, where

$$K_{p,q} = \begin{pmatrix} I_p & 0 & 0 & 0 \\ 0 & -I_q & 0 & 0 \\ 0 & 0 & I_p & 0 \\ 0 & 0 & 0 & -I_q \end{pmatrix}, \quad p+q=n.$$

II. Cartan Decompositions of Noncompact Real Forms of $sp(n,\mathbb{C})$.

1. $\mathfrak{sp}(n,\mathbb{R}) = \mathfrak{k} \oplus \mathfrak{p}$, where

$$\mathfrak{k} = \left\{ \begin{pmatrix} A & B \\ -B & A \end{pmatrix} \,\middle|\, A, B \in \mathbb{R}^{n\times n},\ B = {}^tB,\ A = -\,{}^tA \right\} \cong \mathfrak{u}(n),$$

$$\mathfrak{p} = \left\{ X \in \mathfrak{sp}(n,\mathbb{R}) \,\middle|\, X = {}^tX \right\}.$$

To see that $\mathfrak{k} \cong \mathfrak{u}(n)$, map

$$\begin{pmatrix} A & B \\ -B & A \end{pmatrix} \in \mathfrak{k} \quad \text{to} \quad A + iB \in \mathfrak{u}(n).$$

The Cartan involution is $\theta(X) = -\,{}^tX$.

2. $\mathfrak{sp}(p,q) = \mathfrak{k} \oplus \mathfrak{p}$, where

$$\mathfrak{k} = \left\{ \begin{pmatrix} X_{11} & 0 & X_{13} & 0 \\ 0 & X_{22} & 0 & X_{24} \\ -X_{13} & 0 & X_{11} & 0 \\ 0 & -X_{24} & 0 & X_{22} \end{pmatrix} \,\middle|\, \begin{matrix} X_{11} \in \mathfrak{u}(p),\ X_{22} \in \mathfrak{u}(q) \\ X_{13} \in \mathbb{C}^{p\times p},\ {}^tX_{13} = X_{13} \\ X_{24} \in \mathbb{C}^{q\times q},\ {}^tX_{24} = X_{24} \end{matrix} \right\}$$

$$\cong \mathfrak{sp}(p) \times \mathfrak{sp}(q).$$

The Cartan involution is $\theta(X) = K_{p,q}XK_{p,q}$.

III. The Corresponding Noncompact Symmetric Spaces.

1. $Sp(n,\mathbb{R})/U(n)$.

   Here $G = Sp(n,\mathbb{R})$ is the symplectic group defined in Exercise 2.1.7 while $U(n)$ is really the subgroup $K = G \cap O(2n)$ which is isomorphic to the unitary group

   $$U(n) = \left\{ g \in \mathbb{C}^{n\times n} \;\middle|\; {}^t\overline{g}g = I \right\},$$

   by part (b) of Lemma 2.1.1 below.

   There are two equivalent but rather different ways to view this symmetric space. The first is as the space $\mathcal{P}_n^*$ of **positive symplectic $2n \times 2n$ real matrices**.

   The second version of $Sp(n,\mathbb{R})/U(n)$ is the **Siegel upper half space** $\mathcal{H}_n$ defined by:

   $$\mathcal{H}_n = \left\{ Z \in \mathbb{C}^{n\times n} \;\middle|\; {}^tZ = Z,\ \operatorname{Im} Z \in \mathcal{P}_n \right\}.$$

   This example is the most important one for the rest of this book. We will discuss the various identifications of $Sp(n,\mathbb{R})/U(n)$ below. Sometimes the space $\mathcal{H}_n$ is called the "Siegel upper half plane," despite the fact that it is definitely not two-dimensional for $n > 1$. We must also apologize for our abusive use of the letter H. In Volume I, H was the Poincaré upper half plane. Now it should be $\mathcal{H}_1$. Then there is the Helgason–Fourier transform. Help! I need more alphabets!
2. $Sp(p,q)/Sp(p) \times Sp(q)$.

   Here $Sp(p,q) = \left\{ g \in SL(p+q,\mathbb{C}) \;\middle|\; {}^tgK_{p,q}\overline{g} = K_{p,q} \right\}$, where, as before

   $$K_{p,q} = \begin{pmatrix} -I_p & & & 0 \\ & I_q & & \\ & & -I_p & \\ 0 & & & I_q \end{pmatrix}.$$

Exercise 2.1.10. Check the computations for the type A and C noncompact symmetric space examples above.

This is just about all the examples of symmetric spaces that we shall discuss. In particular, we are avoiding the exceptional Lie groups and their symmetric spaces. Table 2.1 lists some other examples of symmetric spaces. We will also be interested in the symmetric space $SL(2,\mathbb{C})/SU(2)$, which can be identified with the quaternionic upper half space or hyperbolic 3-space. It is considered at the end of this section. It is the symmetric space of a complex Lie group considered as a real group. We do not discuss compact symmetric spaces here, except to note that there is a duality between symmetric spaces U/K' of compact type and symmetric spaces G/K of noncompact type. This duality is obtained on the Lie algebra level by writing

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

and

$$\mathfrak{u} = \mathfrak{k} \oplus i\mathfrak{p},$$

where $\mathfrak{u}$ is a compact real form of the complexification of $\mathfrak{g}$. See Helgason [273, Ch. 5] for more details. Helgason [273, p. 321] gives a global duality result for bounded symmetric domains (which will be defined below) allowing them to be viewed as open submanifolds of a compact Hermitian space. This is the Borel embedding theorem (see Borel [61, 62]). Such results can be applied to compute dimensions of spaces of automorphic forms via the Hirzebruch–Riemann–Roch theorem (see Hirzebruch [297, pp. 162–165] and Section 2.2.3). Healy [266] provides an example of the implications of this duality for harmonic analysis on $SU(2)$ and hyperbolic 3-space.

Example 2.1.4 (The Duality Between Compact and Noncompact Symmetric Spaces).
This example shows that hyperbolic geometry is dual to spherical geometry. We begin with the two Cartan decompositions:

$$\mathfrak{sl}(2,\mathbb{R}) = \mathfrak{so}(2) \oplus \mathfrak{p}_2, \quad \mathfrak{su}(2) = \mathfrak{so}(2) \oplus i\mathfrak{p}_2,$$

where

$$\mathfrak{p}_2 = \left\{ X \in \mathbb{R}^{2\times 2} \;\middle|\; {}^tX = X,\ \mathrm{Tr}\, X = 0 \right\}.$$

The symmetric space $SL(2,\mathbb{R})/SO(2)$ can be viewed as the hyperbolic upper half plane of Chapter 3, Vol. I, while $SU(2)/SO(2)$ can be viewed as the sphere S^2 in $\mathbb{R}^3$, which is the symmetric space considered in Chapter 2, Vol. I.

The Siegel upper half space $\mathcal{H}_n$ is a **Hermitian symmetric space**; i.e., a symmetric space with a complex structure, invariant under each geodesic-reversing symmetry. It turns out that the Hermitian symmetric spaces of the compact or noncompact type have non-semisimple maximal compact subgroups K (see Helgason [273, p. 281] or Loos [408, Vol. II, p. 161]). Such is indeed the case for $G = Sp(n,\mathbb{R})$, $K = U(n)$. It also turns out that the Hermitian symmetric spaces of noncompact type are the **bounded symmetric domains** D in complex n-space. Here **symmetric** means that for every $z \in D$ there is a biholomorphic involutive map on D having z as an isolated fixed point (see Helgason [273, pp. 311–322] or Loos [408, Vol. II, p. 164]). Koecher [361] found a way of constructing all the Hermitian symmetric spaces from Jordan algebras. We shall see in Exercise 2.1.11 that the

Siegel upper half space is identifiable with a bounded symmetric domain, namely the **generalized unit disc**:

$$\mathcal{D}_n = \left\{ W \in \mathbb{C}^{n\times n} \;\middle|\; {}^tW = W,\ I - \overline{W}W \in \mathcal{P}_n \right\}. \tag{2.15}$$

The identification map is the **generalized Cayley transform**:

$$\begin{aligned} \alpha : \mathcal{H}_n &\to \mathcal{D}_n \\ Z &\mapsto (Z - iI)(Z + iI)^{-1}. \end{aligned} \tag{2.16}$$

Exercise 2.1.11 (The Cayley Transform). Show that $W = (Z - iI)(Z + iI)^{-1}$ maps $Z \in \mathcal{H}_n$ into W in the generalized unit disc defined by

$$\mathcal{D}_n = \left\{ W \in \mathbb{C}^{n\times n} \mid {}^tW = W,\ \ I - W\overline{W} \in \mathcal{P}_n \right\}.$$

This mapping allows us to view the symmetric space of the symplectic group as a bounded symmetric domain. What is the image of $\mathcal{P}_n$, viewing $Y \in \mathcal{P}_n$ as the element $iY \in \mathcal{H}_n$?

Cartan proved in 1935 that there are only six types of irreducible homogeneous bounded symmetric domains (see Helgason [278, p. 518]). Here **irreducible** means that the corresponding Lie group is simple. It is possible to generalize many results from analysis and number theory to these classical domains (see Hua [308], Piatetski-Shapiro [485], and Siegel [564], [565, Vol. II, pp. 274–369]).

We could also have differentiated between the three types of symmetric spaces M according to their **sectional curvature**. The sectional curvature is defined as $-g(R(u,v)u,v)$, where g is the Riemannian metric for M, R is the curvature tensor, and u,v are orthonormal tangent vectors in $T_P(M)$, the tangent space to M at a point P. For a symmetric space, the curvature tensor at the origin is:

$$R_o(X,Y)Z = -[[X,Y],Z], \quad \text{for} \quad X,Y,Z \in \mathfrak{p}$$

(see Helgason [273, p. 180]).

Then one has the **classification of types of symmetric spaces M by sectional curvature** (see Helgason [273, p. 205]):

M is of noncompact type $\Leftrightarrow$ the sectional curvature of M is ≤ 0;
M is of compact type $\Leftrightarrow$ the sectional curvature of M is ≥ 0;
M is of Euclidean type $\Leftrightarrow$ the sectional curvature of M is $= 0$.

It is possible to prove the **conjugacy of all maximal compact subgroups of noncompact semisimple real Lie groups** G using Cartan's fixed point theorem, which says that if a compact group K_1 acts on a simply connected Riemannian manifold of negative curvature such as G/K, there must be a fixed point. And $x^{-1}K_1x \subset K$ means xK is fixed. See Helgason [273, p. 75] for more details on Cartan's theorem.

Table 2.1 Irreducible Riemannian symmetric spaces of types I and III for the non-exceptional groups

	Noncompact	Compact
AI	$SL(n,\mathbb{R})/SO(n)$	$SU(n)/SO(n)$
AII	$SU^*(2n)/Sp(n)$	$SU(2n)/Sp(n)$
AIII	$SU(p,q)/S\left(U_p \times U_q\right)$	$SU(p+q)/S\left(U_p \times U_q\right)$
BDI	$SO_o(p,q)/SO(p) \times SO(q)$	$SO(p+q)/SO(p) \times SO(q)$
DIII	$SO^*(2n)/U(n)$	$SO(2n)/U(n)$
CI	$Sp(n,\mathbb{R})/U(n)$	$Sp(n)/U(n)$
CII	$Sp(p,q)/Sp(p) \times Sp(q)$	$Sp(p+q)/Sp(p) \times Sp(q)$

The grand finale of the classification theory is the listing of the four types of irreducible symmetric spaces given in Helgason [273, Ch. 9] and [278, pp. 515–518]. Once more, irreducible means that the corresponding Lie group is simple.
The Four Types of Irreducible Symmetric Spaces are:

I. G/K, where G is a compact connected simple real Lie group and K is the subgroup of points fixed by an involutive automorphism of G.
II. G is a compact, connected simple Lie group provided with a left and right invariant Riemannian structure unique up to constant factor.
III. G/K where G is a connected noncompact simple real Lie group and K is the subgroup of points fixed by an involutive automorphism of G (a maximal compact subgroup).
IV. G/U, where G is a connected Lie group whose Lie algebra is a simple Lie algebra over $\mathbb{C}$ viewed as a real Lie algebra, and U is a maximal compact subgroup of G.

The irreducible symmetric spaces of types I and III which come from non-exceptional Lie groups are in Table 2.1. In this table $SO^*(2n) = \{g \in SO(2n,\mathbb{C}) \mid {}^t\overline{g}J_n g = J_n\}$, where J_n is defined in part (4) of the list of real forms of $\mathfrak{sl}(n,\mathbb{C})$ and $SO_o(p,q)$ is the identity component of $SO(p,q) = \left\{g \in SL(n,\mathbb{R}) \bigm| {}^t g I_{p,q} g = I_{p,q}\right\}$, where $n = p+q$ and $I_{p,q}$ is defined in part (3) of the list of real forms of $\mathfrak{sl}(n,\mathbb{C})$.

2.1.3 Cartan, Iwasawa, and Polar Decompositions, Roots

From now on, our emphasis will be upon the symmetric space $Sp(n,\mathbb{R})/U(n)$. Our first task is to study the various realizations of this space. We begin with the realization as the space of **positive symplectic matrices**:

$$Sp(n,\mathbb{R})/U(n) \cong \mathcal{P}_n^* = \{Y \in Sp(n,\mathbb{R}) \mid Y \in \mathcal{P}_{2n}\}. \tag{2.17}$$

The proof of (2.17) involves the global or group level Cartan decomposition. Let $\mathfrak{g}$ be a noncompact semisimple (real) Lie algebra with Cartan decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$. Suppose that K and G are the corresponding connected Lie groups with Lie algebras $\mathfrak{k}$ and $\mathfrak{g}$, respectively. Then we have the **global Cartan decomposition**:

$$G = KP, \qquad P = \exp \mathfrak{p},$$

and G is diffeomorphic to $K \times \mathfrak{p}$. Note that G and K are Lie groups but P is not. The main idea of the proof of (2.17) is to use the Adjoint representation of G to deduce the Cartan decomposition of G from that for $GL(n, \mathbb{R})$, which is:

$$GL(n, \mathbb{R}) = O(n) \cdot \mathcal{P}_n \tag{2.18}$$

(see Exercise 1.1.5 of Section 1.1.2). Proofs of the general Cartan decomposition can be found in Helgason [273, p. 215] or Loos [408, Vol. I, p. 156]. We shall only consider the special case of interest.

Lemma 2.1.1 (The Cartan Decomposition for the Symplectic Group).

(a) The Cartan decomposition for $G = Sp(n, \mathbb{R})$ comes from the Cartan decomposition (2.18) for $GL(2n, \mathbb{R})$ by taking intersections; i.e.,

$$G = Sp(n, \mathbb{R}) = K \cdot \mathcal{P}_n^*, \quad \text{with } \mathcal{P}_n^* = \mathcal{P}_{2n} \cap G \text{ and } K = O(2n) \cap G.$$

(b) The maximal compact subgroup K of G given in part (a) can be identified with the unitary group $U(n) = \{g \in \mathbb{C}^{n\times n} \mid {}^t\overline{g}g = I\}$. It follows also that the symmetric space $G/K = Sp(n, \mathbb{R})/U(n)$ can be identified with $\mathcal{P}_n^*$.

Proof. (a) See Helgason [273, p. 345] and [278, p. 450].

Observe that (2.18) says that $g \in G$ can be written as $g = up$ with $u \in O(2n)$ and $p \in \mathcal{P}_{2n}$. We need to show that both u and p lie in $Sp(n, \mathbb{R})$. To see this, note that $p^2 = {}^tgg$. Moreover, $g \in G$ implies that ${}^tg^{-1}$ and ${}^tg$ both also lie in G (a situation really brought about by the existence of an involution of G with differential the Cartan involution of $\mathfrak{g}$). Thus $p^2 \in G$.

Now we must show that $p^2 \in G$ implies that p lies in G. To do this, note that G is a pseudoalgebraic group, meaning that there is a finite set of polynomials

$$f_j \in \mathbb{C}[X_1, \ldots, X_{4n^2}]$$

such that a matrix g lies in G if and only if g is a root of all the f_j. Now there is a rotation matrix $k \in O(2n)$ such that

$$k^{-1}p^2k = \begin{pmatrix} e^{h_1} & & 0 \\ & \ddots & \\ 0 & & e^{h_{2n}} \end{pmatrix}.$$

And $k^{-1}Gk$ is also a pseudoalgebraic group. Thus the diagonal matrices

$$k^{-1}p^{2r}k = \begin{pmatrix} e^{rh_1} & & 0 \\ & \ddots & \\ 0 & & e^{rh_{2n}} \end{pmatrix}$$

satisfy a certain set of polynomial equations for any integer r. But if an exponential polynomial

$$F(t) = \sum_{j=1}^{B} c_j \exp(b_j t)$$

vanishes for all integers t, then it must vanish for all real numbers t as well. Thus, in particular, p must lie in the group G, as will all elements

$$p_t = \exp(tX), \quad \text{for} \quad t \in \mathbb{R}, \quad \text{if} \quad p^2 = \exp(2X).$$

But then $p \in G$ implies that $u = gp^{-1} \in G$. This completes the proof of part (a)—except to show the uniqueness of the expression $g = up$ and the fact that G is diffeomorphic to $K \times \mathfrak{p}$. We leave these proofs as an **exercise**.

(b) To see that K is isomorphic to $U(n)$, proceed as follows. First recall that we have $K = G \cap O(2n)$. Thus if $J = J_n$ is as defined in Exercise 2.1.7, then

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in K \Leftrightarrow JM = MJ \text{ and } {}^tMJM = J$$

$$\Leftrightarrow C = -B,\ D = A,\ {}^tAB = {}^tBA \quad \text{and} \quad {}^tAA + {}^tBB = I.$$

The last statement is equivalent to saying that $A + iB \in U(n)$. Thus the identification of K and $U(n)$ on the group level is the same as that on the Lie algebra level which was discussed when we listed the Cartan decompositions corresponding to noncompact real forms of $\mathfrak{sp}(n, \mathbb{C})$. In fact,

$$\sigma \begin{pmatrix} A & B \\ -B & A \end{pmatrix} = A + iB$$

defines a mapping which preserves matrix multiplication as well as addition. The map σ identifies K with $U(n)$. A good reference for these things is Séminaire Cartan [547, Exp 3]. The proof of Lemma 2.1.1 is now complete. ■

Exercise 2.1.12. Fill in all the details in the proof of Lemma 2.1.1.

Note that most calculations are far easier on the Lie algebra level than on the group level. For an example of the difference between the algebra and the group, note that it is clear that $\mathfrak{sp}(n,\mathbb{R})$ is contained in $\mathfrak{sl}(2n,\mathbb{R})$, but it is not obvious that $Sp(n,\mathbb{R})$ is contained in $SL(2n,\mathbb{R})$, though it is true.

Exercise 2.1.13. Prove the last statement.
Hint. Show that $Sp(n,\mathbb{R})$ is connected. See Chevalley [104, p. 36] for the useful result which says that H and G/H connected implies G connected, where H is a closed subgroup of the topological group G.

Next we seek to generalize the **Iwasawa decomposition** from Exercise 1.2.12 in Section 1.1.3:

$$G = GL(n,\mathbb{R}) = KAN, \tag{2.19}$$

where K is the compact group $O(n)$, A is the abelian group of positive diagonal matrices in G, and N is the nilpotent group of upper triangular matrices in G with ones on the diagonal. In order to obtain such an Iwasawa decomposition for any noncompact semisimple (real) Lie group G, one must discuss the **root space decomposition** of the Lie algebra of G. We do not give a detailed discussion of root spaces, except for several examples. The details for the general case can be found in Helgason [273, 278] or Loos [408].

Some definitions are needed to discuss the root space decomposition of the Lie algebra $\mathfrak{g}$ of G. Define $\mathfrak{a}$ to be a **maximal abelian subspace of** $\mathfrak{p}$. Here $\mathfrak{p}$ comes from the Cartan decomposition of $\mathfrak{g}$. Then for any **real linear functional (root)** $\alpha : \mathfrak{a} \to \mathbb{R}$, define the **root space**:

$$\mathfrak{g}_\alpha = \{X \in \mathfrak{g} \mid (\operatorname{ad} H)X = \alpha(H)X, \ \text{ for all } \ H \in \mathfrak{a}\}.$$

If $\mathfrak{g}_\alpha \neq \{0\}$, then we say that the linear functional α is a **restricted root**. Let Λ denote the set of all **nonzero restricted roots**. When we are considering a normal real form such as $\mathfrak{sl}(n,\mathbb{R})$, the **restricted roots** are restrictions of roots of the complexification of $\mathfrak{g}$.

Next set $\mathfrak{m}$ equal to the **centralizer** of $\mathfrak{a}$ in $\mathfrak{k}$, where $\mathfrak{k}$ comes from the Cartan decomposition of $\mathfrak{g}$; i.e.,

$$\mathfrak{m} = \{X \in \mathfrak{k} \mid [X,\mathfrak{a}] = 0\}.$$

In fact, $\mathfrak{m}$ will always be zero for normal or split real forms such as $\mathfrak{sp}(n,\mathbb{R})$.

Finally, **the root space decomposition of the real noncompact semisimple Lie algebra** $\mathfrak{g}$ is:

$$\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{m} \oplus \sum_{\alpha\in\Lambda}{}^{\oplus} \mathfrak{g}_\alpha.$$

To prove the validity of this decomposition, consider the positive definite bilinear form F on $\mathfrak{g}$ defined as follows, using the Killing form B and the Cartan involution θ of $\mathfrak{g}$:

$$F(X,Y) = -B(X,\theta Y), \quad \text{for} \quad X,Y \in \mathfrak{g}. \tag{2.20}$$

If $X \in \mathfrak{p}$, then ad$X$ is symmetric with respect to F and thus is a diagonalizable linear transformation of $\mathfrak{g}$. Therefore the commuting family of all the ad$X$ for $X \in \mathfrak{a}$ is simultaneously diagonalizable with real eigenvalues. It remains to show that the eigenspace corresponding to the zero functional is:

$$\mathfrak{g}_0 = (\mathfrak{g}_0 \cap \mathfrak{k}) \oplus (\mathfrak{g}_0 \cap \mathfrak{p}) \ = \mathfrak{m} \oplus \mathfrak{a}.$$

This comes from the definitions.

Note that if $\mathfrak{g}$ is $\mathfrak{sl}(n,\mathbb{R})$ or $\mathfrak{sp}(n,\mathbb{R})$, then $\mathfrak{m} = \{0\}$ and restricted roots are the same as the roots of the complexifications $\mathfrak{sl}(n,\mathbb{C})$ and $\mathfrak{sp}(n,\mathbb{C})$ restricted to the normal real form.

One can define the set of **positive restricted roots** Λ^+ as a subset of Λ such that Λ is the disjoint union of Λ^+ and $-\Lambda^+$. We will soon see how to find such sets of positive roots in our favorite cases.

We need to use the positive roots to construct a certain nilpotent Lie subalgebra $\mathfrak{n}$ of $\mathfrak{g}$. By definition, a Lie algebra $\mathfrak{n}$ is said to be **nilpotent** if the lower central series $\mathfrak{n}^k$ defined by

$$\mathfrak{n}^0 = \mathfrak{n}, \ \mathfrak{n}^1 = [\mathfrak{n},\mathfrak{n}], \ \mathfrak{n}^{k+1} = [\mathfrak{n},\mathfrak{n}^k]$$

terminates; i.e., $\mathfrak{n}^k = \{0\}$, for some k.

Suppose that Λ^+ denotes the chosen set of positive roots of $\mathfrak{g}$. Define the **nilpotent Lie subalgebra** $\mathfrak{n}$ *of* $\mathfrak{g}$ by:

$$\mathfrak{n} = \sum_{\alpha\in\Lambda^+}{}^{\oplus}\, \mathfrak{g}_\alpha.$$

We can also define the **opposite nilpotent subalgebra** $\overline{\mathfrak{n}}$ of $\mathfrak{g}$ by:

$$\overline{\mathfrak{n}} = \sum_{\alpha\in\Lambda^+}{}^{\oplus}\, \mathfrak{g}_{-\alpha}.$$

To prove that $\mathfrak{n}$ is nilpotent, it suffices to know the following simple facts about roots.

Simple Facts About Roots

(1) Λ^+ is a finite set;
(2) $[\mathfrak{g}_\alpha,\mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta}$;
(3) $\alpha,\beta \in \Lambda^+$ implies that $\alpha+\beta$ is either a positive root or not a root at all.

Furthermore, if θ denotes the Cartan involution of $\mathfrak{g} = \mathfrak{k}\oplus\mathfrak{p}$, then θ interchanges the nilpotent algebra $\mathfrak{n}$ and its opposite $\overline{\mathfrak{n}}$; i.e., $\theta\mathfrak{n} = \overline{\mathfrak{n}}$. To see this, note that $\theta(X) = -X$ for all X in $\mathfrak{a} \subset \mathfrak{p}$ and θ preserves the Lie bracket.

From the preceding considerations, it is easy to obtain the **Iwasawa decomposition of the noncompact real semisimple Lie algebra**:

$$\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{a} \oplus \mathfrak{n}.$$

For clearly, one has $\mathfrak{g} = \overline{\mathfrak{n}} \oplus \mathfrak{m} \oplus \mathfrak{a} \oplus \mathfrak{n}$. And $\mathfrak{k} \oplus \mathfrak{a} \oplus \mathfrak{n}$ is a direct sum, since

$$X + H + Y = 0, \quad \text{for } X \in \mathfrak{k},\ H \in \mathfrak{a},\ Y \in \mathfrak{n},$$

implies that

$$0 = \theta(X + H + Y) = X - H + \theta(Y).$$

Subtract the two equations to see that $2H + Y - \theta(Y) = 0$. This implies that $H = 0$ by the fact that the root space decomposition is a direct sum. Thus $Y = \theta(Y) = 0$ and $X = 0$.

To complete the proof of the Lie algebra Iwasawa decomposition, we need to only show that the dimensions are correct. It suffices to look at the following mapping:

$$\begin{aligned} &\mathfrak{m} \oplus \overline{\mathfrak{n}} \to \mathfrak{k}, && 1-1, \text{ onto} \\ &X + Y \mapsto X + Y + \theta(Y), && \text{for } \ X \in \mathfrak{m},\ Y \in \overline{\mathfrak{n}}. \end{aligned}$$

Next we want to consider three examples: $\mathfrak{sl}(n, \mathbb{R})$, $\mathfrak{sp}(n, \mathbb{R})$, and $\mathfrak{su}(3, 1)$. The first two examples are **split** or **normal**, so that $\mathfrak{m} = \{0\}$, the restricted roots are restrictions of complex roots of the complexification, and all the roots spaces are one-dimensional real vector spaces.

Three Examples of Iwasawa Decompositions of Real Semisimple Lie Algebras

Example 2.1.5 ($\mathfrak{sl}(n, \mathbb{R})$)**.**

Recall that the Cartan decomposition is $\mathfrak{sl}(n, \mathbb{R}) = \mathfrak{k} \oplus \mathfrak{p}_n$, where

$$\begin{aligned} \mathfrak{k} &= \mathfrak{so}(n) = \left\{ X \in \mathbb{R}^{n \times n} \mid {}^t X = -X \right\}, \\ \mathfrak{p}_n &= \left\{ X \in \mathbb{R}^{n \times n} \mid {}^t X = X,\ \operatorname{Tr} X = 0 \right\}. \end{aligned}$$

One can show that a maximal abelian subspace of $\mathfrak{p}_n$ is:

$$\mathfrak{a} = \left\{ H \in \mathbb{R}^{n \times n} \mid H \text{ is diagonal of trace } 0 \right\}.$$

Next let E_{ij} for $1 \leq i,j \leq n$ denote the matrix with 1 in the i,j place and 0's elsewhere. Then set $e_i(H) = h_i$ if H is a diagonal matrix with h_i as its ith diagonal entry. Let $\alpha_{ij} = e_i - e_j$. Then $[H, E_{ij}] = \alpha_{ij}(H)E_{ij}$ and we find the root space decomposition involves the

$$\mathfrak{g}_{\alpha_{ij}} = \mathbb{R}E_{ij} \text{ with } \Lambda^+ = \left\{\alpha_{ij} \mid 1 \leq i < j \leq n\right\}.$$

Thus

$$\mathfrak{n} = \sum_{1 \leq i < j \leq n}^{\oplus} \mathbb{R}E_{ij} = \text{the upper triangular real } n{\times}n \text{ matrices with 0 on the diagonal.}$$

It follows that the Lie algebra analogue of the Iwasawa decomposition of $SL(n, \mathbb{R})$ coming from (2.19) says:

$$\mathfrak{sl}(n, \mathbb{R}) = \mathfrak{so}(n) \oplus \mathfrak{a} \oplus \mathfrak{n},$$

with $\mathfrak{so}(n)$ denoting the skew symmetric $n \times n$ real matrices, $\mathfrak{a}$ equal to the $n \times n$ real diagonal trace zero matrices, and $\mathfrak{n}$ equal to the upper triangular $n \times n$ real matrices with zeros on the diagonal.

Example 2.1.6 ($\mathfrak{sp}(n, \mathbb{R})$).
Recall that $\mathfrak{sp}(n, \mathbb{R})$ consists of matrices

$$(A, B, C) \doteqdot \begin{pmatrix} A & B \\ C & -{}^tA \end{pmatrix},$$

with $A, B, C \in \mathbb{R}^{n \times n}$ and B, C symmetric. We found the Cartan decomposition had:

$$\begin{aligned} \mathfrak{k} &= \{(A, B, -B) \mid B \text{ symmetric, } A \text{ skew symmetric}\}, \\ \mathfrak{p} &= \{(A, B, B) \mid A, B \text{ symmetric}\}. \end{aligned}$$

A calculation shows that a maximal abelian subspace of $\mathfrak{p}$ is:

$$\mathfrak{a} = \{(H, 0, 0) \mid H \text{ is real } n \times n \text{ diagonal}\}.$$

Suppose that the $E_{ij},\ 1 \leq i < j \leq n$, are as in Example 2.1.5. Set $G_{pq} = E_{pq} + E_{qp}$, for $1 \leq p \leq q \leq n$. Then, if we abuse notation and write $H = (H, 0, 0)$, we have

$$\begin{aligned} [H, (E_{ij}, 0, 0)] &= (e_i - e_j)(H)(E_{ij}, 0, 0), \\ [H, (0, G_{pq}, 0)] &= (e_p + e_q)(H)(0, G_{pq}, 0), \\ [H, (0, 0, G_{pq})] &= -(e_p + e_q)(H)(0, 0, G_{pq}). \end{aligned}$$

It follows that we can take $\Lambda^+ = \left\{e_i - e_j \mid 1 \leq i < j \leq n\right\} \cup \left\{e_p + e_q \mid 1 \leq p \leq q \leq n\right\}$. Thus

$$
\begin{aligned}
\mathfrak{n} &= \sum_{1\le i<j\le n}{}^{\oplus}\, \mathbb{R}(E_{ij},0,0) + \sum_{1\le p\le q\le n}{}^{\oplus}\, \mathbb{R}(0,G_{pq},0) \\
&= \{(A,B,0) \mid A \text{ upper triangular, 0 on diagonal, } B \text{ symmetric}\}\,.
\end{aligned}
$$

So the **Iwasawa decomposition** is:

$$\mathfrak{sp}(n,\mathbb{R}) = \mathfrak{k} \oplus \mathfrak{a} \oplus \mathfrak{n},$$

where

$$
\begin{aligned}
\mathfrak{k} &= \{(A,B,-B) \mid B \text{ symmetric, } A \text{ skew-symmetric}\}\,, \\
\mathfrak{a} &= \{(H,0,0) \mid H \text{ diagonal}\}\,, \\
\mathfrak{n} &= \{(A,B,0) \mid A \text{ upper triangular with 0 on the diagonal, } B \text{ symmetric}\}\,.
\end{aligned}
$$

Example 2.1.7 **(** $\mathfrak{su}(3,1)$**).**
First recall that

$$\mathfrak{su}(3,1) = \left\{X \in \mathfrak{sl}(4,\mathbb{C}) \;\middle|\; -I_{3,1}\, {}^t\overline{X}\, I_{3,1} = X\right\}$$

where

$$I_{3,1} = \begin{pmatrix} 1 & & & 0 \\ & 1 & & \\ & & 1 & \\ 0 & & & -1 \end{pmatrix}.$$

The corresponding Lie group is $SU(3,1) = \{g \in SL(4,\mathbb{C}) \mid {}^t\overline{g} I_{3,1} g = I_{3,1}\}$.
One sees easily that

$$\mathfrak{su}(3,1) = \left\{(A,b,c) \;\middle|\; A \in \mathfrak{u}(3),\ b \in \mathbb{R},\ c \in \mathbb{C}^{3\times 1},\ \mathrm{Tr}A + ib = 0\right\},$$

where

$$(A,b,c) = \begin{pmatrix} A & c \\ {}^t\overline{c} & ib \end{pmatrix}.$$

We saw that the Cartan decomposition of $\mathfrak{su}(3,1)$ involves

$$
\begin{aligned}
\mathfrak{k} &= \{(A,b,0) \mid A \in \mathfrak{u}(3),\ b \in \mathbb{R},\ \mathrm{Tr}A + ib = 0\}\,, \\
\mathfrak{p} &= \left\{(0,0,c) \;\middle|\; c \in \mathbb{C}^{3\times 1}\right\}.
\end{aligned}
$$

A maximal abelian subspace of $\mathfrak{p}$ is $\mathfrak{a} = \mathbb{R}(0, 0, e_1)$ where $e_1 = {}^t(1, 0, 0)$. Note that $(0, 0, ie_1)$ does not commute with $(0, 0, e_1)$. You need to multiply matrices to check these things (see Exercise 2.1.14 below). Similarly you can show that:

$$\mathfrak{m} = \left\{ \begin{pmatrix} ib & 0 & 0 & 0 \\ 0 & u_1 & u_2 & 0 \\ 0 & u_3 & u_4 & 0 \\ 0 & 0 & 0 & ib \end{pmatrix} \middle| \, b \in \mathbb{R}, \ U = \begin{pmatrix} u_1 & u_2 \\ u_3 & u_4 \end{pmatrix} \in \mathfrak{u}(2), \ \mathrm{Tr}U + 2ib = 0 \right\}.$$

To prove this, one must show that the matrices of $\mathfrak{m}$ centralize $\mathfrak{a}$ and that nothing else in $\mathfrak{k}$ does the same trick. Once again, this is checked by multiplying matrices. Thus we have come upon an example of a nonzero and rather fat $\mathfrak{m}$. The root space decomposition of $\mathfrak{su}(3, 1)$ is rather complicated. We find roots λ such that 2λ is also a root. Such things cannot happen for complex semisimple Lie algebras. And one finds root spaces $\mathfrak{g}_\lambda$ of dimension greater than one over $\mathbb{R}$.

The positive roots of $\mathfrak{su}(3, 1)$ are $\Lambda^+ = \{\lambda, 2\lambda\}$, where $\lambda((0, 0, e_1)) = 1$. And the root space decomposition of $\mathfrak{su}(3, 1)$ is:

$$\begin{aligned} &\mathfrak{su}(3, 1) = \mathfrak{a} \oplus \mathfrak{m} \oplus \mathfrak{g}_\lambda \oplus \mathfrak{g}_{-\lambda} \oplus \mathfrak{g}_{2\lambda} \oplus \mathfrak{g}_{-2\lambda} \\ &\dim_{\mathbb{R}} \mathfrak{a} = 1, \quad \dim_{\mathbb{R}} \mathfrak{m} = 4, \quad \dim_{\mathbb{R}} \mathfrak{g}_\lambda = \dim_{\mathbb{R}} \mathfrak{g}_{-\lambda} = 4, \\ &\dim_{\mathbb{R}} \mathfrak{g}_{2\lambda} = \dim_{\mathbb{R}} \mathfrak{g}_{-2\lambda} = 1. \end{aligned}$$

To see this, note that

$$\left[\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & a & b & 0 \\ -\overline{a} & 0 & 0 & c \\ -\overline{b} & 0 & 0 & d \\ 0 & \overline{c} & \overline{d} & 0 \end{pmatrix} \right] = \begin{pmatrix} 0 & \overline{c} & \overline{d} & 0 \\ -c & 0 & 0 & \overline{a} \\ -d & 0 & 0 & \overline{b} \\ 0 & a & b & 0 \end{pmatrix} = k \begin{pmatrix} 0 & a & b & 0 \\ -\overline{a} & 0 & 0 & c \\ -\overline{b} & 0 & 0 & d \\ 0 & \overline{c} & \overline{d} & 0 \end{pmatrix}$$

implies that $k = \pm 1$ and that $ka = \overline{c}$, $kb = \overline{d}$. Thus, if $k = 1$, we find that $\mathfrak{g}_\lambda$ is four-dimensional over $\mathbb{R}$:

$$\mathfrak{g}_\lambda = \left\{ \begin{pmatrix} 0 & a & b & 0 \\ -\overline{a} & 0 & 0 & \overline{a} \\ -\overline{b} & 0 & 0 & \overline{b} \\ 0 & a & b & 0 \end{pmatrix} \middle| \, (a, b) \in \mathbb{C}^2 \right\}.$$

If $k = -1$, we find that again $\mathfrak{g}_{-\lambda}$ is four-dimensional over $\mathbb{R}$:

$$\mathfrak{g}_{-\lambda} = \left\{ \begin{pmatrix} 0 & a & b & 0 \\ -\overline{a} & 0 & 0 & -\overline{a} \\ -\overline{b} & 0 & 0 & -\overline{b} \\ 0 & -a & -b & 0 \end{pmatrix} \middle| \, (a, b) \in \mathbb{C}^2 \right\}.$$

Recalling what it means to be in $\mathfrak{su}(3,1)$, we see that it remains to deal with

$$\mathfrak{g}_{2\lambda} = \mathbb{R}\begin{pmatrix} i & 0 & 0 & -i \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ i & 0 & 0 & -i \end{pmatrix} \text{ and } \mathfrak{g}_{-2\lambda} = \mathbb{R}\begin{pmatrix} i & 0 & 0 & i \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -i & 0 & 0 & -i \end{pmatrix}.$$

The nilpotent Lie algebra $\mathfrak{n}$ is then:

$$\mathfrak{n} = \mathfrak{g}_\lambda \oplus \mathfrak{g}_{2\lambda} = \left\{ \begin{pmatrix} ic & a & b & -ic \\ -\overline{a} & 0 & 0 & \overline{a} \\ -\overline{b} & 0 & 0 & \overline{b} \\ ic & a & b & -ic \end{pmatrix} \middle| \; a, b \in \mathbb{C},\ c \in \mathbb{R} \right\}.$$

Exercise 2.1.14. (a) Check the calculations in the preceding three examples.
(b) Perform the analogous calculation to that of part (a) in the case of the Lorentz algebra $\mathfrak{so}(3,1)$.

Our next goal is to understand **the group level Iwasawa decomposition of a noncompact semisimple connected real Lie group G with finite center**:

$$G = KAN,$$

where K, A, N are connected Lie subgroups of G with Lie algebras $\mathfrak{k}, \mathfrak{a}, \mathfrak{n}$, respectively. G is actually diffeomorphic to the product $K \times A \times N$. The exponential maps $\mathfrak{k}$ onto the compact group K. And exp is a diffeomorphism which maps $\mathfrak{a}$ onto the abelian group A while taking addition to multiplication. The exponential is a diffeomorphism of $\mathfrak{n}$ onto the nilpotent group N. Recall that the exponential does not in general map $\mathfrak{g}$ onto G, nor is exp a *diffeomorphism* in general (see Exercise 2.1.6). For a proof that the exponential map is onto for abelian, nilpotent, and compact Lie groups, see Helgason [273, pp. 229, 56–58, 188–189].

In our discussion of the group level Iwasawa decomposition, we shall only consider the special case of the symplectic group. The proof of the global Iwasawa decomposition in the general case uses the Adjoint representation (see Helgason [273, 278] or Loos [408]).

Recall the following definition:

$$Sp(n,\mathbb{R}) = \text{the symplectic group} = \left\{ g \in SL(2n,\mathbb{R}) \;\middle|\; {}^t g J_n g = J_n \right\},$$

where

$$J_n = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}.$$

It follows that

$$Sp(n,\mathbb{R}) = \left\{ \begin{pmatrix} A & B \\ C & D \end{pmatrix} \middle| \ {}^tAC = {}^tCA,\ {}^tBD = {}^tDB,\ {}^tAD - {}^tCB = I_n \right\}. \tag{2.21}$$

And **Lie subgroups** of $Sp(n,\mathbb{R})$ **which correspond to the Lie subalgebras** $\mathfrak{k}, \mathfrak{a}, \mathfrak{n}$ **in the Iwasawa decomposition** $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{a} \oplus \mathfrak{n}$ are:

$$\left.\begin{aligned} K_n^* &= \left\{ \begin{pmatrix} A & B \\ -B & A \end{pmatrix} \middle| \ A + iB \in U(n) \right\}, \\ A_n^* &= \left\{ \begin{pmatrix} H & 0 \\ 0 & H^{-1} \end{pmatrix} \middle| \ H \text{ positive diagonal} \right\}, \\ N_n^* &= \left\{ \begin{pmatrix} A & B \\ 0 & {}^tA^{-1} \end{pmatrix} \middle| \ A \text{ upper triangular; } A\,{}^tB = B\,{}^tA \right\}. \end{aligned}\right\} \tag{2.22}$$

We have seen that the **symmetric space** associated with $Sp(n,\mathbb{R})$ is:

$$Sp(n,\mathbb{R})/K_n^* \cong \mathcal{P}_n^* = \mathcal{P}_{2n} \cap Sp(n,\mathbb{R}) \tag{2.23}$$

(see Lemma 2.1.1). Now we wish to find (along with the Iwasawa decomposition) another realization of this symmetric space—**Siegel's upper half space**:

$$\mathcal{H}_n = \left\{ Z \in \mathbb{C}^{n\times n} \mid {}^tZ = Z,\ \operatorname{Im} Z \in \mathcal{P}_n \right\}. \tag{2.24}$$

Observe that we can define the following **actions** of $G = Sp(n,\mathbb{R})$ on the three versions of the symmetric space:

$$\left.\begin{aligned} &\text{action of } g \in G \text{ on } G/K \text{ is } a_g(xK) = gxK \quad \text{for } x \in G; \\ &\text{action of } g \in G \text{ on } \mathcal{P}_n^* \text{ is } b_g(Y) = Y[g] = {}^tgYg \ \text{ for } Y \in \mathcal{P}_n^*; \\ &\text{action of } g \in G \text{ on } \mathcal{H}_n \text{ is } c_g(Z) = (AZ+B)(CZ+D)^{-1} \\ &\qquad\qquad\qquad\qquad \text{for } g = \begin{pmatrix} A & B \\ C & D \end{pmatrix} Z \in \mathcal{H}_n. \end{aligned}\right\} \tag{2.25}$$

Exercise 2.1.15. (a) Prove formula (2.21).
(b) Check that $c_{hg}(Z) = c_h(c_g(Z))$ in formula (2.25).

The following lemma will allow us to identify all the versions of the symmetric space associated with the symplectic group. To see this, study the following diagram of mappings:

$$\left.\begin{matrix} Sp(n,\mathbb{R})/K_n^* & \to & \mathcal{P}_n^* & \to & \mathcal{H}_n \\ gK_n^* & \mapsto & g\,{}^tg & \mapsto & X + iY. \end{matrix}\right\} \tag{2.26}$$

Here X, Y come from the **partial Iwasawa decomposition** of $S \in \mathcal{P}_n^*$:

$$S = \begin{pmatrix} Y & 0 \\ 0 & Y^{-1} \end{pmatrix} \begin{bmatrix} I & 0 \\ X & I \end{bmatrix}, \quad \text{for } \; X = {}^tX \in \mathbb{R}^{n\times n}, \; Y \in \mathcal{P}_n. \tag{2.27}$$

Lemma 2.1.2 below gives the existence and uniqueness of this decomposition for every positive symplectic matrix. There is an equivalent Iwasawa decomposition obtained by applying matrix inverse to formula (2.27):

$$S = \begin{pmatrix} Y^{-1} & 0 \\ 0 & Y \end{pmatrix} \begin{bmatrix} I & -X \\ 0 & I \end{bmatrix}, \quad \text{for } \; X = {}^tX \in \mathbb{R}^{n\times n}, \; Y \in \mathcal{P}_n. \tag{2.28}$$

Exercise 2.1.16. (a) Show that the composition of the two maps in (2.26) takes gK_n^* with

$$g = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

to $(Ai + B)(Ci + D)^{-1}$ in $\mathcal{H}_n$.

(b) Show also that the maps in (2.26) preserve the group actions in (2.25). More precisely, define $i_1\left({}^tgK_n\right) = I[g]$ for $g \in G$ and define $i_2(S) = X + iY$, for S with partial Iwasawa decomposition (2.27). Prove that

$$i_1 \circ a_g = b_{\,{}^tg} \circ i_1 \;\text{ and }\; i_2 \circ b_{\,{}^tg} = c_g \circ i_2.$$

(c) Suppose that $Z^* = c_g(Z) = (AZ + B)(CZ + D)^{-1}$ for $Z \in \mathcal{H}_n$ and g as in part (a). Show that the imaginary part of Z^* is $Y^* = Y\{(CZ + D)^{-1}\}$, where Y is the imaginary part of Z and $Y\{W\} = {}^t\overline{W}YW$. Here $\overline{W}$ is the matrix obtained from W by complex conjugation of all the entries of W. Then show that $Z^* \in \mathcal{H}_n$.

(d) Show that the Jacobian $|\partial Z^*/\partial Z| = |CZ + D|^{-n-1}$.

Hints. See Maass [426, p. 33].

(a) Note that

$$\begin{aligned} Z &= (Ai + B)(Ci + D)^{-1} \\ &= (Ai + B)\left(-{}^tCi + {}^tD\right)\left(-{}^tCi + {}^tD\right)^{-1}(Ci + D)^{-1}. \end{aligned}$$

(c) Note that $c_g(W) - c_g(Z) = \left(W\,{}^tC + {}^tD\right)^{-1}(W - Z)(CZ + D)^{-1}$. To find Y^*, let $W = \overline{Z}$.

Lemma 2.1.2 (Iwasawa Decomposition for the Symplectic Group). *Here we use the notation (2.21)–(2.28).*

(a) Every positive symplectic matrix has the partial Iwasawa decomposition given in (2.27) or (2.28). Thus the mappings in (2.26) are identifications of the three differentiable manifolds.

(b) The Iwasawa decomposition of $G = Sp(n, \mathbb{R})$ *says that*

$$G = K_n^* A_n^* N_n^*, \quad \textit{with} \quad K_n^*, A_n^*, N_n^* \textit{ as in (2.22).}$$

Proof. (a) We know from (2.19) that we can write $S \in \mathcal{P}_n^*$ as:

$$S = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \begin{bmatrix} I & X \\ 0 & I \end{bmatrix} \quad \text{with } A, B \in \mathcal{P}_n, \; X \in \mathbb{R}^{n \times n}.$$

Since S is symplectic, it follows that for J_n as defined in Exercise 2.1.7, we have $SJ_nS = J_n$. Thus $J_nSJ_n = -S^{-1}$.

The only way for J_nSJ_n to be equal to $-S^{-1}$ when S has the given partial Iwasawa decomposition is that

$$A = B^{-1} \quad \text{and} \quad X = {}^tX.$$

This is easily seen using again the fact that $J_n^2 = -I$. For

$$J_nSJ_n = J_n \begin{pmatrix} I & 0 \\ {}^tX & I \end{pmatrix} J_nJ_n \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} J_nJ_n \begin{pmatrix} I & X \\ 0 & I \end{pmatrix} J_n = -\begin{pmatrix} B & 0 \\ 0 & A \end{pmatrix} \begin{bmatrix} I & 0 \\ -X & I \end{bmatrix}.$$

(b) Use part (a). This allows one to write $S \in \mathcal{P}_n^*$ as

$$S = \begin{pmatrix} A & 0 \\ 0 & A^{-1} \end{pmatrix} \begin{bmatrix} I & B \\ 0 & I \end{bmatrix}, \qquad \text{for } A \in \mathcal{P}_n, \; B = {}^tB \in \mathbb{R}^{n \times n}.$$

Then express A as $A = H[Q]$ with H positive diagonal and Q upper triangular with 1's on the diagonal. This is possible by the Iwasawa decomposition for $GL(n, \mathbb{R})$. Thus

$$Y = \begin{pmatrix} H & 0 \\ 0 & H^{-1} \end{pmatrix} \begin{bmatrix} Q & QB \\ 0 & {}^tQ^{-1} \end{bmatrix},$$

which is the full Iwasawa decomposition of $Y \in \mathcal{P}_n^*$. This translates to the Iwasawa decomposition for an element of the symplectic group using the Cartan decomposition (Lemma 2.1.1), completing the proof of Lemma 2.1.2.

■

Exercise 2.1.17. Show that the Killing form on $\mathfrak{g} = \mathfrak{sp}(n, \mathbb{R})$ is

$$B(X, Y) = 4(n+1)\mathrm{Tr}(XY).$$

Hint. Use the root space decomposition of $\mathfrak{g}$.

The Riemannian metric on $Sp(n,\mathbb{R})/K_n^* \cong \mathcal{P}_n^*$ is given by:

$$Q_Y(u,v) = \mathrm{Tr}(Y^{-1}uY^{-1}v),$$

for $Y \in \mathcal{P}_n^*,\ u,v \in T_Y(\mathcal{P}_n^*) =$ the tangent space to $\mathcal{P}_n^*$ at Y (see (2.14) and the analogous result for $SL(n,\mathbb{R})$). Here we have dropped the constant in the Killing form of Exercise 2.1.17. Using the notation of formula (1.11) in Section 1.1.3, if $dY = (dy_{ij}) \in T_Y(\mathcal{P}_n^*),\ Y \in \mathcal{P}_n^*$, then the **arc length** on $\mathcal{P}_n^*$ is

$$ds^2 = Tr(Y^{-1}dY\, Y^{-1}dY). \tag{2.29}$$

We want to show that the geodesics for this metric come from matrix exp and thus that $\mathcal{P}_n^*$ is a totally geodesic submanifold of $\mathcal{P}_{2n}$. We can use partial Iwasawa coordinates from Lemma 2.1.2 for this purpose. Now $W \in \mathcal{P}_n^*$ has partial Iwasawa decomposition

$$W = \begin{pmatrix} V & 0 \\ 0 & V^{-1} \end{pmatrix} \begin{bmatrix} I & X \\ 0 & I \end{bmatrix}, \quad V \in \mathcal{P}_n, \quad X = {}^tX \in \mathbb{R}^n. \tag{2.30}$$

Just as in Exercise 1.1.14 of Section 1.1.3, we obtain the following formula for the **arc length** on $\mathcal{P}_n^*$ **in partial Iwasawa coordinates** (2.30):

$$ds^2 = \mathrm{Tr}\left(\left(V^{-1}dV\right)^2 + \left(V\,d(V^{-1})\right)^2 + 2\left(V^{-1}\right)\,\left({}^t dX\right)\,V^{-1}\,dX\right). \tag{2.31}$$

Exercise 2.1.18. Prove formula (2.31). Then note that $d(V^{-1}) = -V^{-1}\,dV\,V^{-1}$ and use this to show that the arc length on $\mathcal{P}_n^*$ can be expressed as follows using partial Iwasawa coordinates (2.30):

$$ds^2 = 2\,\mathrm{Tr}\left(\left(V^{-1}dV\right)^2 + \left(V^{-1}\,dX\right)^2\right).$$

Show that the action of $G = Sp(n,\mathbb{R})$ on $\mathcal{P}_n^*$ leaves the arc length invariant.

Using Exercise 2.1.18 we find that the **arc length on the Siegel upper half space** $\mathcal{H}_n$ is:

$$ds^2 = 2\,\mathrm{Tr}\left(V^{-1}dZ\,V^{-1}d\overline{Z}\right), \quad \text{if } Z = U + iV \in \mathcal{H}_n. \tag{2.32}$$

This is indeed the arc length considered by Siegel [565, Vol. II, p. 276].

Before proceeding to the study of geodesics in $\mathcal{P}_n^*$ or $\mathcal{H}_n$, we need to consider the analogue of polar coordinates in these spaces.

Lemma 2.1.3 (The Polar Decomposition of a Noncompact Semisimple Real Lie Group). *Let G be a noncompact connected real semisimple Lie group with connected Lie subgroups K and A, as in the Cartan and Iwasawa decompositions. Then G has the polar decomposition:*

$$G = KAK.$$

Proof. First we show that if the Lie algebra $\mathfrak{g}$ of G has Cartan decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ and Iwasawa decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{a} \oplus \mathfrak{n}$, then

$$\mathfrak{p} = \mathrm{Ad}(K)\mathfrak{a}, \tag{2.33}$$

where the Adjoint representation Ad is defined in formula (2.11). To prove (2.33), choose H in $\mathfrak{a}$ so that its centralizer in $\mathfrak{p}$ is $\mathfrak{a}$; i.e., take $H \in \mathfrak{a}$ such that $\alpha(H) \neq 0$ for all roots $\alpha \in \Lambda$. Set K^* equal to $\mathrm{Ad}_G K$ and suppose that X is in $\mathfrak{p}$. Now there is an element k_0 in K^* such that

$$B(H, \mathrm{Ad}(k_0)X) = \mathrm{Min}\left\{B(H, \mathrm{Ad}(k)X) \mid k \in K^*\right\}.$$

Suppose that $T \in \mathfrak{k}$. Then the derivative at $t = 0$ of the following function $f(t)$ of the real variable t must be 0 by the first derivative test:

$$f(t) = B(H, \mathrm{Ad}(\exp tT)\ \mathrm{Ad}(k_0)X).$$

This implies using the fact that the derivative of Ad is ad:

$$B(H, (\mathrm{ad}T)(\mathrm{Ad}(k_0)X)) = 0 \text{ for all } T \text{ in } \mathfrak{k}.$$

Thus (by Exercise 2.1.3)

$$B(T, [H, \mathrm{Ad}(k_0)X]) = 0 \text{ for all } T \text{ in } \mathfrak{k}.$$

Since $[\mathfrak{p}, \mathfrak{p}] \subset \mathfrak{k}$, and B is negative definite on $\mathfrak{k}$, it follows that $[H, \mathrm{Ad}(k_0)X] = 0$ which says that $\mathrm{Ad}(k_0)X \in \mathfrak{a}$, by the definition of H. The proof of Lemma 2.1.3 is completed by observing that (2.33) implies

$$\exp \mathfrak{p} = \exp(\mathrm{Ad}(K))\mathfrak{a} = \mathrm{Int}(K)(\exp \mathfrak{a}) = \bigcup_{k \in K} kAk^{-1}.$$

Lemma 2.1.3 follows from this equality and Lemma 2.1.1. ∎

Next we consider some examples.

Examples of the Polar Decomposition

Example 2.1.8 ($GL(n, \mathbb{R}) = O(n)A_nO(n)$, Where A_n Consists of All Positive Diagonal Matrices).

This is equivalent (via the Cartan decomposition) to saying that for any positive matrix Y in $\mathcal{P}_n$, there is an orthogonal matrix k in $O(n)$ and a positive diagonal matrix a in A_n such that $Y = k^{-1}ak$. Thus the polar decomposition is just the **spectral theorem** for positive definite symmetric matrices, as we noted already in formula (1.22) of Section 1.1.4.

The next question is: **How unique are the a and k in the polar decomposition** of Y in $\mathcal{P}_n$? We saw in the paragraph after Exercise 1.1.24 of Section 1.1.4 that these coordinates give a $(2^n n!)$-fold covering of $\mathcal{P}_n$, since the entries of a are unique up to the action of the Weyl group of permutations of the diagonal entries and the matrices in $O(n)$ that commute with all the diagonal matrices must themselves be diagonal with entries ± 1.

Example 2.1.9 (Euler Angle Decomposition of the Compact Group $SO(3)$). A reference is Hermann [289, pp. 30–39]. Set $G = SO(3)$,

$$k = \begin{pmatrix} SO(2) & 0 \\ 0 & 1 \end{pmatrix}.$$

The Cartan decomposition of $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ is:

$$\mathfrak{so}(3) = \begin{pmatrix} \mathfrak{so}(2) & 0 \\ 0 & 0 \end{pmatrix} \oplus \left\{ \begin{pmatrix} 0 & c \\ -{}^t c & 0 \end{pmatrix} \middle| \, c \in \mathbb{R}^2 \right\}.$$

Then we can take the maximal abelian subspace of $\mathfrak{p}$ to be

$$\mathfrak{a} = \mathbb{R} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}.$$

And

$$\exp \left\{ t \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \right\} = \begin{pmatrix} \cos t & 0 & \sin t \\ 0 & 1 & 0 \\ -\sin t & 0 & \cos t \end{pmatrix}.$$

which is easily seen by writing out the series for the matrix exponential. Thus the **Euler angle decomposition** of g in $SO(3)$ is:

$$\begin{pmatrix} \cos u & \sin u & 0 \\ -\sin u & \cos u & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos t & 0 & \sin t \\ 0 & 1 & 0 \\ -\sin t & 0 & \cos t \end{pmatrix} \begin{pmatrix} \cos v & \sin v & 0 \\ -\sin v & \cos v & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The three Euler angles are u, t, v. Thus any rotation in 3-space is a product of three rotations about two axes.

Example 2.1.10 (The Dual Noncompact Group to $SO(3)$ Is the Lorentz Group $SO(2,1)$).

The group $SO(2,1)$ again has an Euler angle decomposition that is well known to physicists. You need two angular variables and one real variable. One finds that the maximal abelian subalgebra $\mathfrak{a}$ of $\mathfrak{p}$ is:

$$\mathfrak{a} = \mathbb{R}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Then

$$\exp\left\{t\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}\right\} = \begin{pmatrix} \cosh t & 0 & \sinh t \\ 0 & 1 & 0 \\ \sinh t & 0 & \cosh t \end{pmatrix}.$$

For $SO(3,1)$, these matrices are called "Lorentz boosts" (see Misner et al. [454, p. 67]). The A-part of this group does not get wound up like the A-part of the compact group in Example 2.1.9.

Example 2.1.11 (Euler Angles for $U(3,1)$).
The physicist Wigner [668] considers this example, for which KAK is:

$$\begin{pmatrix} A & 0 \\ 0 & u \end{pmatrix}\begin{pmatrix} \cosh t & 0 & 0 & \sinh t \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \sinh t & 0 & 0 & \cosh t \end{pmatrix}\begin{pmatrix} B & 0 \\ 0 & v \end{pmatrix},$$

for A, B in $U(3)$ and u, b in $i\mathbb{R}$.

Example 2.1.12 ($SU(2)$ Has Euler Angles: $(0 \leq \theta \leq \pi,\ 0 \leq \varphi \leq 2\pi,\ 0 \leq \psi \leq 4\pi)$).

$$\begin{pmatrix} \exp(i\varphi/2) & 0 \\ 0 & \exp(-i\varphi/2) \end{pmatrix}\begin{pmatrix} \cos(\theta/2) & \sin(\theta/2) \\ -\sin(\theta/2) & \cos(\theta/2) \end{pmatrix}\begin{pmatrix} \exp(i\psi/2) & 0 \\ 0 & \exp(-i\psi/2) \end{pmatrix}.$$

Exercise 2.1.19. Fill in the details of the derivations of polar decompositions in Exercises 2.1.9–2.1.12. How unique are these decompositions?

Example 2.1.13 (The Symplectic Group $Sp(n,\mathbb{R})$).
The polar decomposition of $Sp(n,\mathbb{R})$ says:

$$Sp(n,\mathbb{R}) = K_n^*\, A_n^*\, K_n^*,$$

where

$$K_n^* = O(2n) \cap Sp(n, \mathbb{R}) \cong U(n),$$
$$A_n^* = A_{2n} \cap Sp(n, \mathbb{R}) = \{\text{positive diagonal symplectic matrices}\}.$$

How unique is this decomposition? This time it is not legal to permute all the $2n$ entries of the diagonal matrix in A_n^* because the matrix has to remain symplectic. The matrix looks therefore like

$$\begin{pmatrix} H & 0 \\ 0 & H^{-1} \end{pmatrix} \text{ with } H = \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix}, \quad a_j \text{ positive.}$$

Certainly it is legal to permute all the a_j. One can also send a_j to a_j^{-1}. The group generated by such transformations of the elements of A_n^* is the *Weyl group* of $Sp(n, \mathbb{R})$, which has order $n!2^n$. If we define A_n^{*+} to be the set of diagonal matrices of the form:

$$\begin{pmatrix} H & 0 \\ 0 & H^{-1} \end{pmatrix} \text{ with } H = \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix}$$

such that $1 \leq a_1 \leq a_2 \leq \cdots \leq a_n$, then the polar decomposition

$$\mathcal{P}_n^* = A_n^{*+}[K_n^*]$$

is *unique*, up to the action of M_n^* = the centralizer of A_n^* in K_n^*, which has order 2^n.

2.1.4 Geodesics and the Weyl Group

In order to discuss the uniqueness of the general polar decomposition, one needs to discuss the Weyl group for a general semisimple noncompact real Lie group. However, let us postpone this until we have obtained the geodesics in the symmetric space for the symplectic group.

Theorem 2.1.1 (Geodesics in the Symmetric Space of the Symplectic Group).

(a) *A geodesic segment in $\mathcal{P}_n^*$ of the form $T(t)$, for $0 \leq t \leq 1$, with $T(0) = I$ and $T(1) = Y \in \mathcal{P}_n^*$ has the expression:*

$$T(t) = \exp\{tB[U]\}, \quad \textit{for } 0 \leq t \leq 1,$$

provided that Y has polar decomposition from Lemma 2.1.3 with

$$Y = \exp B[U], \ \textit{for } U \in O(n) \text{ and}$$

$$B = \begin{pmatrix} H & 0 \\ 0 & -H \end{pmatrix} \textit{ with } H = \begin{pmatrix} h_1 & & 0 \\ & \ddots & \\ 0 & & h_n \end{pmatrix}, \ h_j \in \mathbb{R}, \ 1 \le j \le n.$$

The length of the geodesic segment is:

$$\left(2 \sum_{j=1}^{n} h_j^2 \right)^{1/2}.$$

(b) *Consider the geodesic through Z_0 and Z_1 in $\mathcal{H}_n$. Set*

$$\rho(Z_1, Z_0) = (Z_1 - Z_0)(\overline{Z}_1 - Z_0)^{-1}(\overline{Z}_1 - \overline{Z}_0)(Z_1 - \overline{Z}_0)^{-1}.$$

A given pair of points Z_0, Z_1 in $\mathcal{H}_n$ can be transformed by the same matrix $M \in Sp(n, \mathbb{R})$ into another pair of points W_0, W_1 in $\mathcal{H}_n$ if and only if the matrices $\rho(Z_0, Z_1)$ and $\rho(W_0, W_1)$ have the same eigenvalues.

If $r_1, \ldots, r_n$ are the eigenvalues of the matrix $\rho(Z_0, Z_1)$, then the symplectic distance between Z_1 and Z_0 is:

$$s(Z_0, Z_1) = \sqrt{2} \left(\sum_{j=1}^{n} \log^2 \frac{1 + \sqrt{r_j}}{1 - \sqrt{r_j}} \right)^{1/2}.$$

Proof. See Maass [426, p. 39].

(a) The proof proceeds exactly as in the proof of Theorem 1.3.1 of Section 1.1.3. In the partial Iwasawa decomposition (2.30) of $T(t)$, we only decrease the arc length by taking X to be identically zero. Then by Exercise 2.1.18,

$$ds^2 = 2\mathrm{Tr}\left(\left(V^{-1}\, dV \right)^2 \right),$$

and we know from Theorem 1.3.1 of Section 1.1.3 that this arc length is minimized by taking V to be diagonal. The rest of part (a) is immediate.

(b) Note that if $W_j = \left(AZ_j + B \right) \left(CZ_j + D \right)^{-1}$, for $j = 0, 1$,

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \in Sp(n, \mathbb{R}),$$

then

$$\rho(W_1, W_0) = \left(Z_1\, {}^tC + {}^tD\right)^{-1} \rho(Z_1, Z_0) \left(Z_1\, {}^tC + {}^tD\right). \tag{2.34}$$

Using part (a), we need to only observe that with H as in part (a), we have

$$\rho(iH, iI) = (H - I)^2 (H + I)^{-2}.$$

The eigenvalues of $\rho(iH, iI)$ are $r_j = (h_j - 1)^2 (h_j + 1)^{-2}$. Thus

$$h_i = \frac{1 + \sqrt{r_j}}{1 - \sqrt{r_j}}.$$

This completes the proof of Theorem 2.1.1.

■

It is possible to generalize Theorem 2.1.1 to all noncompact real symmetric spaces.

Exercise 2.1.20. Prove formula (2.34) which was used in the proof of part (b) of Theorem 2.1.1.
Hint. First show that

$$W_1 - W_0 = \left(Z_1\, {}^tC + {}^tD\right)^{-1} (Z_1 - Z_0)\, (CZ_0 + D)^{-1}.$$

Theorem 2.1.2. *Suppose that G is a connected noncompact real semisimple Lie group.*

(a) A geodesic in G/K which passes through gK has the form:

$$\gamma_X(t) = g \exp(tX)K, \quad \text{for some } X \in \mathfrak{p}, \quad \text{with } t \in \mathbb{R}.$$

Here the Cartan decomposition of the Lie algebra $\mathfrak{g}$ of G is $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$, and K is a connected Lie subgroup of G with Lie algebra $\mathfrak{k}$.

(b) Geodesics of G/K have the form $\gamma_X(t)$, for all $t \in \mathbb{R}$, using the notation of part (a). This means that G/K is a complete Riemannian manifold. Moreover, any two points of G/K can be joined by a geodesic segment of length equal to the Riemannian distance between the points.

(c) A geodesic through the origin K in G/K has the form

$$\gamma(t) = k \exp(tX)K \quad \text{for some } k \in K,\ X \in \mathfrak{a}, \quad \text{with } t \in \mathbb{R}.$$

Here the Iwasawa decomposition of $\mathfrak{g}$ is $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{a} \oplus \mathfrak{n}$.

Proof. (a) Let g_0K and g_1K be two cosets in G/K. Apply the transformation $a_{g_o^{-1}}$ from (2.25)to transform these cosets to K and $(g_0^{-1}g_1)K$. Now we have the polar decomposition (Lemma 2.1.3): $g_0^{-1}g_1 = k_1ak_2$, with $k_i \in K$ and $a \in A$. So the transformation $a_{k_1^{-1}}$ sends these cosets to K and aK. Thus we have reduced the proof to the case that $\gamma(t)$ is a geodesic with $\gamma(0) = K$ and $\gamma(1) = aK$, with $a \in A$. Write

$$\gamma(t) = a(t)n(t)K, \text{ using the Iwasawa decomposition.}$$

We want to show that $n(t) = e$, the identity in G. Then we would be reduced to the known result that straight lines in the Euclidean space $\mathfrak{a}$ are the geodesics.

The Riemannian structure on G/K comes from the Killing form B of formula (2.14). If $\pi : G \to G/K$ with $\pi(g) = gK$ and $\gamma(t) = w(t)K$, with $w(t) \in G,\ w(t) = a(t)n(t)$, then

$$Q_{\gamma(t)}(\gamma'(t), \gamma'(t))=B\left(\left((d\pi)_{w(t)}(dL_{w(t)})_e\right)^{-1}\gamma'(t), \left((d\pi)_{w(t)}(dL_{w(t)})_e\right)^{-1}\gamma'(t)\right),$$

with $L_g(x) = gx$. Since $\gamma = \pi \circ w$, we have

$$Q_{\gamma(t)}(\gamma'(t), \gamma'(t)) = B\left(dL_{w(t)}^{-1}w'(t), dL_{w(t)}^{-1}w'(t)\right).$$

Now, we can calculate the **differential of multiplication** $w(t) = a(t)n(t)$ as follows—using Exercise 2.1.21 below:

$$w'(t) = (dL)_{w(t)}\left(\mathrm{Ad}(n(t))^{-1}dL_{a(t)}^{-1}\left(a'(t)\right) + dL_{n(t)}^{-1}\left(n'(t)\right)\right). \tag{2.35}$$

Thus

$$\begin{aligned} Q_{\gamma(t)}(\gamma'(t), \gamma'(t)) &= B\left(\mathrm{Ad}(n(t))^{-1}dL_{a(t)}^{-1}(a'(t)) + dL_{n(t)}^{-1}(n'(t)),\ \text{same}\right) \\ &= B\left(dL_{a(t)}^{-1}(a'(t)) + \mathrm{Ad}(n(t))dL_{n(t)}^{-1}(n'(t)),\ \text{same}\right) \\ &= Q_{a(t)}\left(a'(t), a'(t)\right) + Q_{n(t)}(n'(t), n'(t)), \end{aligned}$$

since

$$0 = 2B\left(dL_{a(t)}^{-1}(a'(t)), \mathrm{Ad}(n(t))dL_{n(t)}^{-1}(n'(t))\right),$$

because the first argument of the Killing form lies in $\mathfrak{a}$ and the second lies in $\mathfrak{n}$. See Exercise 2.1.22 below. Thus the distance is only made smaller by setting $n(t) = e$. It follows that we are reduced to the computation of the geodesics in the space $\mathfrak{a}$. The Killing form gives a metric on $\mathfrak{a}$ which is equivalent to the usual Euclidean metric. So the geodesics in $\mathfrak{a}$ are straight lines and the geodesics in $A = \exp \mathfrak{a}$ through e are of the form $\exp(tX),\ t \in \mathbb{R}$, for some $X \in \mathfrak{a}$. This completes the proof of part (a) of Theorem 2.1.2.

(b) This is proved in Helgason [273, p. 56] using part (a).
(c) This follows from part (1) and the polar decomposition (Lemma 2.1.3).
■

Exercise 2.1.21 (The Differential of Multiplication). Suppose that the Lie algebra $\mathfrak{g}$ can be decomposed into a direct sum of subalgebras

$$\mathfrak{g} = \mathfrak{m} \oplus \mathfrak{h}$$

and let $G \supset M, H$ be the corresponding connected Lie subgroups. If the map $\alpha : M \times H \to G$ is defined by $\alpha(m,h) = mh$, show that the differential is:

$$(d\alpha)_{(m,h)}\,(dL_m X, dL_h Y) = (dL_{mh})\left(\mathrm{Ad}\left(h^{-1}\right)X + Y\right), \text{ for } X \in \mathfrak{m},\ Y \in \mathfrak{h}.$$

Hint. Define $L_m \times L_h : M \times H \to M \times H$ by $(L_m \times L_h)(x,y) = (mx, hy)$, for $x \in M,\ y \in H$. Then

$$\alpha \circ (L_m \times L_h) = L_{mh} \circ \alpha \circ (\mathrm{Int}(h^{-1}) \times I).$$

Thus you can use formula (2.7) relating multiplication on G and Lie bracket on $\mathfrak{g}$ to show that:

$$(d\alpha)_{(e,e)}(X,Y)f = \frac{df}{dt}(\exp tX \exp tY)\bigg|_{t=0} = (X+Y)f.$$

Exercise 2.1.22. Suppose that $\mathfrak{g}$ has the Iwasawa decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{a} \oplus \mathfrak{n}$ and Cartan involution θ. Consider the form $F(X,Y) = -B(X, \theta Y)$ for $X, Y \in \mathfrak{g}$ from formula (2.20) in Section 2.1.3. Then F is a positive definite bilinear form on $\mathfrak{g}$. Show that

$$X \in \mathfrak{k} \Rightarrow \text{ad}\, X \text{ is skew symmetric;}$$

$$X \in \mathfrak{a} \Rightarrow \text{ad}\, X \text{ is diagonal;}$$

$$X \in \mathfrak{n} \Rightarrow \text{ad}\, X \text{ is upper triangular with 0 on the diagonal.}$$

Hint. (See Wallach [651, p. 166] or Helgason [273, p. 223].) You need to take an ordered set of positive roots: $\alpha_1, \alpha_2, \ldots, \alpha_m$. Then form an orthonormal basis of $\mathfrak{g}$ by taking orthonormal bases of

$$\mathfrak{g}_{\alpha_m}, \ldots, \mathfrak{g}_{\alpha_1},\ \mathfrak{a} \oplus \mathfrak{m},\ \theta(\mathfrak{g}_{\alpha_1}), \ldots, \theta(\mathfrak{g}_{\alpha_m}).$$

You need to use properties of the roots such as the fact that: $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta}$.

Next we consider the Weyl group of the symmetric space. See (1.296) of Section 1.5.3 for the definition in the case of $GL(n, \mathbb{R})$. As usual, suppose that G

is a noncompact real semisimple Lie group with the standard definitions of $K, \mathfrak{a}$, etc. Define the following subgroups of K:

$$\left.\begin{array}{l} M = \text{the } \textbf{centralizer} \text{ of } \mathfrak{a} \text{ in } K = \{k \in K \mid \mathrm{Ad}(k)|_{\mathfrak{a}} = \text{identity}\}\,, \\ M' = \text{the } \textbf{normalizer} \text{ of } \mathfrak{a} \text{ in } K = \{k \in K \mid \mathrm{Ad}(k)\mathfrak{a} \subset \mathfrak{a}\}\,. \end{array}\right\} \tag{2.36}$$

Both M and M' are closed subgroups of K. The **Weyl group** of G/K is defined to be $W = M'/M$. Note that W is independent of the choice of $\mathfrak{a}$, by the conjugacy of all maximal abelian subspaces of $\mathfrak{p}$ (see the proof of Lemma 2.1.3). These definitions can also be made in the case that G is compact (see Helgason [273, p. 244]).

Theorem 2.1.3 (The Weyl Group).

(1) The Weyl group is a finite group contained in the orthogonal group in $GL(\mathfrak{a})$ with respect to the inner product on $\mathfrak{a}$ defined by the Killing form of $\mathfrak{g}$.

(2) The Weyl group permutes the restricted roots. Define a ***Weyl chamber*** *to be a connected component of*

$$\left(\mathfrak{a} - \bigcup_{\alpha \in \Lambda} \alpha^{-1}(0)\right), \quad \text{for } \alpha \in \Lambda.$$

Note that $\alpha^{-1}(0)$ is a hyperplane in $\mathfrak{a}$. The Weyl group also permutes the Weyl chambers. Moreover the action of the Weyl group on the Weyl chambers is simply transitive.

(3) For $\lambda \in \Lambda$, define $s_\lambda : \mathfrak{a} \to \mathfrak{a}$ by

$$s_\lambda(H) = H - 2(\lambda(H)/\lambda(H_\lambda))H_\lambda, \quad \text{where} \quad H_\lambda \in \mathfrak{a}$$

is defined by

$$B(H, H_\lambda) = \lambda(H), \textit{ for all } H \in \mathfrak{a}.$$

Then s_λ is the reflection in the hyperplane $\lambda^{-1}(0)$. The Weyl group is generated by these reflections s_λ, for $\lambda \in \Lambda$.

Proof. We shall only prove part (1). For the other parts of the theorem, see Helgason [273, Ch. 7] or Wallach [651, pp. 77, 168]. By the Lie group/Lie algebra dictionary,

$$\text{the Lie algebra of } M = \mathrm{Lie}(M) = \mathfrak{m} = \{X \in \mathfrak{k} \mid \mathrm{ad}\, X|_{\mathfrak{a}} = 0\}\,.$$

If we can show that M' has the same Lie algebra as M, then it will follow that the quotient M'/M is both discrete and compact (thus finite). Suppose that T is in the Lie algebra of M'. Then write out the root space decomposition of T:

$$T = Y + \sum_{\lambda\in\Lambda} X_\lambda, \quad \text{for } Y \in \mathfrak{m} \oplus \mathfrak{a},\ X_\lambda \in \mathfrak{g}_\lambda.$$

It follows that for all $H \in \mathfrak{a}$

$$[H,T] = \sum \lambda(H) X_\lambda \in \mathfrak{a} \text{ implies that } [H,T] = 0.$$

Here we have used the fact that the sum in the root space decomposition is direct.

To see that the group M'/M permutes the restricted roots is easy. To see that the reflections s_λ come from some $\mathrm{Ad}(k)$, $k \in K$, is harder. To see that the s_λ generate the Weyl group is even harder. Note that we cannot claim that the s_λ, with λ from a system of simple roots, generate the Weyl group. A system of simple roots has the property that any root is a linear combination of simple roots with integer coefficients that are either all positive or all negative (with $r = \dim \mathfrak{a}$ elements). Such simple root systems *do* give generators of the Weyl group in the case of *complex* semisimple Lie algebras. However, real Lie algebras are somewhat different, as we will see in the following examples. ■

Exercise 2.1.23. (a) Why is it reasonable to call a Lie algebra "semisimple" if the Killing form is nondegenerate? What is the connection with the standard notion that an algebraic object is semisimple if it is a direct sum of simple objects?

(b) Why do we call a semisimple Lie algebra "compact" if its Killing form is negative definite? What is the connection with compact Lie groups? Can we drop the hypothesis that the Lie algebra be semisimple?

(c) Recall that we said a Lie algebra is "nilpotent" if all sufficiently long brackets must vanish. What is the connection with the usual idea of a nilpotent linear transformation (such as $\mathrm{ad}X$)?

Hints.

(a) See Helgason [273, pp. 121–122].
(b) See Helgason [273, pp. 120–122]. Think about $\mathbb{R}$ and $\mathbb{R}/\mathbb{Z}$.
(c) See Helgason [273, pp. 135–137].

Examples of Weyl Groups

Example 2.1.14 **($GL(n,\mathbb{R})$).**

Since $\mathrm{Ad} = \mathrm{Int}$ for matrix groups, we have:

$$\begin{aligned} M &= \left\{k \in O(n) \;\middle|\; kXk^{-1} = X, \text{ for any diagonal matrix } X\right\}, \\ M' &= \left\{k \in O(n) \;\middle|\; kXk^{-1} \text{ is diagonal, for any diagonal matrix } X\right\}. \end{aligned}$$

It follows that

$$\begin{aligned} M &= \{\text{diagonal matrices with entries } +1 \text{ or } -1\}, \\ M' &= \{\text{matrices with each row or column having exactly one non} - 0 \text{ entry of } \pm 1\}. \end{aligned}$$

Thus the Weyl group of $GL(n,\mathbb{R})$ is the group of all permutations of n objects as we also saw in Exercise 1.5.30 of Section 1.5.3.

Example 2.1.15 ($Sp(n,\mathbb{R})$**).** Here

$$
\begin{aligned}
K &= \left\{ \begin{pmatrix} A & B \\ -B & A \end{pmatrix} \,\middle|\, A+iB \in U(n) \right\}, \\
\mathfrak{a} &= \left\{ \begin{pmatrix} H & 0 \\ 0 & -H \end{pmatrix} \,\middle|\, H \;\; n\times n \text{ real diagonal} \right\}, \\
M &= \left\{ \begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix} \,\middle|\, A \text{ diagonal } n\times n,\ \text{ entries } \pm 1 \right\}, \\
M' &= \left\{ \begin{pmatrix} A & B \\ -B & A \end{pmatrix} \,\middle|\, A+B \text{ is in the } M' \text{ for } GL(n,\mathbb{R}) \right\}.
\end{aligned}
$$

It follows that the Weyl group $W = M'/M$ contains all permutations of entries of H in

$$\begin{pmatrix} H & 0 \\ 0 & -H \end{pmatrix} \text{ in } \mathfrak{a},$$

as well as all possible changes of sign. So it has 2^n times $n!$ elements. For example, let $n = 3$ and

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},\ B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},\ k = \begin{pmatrix} A & B \\ -B & A \end{pmatrix},\ H = \begin{pmatrix} h_1 & 0 & 0 \\ 0 & h_2 & 0 \\ 0 & 0 & h_3 \end{pmatrix},$$

$$a = \begin{pmatrix} H & 0 \\ 0 & -H \end{pmatrix}, \quad \text{then} \quad \mathrm{Ad}(k)a = \begin{pmatrix} H' & 0 \\ 0 & -H' \end{pmatrix},$$

where

$$H' = \begin{pmatrix} h_2 & 0 & 0 \\ 0 & h_1 & 0 \\ 0 & 0 & -h_3 \end{pmatrix}.$$

Exercise 2.1.24. Check the results stated in Example 2.1.15 above for $Sp(n,\mathbb{R})$.

Example 2.1.16 ($SU(2,1)$)**.**
For this example,

$$K = \left\{ \begin{pmatrix} U & 0 \\ 0 & t \end{pmatrix} \,\middle|\, U \in U(2),\ t = (\det U)^{-1} \right\},$$

$$\mathfrak{a} = \mathbb{R}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix},$$

$$M = \left\{ k = \begin{pmatrix} e^{i\alpha} & 0 & 0 \\ 0 & e^{i\beta} & 0 \\ 0 & 0 & e^{i\alpha} \end{pmatrix} \middle| \det k = 1 \right\},$$

$$M' = \left\{ k = \begin{pmatrix} e^{i\alpha} & 0 & 0 \\ 0 & e^{i\beta} & 0 \\ 0 & 0 & \pm e^{i\alpha} \end{pmatrix} \middle| \det k = 1 \right\}.$$

So the Weyl group has only two elements. The entries of the diagonal matrices in M' are supposed to be of complex norm 1.

Exercise 2.1.25. Verify the results stated in Example 2.1.16 for $SU(2,1)$.

Now that we have described the Weyl group, it is possible to discuss the **nonuniqueness of the polar decomposition**. The precise result is that if we set $\mathfrak{a}' = \{H \in a \mid \lambda(H) \neq 0, \text{ for all } \lambda \in \Lambda\}$, $A' = \exp \mathfrak{a}'$, and define the map $f : (K/M) \times A' \to G/K$ by $f(kM, a) = (ka)K$, then the map f is $\#W$ to 1, regular, and onto an open submanifold of G/K whose complement in G/K has lower dimension (see Helgason [273, p. 381] or [278, p. 402] or Wallach [651]).

One should also consider the relation between the structure theory for $\mathfrak{g}$ a noncompact semisimple real Lie algebra and that for the complexification $\mathfrak{g}^c = \mathfrak{g} \otimes_{\mathbb{R}} \mathbb{C}$. The same question could be asked for the compact real form of $\mathfrak{g}^c$. As an example, consider $SU(2,1)$ again. The **Cartan subalgebra** or maximal abelian subalgebra $\mathfrak{h}$ of $\mathfrak{su}(2,1)$ containing $\mathfrak{a}$ is:

$$\mathfrak{h} = \left\{ \begin{pmatrix} a & 0 & b \\ 0 & c & 0 \\ b & 0 & a \end{pmatrix} \middle| a, c \in i\mathbb{R},\ b \in \mathbb{R} \right\}.$$

Clearly the complexification of $\mathfrak{h}$ is a Cartan subalgebra of the complexification of $\mathfrak{g}$. This shows that much is missing from the complexification of $\mathfrak{a}$. One can show that the restricted roots are really restrictions of roots of the complexified Lie algebra (see Helgason [273, Ch. 6]). Again, some roots from the complexification may be missing in the real version of the Lie algebra.

2.1.5 Integral Formulas

Our next topic is integral formulas for noncompact semisimple real Lie groups. First perhaps we should discuss the Haar measures in G, A, N, and K. See our earlier comments on this subject in Chapter 2 of Vol. I and Chapter 1 of this Volume. More details about Haar measures can be found in Helgason [273, Chapter 10]. Because Haar measure is unique up to a positive scalar multiple, we can define the **modular function** $\delta : G \to \mathbb{R}^+$ by the formula (assuming dg = left Haar measure):

$$\int f(gs^{-1})dg = \delta(s) \int f(g)dg. \tag{2.37}$$

For the left-hand side of the equality is a left G-invariant integral for fixed s. Thus it must be a positive constant times the Haar integral of f. It follows easily that δ is continuous, $\delta(st) = \delta(s)\delta(t)$, and $d(gs) = \delta(s)dg$. Thus the modular function relates right and left Haar measure. By definition, a **unimodular group** has $\delta = 1$ identically. Furthermore, it is easy to see that $d(g^{-1}) = \delta(g^{-1})dg$. If G is a Lie group, one also has $d(s^{-1}gs) = d(gs) = \delta(s)dg$. Thus $\det(\mathrm{Ad}(s^{-1})) = \delta(s)$, for all s in G.

We prove that compact, semisimple, and nilpotent Lie groups are all unimodular. Suppose first that K is compact. Then δ maps K onto a compact subgroup of $\mathbb{R}^+$ which must contain only one element, since otherwise powers would approach 0 or infinity. Suppose next that G is semisimple. Then Ad (s) leaves the Killing form invariant for s in G. But the Killing form of a semisimple group is nondegenerate and thus equivalent to

$$I_{p,q} = \begin{pmatrix} I_p & 0 \\ 0 & -I_q \end{pmatrix}, \quad \text{for some} \quad p, q.$$

If ${}^t g I_{p,q} g = I_{p,q}$, the determinant of g must have absolute value 1. Finally suppose that N is nilpotent and connected. Then

$$\det(\mathrm{Ad}(n)) = \exp(\mathrm{Tr}(\mathrm{ad}(\log\ n))) = 1,$$

for $n \in N$, since ad $(\log\ n)$ is a nilpotent linear transformation.

Proposition 2.1.1 (The Integral Formula for the Iwasawa Decomposition). *Define* $m_\lambda = \dim_{\mathbb{R}} \mathfrak{g}_\lambda$ *and*

$$J(a) = \prod_{0<\lambda\in\Lambda^+} \exp(m_\lambda \lambda(\log a)),$$

for $a \in A$. *Then*

$$\int_A \int_N \int_K f(ank)\, da\, dn\, dk = \int_G f(g)\, dg,$$

where all the measures are left-invariant (and thus right-invariant) Haar measures on G, A, N, K. *However, changing the order gives:*

$$\int_K \int_A \int_N f(kan) J(a)\, dk\, da\, dn = \int_G f(g)\, dg.$$

Proof. In order to compute the Jacobian of the Iwasawa decomposition, we proceed as in Exercise 1.1.20 of Section 1.1.4. Thus we need the differential of $\mathrm{Int}(a)n = ana^{-1}$, for $n \in N$ and $a \in A$. We know that the differential of $\mathrm{Int}(a)$ is $\mathrm{Ad}(a)$, by definition. Thus if $a = \exp H$ for $H \in \mathfrak{a}$, we find that:

$$\begin{aligned}\det(\mathrm{Ad}(a)) &= \det(\exp(\mathrm{ad} H)) = \exp\left(\mathrm{Tr}(\mathrm{ad} H)\right) \\ &= \exp\Big(\sum_{\lambda\in\Lambda^+} m_\lambda \lambda(H)\Big) = \prod_{\lambda\in\Lambda^+} \exp\left(m_\lambda \lambda(H)\right),\end{aligned}$$

which is simply $J(a)$, as defined in the proposition, since $H = \log a$. Here we have used the fact that:

$$\mathfrak{n} = \sum_{\lambda\in\Lambda^+}{}^{\oplus}\, \mathfrak{g}_\lambda, \qquad \mathfrak{g}_\lambda = \{X \in \mathfrak{g} \mid \mathrm{ad} H(X) = \lambda(H)X, \quad \text{for all} \quad H \in \mathfrak{a}\}.$$

Note that $\mathrm{Int}(a) : N \to N$ for any $a \in A$.

The rest of the argument is really the same as that of Exercise 1.1.20 in Section 1.1.4, but we shall repeat it for completeness. First observe that the left Haar measures on G and K can be normalized so that if $d\overline{g}$ denotes the G-invariant measure on the symmetric space G/K, then the following equality prevails:

$$\int_G f(g)\, dg = \int_{\overline{g}=gK\in G/K} \int_{k\in K} f(gk)\, dk\, d\overline{g}.$$

Now G/K can be identified with AN. Thus we need to only show

$$\int_A \int_N f(an)\, dn\, da$$

gives a left AN-invariant integral on AN. Let $a_1 \in A$ and $n_1 \in N$. Then we have

$$\int_A \int_N f(a_1 n_1 an)\, dn\, da = \int_A \int_N f(a_1 a n_2 n)\, dn\, da, \quad \text{if} \;\; n_2 = a^{-1} n_1 a.$$

Since both da and dn are left invariant, the last integral is just

$$\int_A \int_N f(an)\, dn\, da.$$

This completes the proof of the first integral formula in the proposition.

Now we are ready to prove the second version of the integral formula for the Iwasawa decomposition. Using the differential of Int(a) and the first integral formula, we get:

$$\int_G f(g)\,dg = \int_N \int_A \int_K f(nak)J(a)^{-1}\,dk\,da\,dn.$$

Now replace $f(g)$ by $f(g^{-1})$. This will reverse orders on the right-hand side and produce

$$\int_N \int_A \int_K f(k^{-1}a^{-1}n^{-1})J(a)^{-1}\,dk\,da\,dn.$$

Finally the fact that G, N, A, K are all unimodular leads to the second integral formula for the Iwasawa decomposition. ■

Examples

(1) $\boldsymbol{G = GL(n,\mathbb{R})}$.

$$J(a) = \prod_{1\le i<j\le n} \frac{a_i}{a_j} = \prod_{i=1}^{n} a_i^{n-2i+1}.$$

(2) $\boldsymbol{G = Sp(n,\mathbb{R})}$.

$$J(a) = \prod_{1\le i<j\le n} \frac{a_i}{a_j} \prod_{1\le i\le j\le n} a_i a_j = \prod_{i=1}^{n} a_i^{2(n+1-i)}.$$

In order to be more precise, we need to fix the invariant volumes on the symmetric spaces.

Invariant Volume Elements on the Symmetric Spaces of $\boldsymbol{GL(n,\mathbb{R})}$ and $\boldsymbol{Sp(n,\mathbb{R})}$

(1) We found in formula (1.16) of Section 1.1.4 that the $GL(n,\mathbb{R})$-invariant volume element on $\mathcal{P}_n$ is:

$$d\mu_n = |Y|^{-(n+1)/2} \prod_{1\le i\le j\le n} dy_{ij}, \quad \text{if} \quad Y = (y_{ij}) \in \mathcal{P}_n.$$

(2) Next we want to find the invariant volume element on the Siegel upper half space $\mathcal{H}_n$. The argument following (1.16) of Section 1.1.4 can be imitated to show that the $Sp(n,\mathbb{R})$-invariant volume on $\mathcal{H}_n$ is:

$$d\mu_n^*(Z) = |Y|^{-(n+1)} \prod_{1\le i\le j\le n} dx_{ij}\, dy_{ij}, \quad \text{if} \quad Z = X + iY \in \mathcal{H}_n,$$

with

$$X = (x_{ij}) \quad \text{and} \quad Y = (y_{ij}).$$

Let us prove this last formula. As in the case of $GL(n,\mathbb{R})$, it suffices to find the Jacobian of the action of a diagonal symplectic matrix on $\mathcal{H}_n$. So observe that the image of $X + iY \in \mathcal{H}_n$ under the matrix

$$\begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix} \in Sp(n,\mathbb{R}), \qquad a = \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix},$$

is $aZa = Z[a] = X[a] + iY[a]$, according to (2.25). Thus the Jacobian of the transformation is:

$$\prod_{1\le i\le j\le n} a_i a_j \prod_{1\le i\le j\le n} a_i a_j = |a|^{2(n+1)}.$$

This shows that the measure $d\mu_n^*$ is $Sp(n,\mathbb{R})$-invariant.

Our next goal is to work out the integral formula for polar coordinates. First we need a Lemma.

Lemma 2.1.4 (The Integral Formula for Exp Restricted to $\mathfrak{p}$ in the Cartan Decomposition). *There is a positive constant c such that:*

$$\int_{G/K} f(x)\, dx = c \int_{\mathfrak{p}} f(\exp Y) J(Y)\, dY,$$

where

$$J(X) = \det\left(\left.\frac{\sinh \operatorname{ad} X}{\operatorname{ad} X}\right|_{\mathfrak{p}} \right), \quad \text{for} \quad X \in \mathfrak{p}.$$

Proof. First recall our calculation of the differential of exp in formula (2.8) or see Helgason [273, p. 95] for the general result. Observe that if $X, Y \in \mathfrak{p}$, then

$$(d\exp)_X Y = (dL_{\exp X})_e \circ \left\{ \sum_{n\geq 0} \frac{1}{(2n+1)!} (-\mathrm{ad}X)^{2n} \right\} Y.$$

For $[\mathfrak{k}, \mathfrak{p}] \subset \mathfrak{p}$ and $[\mathfrak{p}, \mathfrak{p}] \subset \mathfrak{k}$ follow from the properties of the Cartan involution. Therefore

$$(\mathrm{ad}X)^{2n+1} Y \in \mathfrak{k} \quad \text{and} \quad (\mathrm{ad}X)^{2n} Y \in \mathfrak{p}.$$

Since $\mathfrak{k} \cap \mathfrak{p} = 0$, we have the vanishing of the sum of the odd powers of ad$X$ in the series expression for the differential of exp.

We can write $X \in \mathfrak{p}$ in the form

$$X = Y + \sum_{\lambda \in \Lambda^+} (X_\lambda - \theta X_\lambda), \qquad Y \in \mathfrak{a},$$

for X_λ in a basis for $\mathfrak{g}_\lambda$. Note also that $H \in \mathfrak{a},\ X_\lambda \in \mathfrak{g}_\lambda$ implies that

$$(\mathrm{ad}H)\,(X_\lambda - \theta X_\lambda) = \lambda(H)\,(X_\lambda + \theta X_\lambda)\,, \ \text{ if } \ X_\lambda \in \mathfrak{g}_\lambda.$$

It follows that

$$(\mathrm{ad}H)^2\,(X_\lambda - \theta X_\lambda) = (\lambda(H))^2\,(X_\lambda - \theta X_\lambda)\,.$$

Thus the differential at H has determinant

$$\left|(d\exp|_{\mathfrak{p}})_H\right| = \left|e^H\right| \prod_{0<\lambda\in\Lambda} \frac{\sinh \lambda(H)}{\lambda(H)},$$

proving Lemma 2.1.4. This shows, in particular, that exp is a diffeomorphism when restricted to $\mathfrak{p}$. ■

Proposition 2.1.2 (The Integral Formula for Polar Coordinates). *Suppose that the root space $\mathfrak{g}_\lambda$ has dimension m_λ for any restricted root λ. Then there is a positive constant c such that if dg denotes the Haar measure on G, then*

$$\int_G f(g)\, dg = c \int_K \int_A \int_K f(k_1 a k_2) D(a)\, dk_1\, da\, dk_2,$$

where

$$D(a) = \prod_{0<\lambda\in\Lambda} |\sinh(\lambda(\log a))|^{m_\lambda}\,, \quad \text{for} \quad a \in A.$$

Proof. A reference for the proof is Helgason [273, Ch. 10]. The main step is the preceding lemma. Then one uses the fact that $\mathfrak{p}$ =Ad$(K)\mathfrak{a}$ from Lemma 2.1.3.

Observe first that $X \in \mathfrak{k}$ has the representation:

$$X = Y + \sum_{\lambda \in \Lambda^+} (X_\lambda + \theta X_\lambda), \quad \text{with} \quad Y \in \mathfrak{m},\ X_\lambda \in \mathfrak{g}_\lambda,$$

where $\mathfrak{m}$ and $\mathfrak{g}_\lambda$ are from the root space decomposition of $\mathfrak{g}$ (see Helgason [273, p. 224]).

Define $f : K \times A \to P$ by $f(k,a) = kak^{-1} = p$. Suppose that $Y \in \mathfrak{k},\ H \in \mathfrak{a}$. Then

$$\begin{aligned}(df)_{(k,a)}(Y,H) &= \lim_{t\to 0} \frac{1}{t} \left\{ f(ke^{tY}, ae^{tH}) - f(k,a) \right\} \\ &= \lim_{t\to 0} \frac{1}{t} \left\{ ke^{tY} ae^{tH} e^{-tY} k^{-1} - kak^{-1} \right\} \\ &= k \left\{ \lim_{t\to 0} \frac{1}{t} \left(e^{tY} ae^{tH} e^{-tY} - a \right) \right\} k^{-1}.\end{aligned}$$

Now suppose that $a = \exp(H_0)$ for $H_0 \in \mathfrak{a}$. Then the object inside the last limit is:

$$\begin{aligned}e^{tY} e^{H_0+tH} e^{-tY} - e^{H_0} &= \exp\left(H_0 + tH + t[Y, H_0 + tH] + o(t^2)\right) - \exp H_0 \\ &= \exp\left(H_0 + t(H + [Y, H_0]) + o(t^2)\right) - \exp H_0.\end{aligned}$$

Use the chain rule to evaluate the derivative of the preceding quantity with respect to t at $t = 0$ and obtain:

$$(d\exp)_{H_0}\left(H + [Y, H_0]\right).$$

Take a basis of $\mathfrak{p}$ coming from $\mathfrak{a}$ and vectors $X_\lambda - \theta X_\lambda$ and a basis of $\mathfrak{k}$ coming from $\mathfrak{m}$ and vectors $X_\lambda + \theta X_\lambda$, with X_λ in the root spaces $\mathfrak{g}_\lambda$. One sees that for $Y = X_\lambda + \theta X_\lambda$, the preceding is:

$$(d\exp)_{H_0}\left(H - \lambda(H_0)(X_\lambda - \theta X_\lambda)\right).$$

Finally use the formula for the differential of exp, along with the fact that the odd powers of $(\mathrm{ad} H_0)$ vanish, once again. This yields:

$$\begin{aligned}&(d\exp)_{H_0}\left(H + [X_\lambda + \theta X_\lambda, H_0]\right) \\ &= e^{H_0} \left\{ H - \lambda(H_0) \sum_{n \geq 0} \frac{1}{(2n+1)!} \lambda(H_0)^{2n} (X_\lambda - \theta X_\lambda) \right\} \\ &= e^{H_0} \left\{ H - \sinh \lambda(H_0)(X_\lambda - \theta X_\lambda) \right\}.\end{aligned}$$

This completes our discussion of Proposition 2.1.2. ■

If G is a real noncompact semisimple Lie group, K a compact subgroup coming from the Cartan decomposition of G, the **boundary** of G/K can be defined as K/M. The group M was defined in (2.36). And we can identify this boundary with G/B, if B is the Borel or minimal parabolic subgroup $B = MAN$, as in (1.19) of Section 1.1.4. Furstenberg [194] and Moore [459] show that G/B is a "maximal boundary" in a certain probabilistic sense.

Example 2.1.17 **($G = SL(n, \mathbb{R})$).**
Here $B = MAN$ consists of all upper triangular matrices of determinant one. We can identify G/B as the **flag manifold**:

$$F_n = \{(V_1, \ldots, V_{n-1}) \mid V_i \text{ is a vector subspace of } \mathbb{R}^n,\ \dim_{\mathbb{R}} V_i = i,\ V_i \subset V_{i+1}\}.$$

The action of $g \in G$ on F_n is $g(V_1, \ldots, V_{n-1}) = (gV_1, \ldots, gV_{n-1})$. This action is easily seen to be transitive. To calculate the stability group of a point, let $e_i \in \mathbb{R}^n$ denote the column vector with ith coordinate one and the rest zero. Then set $V_i^0 = \mathbb{R}e_1 \oplus \cdots \oplus \mathbb{R}e_i$. Then g fixes V_i^0, for all i, means $g \in B$. See Exercise 1.3.11 of Section 1.3.6.

Exercise 2.1.26. Show that the Jacobian of the action of $g \in G$ on the boundary $G/MAN \cong K/M$ is given by the following integral formula:

$$\int_{K/M} f(\overline{k})\, d\overline{k} = \int_{\overline{k}=kM \in K/M} f(g(\overline{k}))J(a(gk))^{-1}\, d\overline{k},$$

where $a(g)$ is the A-part of the KAN-Iwasawa decomposition of g, and $J(a)$ is the Jacobian of the Iwasawa decomposition in Proposition 2.1.1. Here $d\overline{k}$ is any K-invariant measure on K/M.

It is also possible to show (using the Bruhat decomposition of G described in Section 1.5.3 for $GL(n)$) that if $\overline{N}$ denotes the opposite nilpotent subgroup corresponding to the Lie subalgebra $\overline{\mathfrak{n}}$,

$$\overline{\mathfrak{n}} = \sum_{0>\alpha\in\Lambda} \mathfrak{g}_\alpha,$$

then $\overline{N}MAN$ is an open subset of G with lower dimensional complement. See Lemma 1.3.2 of Section 1.3.3 for a proof of this result when $G = SL(n, \mathbb{R})$. This allows us to identify K/M with $\overline{N}$ as far as integration is concerned. In Section 1.3, we applied such a result to obtain the asymptotics of spherical functions. For more information on boundaries and compactifications of symmetric spaces, see Gérardin [216], Helgason [273–282], Koranyi [364], and the references mentioned when we defined the boundary of $\mathcal{P}_n$.

This concludes our discussion of the basic integral formulas for symmetric spaces. Next let us consider differential operators on the symmetric space G/K when G is a noncompact real semisimple Lie group.

2.1.6 Invariant Differential Operators

Let φ be a diffeomorphism of a manifold M. We say that a differential operator D on M is **invariant** under φ if D commutes with φ; i.e., if $D(f \circ \varphi) = (Df) \circ \varphi$, for all infinitely differentiable functions f on M. For each $g \in G$, we have a diffeomorphism a_g of the symmetric space G/K defined by $a_g(xK) = (gx)K$, for $g, x \in G$. Define $D(G/K)$ to be the set of all differential operators on G/K which are a_g-invariant for all $g \in G$. So $\boldsymbol{D(G/K)}$ is the **algebra of invariant differential operators on $\boldsymbol{G/K}$**. The Laplacian will, of course, be such an operator. In general, however, there will be invariant differential operators on G/K which are not polynomials in the Laplacian, just as for $\mathcal{P}_n$ (see Theorem 1.1.2 of Section 1.1.5). The following theorem is proved in Helgason [273, p. 432]). We will not prove it here.

Theorem 2.1.4 (Harish-Chandra and Chevalley). *Suppose that G is a noncompact real semisimple Lie group with* $\dim_{\mathbb{R}} \mathfrak{a} = r =$***rank of $\boldsymbol{G/K}$***. *Then the algebra $D(G/K)$ of all invariant differential operators on G/K is a commutative algebra. In fact, it is a polynomial ring with r algebraically independent generators.*

There is a close relation between $D(G/K)$ and $D(G)$ = the left-invariant differential operators on G or the **universal enveloping algebra** of G (see Helgason [273, Ch. 10]).

Question. Can one relate the G-invariant differential operators $D(G/K)$ for the following chain of inclusions of totally geodesic submanifolds?

$$\begin{array}{lllll} \mathcal{P}_n & \to \mathcal{H}_n & \to \mathcal{P}_{2n}^* & & \subset \mathcal{P}_{2n}, \\ Y & \mapsto iY & \mapsto \begin{pmatrix} Y & 0 \\ 0 & Y^{-1} \end{pmatrix}. & & \end{array}$$

We know, for example, that the arc length on $\mathcal{H}_n$ is given by:

$$ds^2 = 2\mathrm{Tr}\left(\left(Y^{-1}dY\right)^2 + \left(Y^{-1}dX\right)^2\right), \quad \text{for} \quad Z = X + iY \in \mathcal{H}_n.$$

Therefore the Laplacian on $\mathcal{H}_n$ must be a sum of the Laplacian on $\mathcal{P}_n$ plus a term involving only differentiation with respect to X-variables. If follows that for functions of the Y-variable alone, the Laplacian on $\mathcal{H}_n$ coincides with that on $\mathcal{P}_n$, disregarding constants.

Let G be a noncompact real semisimple Lie group, as usual. A function $u : G/K \to \mathbb{C}$ is called **harmonic** if $Du = 0$ for any operator $D \in D(G/K)$

such that D annihilates constants. This definition was made by Godement [223]. Furstenberg [194] shows that, in fact, a bounded solution of $\Delta u = 0$ on G/K is automatically harmonic. Other references for potential theory (i.e., the study of harmonic functions) on symmetric spaces are Helgason [275] and Koranyi [364].

Theorem 2.1.5 (Godement's Mean Value Theorem). *Suppose $u : G/K \to \mathbb{C}$ is infinitely differentiable. Then u is harmonic if and only if*

$$\int_{k\in K} u(gkhK)\,dk = u(gK), \quad \text{for all} \quad g,h \in G.$$

Proof (Helgason [275, pp. 42–43]).
$\Rightarrow$ Let u be harmonic and

$$F(h) = \int_K u(gkhK)\,dk.$$

We want to show that $F(h) = F(e) = u(gK)$, $e =$ identity of G. Since F satisfies an elliptic partial differential equation with analytic coefficients, it follows, by a theorem of Bernstein, that F is analytic (see John [332, pp. 57, 142]). Now it suffices to show that

$$(DF)(e) = 0,$$

for every left invariant differential operator D on the Lie group G such that D annihilates the constants.

To show that DF vanishes at the identity, we must merely relate differential operators in $D(G/K)$ with those in $D(G)$. This is done in detail in Helgason [273, Ch. 10]. We merely sketch the process. Let us use the following notation for a diffeomorphism φ of G:

$$D^{\varphi} f = D(f \circ \varphi) \circ \varphi^{-1}, \quad \text{if} \quad D \in D(G).$$

Then for $D \in D(G)$ write

$$D^{\#} f = \int_K D^{R_k} f\,dk, \quad \text{if} \quad R_k x = xk \quad \text{for} \quad x \in G.$$

Now it can be shown that $D^{\#}$ is a differential operator which is invariant under all the R_k, $k \in K$ and thus gives rise to an operator $\tilde{D}$ in $D(G/K)$. And we find that by hypothesis:

$$(DF)(e) = (D^{\#}F)(e) = \int_K (\tilde{D}u)(gkK)\,dk = 0.$$

$\Leftarrow$ Assume that u has the mean value property stated in the theorem and that $D \in D(G/K)$ annihilates constants. As usual, set $a_g(hK) = ghK$, for g, h in G. Thus

$$\int_{k\in K} u(a_{gk}(x))\, dk = u(gK), \quad \text{if} \quad x \in G/K.$$

Apply D to both sides of this equation considered as functions of $x \in G/K$ to obtain

$$\int_{k\in K} (Du)(a_{gk}(x))dk = 0,$$

since D and a_g commute. Take x to be the coset K in G/K to see that $Du = 0$. ∎

Theorem 2.1.6 ((Furstenberg) (Poisson Integral Formula)). *Let us suppose u: $G/K \to \mathbb{C}$ is a bounded harmonic function. Then there is a bounded measurable function $\widehat{u} : K/M \to \mathbb{C}$ such that*

$$u(gK) = \int_{\overline{k}\in K/M} \widehat{u}(g(\overline{k}))\, d\overline{k}, \quad \text{for all} \quad g \in G. \tag{2.38}$$

Here $d\overline{k}$ is the unique K-invariant measure on the boundary K/M such that

$$\int_{K/M} d\overline{k} = 1.$$

And conversely, given $\widehat{u}$, as above, the function u on G/K defined by (2.38) is harmonic. We can rewrite formula (2.38) as:

$$u(x) = \int_{\overline{k}\in K/M} \widehat{u}(\overline{k}) P(x, \overline{k})\, d\overline{k},$$

*where **Poisson's kernel** $P(x, \overline{k})$ is:*

$$P(gK, kM) = d\left(g^{-1}\left(\overline{k}\right)\right)/d\overline{k} = J^{-1}\left(a\left(g^{-1}k\right)\right). \tag{2.39}$$

Here J denotes the Jacobian of the Iwasawa decomposition from Proposition 2.1.1 and $a(g)$ is the A-part of the KAN Iwasawa decomposition of $g \in G$.

Proof. Sketch (See Helgason [275, pp. 42–52].)
$\Rightarrow$ We need to know that the Borel subgroup $B = MAN$ has the following **fixed point property**. Suppose that B acts continuously on a locally convex topological vector space by linear transformations leaving a nonempty compact convex set invariant. Then B has a fixed point in the convex set. Assuming this result, suppose $u : G/K \to \mathbb{C}$ is bounded and harmonic. Define the set

$$Q_u = \left\{ w \in L^\infty(G) \;\middle|\; \begin{array}{l} \|w\|_\infty = \text{l.u.b.}\,\{|w(h)| \mid h \in G\} \le \|u\|_\infty \\[2ex] u(gK) = \int_K w(gkh)dk, \quad \text{for all } \; g,h \in G \end{array} \right\}.$$

By Godement's Mean Value Theorem, $u \circ \pi = \tilde{u} \in Q_u$, where $\pi : G \to G/K$ is defined by $\pi(g) = gK$.

Suppose that MAN leaves u_1 in Q_u fixed. Set $\widehat{u}(gMAN) = u_1(g)$, for all $g \in G$. Then $\widehat{u}$ has the required property.

$\Leftarrow$ If $\widehat{u}$ is as described in the theorem, then u defined by (2.38) is easily shown to have the mean value property. ■

Exercise 2.1.27. Prove this last statement; i.e., the $\Leftarrow$ of Theorem 2.1.6.

Another standard result in potential theory generalizes as follows.

Theorem 2.1.7. *Suppose that F is continuous on the boundary of G/K and that $P(x,b)$ =Poisson's kernel from (2.39). Set*

$$u(gK) = \int_{K/M} P(gK, \overline{k})F(\overline{k})\, d\overline{k}, \qquad \text{for } g \in G.$$

Then u has boundary values given by F; i.e.,

$$\lim_{t\to\infty} u\left((k \exp tH)K\right) = F(kM), \qquad \text{for } k \in K,\; H \in \mathfrak{a}^+,$$

where $\mathfrak{a}^+$ is a fixed Weyl chamber in $\mathfrak{a}$ (from the Iwasawa decomposition).

Proof (Helgason [275, pp. 47–48]). First one must identify the boundary K/M (up to set of measure 0) with $\overline{N}$ the Lie subgroup of G corresponding to the Lie subalgebra

$$\overline{\mathfrak{n}} = \sum_{\alpha \in \Lambda^-}{}^{\oplus}\, \mathfrak{g}_\alpha,$$

where the Weyl chamber is

$$\mathfrak{a}^+ = \{H \in \mathfrak{a} \mid \alpha(H) < 0 \text{ if } \alpha \in \Lambda^-\}.$$

Set $a_t = \exp(tH)$, for $t \in \mathbb{R}$. Write $k(g)$ = the K-part in the KAN Iwasawa decomposition of $g \in G$ and obtain:

$$\int_{K/M} F(a_t(\overline{k}))\, d\overline{k} = \int_{\overline{N}} F(k(\mathrm{Int}(a_t)\overline{n})M)\, \frac{d\overline{k}}{d\overline{n}}\, d\overline{n},$$

since $a_t\overline{n}MAN = a_t\overline{n}a_t^{-1}MAN$. Set $\overline{n} = \exp\sum_{\alpha<0} X_\alpha$, for $X_\alpha \in \mathfrak{g}_\alpha$. Then

$$\begin{aligned}\operatorname{Int}(\exp tH)\,\overline{n} &= \exp\left(\operatorname{Ad}(\exp tH)\sum_{\alpha<0} X_\alpha\right) = \exp\left(e^{\operatorname{ad} tH}\sum_{\alpha<0} X_\alpha\right)\\ &= \exp\left(\sum_{\alpha<0} e^{t\alpha(H)}\right) X_\alpha \to e, \quad \text{as} \quad t\to\infty,\end{aligned}$$

where e denotes the identity, because $\alpha(H) < 0$ if $H \in \mathfrak{a}^+$. ■

2.1.7 Special Functions and Harmonic Analysis on Symmetric Spaces

It is now possible to discuss various types of special functions on the symmetric space $K\backslash G$ of a noncompact real semisimple Lie group G. We shall view G as acting on the right in order to remain close to the notation that we used in Section 1.2. The basic eigenfunction of the invariant differential operators on $K\backslash G$ is the **power function** $p(Kg)$ defined as follows. Let $\lambda : \mathfrak{a} \to \mathbb{C}$ be a linear functional over $\mathbb{R}$; i.e., $\lambda \in \mathfrak{a}^*$. For $g \in G$, with Iwasawa decomposition $g = kan$, write $H(g) = \log a \in \mathfrak{a}$. Then define the **power function**

$$p_\lambda(Kg) = \exp(\lambda(H(g))). \tag{2.40}$$

The power function is indeed an eigenfunction for all the G-invariant differential operators $D \in D(K\backslash G)$. The proof is the same as that for Proposition 1.2.1 of Section 1.2.1. We know that if $t = a_1n_1$, for $a_1 \in A,\ n_1 \in N$, we have

$$p_\lambda((Kx)t) = p_\lambda(Kx)p_\lambda(Kt),$$

since $x = kan$ implies that $xa_1 = kana_1 = kaa_1(a_1^{-1}na_1)$ with $a_1^{-1}na_1 \in N$. Then, if $D \in D(K\backslash G),\ t \in AN$, and $Ky = (Kx)t$,

$$Dp_\lambda(Ky) = (Dp_\lambda)(Kxt) = (Dp_\lambda(Kx))p_\lambda(Kt).$$

Set $x = e =$ the identity, to complete the proof that the power function is indeed an eigenfunction for all the invariant differential operators on $K\backslash G$.

Define a **spherical function** of $K\backslash G$ to be a function $f : K\backslash G \to \mathbb{C}$ such that $f(K) = 1$ and such that f is a K-invariant eigenfunction of all the invariant differential operators in $D(K\backslash G)$.

Spherical functions can be built up out of power functions as in Theorem 1.2.3 of Section 1.2.3. The following theorem is proved in Helgason [273, Ch. 10]. In fact, the proof that we gave in Section 1.2.3 generalizes. It is also possible to extend the rest of the Theorem 1.2.3 of Section 1.2.3 to $K\backslash G$.

Theorem 2.1.8 (Harish-Chandra). *A spherical function has the form*

$$h_\lambda(Kg) = \int_K p_\lambda(Kgk)\, dk,$$

where $\lambda = i\mu - p$, $\rho = \frac{1}{2}\sum_{\alpha>0}\alpha$. *Moreover spherical functions* $h_{i\mu-\rho}$ *are invariant under the Weyl group acting on the* μ*-variable. Here* $\mu \in \mathfrak{a}^* =$ *the dual vector space of* $\mathfrak{a}$.

Harish-Chandra [263] obtained the asymptotics of the spherical function:

$$\left.\begin{aligned} &h_{i\mu-\rho}(\exp H) \sim e^{-\rho(H)} \sum_{s\in W} c(s\mu) e^{is\mu(H)}, \quad \text{as} \quad H \to \infty, \quad H \in \mathfrak{a}^+, \\ &c(\mu) = \int\limits_{\overline{N}} \exp\{(-i\mu - \rho)(H(\overline{n}))\}\, d\overline{n}. \end{aligned}\right\} \tag{2.41}$$

Gindikin and Karpelevic [220] obtained the explicit formula for the c-function:

$$c(\mu) = I(i\mu)/I(\rho), \quad \text{where } I(\nu) = \prod_{\alpha>0} B\left(\frac{1}{2}m_\alpha, \frac{1}{4}m_{\alpha/2} + \frac{(\nu,\alpha)}{(\alpha,\alpha)}\right), \quad \nu \in \mathfrak{a}^*, \tag{2.42}$$

where B is the beta function (not to be confused with the Killing form), $m_\alpha = \dim_{\mathbb{R}} \mathfrak{g}_\alpha$, $\rho = \frac{1}{2}\sum_{\alpha>0}\alpha$. Here (ν,α) denotes the inner product on the dual space $\mathfrak{a}^*$ induced by the Killing form of $\mathfrak{g}$ restricted to $\mathfrak{a}$ (a form which is automatically positive definite).

This concludes what we have to say about spherical functions. It would be interesting to look at analogues of Bessel and Whittaker functions for general symmetric spaces, but we will not do this here.

The asymptotics and functional equations of the spherical functions h_λ are sufficient to study the **Helgason–Fourier transform** of $f : K\backslash G \to \mathbb{C}$ defined by:

$$\mathcal{H}f(\lambda, \overline{k}) = \int\limits_{x\in K\backslash G} f(x)\overline{p_\lambda(xk)}\, dx. \tag{2.43}$$

Here $\lambda \in (\mathfrak{a}^*)^c$, the complexification of the dual vector space to $\mathfrak{a}$, $\overline{k} = kM \in K/M$, xg, for $x \in K\backslash G$ and $g \in G$, denotes the right action given by $(Kh)g = K(hg)$, and dx denotes the G-invariant volume on the symmetric space $K\backslash G$.

The **inversion formula** for this transform is due to Harish-Chandra and Helgason (see Helgason [273–282]) and says that

$$f(x) = \int\limits_{\mu\in\mathfrak{a}^*} \int\limits_{B=K/M} \mathcal{H}f(i\mu+\rho,\overline{k})\, p_{i\mu+\rho}(xk)\, |c(\mu)|^{-2}\, d\mu\, d\overline{k}, \tag{2.44}$$

with a suitable normalization of the Euclidean measure on the real vector space $\mathfrak{a}^*$, which is the dual space to $\mathfrak{a}$. The proof of (2.44) is analogous to that of Theorem 1.3.1 in Section 1.3. Helgason [282, Ch. IV] gives a detailed account of the transform for K bi-invariant functions on G. Information on the history of the subject can be found in the same place.

We leave it to the reader to note the remaining properties of the Helgason transform, analogous to those listed in Theorem 1.3.1 in Section 1.3.

Example 2.1.18 ($G = Sp(n,\mathbb{R})$).
Recall our identifications in formula (2.26) in Section 2.1.3 of $K\backslash G$ and $\mathcal{H}_n$ with $\mathcal{P}_n^*$ via:

$$W = \begin{pmatrix} Y & 0 \\ 0 & Y^{-1} \end{pmatrix} \begin{bmatrix} I & X \\ 0 & I \end{bmatrix}, \qquad {}^tX = X \in \mathbb{R}^{n\times n},\ Y \in \mathcal{P}_n.$$

Thus, the power function is:

$$p_s(W) = \prod_{j=1}^{n} |Y_j|^{s_j}, \quad \text{for } W, Y \text{ as above, } s \in \mathbb{C}^n.$$

Viewing $\mathcal{P}_n$ as the subset of $\mathcal{P}_n^*$ consisting of the W with $X = 0$ in the partial Iwasawa decomposition above, it follows that the power function on $\mathcal{P}_n^*$ restricts to the power function on $\mathcal{P}_n$ which was defined in Equation (1.41) of Section 1.2.1.

A possible analogue of the **gamma function** for $\mathcal{P}_n^*$ is the Helgason–Fourier transform of $\exp(-\mathrm{Tr}(W))$:

$$\begin{aligned}
&\int\limits_{\mathcal{P}_n^*} \exp(-\mathrm{Tr}(W)) p_s(W)\, d\mu_n^*(W) \\
&\quad = \int\limits_{Y\in\mathcal{P}_n} \int\limits_{\substack{X\,\in\mathbb{R}^{n\times n}\\ X = {}^tX}} \exp\left\{-\mathrm{Tr}\left(Y + Y^{-1} + Y^{-1}\left[X\right]\right)\right\}\, p_s(Y)|Y|^{-(n+1)/2}\, d\mu_n(Y)\, dX \\
&\quad = \pi^{n(n+1)/4} K_n(s^{\#}|I,I), \qquad s^{\#} = s - \left(0,\ldots,0,\tfrac{1}{2}\right),
\end{aligned}$$

where K_n denotes the K-Bessel function for $\mathcal{P}_n$ defined by formula (1.61) of Section 1.2.2. We saw the case $n = 1$ of this result in formula (3.141) in Section 3.7 of Volume I. The present formula should allow one to generalize (3.142) of Section 3.7, Vol. I, to $Sp(n,\mathbb{Z})$ using the spectral resolution of the G-invariant differential operators on $L^2(\mathcal{H}_n/Sp(n,\mathbb{Z}))$.

Example 2.1.19 ((The Heat Equation on $K \setminus G$) (Gangolli [197, pp. 108–109])). We want to find $u(Kx, t)$ such that

$$\begin{cases} \Delta u = u_t, & \text{where} \quad \Delta = \text{ the Laplacian for } K \setminus G, \\ u(Kx, 0) = f(Kx), & \text{for some given } K\text{-invariant function } f \text{ on } K \setminus G. \end{cases}$$

Now, it can be shown that

$$\Delta p_{i\mu - p} = -\{(\mu, \mu) + (\rho, \rho)\}\, p_{i\mu-\rho}.$$

Thus the same sort of argument that worked in § 1.3.4 shows that **the heat kernel** is

$$G_t(Kx) = \int\limits_{\mu \in \mathfrak{a}^*} \exp\left(-\{(\mu, \mu) + (\rho, \rho)\}t\right) h_{i\mu-\rho}(Kx)\, |c(\mu)|^{-2}\, d\mu$$

and

$$u(t, Kx) = G_t * f, \quad \text{where the convolution is over } G.$$

Recall that convolution was defined in formula (1.24). Gangolli [197] shows that $G_t(Kx)$ has the standard properties of the fundamental solution of the heat equation, just as we saw for $\mathcal{P}_n$ in Exercise 1.3.8 of Section 1.3.4.

As we noted earlier, Ólafsson and Schlichtkrull [479] consider the holomorphic extension of the heat transform $G_t * f$, for f an L^2 function on a general symmetric space. The extension is to the complex crown of the symmetric space. Helgason's conjecture on eigenfunctions of the invariant differential operators being reconstructible from their hyperfunction boundary values can be considered using the crown of the symmetric space. See Gindikin [219].

Helgason [275, pp. 67–68] solves the wave equation on a symmetric space using the Radon transform on $K \setminus G$. He also discusses Huyghen's principle for a symmetric space. It is also shown by Helgason that eigenfunctions of $D(G/K)$ can be expressed as a Poisson integral over the boundary of the symmetric space.

2.1.8 An Example of a Symmetric Space of Type IV: The Quaternionic Upper Half 3-Space

References for this example include Belinfante and Kolman [41], Bougerol [73], Elstrodt et al. [165–168], Jauch [331], Kubota [373–375], Maass [426, Ch 1], Mennicke [442, 443], Sarnak [526], and Marie-France Vignéras [633].

First we need a brief review of **quaternions**. See also Volume I, p. 218. The quaternions, denoted $\mathbb{H}$ for Hamilton, form a division ring or noncommutative field:

$$\mathbb{H} = \mathbb{R} \oplus \mathbb{R}i \oplus \mathbb{R}j \oplus \mathbb{R}k,$$

where $ij = k = -ji$, $jk = i = -kj$, $ki = j = -ik$, $i^2 = j^2 = k^2 = -1$. The **reduced norm** of a quaternion $q = a + bi + cj + dk$, with a, b, c, d real is $\mathrm{Nrd}\,(q) = qq^c$, with the **conjugate** $q^c = a - bi - cj - dk$. The Euclidean length of the quaternion thought of as a vector in $\mathbb{R}^4$ is:

$$\|q\| = \sqrt{a^2 + b^2 + c^2 + d^2} = \sqrt{\mathrm{Nrd}\,(q)},$$

All goes very much as with the complex numbers except that things do not commute.

It is possible to represent quaternions by complex 2×2 matrices via:

$$\begin{aligned} 1 &\mapsto \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ i &\mapsto \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix} = -i\sigma_1 \\ j &\mapsto \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = -i\sigma_2 \\ k &\mapsto \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} = -i\sigma_3. \end{aligned}$$

These matrices are $-i$ times the Pauli matrices $\sigma_1, \sigma_2, \sigma_3$ from quantum mechanics.

One can view $SU(2)$ as the unit quaternions via such an identification. Thus $SU(2)$ is simply connected. Call the preceding map from quaternions to matrices f. Then we claim $SU(2) \cong \{f(q) \mid \|q\| = 1\}$.

The group $K = SU(2)$ can be mapped onto $SO(3, \mathbb{R})$ via a homomorphism of fundamental importance in the Dirac theory of electron spin. The map is given by taking Q in $SU(2)$ to $A = (a_{ij})$ in $\mathbb{R}^{3\times3}$ via $a_{ij} = \mathrm{Tr}(Q\sigma_i {}^t\overline{Q}\sigma_j)/2$. The map is onto with kernel the center of $SU(2)$.

After this brief discussion of quaternions, we can give various descriptions of a symmetric space that has been of interest to number theorists and physicists. The space is

$$SL(2, \mathbb{C})/SU(2).$$

It fits into type IV of Cartan's classification of symmetric spaces. We can identify this space as the space of positive Hermitian matrices of determinant one:

$$\mathcal{SP}_2^c = \left\{ Y \in \mathbb{C}^{2\times2} \;\middle|\; Y = {}^t\overline{Y},\ Y \text{ positive},\ \ |Y| = 1 \right\}.$$

The identification is:

$$\begin{aligned} SL(2, \mathbb{C})/SU(2) &\to \mathcal{SP}_2^c, \\ gSU(2) &\mapsto g\,{}^t\overline{g}. \end{aligned} \tag{2.45}$$

Here, we define a **positive Hermitian matrix** Y to be a Hermitian matrix $Y \in \mathbb{C}^{n\times n}$ such that $Y\{x\} = {}^t\bar{x}Yx > 0$ for all $x \in \mathbb{C}^n - 0$. These matrices are quite analogous to ordinary positive matrices. We could rewrite Chapter 3 of Vol. I and Chapter 1 of this Volume in the Hermitian case, if we had the time. Mercifully Elstrodt et al. [168] have done this and more.

By generalizing the Iwasawa decomposition, one sees that the coset representatives $g \in SL(2,\mathbb{C})$ for $SL(2,\mathbb{C})/SU(2)$ can be chosen to have the form:

$$g = \begin{pmatrix} \sqrt{t} & z/\sqrt{t} \\ 0 & 1/\sqrt{t} \end{pmatrix},\ z \in \mathbb{C},\ t > 0,\ z = x + iy. \tag{2.46}$$

This allows us to identify $SL(2,\mathbb{C})/SU(2)$ with **the quaternionic upper half space**:

$$\mathcal{H}^c = \{z + kt = x + iy + kt \mid x, y \in \mathbb{R},\ t > 0\}\,. \tag{2.47}$$

Thus the elements of $\mathcal{H}^c$ are quaternions with j-coordinate equal to zero and positive k-coordinate. The mapping from $SL(2,\mathbb{C})/SU(2)$ to $\mathcal{H}^c$ sends $gSU(2)$ with g given by (2.46) to $z + kt$.

The **action** of a matrix

$$g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2,\mathbb{C})$$

on an element q of the quaternionic upper half plane is:

$$g(q) = (aq + b)(cq + d)^{-1} = q^* \ \text{ with } \ t^* = t\|cq + d\|^{-2}. \tag{2.48}$$

Recall that it is all right to divide by quaternions (on one side or the other), but it is not all right to interchange the order of multiplication.

The **action** of $g \in SL(2,\mathbb{C})$ on $Y \in \mathcal{SP}_2^c$ is

$$Y \mapsto Y\{g\} = {}^t\overline{g}Yg. \tag{2.49}$$

Using this action we identify our symmetric space as $SU(2)\backslash SL(2,\mathbb{C})$; i.e., we consider left rather than right cosets.

Exercise 2.1.28. Check that the three group actions are preserved in our identifications of $SL(2,\mathbb{C})/SU(2)$ with $\mathcal{H}^c$ and $\mathcal{SP}_2^c$.

Exercise 2.1.29. Show that the invariant arc length, volume element, and Laplacian on $\mathcal{H}^c$ are given by:

$$\begin{aligned} ds^2 &= t^{-2}(dx^2 + dy^2 + dt^2); \\ d\mu &= t^{-3}\, dx\, dy\, dt; \\ \Delta &= t^2(\partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial t^2) - t\partial/\partial t. \end{aligned}$$

Exercise 2.1.30. (a) Show that a spherical function on $\mathcal{SP}_2^c$ has the form

$$h_\lambda(Y) = \tfrac{2\sin(\lambda r/2)}{\lambda \sinh r}, \quad \text{if } \; Y = a_r[k],$$
$$a_r = \begin{pmatrix} \exp(r/2) & 0 \\ 0 & \exp(-r/2) \end{pmatrix}, \qquad k \in SU(2) = K.$$

Here r is the geodesic radial coordinate in the polar coordinate decomposition of Y in $\mathcal{SP}_2^c$.

(b) Show that if f is in $L^1(\mathcal{SP}_2^c/K)$, then

$$\int\limits_{\mathcal{SP}_2^c} f(Y)d\mu = \int\limits_{\mathbb{R}} f(a_r)\sinh^2 r\, dr.$$

(c) Use part (b) to show that the Helgason–Fourier transform for K-invariant functions on $\mathcal{SP}_2^c/K$ has the form:

$$\widehat{f}(\lambda) = \frac{2}{\lambda i}\int\limits_{\mathbb{R}} \exp(i\lambda t/2) f(a_t)\sinh t\, dt.$$

(d) Use the inversion formula for the ordinary Euclidean Fourier transform from Section 1.2 of Volume I to show that the spectral measure for Fourier inversion on $\mathcal{SP}_2^c$ is:

$$\frac{|\lambda|^2}{16}d\lambda, \quad \text{where} \quad d\lambda = \text{Lebesgue measure on } \mathbb{R}.$$

(e) Find the fundamental solution for the heat equation on $\mathcal{SP}_2^c$.

Hint. See Bougerol [73], Karpelevich et al. [340], or Burridge and Papanicolaou [92].

Exercise 2.1.30 shows that harmonic analysis on $SL(2,\mathbb{C})$ is far simpler than that on $SL(2,\mathbb{R})$. This is an example of a general phenomenon (see Helgason [276, p. 31] for the generalization of part (a) of Exercise 2.1.30). It would also be nice to consider analogues of Bessel and Whittaker functions for $\mathcal{H}^c$.

This completes our brief sketch of the theory of harmonic analysis on general symmetric spaces. There are many applications, other than those mentioned so far. For example, Resnikoff [506] considers the consequences of using the geometries of the spaces $(\mathbb{R}^+)^3$ or $(\mathbb{R}^+ \times SL(2,\mathbb{R})/SO(2))$ as models for color perception. An experiment is posed for distinguishing which geometry gives a more accurate model. Other references for the general theory are Gurarie [254] and Wawrzyńczyk [657].

2.2 Geometry and Analysis on $\Gamma\backslash G/K$

"Say what you know, do what you must, and whatever will be, will be."

Sofya Kovalevskaya's maxim from her paper [368] quoted in Pelageya Kochina [358, p. 168]

2.2.1 Fundamental Domains

Our goal for the remainder of this volume is to give a very brief sketch of parts of the story of harmonic analysis on $\Gamma\backslash G/K$, and automorphic forms for certain subgroups Γ of G **acting discontinuously** on the symmetric space $X = G/K$. This means that for each $x \in X$, the set of images of x under Γ has no limit point in X. We will concentrate on two specific discontinuous groups: $GL(n, \mathfrak{O}_K)$, where $\mathfrak{O}_K$ is the ring of integers in an algebraic number field K,[2] and the **Siegel modular group** $Sp(n, \mathbb{Z})$. Here $GL(n, \mathfrak{O}_K)$ is the **modular group over an algebraic number field** which consists of $n \times n$ matrices γ such that both γ and γ^{-1} have entries in $\mathfrak{O}_K$. See Section 1.4 of Volume I for the necessary definitions from algebraic number theory. The group $Sp(n, \mathbb{Z})$ consists of all symplectic $2n \times 2n$ integral matrices.

There are many reasons to study such discontinuous groups. Of course knowledge of $GL(n, \mathfrak{O}_K)$ and related groups leads to greater understanding of the arithmetic of K itself. For example, there are applications to explicit class field theory, distribution of Gauss sums, values of Dedekind zeta functions and L-functions, asymptotics of units, elliptic curves, quadratic forms, and abelian varieties. References for these subjects include: Andrianov [13, 14], Bruinier et al. [82], Borel and Casselman [66], Borel and Mostow [68], Elstrodt et al. [165–168], Freitag [185, 186], Gelbart [208, 209], Goldfeld et al. [231], Heath-Brown and Patterson [267], Hecke [268, pp. 21–114], Jacquet and Langlands [324], Klingen [355], Kubota [373–376], Langlands [394], Maass [426], Mennicke [442, 443], Saito [525], Sarnak [527], Séminaire Cartan [547], Shimura [555], Shintani [558–560], Siegel [563, 565], Tunnell [620, 621], and Weil [660].

The Siegel modular group $Sp(n, \mathbb{Z})$ and kindred groups appear in many diverse areas of physics, often via the connections with abelian integrals and Riemann theta functions which arise in many theories from boson fields to solitons. See Cartier's article in Borel and Mostow [68, pp. 361–386], Cooke [124], Dubrovin et al. [143], Linda Keen [345], Pelageya Kochina [358], Sofya Kovalevskaya [368], Gérard Lion and Michèle Vergne [406], Lonngren and Scott [407], McKean and Trubowitz [440], Monastyrsky and Perelomov [457], Mumford [471], Novikov [474], Perelomov [484], Shale [551], and Wallach [652].

---

[2] Hopefully the beleaguered reader will not be too confused by our use of K for the maximal compact subgroup of G as well as an algebraic number field.

Theta functions also play a major role in the analytic theory of quadratic forms. See Siegel [565, Vol. I, pp. 326–405, 410–443, 469–548] and Weil [662, Vol. 2, pp. 1–157].

Here we seek to outline the foundations of a building which would ultimately encompass the generalization of Sections 3.3–3.7 of Volume I to these new discontinuous groups Γ. Our achievements will be pitiful compared with what is required. In particular, we will not say much about extensions of Section 3.7 of Volume I; i.e., the analogues of the non-Euclidean Poisson summation formula and the Selberg trace formula. Such results have already found various arithmetic and geometric applications, e.g., in computing dimensions of spaces of holomorphic automorphic forms. There are also results on units in number fields over imaginary quadratic fields and elliptic curves over imaginary quadratic fields. References for such work include: Christian [109], Eie [160], Efrat [152], Elstrodt et al. [165–168], Hashimoto's article in Hejhal et al. [272, pp. 253–276], Hashimoto [264], Langlands [389–395], Mennicke [442, 443], Morita [463], Müller [469], Petra Ploch [489], Sarnak [527], Tanigawa [591], Marie-France Vignéras [630, 631], Yamazaki [673], and Zograf [677].

General references for this section include: Andrianov [9–14], Baily [32], Hel Braun [74–76], Bruinier et al. [82], Christian [108], Elstrodt et al. [168], Freitag [185, 186], Gelfand, Graev, and Piatetski-Shapiro [214], Hecke [268], Hirzebruch and Van der Geer [299], Klingen [355], Maass [426], Mennicke [442], Séminaire Cartan [547], Shimura [554], Siegel [563–565], and Weil [658, 660, 662].

The following quote is from Van der Geer in [82, p. 182]:

> The general theory of automorphic representations provides a generalization of the theory of elliptic modular forms. But despite the obvious merits of this approach some of the attractive explicit features of the $g = 1$ [i.e., $SL(2,\mathbb{R})$] theory are lost in the generalization.

Our first topic is fundamental domains $K\backslash G/\Gamma$, for our favorite examples.

Example 2.2.1 (The Picard Modular Group).

Let K be the number field $\mathbb{Q}(i)$ with ring of Gaussian integers

$$\mathfrak{O}_K = \mathbb{Z}[i] = \{x + iy \mid x, y \in \mathbb{Z}\}.$$

Here $i = \sqrt{-1}$. The **Picard modular group** is defined to be

$$\Gamma = SL(2,\mathfrak{O}_K) = \left\{ \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \middle| \; a,b,c,d \in \mathfrak{O}_K, \;\; \det\gamma = 1 \right\}.$$

The group $SL(2,\mathfrak{O}_K)$ acts discontinuously on the quaternionic upper half space $\mathcal{H}^c$ by fractional linear transformation defined by formula (2.48) in Section 2.1.8. An equivalent version of this action from formula (2.49) in the same subsection gives the action of $\gamma \in SL(2,\mathfrak{O}_K)$ on a positive determinant one Hermitian matrix $Y \in \mathcal{SP}_2^c$ via:

$$Y \mapsto Y\{\gamma\} = {}^t\overline{\gamma} Y \gamma.$$

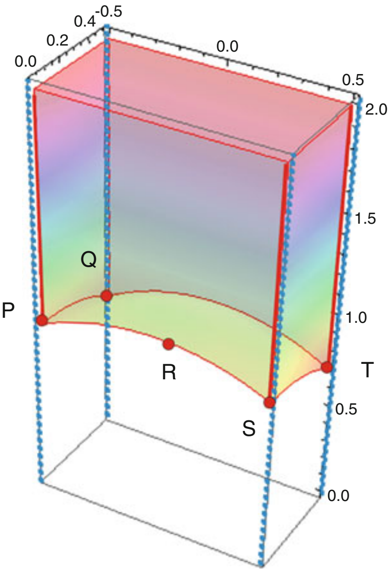

Fig. 2.1 A fundamental domain for $SL(2,\mathbb{Z}[i])$ in the quaternionic upper half plane:

$$D=\left\{ x+iy+kt \;\middle|\; \begin{array}{c} |x|\le .5,\ t>0 \\ 0\le y\le .5 \\ x^2+y^2+t^2\ge 1 \end{array} \right\}.$$

The labeled points are:

$$P=\left(-\frac{1}{2},0,\sqrt{3}/2\right), Q=\left(-\frac{1}{2},\frac{1}{2},1/\sqrt{2}\right),$$

$$R=(0,0,1), S=\left(\frac{1}{2},0,\sqrt{3}/2\right), T=\left(\frac{1}{2},\frac{1}{2},1/\sqrt{2}\right)$$

A fundamental domain for the action of $SL(2,\mathfrak{O}_K)$ on $\mathcal{H}^c$ was determined by Picard [488] and is pictured in Figure 2.1. Various views of a tessellation of hyperbolic 3-space $\mathcal{H}^c$ obtained by transforming this fundamental domain by elements of $\Gamma = SL(2,\mathfrak{O}_K)$ are shown in Figures 2.2, 2.3, 2.4, 2.5, and 2.6. Figures 2.7 and 2.8 show Cayley transforms of this tessellation which are inside of the unit sphere in 3-space. The figures are shown in stereo. If you stare at the two versions of the picture, one for each eye, you should be able to see a 3D tessellation. All of the tessellations were created by Mark Eggert using one of UCSD's VAX computers in the 1980s. Part of a tessellation obtained by taking a union of the fundamental domain in Figure 2.1 with its image under $y\mapsto -y$ can be found just before the index of this volume.

As for $SL(2,\mathbb{Z})$ (see Exercise 3.3.1, Volume I), the sides of the fundamental domain are mapped to each other by generators of Γ, which are in this case:

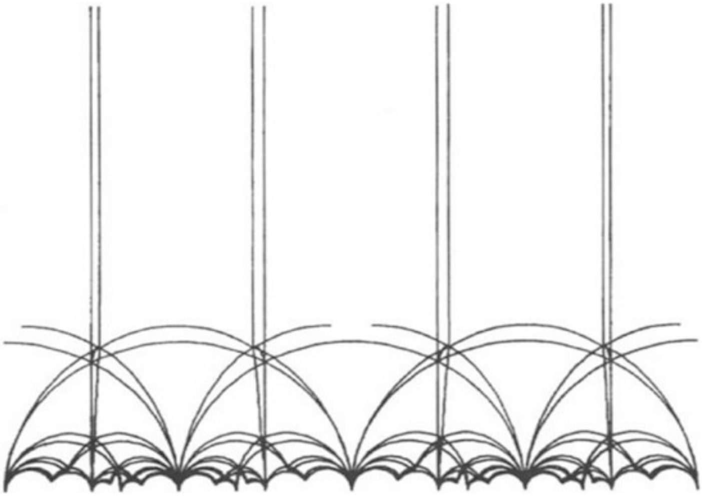

Fig. 2.2 Tessellation of the quaternionic upper halfplane from $SL(2, \mathbb{Z}[i])$ in stereo. It may help to put a division between the two halves of this figure and those that follow, in order to produce the 3D effect. The figure was created by Mark Eggert using the UCSD VAX computer

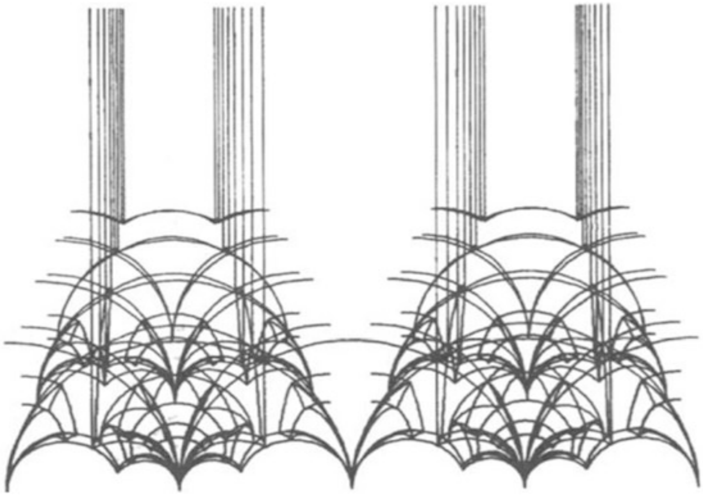

Fig. 2.3 Tessellation of the quaternionic upper half plane from $SL(2, \mathbb{Z}[i])$ in stereo. The figure was created by Mark Eggert using the UCSD VAX computer

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & i \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Given any imaginary quadratic number field $K = Q(\sqrt{D})$, of discriminant $D > 0$, one can consider $SL(2, \mathfrak{O}_K)$, $\mathfrak{O}_K =$ the ring of algebraic integers in K,

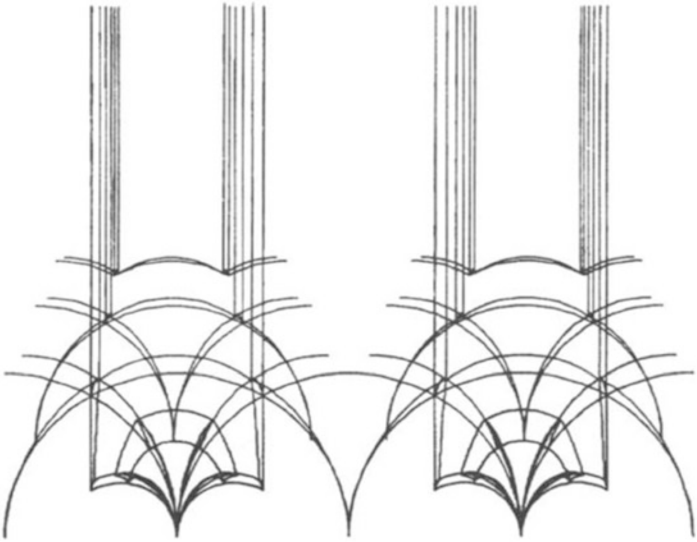

Fig. 2.4 Tessellation of the quaternionic upper half plane from $SL(2, \mathbb{Z}[i])$ in stereo. The figure was created by Mark Eggert using the UCSD VAX computer

Fig. 2.5 Tessellation of the quaternionic upper half plane from $SL(2, \mathbb{Z}[i])$ in stereo. The figure was created by Mark Eggert using the UCSD VAX computer

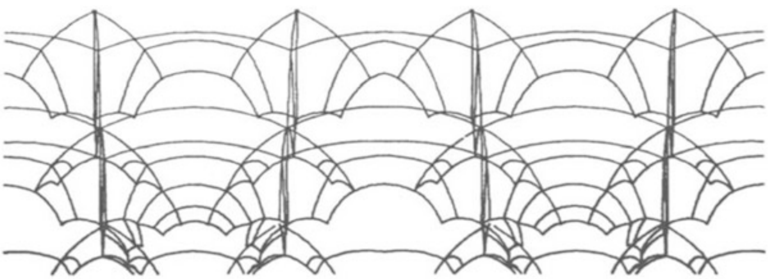

Fig. 2.6 Tessellation of the quaternionic upper half plane from $SL(2, \mathbb{Z}[i])$ in stereo. The figure was created by Mark Eggert using the UCSD VAX computer

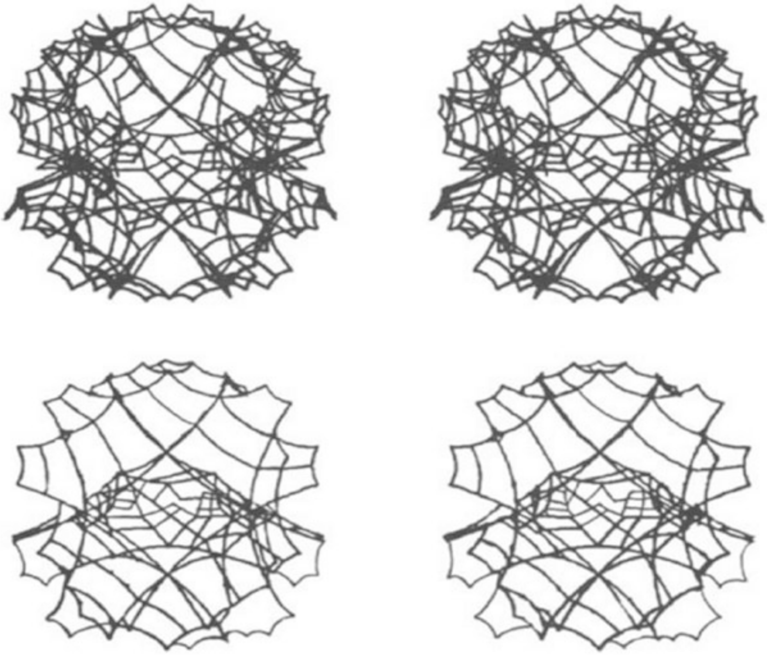

Fig. 2.7 Stereo tessellation of the unit sphere obtained by mapping the preceding tessellations of the quaternionic upper half plane into the unit sphere by a Cayley transform. The mapping for this figure and Figure 2.8 is

$$q \mapsto (q-k)(-kq+1)^{-1}.$$

The figure was created by Mark Eggert using the UCSD VAX computer

$$SL(2,\mathfrak{O}_K) = \left\{ \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \;\middle|\; a,b,c,d \in \mathfrak{O}_K,\ \det\gamma = 1 \right\}.$$

This investigation was begun by Bianchi [50]. Humbert [310] showed that the **volume of the fundamental domain** $SL(2,\mathfrak{O}_K)\backslash\mathcal{H}^c$ is

$$\frac{|D_K|^{3/2}\zeta_K(2)}{4\pi^2}, \quad \text{where} \quad \zeta_K(s) = \ \text{the Dedekind zeta function of} \ K$$

(see Section 1.4 of Volume I for the definitions of ζ_K and the discriminant D_K). The geometry of the fundamental domain is thus closely associated with the arithmetic of the number field K. In particular, the number of cusps of the fundamental domain is equal to the class number of K, which was defined in Section 1.4 of Vol. I. We will demonstrate this fact in Proposition 2.1.1.

Siegel gives two methods to prove formulas for the volume of fundamental domains of this sort. See Siegel [565, Vol. I, pp. 464–465, Vol. II, pp. 330–331, and Vol. III, pp. 39–46, 328–333]. One of Siegel's methods is the one we used in Theorem 1.4.4 of Section 1.4.4 to find the volume of the fundamental domain

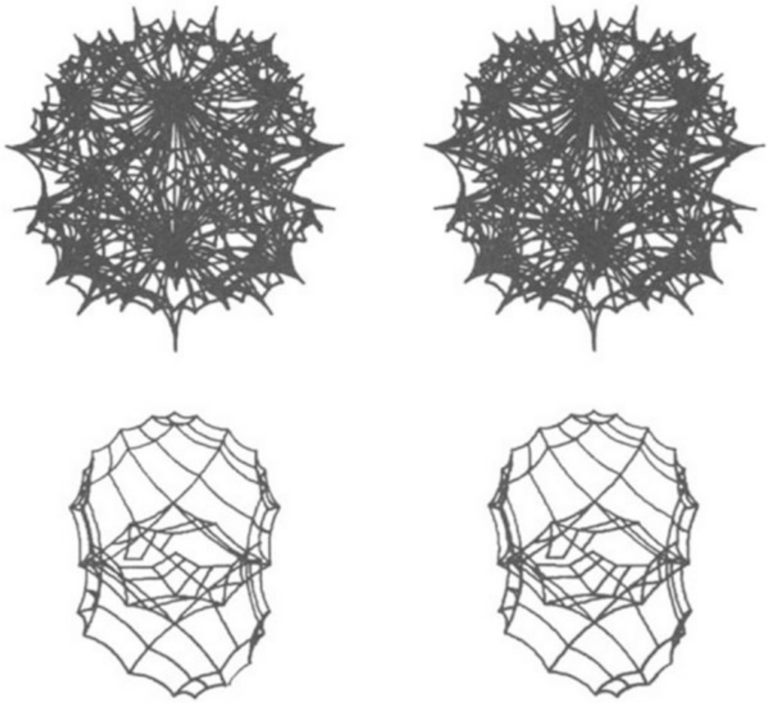

Fig. 2.8 Stereo tessellation of the unit sphere obtained by mapping the preceding tessellations of the quaternionic upper half plane into the unit sphere by a Cayley transform. The figure was created by Mark Eggert using the UCSD VAX computer

for $GL(n,\mathbb{Z})$ via Siegel's integral formula. The other method involves finding the residues of Eisenstein series like (1.174) in Section 1.4.1, using the method of theta functions.

References for fundamental domains in quaternionic upper half space include: Ahlfors [2], Elstrodt et al. [165–168], Humbert [309, 310], Kubota [373–376], Mennicke [442, 443], Milnor [450], Sarnak [527], and Stark [575].

The next example to be considered is the analogue of Example 2.2.1 for real quadratic fields K.

Example 2.2.2 (The Hilbert Modular Group).

Suppose that K is a real quadratic number field and $K = \mathbb{Q}(\sqrt{D})$ has positive discriminant D; e.g., $K = \mathbb{Q}(\sqrt{2})$. Such a field has, as mentioned in Section 1.4 of Volume I, two conjugations mapping K into the field of real numbers and denoted $x^{(1)}, x^{(2)}$, for $x \in K$. If $x = a + b\sqrt{d} \in K$, with $a, b \in \mathbb{Q}$, then $x^{(1)} = x$ and $x^{(2)} = a - b\sqrt{D}$. Form the group

$$\Gamma = SL(2,\mathfrak{O}_K) = \left\{ \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \;\middle|\; a,b,c,d \in \mathfrak{O}_K,\ \det\gamma = 1 \right\}.$$

This is called the **Hilbert modular group** for the field K. Then Γ acts discontinuously on the product $\mathcal{H}^2$ of two ordinary upper half planes via

$$\gamma(z^{(1)}, z^{(2)}) = (\gamma^{(1)} z^{(1)}, \gamma^{(2)} z^{(2)}), \quad \text{for} \quad z^{(j)} \in \mathcal{H}, \tag{2.50}$$

and $\gamma^{(j)}$ denoting the matrix each of whose entries is obtained by taking the jth conjugate of the corresponding entry of $\gamma \in \Gamma$, $j = 1, 2$. Here $\gamma^{(j)}$ acts on $z^{(j)}$ by fractional linear transformation as in Chapter 3 of Volume I.

Exercise 2.2.1. (a) Show that the action of $\Gamma = SL(2, \mathfrak{O}_K)$, for a real quadratic field K, on the ordinary upper half plane H via $z \mapsto \gamma z$, for $z \in H$, is not discontinuous.

(b) Show that the action of $\Gamma = SL(2, \mathfrak{O}_K)$, for a real quadratic field K, on $\mathcal{H}^2$ defined by (2.50) is discontinuous.

Hint. (a) Make use of the units in K; that is $x \in \mathfrak{O}_K$ such that $x^{-1} \in \mathfrak{O}_K$.

It is an easy matter to generalize to $SL(2, \mathfrak{O}_K)$ where K is any **totally real algebraic number field** K; i.e., a number field K such that every conjugation is real-valued. Then the Hilbert modular group $SL(2, \mathfrak{O}_K)$ acts discontinuously on $\mathcal{H}^m$, where m is the degree of K over $\mathbb{Q}$.

The fundamental domains $\Gamma\backslash\mathcal{H}^m$ have been rather intensively studied. They are $2m$-dimensional and complicated by the existence of units of infinite order in $\mathfrak{O}_K$. The formula for the **volume of the fundamental domain** is

$$2(-2\pi)^m \zeta_K(-1) = 2\pi^{-m} D_K^{3/2} \zeta_K(2),$$

where D_K is the absolute value of the discriminant of K and $\zeta_K(s)$ is Dedekind's zeta function of K. See Klingen [351] for a proof of a much more general result.

References for these results include: Blumenthal [52], Harvey Cohn [116], Freitag [186], Giraud [221], Gundlach [253], Hammond [259], Hirzebruch [296], Hirzebruch and Van der Geer [299], Hirzebruch and Zagier [300, 301], Humbert [309], Klingen [350–352], Maass [415, 416], Resnikoff [505], Shimizu [553], Shimura [554], Siegel [563, 565], Thomas and Vasquez [613], and Weisser [665]. See our earlier comments on volumes of fundamental domains for the Picard type groups.

Example 2.2.3 (The Modular Group over any Number Field).

Suppose that K is any number field, $\mathfrak{O}_K$ its ring of integers, and m its degree over $\mathbb{Q}$. Then, as in Section 1.4 of Volume I, K must be equal to $\mathbb{Q}(a)$ for some complex number a with minimal polynomial $f(x)$ and K is isomorphic to the quotient $\mathbb{Q}[x] / ((f(x))$. So

$$\left.\begin{aligned} &K \otimes_{\mathbb{Q}} \mathbb{R} \cong \mathbb{R}[x]/(f(x)) \cong \sum_{j=1}^{r_1+r_2}{}^{\oplus} E_j, \\ &E_j \cong \begin{cases} \mathbb{R}, j = 1, \ldots, r_1; \\ \mathbb{C}, j = r_1 + 1, \ldots, r_1 + r_2. \end{cases} \end{aligned}\right\} \tag{2.51}$$

Therefore we have m conjugations sending K into E_j by mapping x to $x^{(j)}$, for $j = 1, \ldots, r_1$, and mapping x to $x^{(j)}$ or $\overline{x^{(j)}}$, for $j = r_1 + 1, \ldots, r_1 + r_2$.

What is a positive matrix over a number field? We are actually seeking the "infinite prime part" of an adelic symmetric space (see Cassels and Fröhlich [101], Gelbart [208], Gelfand et al. [214], and Weil [658, 660, 662]).

Define a **positive quadratic form** Y **over the number field** K to be a vector

$$Y = \left(Y^{(1)}, \ldots, Y^{(r_1+r_2)}\right),$$

with $Y^{(j)} \in \mathcal{P}_n$, for $j = 1, \ldots, r_1$ and $Y^{(j)} \in \mathcal{P}_n^c$, $j = r_1 + 1, \ldots, r_1 + r_2$. Here $\mathcal{P}_n$ is the symmetric space of positive real $n \times n$ matrices studied in Chapter 1 of this volume, while $\mathcal{P}_n^c$ is the symmetric space of positive $n \times n$ Hermitian complex matrices; i.e.,

$$\mathcal{P}_n^c = \left\{ Y \in \mathbb{C}^{n \times n} |\ {}^t\overline{Y} = Y,\ Y \text{ positive} \right\} \cong U(n)\backslash GL(n, \mathbb{C}).$$

A complex Hermitian matrix Y is called **positive** if $Y\{x\} = {}^t\overline{x}Yx > 0$ for every column vector $x \in \mathbb{C}^n - 0$. Set $\mathcal{P}_n^K$ = the **space of positive quadratic forms** over K. Clearly this symmetric space will generalize the two preceding examples, if we restrict to the **determinant one subspace**

$$\mathcal{SP}_n^K = \left\{ Y \in \mathcal{P}_n^K \;\middle|\; |Y^{(j)}| = 1,\ j = 1, \ldots, r_1 + r_2 \right\}.$$

Set

$$\Gamma = GL(n, \mathfrak{O}_K) = \left\{ \gamma \in \mathfrak{O}_K^{n \times n} \mid \gamma^{-1} \in \mathfrak{O}_K^{n \times n} \right\}.$$

The **action of the modular group** $GL(n, \mathfrak{O}_K)$ on $Y \in \mathcal{P}_n^K$ is given by:

$$Y \mapsto Y\{A\}, \quad \text{with} \ \ (Y\{A\})^{(j)} = \left(Y^{(j)}\right)\left\{A^{(j)}\right\} = {}^t\overline{A^{(j)}}Y^{(j)}A^{(j)},$$

for $j = 1, \ldots, r_1 + r_2$.Here $A^{(j)}$ denotes the matrix all of whose entries are the jth conjugate of the corresponding entries in A.

There is a long history of looking only at the "infinite prime" part of the symmetric space rather than the adelic version which includes an infinite number of p-adic components, one for each finite prime p—a much more recent construct. Some references are Hecke [268, pp. 21–55], Humbert [309, 310], Klingen [350–352], Ramanathan [497, 498], Siegel [563], Weil [658], and Weyl [666, Vol. IV, pp. 232–264]. Much of our discussion here was inspired by working with John Hunter who considered number-theoretic applications of analogues of Siegel's integral formula (Proposition 2.1.2 of Section 1.4.4) for $SL(n, \mathfrak{O}_K)$, K imaginary quadratic. I regret that John's death prevents publication of his thesis work (see Hunter [311]).

Exercise 2.2.2. Show that $GL(n, \mathfrak{O}_K)$ consists of all matrices in $\mathfrak{O}_K^{n\times n}$ whose determinant is a unit in $\mathfrak{O}_K$.

Fundamental domains for $\mathcal{P}_n^K/GL(n, \mathfrak{O}_K)$ were discussed by Humbert [309], who generalized many of the results that we presented in Section 1.4 for the case that K is the field of rational numbers. Siegel [565] obtains analogues of many of the results of Section 1.4 in various places. For example, Siegel [565, Vol. I, p. 475] gives an analogue of Lemma 1.4.2 of Section 1.4.2. And Siegel [565, pp. 464–465] obtains a formula for the volume of the fundamental domain, in a paper which was to be corrected later [565, Vol. III, pp. 328–333].

We choose not to rewrite all of Section 1.4 in this case. Instead we take up some aspects of the theory when $\Gamma = SL(2, \mathfrak{O}_K)$. In particular, we discuss a result of Maass [416] correcting an error of Blumenthal [52]. This error also appears in Hecke's first paper (see Hecke [268, pp. 21–55 and the notes at the end of the volume]). We want to show that the cusps of the fundamental domain $\mathcal{SP}_2^K/SL(2, \mathfrak{O}_K)$ are in one-to-one correspondence with the ideal classes of K. The cusps are the points of the fundamental domain which are equivalent to infinity under the action of $SL(2, K)$. Thus they are elements of $\widehat{K} = K \cup \{\infty\}$.

Proposition 2.2.1 (The Cusp-Ideal Class Correspondence). *The cusps of the fundamental domain for $\mathcal{SP}_2^K/SL(2, \mathfrak{O}_K)$, K any number field, are in one-to-one correspondence with the ideal class group I_K of K.*

Proof. See Siegel [563, p. 242]. Let h denote the class number of K. Choose fixed integral ideals $\mathfrak{a}_1, \ldots, \mathfrak{a}_h$ representing the ideal class group I_K. We want to show that the elements of $\widehat{K} = K \cup \{\infty\}$ are divided into h equivalence classes by the action of

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, \mathfrak{O}_K) \ \text{ on } \ x \in \widehat{K} \ \text{ defined by } \ \gamma(x) = \frac{ax+b}{cx+d}, \ \gamma(\infty) = a/c.$$

Suppose that $x = p/s$ with $p, s \in \mathfrak{O}_K$. Here we write $\infty = 1/0$. Then define $f(x)$ to be the integral ideal (p, s) which is generated by p and s. Note that the ideal class of $f(x)$ is well defined. For if $x = p_1/s_1 = p_2/s_2$, then $\mathfrak{a}_2 = k\mathfrak{a}_1$, for $k = p_2/p_1$, where $\mathfrak{a}_i = f(p_i/s_i)$.

So there is an induced map $\overline{f} : \widehat{K}/SL(2, \mathfrak{O}_K) \to I_K$, since if $\gamma \in SL(2, \mathfrak{O}_K)$ and

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \ \text{ then } \ f(\gamma(p/s)) = (ap + bs, cp + ds) \subset (p, s).$$

The reverse inclusion must hold as well because the determinant of γ is one.

The map $\overline{f}$ is onto, since every ideal in $\mathfrak{O}_K$ has at most two generators (see Pollard [490]).

In order to show that $\overline{f}$ is one-to-one, you will probably first think of the following argument. Suppose that $f(p_1/s_1) = kf(p_2/s_2)$ for some $k \in K$. If $k = \omega/\tau$, for $\omega, \tau \in \mathfrak{O}_K$, then we see that

$$\tau(ap_1 + bs_1) = \omega p_2, \ \text{ and } \ \tau(cp_1 + ds_1) = \omega s_2,$$

for some

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL(2, \mathfrak{O}_K).$$

It follows that

$$\frac{ap_1 + bs_1}{cp_1 + ds_1} = \frac{p_2}{s_2}.$$

This says that p_1/s_1 and p_2/s_2 are indeed equivalent modulo $GL(2, \mathfrak{O}_K)$. But unfortunately we need to know that they are equivalent modulo $SL(2, \mathfrak{O}_K)$. The difference between special and general linear groups over $\mathfrak{O}_K$ can be rather large, thanks to the presence of lots of units.

Since our boat seems to have stopped moving, we take a different tack. This time we follow Siegel's argument. Suppose that $\mathfrak{a}^{-1}$ denotes the inverse ideal to $\mathfrak{a}$. Then $\mathfrak{a}\mathfrak{a}^{-1} = \mathfrak{O}_K = (1)$ and thus

$$p_1 v_1 - s_1 u_1 = 1 \text{ and } p_2 v_2 - s_2 u_2 = 1 \text{ for some } u_i, v_i \in \mathfrak{a}_i^{-1}.$$

Set

$$A_i = \begin{pmatrix} p_i & u_i \\ s_i & v_i \end{pmatrix}, \ i = 1, 2,$$

and note that although A_i only has entries in K, the product $A_2 A_1^{-1}$ is actually in $SL(2, \mathfrak{O}_K)$. For we have assumed that $\bar{f}(p_i/s_i)$ both equal $\mathfrak{a}_1$, say. Thus if you compute the first entry of $A_2 A_1^{-1}$, for example, you find it is $p_2 v_1 - u_2 s_1$, which is in the ideal $\mathfrak{a}_2 \mathfrak{a}_1^{-1} + \mathfrak{a}_1 a_2^{-1} = \mathfrak{O}_K$, since $\mathfrak{a}_1 = \mathfrak{a}_2$. It is important that we have chosen a fixed set of representatives of our ideal classes. But note here that $f(p_1/s_1) = kf(p_2/s_2)$, for $k \in K$, implies that we can assume $k = 1$ by replacing p_2/s_2 by $(kp_2)/(ks_2)$. To see that $A_2 A_1^{-1}$ does indeed take p_1/s_1 to p_2/s_2, note that A_i maps ∞ to p_i/s_i (acting by fractional linear transformation).

This completes the proof of Proposition 2.2.1. ■

Lemma 2.2.1. *The stabilizer in $SL(2, \mathfrak{O}_K)$ of a cusp $x_i = p_i/s_i$ for the fundamental domain $\mathcal{SP}_2^K/SL(2, \mathfrak{O}_K)$ is defined by:*

$$\Gamma_{x_i} = \{\gamma \in SL(2, \mathfrak{O}_K) \mid \gamma x_i = x_i\}.$$

Suppose that $x_i = A_i \infty$ and that

$$A_i = \begin{pmatrix} p_i & u_i \\ s_i & v_i \end{pmatrix} \in SL(2, K), \quad \textit{with } \mathfrak{a}_i = (p_i, s_i), \ u_i, v_i \in \mathfrak{a}_i^{-1},$$

for $i = 1, 2, \ldots, h$. Here the ideals $\mathfrak{a}_1, \ldots, \mathfrak{a}_h$ represent the ideal class group I_K. And $\mathfrak{a}^{-1}$ is the inverse ideal to $\mathfrak{a}$. Let U_K denote the group of units in $\mathfrak{O}_K$. Then the stabilizer of a cusp has the form:

$$\Gamma_{x_i} = \left\{ A_i \begin{pmatrix} w & z \\ 0 & w^{-1} \end{pmatrix} A_i^{-1} \,\middle|\, z \in \mathfrak{a}_i^{-2},\ w \in U_K \right\}.$$

Proof. See Exercise 2.2.3 below. The result is clearly true for the infinite cusp. ■

Exercise 2.2.3. Prove the formula for the stabilizer of the infinite cusp in Lemma 2.2.1 above. Then deduce the result for an arbitrary cusp.
Hint. You can find the details in Siegel [563, p. 245]. If γ stabilizes x_i, then $A_i^{-1}\gamma A_i$ stabilizes infinity and therefore

$$A_i^{-1}\gamma A_i = \begin{pmatrix} w & z \\ 0 & w^{-1} \end{pmatrix}.$$

To see that z must lie in $\mathfrak{a}_i^{-2}$, just multiply out the matrices. And note that

$$(ap_i + bs_i)s_i = (cp_i + ds_i)p_i \ \text{ if } \ \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

But then it follows that w must be a unit, since, after division by $\mathfrak{a}_i^2$, we see that:

$$\frac{(ap_i + bs_i)}{\mathfrak{a}_i} = \frac{(p_i)}{\mathfrak{a}_i}, \qquad \frac{(cp_i + ds_i)}{\mathfrak{a}_i} = \frac{(s_i)}{\mathfrak{a}_i}.$$

You also have to multiply out the following matrices:

$$\begin{pmatrix} p_i & u_i \\ s_i & v_i \end{pmatrix}\begin{pmatrix} w & z \\ 0 & w^{-1} \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} p_i & u_i \\ s_i & v_i \end{pmatrix},$$

to see that w is a unit.

Example 2.2.4 (The Siegel Modular Group).

The Siegel modular group $Sp(n, \mathbb{Z})$ is the group of all symplectic matrices with integer entries. It acts discontinuously on Siegel's upper half space $\mathcal{H}_n$ (or on the space of positive symplectic matrices $\mathcal{P}_n^*$) considered in Section 2.1.

Siegel [565, Vol. II, pp. 300–301] shows that a fundamental domain for $Sp(n, \mathbb{Z})\backslash\mathcal{H}_n$ can be obtained by generalizing the method of perpendicular bisectors from Exercise 3.3.6 of Volume I. That is, taking $\Gamma = Sp(n, \mathbb{Z})$, we consider the domain:

$$\mathcal{D}_n = \{Z \in \mathcal{H}_n \mid d(Z, W) \leq d(Z, \gamma W) \ \text{ for all } \ \gamma \in \Gamma\}.$$

Here $W \in \mathcal{H}_n$ is chosen to be a point not fixed by Γ.

As we quoted at the end of Section 1.4.3, Siegel [565, Vol. II, p. 309] said: "The application of the general method [stated above]... would lead to a rather complicated shape of the frontier [boundary] of F." So Siegel goes on to consider another fundamental domain for $Sp(n,\mathbb{Z})\backslash\mathcal{H}_n$. Before we can say what this new domain is, we need to make a definition. Call $C, D \in \mathbb{Z}^{n\times n}$ a **coprime symmetric pair** if $C\,{}^tD = D\,{}^tC$ and the matrices C and D are relatively prime in the sense that: if for any matrix $G \in \mathbb{Q}^{n\times n}$, the matrices GC and GD are both integral, then G must be integral; i.e., in $\mathbb{Z}^{n\times n}$.

Exercise 2.2.4. Show that $C, D \in \mathbb{Z}^{n\times n}$ are a coprime symmetric pair if and only if $(C \;\; D)$ can be completed to a matrix

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \in Sp(n,\mathbb{Z}).$$

Now we can obtain Siegel's fundamental domain $\mathcal{D}_n^S$ for $Sp(n,\mathbb{Z})\backslash\mathcal{H}_n$, as in Siegel [565, Vol. II, p. 108], by finding an analogue of the highest point method from Exercise 3.3.1 of Vol. I. For this, we need to recall from Exercise 2.1.16 of Section 2.1 that if $W = (AZ+B)(CZ+D)^{-1}$, then the imaginary part of W is $Y\left\{(CZ+D)^{-1}\right\}$ if Y is the imaginary part of Z and $Y\{A\} = {}^t\overline{A}YA$. We will take the **height** of $Z = X + iY \in \mathcal{H}_n$ to be the determinant $|Y|$. So we see that $|W| = |Y| \; \|CZ+D\|^{-2}$. This concept of height leads to the following construction for a fundamental domain by a highest point method.

The **Siegel fundamental domain** $\mathcal{D}_n^S$ for $Sp(n,\mathbb{Z})\backslash\mathcal{H}_n$ is the set of $Z \in \mathcal{H}_n$ such that the following three statements hold (if we ignore boundary identifications):

$$\left.\begin{array}{l} \textbf{(1)}\ \|CZ+D\| \geq 1, \text{ for all coprime symmetric pairs } C, D \text{ with } C \neq 0; \\ \textbf{(2)}\ Y = \operatorname{Im} Z \in \mathcal{M}_n = \text{ Minkowski's fundamental domain for } \mathcal{P}_n\backslash GL(n,\mathbb{Z}); \\ \textbf{(3)}\ X = \operatorname{Re} Z,\ X = \left(x_{ij}\right),\ \text{with } \left|x_{ij}\right| \leq 1/2,\ \ 1 \leq i,j \leq n. \end{array}\right\} \tag{2.52}$$

Again there is a certain relation between Siegel's fundamental domain and matrices in $Sp(n,\mathbb{Z})$:

(1) $\begin{pmatrix} A & B \\ C & D \end{pmatrix} \in Sp(n,\mathbb{Z})$ with $C \neq 0$;

(2) $\begin{pmatrix} U & 0 \\ 0 & {}^tU^{-1} \end{pmatrix}, \quad U \in GL(n,\mathbb{Z})$;

(3) $\begin{pmatrix} I & N \\ 0 & I \end{pmatrix}, \quad N = {}^tN \in \mathbb{Z}^{n\times n}$.

In fact, Maass [426, § 11] shows that $Sp(n,\mathbb{Z})$ can be generated by matrices of the form

$$\begin{pmatrix} I & N \\ 0 & I \end{pmatrix}, \; N = {}^tN \in \mathbb{Z}^{n\times n} \quad \text{and} \quad J = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}.$$

The Siegel fundamental domain $\mathcal{D}_n^S$ can be shown to be closed, connected, and bounded by finitely many algebraic hypersurfaces. Compactifications have been studied and their singularities resolved (see Ash et al. [30], Chai [103], Van der Geer in Bruinier et al. [82] and Namikawa [472]). Gottschling [238] found the explicit list of (28) inequalities defining the fundamental domain $\mathcal{D}_2^S$ for $Sp(2,\mathbb{Z})$. As far as I know, no one has written down the explicit inequalities for $\mathcal{D}_n^S$, when n is larger than 2. Other references for related facts are Christian [108], Freitag [185], Maass [426], Séminaire H. Cartan [547], and Siegel [564, 565]. Van der Geer shows that there is a canonical 1–1 correspondence between the set of isomorphism classes of principally polarized abelian varieties of dimension n and the orbit space $Sp(n,\mathbb{Z})\backslash\mathcal{H}_n$ (see [82, p. 202]).

The symplectic **volume of the fundamental domain** $Sp(n,\mathbb{Z})\backslash\mathcal{H}_n$ was computed by Siegel [565, Vol. II, p. 279] to be:

$$2\prod_{k=1}^{n}\Lambda(k), \quad \text{if } \Lambda(s) = \pi^{-s}\Gamma(s)\zeta(2s).$$

Setting $V_n = \mathrm{Vol}(\mathcal{D}_n^S)$, it follows that

$$V_1 = \frac{\pi}{3}, \quad V_2 = \frac{\pi^3}{270}, \quad V_3 = \frac{\pi^6}{127575}, \quad V_4 = \frac{\pi^{10}}{200930625}.$$

Klingen [351] generalized this result to the **Hilbert–Siegel modular group** which is defined for any totally real algebraic number field K by:

$$Sp(n,\mathfrak{O}_K) = \left\{\gamma \in \mathfrak{O}_K^{n\times n} \mid {}^t\gamma J\gamma = J\right\}, \quad J = \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix}.$$

It is possible to connect the volume of the fundamental domain for $Sp(n,\mathbb{Z})$ with the Euler characteristic of the fundamental domain via the Gauss–Bonnet theorem (see Siegel [565, Vol. II, p. 277, 331], Harder [261] and Klingen [351]). Harder's result is very general. However, the symmetric spaces involved do not include those without complex structure like $\mathcal{SP}_n$ for $n > 2$ or the quaternionic upper half plane.

Mathematicians have studied much more general arithmetic groups Γ and their fundamental domains, including that for $GL(n)$ over a simple associative algebra. Some references are: Borel [65], Hel Braun [75, 76], L. Cohn [120], Feit [176], Krieg [370], Ramanathan [497, 498], Resnikoff and Tai [509], Siegel [565, Vol. III, pp. 143–153; Vol. II, pp. 390–405], Weyl [666, Vol. IV, pp. 232–264], and Weil [658]. Margulis [433] has characterized arithmetic subgroups of connected noncompact Lie groups G of the sort we consider.

Before we leave the subject of fundamental domains, let us give an example of a discrete subgroup Γ of isometries of G/K such that $\Gamma\backslash G/K$ is compact. This example comes from notes of D. Sullivan. Take

$$G = \left\{g \in GL(n+1,\mathbb{R}) \;\middle|\; {}^t g\varphi g = \varphi\right\},$$

where

$$\varphi = \begin{pmatrix} 1 & & 0 & \\ & \ddots & & \\ 0 & & 1 & \\ & & & -\sqrt{2} \end{pmatrix}.$$

Let $\Gamma = G \cap GL(n+1, \mathfrak{O}_L)$ for $L = \mathbb{Q}\left(\sqrt{2}\right)$. Now Γ is a discrete group of isometries of G/K. We can identify G with the Lorentz-type group $O(n, 1)$ and G/K with one sheet of the hyperboloid consisting of the set of points $x \in \mathbb{R}^{n+1}$ such that $\varphi(x) = -1$.

It can be shown that the quotient $\Gamma\backslash G/K$ is compact. We sketch the proof. Note that when $L = \mathbb{Q}\left(\sqrt{2}\right)$, $\mathfrak{O}_L = \mathbb{Z}\left[\sqrt{2}\right]$ which is not discrete in $\mathbb{R}$. Thus we must work harder than usual to show that Γ is actually a discrete subgroup of G. Note the conjugations map:

$$\begin{aligned} \mathfrak{O}_L &\to \mathbb{R}^2, \\ m + n\sqrt{2} &\mapsto \left(m + n\sqrt{2}, m - n\sqrt{2}\right) \quad \text{for } m, n \in \mathbb{Z}. \end{aligned}$$

This induces a mapping which sends Γ discretely into $GL(n+1,\mathbb{R})^2$. Now $\gamma \in \Gamma$ leaves φ invariant and thus γ' leaves φ' invariant where γ' denotes the matrix formed by conjugating all entries of γ; i.e., sending $m + n\sqrt{2}$ to $m - n\sqrt{2}$. Now

$$\varphi' = \begin{pmatrix} 1 & & 0 & \\ & \ddots & & \\ 0 & & 1 & \\ & & & \sqrt{2} \end{pmatrix}$$

and thus the matrices leaving φ' invariant form a compact group. It follows that Γ is a discrete subgroup of G. Why?

To see that $K\backslash G/\Gamma$ is compact, one need to use some version of the Hermite–Mahler compactness theorem in Exercise 1.4.14, Section 1.4.2.

A similar example is $G = \{g \in GL(4,\mathbb{R}) \mid {}^t g\chi g = \chi\}$, where

$$\chi = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -7 \end{pmatrix}.$$

Let $\Gamma = G \cap GL(4,\mathbb{Z})$. It is clear that Γ is discrete. To see that the fundamental domain $K\backslash G/\Gamma$ is compact, identify $K\backslash G/\Gamma$ with a subset S of $\mathcal{P}_4/GL(4,\mathbb{Z})$ by mapping Kg to $I[g]$, as usual. Now to see S has compact closure, we need to only show that $Y \in S$ implies $|Y|$ bounded above and $m_Y = \min\left\{Y[a] \mid a \in \mathbb{Z}^4 - 0\right\}$ bounded below. The fact that m_Y is bounded from below comes from the fact that $\chi[x] = 0$ has no solution $x \in \mathbb{Z}^4 - 0$. For the neighborhood U of 0 defined by

$$U = \left\{x \in \mathbb{R}^4 \mid \chi[x] \le 1/2\right\}$$

contains no point of $g(\mathbb{Z}^4 - 0)$, for $g \in G$.

To see that $\chi[x] \neq 0$ for $x \in \mathbb{Z}^4 - 0$, one looks at

$$x_1^2 + x_2^2 + x_3^2 - 7t^2 \equiv 0 \pmod 8 .$$

Since any integer can be written as a sum of 4 squares, the quadratic form

$$\varphi = x_1^2 + \cdots + x_n^2 - 7t^2$$

does vanish on $\mathbb{Z}^n - 0$ when n is larger than 3. That's why we took the form φ for general n.

There are other references for examples in which $K\backslash G/\Gamma$ is compact; e.g., Borel [65, p. 57]. Borel notes that if G is the orthogonal group of a quadratic form over $\mathbb{Q}$ in n variables and Γ is a group of units of a lattice in $\mathbb{Q}^n$, then $G_{\mathbb{R}}/\Gamma$ is compact iff the form F does not represent 0 over $\mathbb{Q}$. Here Borel uses another sort of compactness criterion. See also Borel [64], Mostow [466], and Mostow and Tamagawa [467]. The last reference proves that if G is a semisimple algebraic matrix group defined over the field $\mathbb{Q}$ and having no unipotent or parabolic elements other than the identity, then $G/G_{\mathbb{Z}}$ is compact. Elstrodt et al. [168] give many examples of $\Gamma \subset G = SL(2,\mathbb{C})$.

2.2.2 Automorphic Forms

Having considered the analogue of Section 1.4 for the types of arithmetic groups in Examples 2.2.1–2.2.4 above, we next begin the study of automorphic forms for these arithmetic groups. The theory of holomorphic forms has received the most attention.

This restricts us here to consideration of automorphic forms for the Hilbert modular group $SL(2,\mathfrak{O}_K)$, for a totally real field K, or the Siegel modular group $Sp(n,\mathbb{Z})$, or the Hilbert–Siegel modular group. It is also possible to discuss non-holomorphic automorphic forms on these symmetric spaces. Such forms satisfy some sort of differential equation involving the invariant differential operators on the symmetric space. Harish-Chandra made a very general definition of automorphic form on a Lie group (see Borel's article in Borel and Mostow [68, p. 199]). Let us begin with a brief sketch of the holomorphic theory which models itself on that from Sections 3.4 and 3.6 of Volume I.

Example 2.2.5 (Holomorphic Hilbert Modular Forms).

Some references for holomorphic Hilbert modular forms are Blumenthal [52], Bruinier's article in [82], Freitag [186], Gundlach [253], Herrmann [292], Hirzebruch [296, 298], Maass [415, 416], Resnikoff [505], Shimura [554], Siegel [563], Marie-France Vignéras [630–632], Van der Geer and Zagier [206], and Zagier [674, 675].

Suppose that K is a totally real algebraic number field of degree m with ring of integers $\mathfrak{O}_K$. We say that a function $f:\mathcal{H}^m\rightarrow\mathbb{C}$ is an (entire) **Hilbert modular form of weight *k*** belonging to $\Gamma=SL(2,\mathfrak{O}_K)$ if f has the following two properties, **assuming that *m* is greater than one**:

(1) $f(z)$ is holomorphic on $\mathcal{H}^m$;
(2) $f(\gamma z)=\mathrm{N}(cz+d)^k f(z)$, for all

$$\gamma=\begin{pmatrix} a & b\\ c & d\end{pmatrix}\in SL(2,\mathfrak{O}_K),\ z\in\mathcal{H}^m.$$

When f satisfies (1) and (2) we say that f is in $\mathcal{M}(SL(2,\mathfrak{O}_K),k)$. The notation in (2) means that for $z=(z^{(1)},\dots,z^{(m)})$, we have

$$\gamma(z)=\left(\gamma^{(1)}z^{(1)},\dots,\gamma^{(m)}z^{(m)}\right),$$

where $\gamma^{(j)}$ denotes the matrix obtained from γ by conjugating each entry of γ by the jth conjugation of the number field K. Here $\gamma^{(j)}$ acts on $z^{(j)}$ by fractional linear transformation. And the **norm** Nz is defined by:

$$\mathrm{N}\,(z)=\prod_{j=1}^{m}z^{(j)},\quad\text{for }\ z\in\mathcal{H}^m. \tag{2.53}$$

One can also look at forms with different weights for each conjugate. This is considered in some of the references listed above but we will not go there.

In analogy to our definition of the norm of $z \in \mathcal{H}^m$, we also define the **trace** as follows:

$$\operatorname{Tr}(z) = \sum_{j=1}^{m} z^{(j)}, \quad \text{for} \quad z \in \mathcal{H}^m. \tag{2.54}$$

Note that if $z \in K$, instead of $\mathcal{H}^m$, the norm and trace are the usual ones for the field extension $K/\mathbb{Q}$.

Of course, one might expect that, as in the case that $m = [K : \mathbb{Q}] = 1$, we should also require a Hilbert modular form to satisfy certain growth conditions at the cusps of the fundamental domain. For the cusp at infinity, one would just require that $f(z)$ be holomorphic at infinity. It was proved by Götzky [239] that this growth condition is unnecessary when $K = \mathbb{Q}(\sqrt{5})$. Gundlach generalized Götzky's result to any totally real number field (see Siegel [563]).

Siegel [563] is a good reference for the basic facts about Hilbert modular forms. For example, Siegel [563, p. 215] shows that a Hilbert modular form of weight $k < 0$ is identically zero, while a Hilbert modular form of weight 0 must be a constant. The argument is analogous to that of Hecke for $SL(2, \mathbb{Z})$ given in Volume I.

Exercise 2.2.5. Show that mk must be an even integer if $\mathcal{M}(SL(2, \mathfrak{O}_K), k)$ is nonzero.

Now to discuss Fourier expansions of Hilbert modular forms, we need to recall the concept of different $\mathfrak{d}_K$ of a number field K (see Section 1.4 of Volume I). The inverse different is the dual lattice to the lattice $\mathfrak{O}_K$ of the number field (cf. pages 80–81 of Volume I). With the trace defined by (2.54), the duality is with respect to the form $\operatorname{Tr}(\alpha\beta)$, and $\mathfrak{d}^{-1} = \{\beta \in K \mid \operatorname{Tr}(\alpha\beta) \in \mathbb{Z} \text{ for all } \alpha \in \mathfrak{O}_K\}$. Fourier expansions are sums running over this dual lattice. Suppose now that f is in $\mathcal{M}\,(SL(2, \mathfrak{O}_K), k)$. Then $f(z)$ is periodic under translations from elements of the lattice $\mathfrak{O}_K$, since matrices

$$\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \text{ are in } SL(2, \mathfrak{O}_K) \text{ when } a \in \mathfrak{O}_K.$$

It follows that $f(z)$ has a **Fourier expansion at the infinite cusp** of the form:

$$f(z) = c(0) + \sum_{0 \ll b \in \mathfrak{d}_K^{-1}} c(b) \, \exp\{2\pi i \operatorname{Tr}(zb)\},$$

where, for $z \in \mathcal{H}^m$ and b a totally positive element of K, we define zb to be the element of $\mathcal{H}^m$ with jth coordinate $z^{(j)} b^{(j)}$. The sum is over b in the inverse different and such that $0 \ll b$, which means that b is **totally positive**; i.e., all conjugates $b^{(j)}$ are positive for $j = 1, \ldots, m$.

Fourier expansions at other cusps x_j can be described using the matrices A_j such that $A_j \infty = x_j$. One must also make use of the formula for the stabilizer of a cusp in Lemma 2.2.1.

Exercise 2.2.6. Show how to obtain the Fourier expansion at infinity of a Hilbert modular form $f(z)$, making use of the Cauchy–Riemann equations to show that the coefficients have the form $c(b)\ \exp\{-2\pi\ \mathrm{Tr}(yb)\}$, if $z = x + iy \in \mathcal{H}^m$.

One example of a Hilbert modular form is the **Eisenstein series** corresponding to an integral ideal $\mathfrak{a}$ in K defined by:

$$E_k(\mathfrak{a}, z) = \sum_{\substack{c,d/U_K \\ (c,d)=\mathfrak{a}}} N(cz+d)^{-k}, \quad k > 2.$$

The sum is over a complete system of representatives for pairs c, d which generate the ideal $\mathfrak{a} = (c, d)$ under the equivalence relation:

$$(c, d) \sim (uc, ud) \text{ for a unit } u \in U_K, \text{ with } Nu = +1.$$

The norm of $cz + d$ is defined by (2.53).

In fact, the Eisenstein series E_k will vanish identically if k is odd and K has a unit of norm -1. It is possible to use an integral test to obtain the convergence of E_k for $k > 2$ (see Siegel [563]). Moreover, Siegel [563, p. 292] proves the vanishing of the lead coefficient or constant term of the Fourier expansion of $E_k(\mathfrak{a}, z)$ with respect to the cusp corresponding to an ideal $\mathfrak{b}$ if $\mathfrak{b}$ is not in the same ideal class as $\mathfrak{a}$. This is to be expected when we recall that Proposition 2.2.1 gave the cusp-ideal class correspondence. Thus one demonstrates the linear independence of the Eisenstein series corresponding to ideals $\mathfrak{a}_1, \ldots, \mathfrak{a}_h$ representing the ideal classes in the ideal class group I_K.

Define a **cusp form** to be a Hilbert modular form f such that $f(z)$ approaches zero as z approaches any cusp of the fundamental domain. Let $\mathcal{S}\,(SL(2, \mathfrak{O}_K), k)$ be the **vector space of cusp forms of weight *k*** for $SL(2, \mathfrak{O}_K)$. Suppose that the ideals $\mathfrak{a}_1, \ldots, \mathfrak{a}_h$ represent the ideal classes in the ideal class group of K. Then we have the direct sum decomposition:

$$\mathcal{M}(SL(2, \mathfrak{O}_K), k) = \left(\sum_{i=1}^{h} {}^{\oplus}\ \mathbb{C}E_k(\mathfrak{a}_i, z) \right) \oplus \mathcal{S}(SL(2, \mathfrak{O}_K), k)$$

(see Siegel [563, p. 294]).

It is also possible to define Poincaré series as in Vol. I, Section 3.4.7. Look at the cusp p/s corresponding to the ideal $\mathfrak{a}$ via Proposition 2.2.1 and let $A \in SL(2, K)$ have the property that $A(\infty) = p/s$. Define a **Poincaré series** by:

$$f_k(\mathfrak{a}, \lambda, z) = \sum_{(c,d)=\mathfrak{a}} N(cz+d)^{-k}\ \exp\left\{2\pi i\, \mathrm{Tr}\,(\lambda \gamma z)\right\}.$$

Here λ is a totally positive element of the ideal $\mathfrak{a}^2\mathfrak{d}_K^{-1}$, where $\mathfrak{d}_K$ is the different of K. The sum is over pairs of generators c, d of the ideal $\mathfrak{a}$ such that

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = A^{-1}\tau, \quad \text{for some } \tau \in SL(2, \mathfrak{O}_K), \quad \text{with} \quad A(\infty) = p/q.$$

See Siegel [563, p. 230]). It can be shown that Poincaré series are cusp forms of weight k. Maass has proved that the Poincaré series and Eisenstein series generate the space $\mathcal{M}\,(SL(2, \mathfrak{O}_K), k)$, for $k \geq 2$. The Poincaré series can vanish identically, but not for large enough weights. See Siegel [563] for more information on the subject, also Bruinier's article in [82] and Freitag [186].

When $m = [K : \mathbb{Q}]$ is larger than one, a function $f(z)$ which is meromorphic on $\mathcal{H}^m$ is called a **Hilbert modular function** if $f(\gamma z) = f(z)$ for all $\gamma \in SL(2, \mathfrak{O}_K)$ and $z \in \mathcal{H}^m$. In the case that $m = 1$ and $K = \mathbb{Q}$, we would add a further requirement that $f(z)$ have at most a pole at the infinite cusp. This need not be assumed when $m \geq 2$. A Hilbert modular function which is holomorphic in $\mathcal{H}^m$ is automatically holomorphic at the cusps and thus must be a modular form of weight zero and therefore a constant. For there are no isolated singularities in several complex variables. Thus, when $m \geq 2$ there does not exist an analogue of the elliptic modular invariant $J(z)$ from Section 3.4.3 of Volume I. This fact was first noted by Götzky [239]. There are errors in Hecke's early papers due to the lack of knowledge of Götzky's result. These early Hecke papers seek to solve Hilbert's 12th problem which asks for an explicit construction of class fields (extension fields having abelian Galois groups) over arbitrary algebraic number fields using automorphic forms (see Hecke [268, p. 942]). Siegel [563] gives a proof that the Hilbert modular functions form an algebraic function field of n variables.

Herrmann [292] investigated the theory of Hecke operators for Hilbert modular forms. This theory has been extended to Picard modular groups by Styer [584]. Shimura [554] describes a very general theory of Hecke operators. Adelic versions of Hecke theory also exist for quite general groups (see Jacquet and Langlands [324], Gelbart [208], and Weil [660]). We leave it to the beleaguered reader to define the Petersson inner product of two cuspidal Hilbert modular forms of weight k and to obtain an analogue of Theorem 3.6.3 in Volume I for Hilbert modular forms.

The correspondence between Hilbert modular forms and Dirichlet series has been much studied (see the preceding references and Stark [575]). The situation is much like that investigated by Weil [662, Vol. III, pp. 165–172] for congruence subgroups of $SL(2, \mathbb{Z})$. One must have functional equations for L-functions that have been "twisted" by Hecke grossencharacters for K, in order to know that the corresponding function $f(z), z \in \mathcal{H}^m$, is a Hilbert modular form for $SL(2, \mathfrak{O}_K)$.

Let us just consider the simplest example of Hecke theory over number fields. Define the **theta function** corresponding to an ideal $\mathfrak{a}$ of a totally real number field K by:

$$\theta(\mathfrak{a}, z) = \sum_{a \in \mathfrak{a}} \exp\left\{-\pi \operatorname{Tr}\left(za^2\right)\right\}.$$

By slightly altering the proof of Theorem 1.4.2 in Volume I, we can view the ideal class zeta functions (which occur as partial sums of the Dedekind zeta function) of a totally real number field as Mellin transforms of this theta function (see Hecke [268, p. 227]). Similarly Hecke obtained the analytic continuation of his L-functions by showing them to be Mellin transforms of theta functions. Generalizations of this theta function are considered by Eichler [159], Kloosterman [356], and Schoeneberg [537]. These authors also look at the effect of Hecke operators on such theta functions.

If Γ is a subgroup of $SL(2, \mathfrak{O}_K)$ without elliptic fixed points, either the Selberg trace formula or the Hirzebruch–Riemann–Roch theorem can be used to compute the dimension of the space of Hilbert modular forms (see Ash et al. [30], Hirzebruch [296–298], Langlands [389], and Shimizu [553], as well as Section 2.2.3). If, for example, K is real quadratic, Γ of index a in $SL(2, \mathfrak{O}_K)/\pm I$, and Γ acts freely on $\mathcal{H}^2$, then, for $k \geq 3$, we have:

$$\dim \mathcal{S}(\Gamma, k) = \frac{k(k-2)}{2} \zeta_K(-1) a + \chi, \quad \chi = 1 + \dim \mathcal{S}(\Gamma, 2).$$

We will say more about the use of trace formulas to compute such dimensions at the end of this chapter. See also Freitag [186].

Example 2.2.6 (Holomorphic Siegel Modular Forms).

Some references for this section are Andrianov [9–14], Baily [32], Böcherer [54, 55], Hel Braun [74], Christian [108, 110], Eichler [157–159], Feit [175, 176], Freitag [185], Garrett [202], Hoobler and Resnikoff [304], Igusa [315, 316], Kaori Imai (Ota) [317], Kalinin [338], Karel [339], Klingen [350, 353–355], Maass [414, 419, 426], Morita [463], Resnikoff [507, 508], Shimura [556], Siegel [564, 565], Tsao [616], Van der Geer's article in Bruinier et al. [82], Weissauer [663, 664], and Yamazaki [673].

We will say that a function f on Siegel's upper half space $\mathcal{H}_n$, $n > 1$, is a holomorphic **Siegel modular form of weight k** and write $f \in \mathcal{M}(Sp(n, \mathbb{Z}), k)$ if it satisfies the following two conditions:

(1) f is holomorphic on $\mathcal{H}_n$;
(2) $f(\gamma Z) = |CZ + D|^k f(Z)$, for all

$$\gamma = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \text{ in } Sp(n, \mathbb{Z}), \quad Z \in \mathcal{H}_n.$$

Some people say that $f \in \mathcal{M}(Sp(n, \mathbb{Z}), k)$ is a Siegel modular form of **genus** n, while others say f is a Siegel modular form of **degree** n. I have decided not to take sides, probably because of a memory of the time that I looked up the German word for genus in my dictionary.

There are generalizations to vector valued Siegel modular forms transforming according to a finite dimensional representation of $GL(n,\mathbb{R})$. See van der Geer's article in Bruinier et al. [82, p. 187].

One might expect to add a third condition that f must be bounded in the region $\mathrm{Im}Z = Y \geq Y_0 > 0$, where $Y \geq Y_0$ means that $Y - Y_0$ lies in the closure of $\mathcal{P}_n$. Koecher shows that this third condition is unnecessary when n is bigger than one (see Proposition 2.2.2 below and Maass [426, Section 13]).

Thus we have another analogue of the space of ordinary modular forms which was studied in Section 3.4 of Volume I. It can be shown that when k is larger than one, $\mathcal{M}\,(Sp(n,\mathbb{Z}),k) \neq 0$ implies $k \in \mathbb{Z}$ and $nk \in 2\,\mathbb{Z}$. It can also be proved that Siegel modular forms of negative weight must vanish while those of weight zero must be constant.

Exercise 2.2.7. Prove the last statements.
Hint. See Klingen [355].

Next we want to consider Fourier expansions of Siegel modular forms. First note that if X lies in the lattice of integral $n \times n$ symmetric matrices, then the dual lattice with respect to the form $\mathrm{Tr}(TX)$ consists of the $n \times n$ semi-integral symmetric matrices $T = (t_{ij})$. Here "semi-integral" means that $t_{jj} \in \mathbb{Z}$ and $t_{ij} \in \frac{1}{2}\mathbb{Z}$, when $i \neq j$.

If f is a Siegel modular form of weight k, then $f(X + iY)$ has period one in each entry of the symmetric matrix X. This implies that f has a **Fourier expansion**:

$$f(Z) = \sum_{\substack{0\,\leq T =\, {}^tT \\ T \text{ semi-integral}}} a(T)\exp\left\{2\pi i\,\mathrm{Tr}(TZ)\right\}. \tag{2.55}$$

Here $T \geq 0$ means that $T[x] \geq 0$ for all $x \in \mathbb{R}^n$.

To see that the Fourier coefficient has the form $a(T)\exp\{2\pi i\,\mathrm{Tr}(TZ)\}$, one must use the fact $f(Z)$ satisfies the Cauchy–Riemann equations in each variable. The sum is over semi-integral matrices because that is the dual lattice to $\mathbb{Z}^{n\times n}$ with respect to the form $\mathrm{Tr}(TX)$.

There is also an expansion known as the Fourier–Jacobi expansion of Piatetski-Shapiro which has proved to be useful. See van der Geer's article in Bruinier et al. [82, p. 196] and Freitag [184, p. 101].

Exercise 2.2.8. Prove what we just said about (2.55)
Hint. The sum is over nonnegative matrices T by Lemma 2.2.2 below.

Lemma 2.2.2. *Let $f \in \mathcal{M}(Sp(n,\mathbb{Z}),k)$ have Fourier coefficients $a(T)$ as in (2.55) above. Then $f(Z)$ is bounded in every domain $Y \geq Y_0 > 0$ (for fixed Y_0) if and only if the Fourier coefficient $a(T) \neq 0$ implies that $T \geq 0$.*

Proof. See Maass [426, pp. 183–184].
$\Rightarrow$ Suppose that $|f(Z)| \leq C(Y_0)$ in $Y \geq Y_0 > 0$. Then

$$a(T)\exp\{-2\pi\,\mathrm{Tr}(TY)\} = \int\limits_{{}^tX=X\in[0,1]^{n\times n}} f(X+iY)\ \exp\{-2\pi i\,\mathrm{Tr}(TX)\}\,dX.$$

In order to show that $a(T)\neq 0$ implies that $T\geq 0$, use the above equality to obtain the bound:

$$|a(T)|\leq C(Y_0)\ \exp\{2\pi\ \mathrm{Tr}(TY)\}.$$

If T is not ≥ 0, then one can show that the right-hand side of this inequality can be made arbitrarily small. Thus $a(T)$ must vanish when T is not ≥ 0.
$\Leftarrow$ Suppose the Fourier expansion of $f(Z)$ has the form

$$f(Z) = \sum_{\substack{{}^tT=T\geq 0\\ \text{semi-integral}}} a(T)\ \exp\{2\pi i\,\mathrm{Tr}(TZ)\}.$$

The convergence of the Fourier series implies that if $Y\geq Y_0\geq aI>0$

$$|f(Z)|\leq C\left(\frac{1}{2}Y_0\right)\sum_{T\geq 0}\exp\{-\pi a\,\mathrm{Tr}(T)\}. \tag{2.56}$$

The series on the right in (2.56) is easily seen to converge. ■

Exercise 2.2.9. Fill in the details in the proof of Lemma 2.2.2 above.

Proposition 2.2.2 (Koecher). *If $f\in\mathcal{M}(Sp(n,\mathbb{Z}),k)$, then f is bounded in any domain $Y\geq Y_0>0$. In particular, f is bounded in the fundamental domain $\mathcal{D}_n^S$ for $Sp(n,\mathbb{Z})$, which is defined in formula (2.52).*

Proof. See Maass [426, pp. 185–187]. For $n=1$, the result is part of the definition of modular form. When n is larger than one, consider the Fourier expansion (2.55) of f. Note that $a\,(T[U])=a(T)$ for all $U\in GL(n,\mathbb{Z})$. The main step in the proof of Proposition 2.2.2 is the proof of the claim that when T is not ≥ 0, then the number

$$c(T,-m)=\#\{\,T[U]\mid\ U\in SL(n,\mathbb{Z}),\ \mathrm{Tr}(T[U])=-m\}$$

is greater than or equal to one for infinitely many $m\geq 1$. Therefore $a(T)$ must vanish in this case.

To prove this claim, observe that, upon setting $U=I+b\,{}^td$, with $b,d\in\mathbb{Z}^n$, such that ${}^tbd=0$, we have

$$|U|=1\ \text{ and }\ \mathrm{Tr}\,(T\,[U])=\mathrm{Tr}\,(T)+2\mathrm{Tr}\left(T\,b\,{}^td\right)+T[b]\ \left({}^tdd\right).$$

If T were not ≥ 0, then we could choose $b \in \mathbb{Z}^n$ so that $T[b] < 0$ and ${}^t bd = 0$, for any $d \in \mathbb{Z}^n$. And then $c(T, -m)$ must indeed be ≥ 1 for infinitely many m. The proof of Proposition 2.2.2 is completed using Lemma 2.2.2. ■

An example of a Siegel modular form of even weight k in $\mathcal{M}(Sp(n, \mathbb{Z}), k)$ is given by the **Eisenstein series**:

$$E_k(Z) = \sum_{C,D} |CZ + D|^{-k}, \text{ for even } k > n + 1,$$

where the sum is over coprime symmetric pairs C, D of matrices in $\mathbb{Z}^{n \times n}$ modulo the equivalence relation

$$(C\ D) \sim (UC \ \ UD), \quad \text{for } U \in GL(n, \mathbb{Z}).$$

Hel Braun [74] proved the convergence of the Eisenstein series in the stated region. Hel Braun was a student of Siegel. She was one of the few women mathematics professors in Germany during her life. I had the opportunity to meet her at an Oberwohlfach meeting. She had much to say about Siegel. It would be interesting to know more of her life. She only discusses the beginnings in her book [77].

See Freitag [185, pp. 66–67] for a convergence proof using a sort of integral test. Freitag [185, p. 67] and Maass [426, Section 14] consider more general Eisenstein series involving modular forms for $Sp(r, \mathbb{Z})$, $r < n$, which were introduced by Klingen [353]. See Van der Geer's article in Bruinier et al. [82, p. 194] for a theorem of Hel Braun on the Klingen Eisenstein series and their behavior under the Siegel Φ-operator defined below in equation (2.57). Another reference is Klingen [355, p. 68], who notes [355, p. 63] that "There is no reasonable way to introduce Eisenstein series for odd weights."

Siegel [565, Vol. II, p. 133] gives the Fourier expansion of the Eisenstein series E_k (k even and $> n + 1$) for $Sp(n, \mathbb{Z})$ (see also Baily [32]). Let

$$T[U] = \begin{pmatrix} T_1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{for} \quad U \in GL(n, \mathbb{Z}),\ T_1 \in \mathbb{Z}^{r \times r},$$

and set $D(T) = |T_1|$. The term corresponding to T in the Fourier expansion of $E_k(Z)$ is:

$$(-1)^{rk/2} 2^{r(k-(r-1)/2)} \prod_{j=0}^{r-1} \frac{\pi^{k-j/2}}{\Gamma(k-j/2)} D(T)^{k-(r+1)/2} \sum_{{}^t R = R \in (\mathbb{Q}/\mathbb{Z})^{r \times r}} e^{2\pi i \mathrm{Tr}(T_1 R)}\ \nu(R)^{-k}.$$

Here $\nu(R)$ is the product of the reduced denominators of the elementary divisors of R. An analogous result for Eisenstein series for $GL(n, \mathbb{Z})$ was discussed in Exercise 1.5.36 of Section 1.5.3 (see also Terras [596]). Siegel used his main theorem on quadratic forms to deduce the rationality of the coefficients of the

Eisenstein series E_k for $Sp(n,\mathbb{Z})$. Baily [32, Ch. 12] gives another derivation. Similar and more general results of this type are obtained by Karel [339] and Tsao [616]. Kaufhold [343] finds even more explicit results for Fourier coefficients of Eisenstein series for $Sp(2,\mathbb{Z})$. When the Eisenstein series have rational coefficients and the Eisenstein series generate the full field of automorphic or Siegel modular functions (i.e., meromorphic functions satisfying condition (2) in the definition of Siegel modular form and having weight $k = 0$) for $Sp(n,\mathbb{Z})$, then the algebraic variety which is the Satake compactification of the fundamental domain $Sp(n,\mathbb{Z})\backslash\mathcal{H}_n$ is a variety defined over $\mathbb{Q}$ (see Baily [32, p. 238]). More information on Eisenstein series can be found in the references mentioned at the beginning of this discussion of holomorphic Siegel modular forms. Some of the references consider non-holomorphic Eisenstein series as well.

An example of a modular form for a congruence subgroup of $Sp(n,\mathbb{Z})$ is the **theta function** defined for $Q \in \mathcal{P}_n \cap \mathbb{Z}^{n\times n}$, $Z \in \mathcal{H}_n$, by:

$$\theta(Z) = \sum_{A\in\mathbb{Z}^{k\times k}} \exp\{\pi i \mathrm{Tr}\,(ZQ\,[A])\}\,.$$

The theta function in formula (1.172) of Section 1.4.1. is simply a restriction of this symplectic theta function to $Z = iY$, $Y \in \mathcal{P}_n$. Eichler [156] shows that in the special case $n = 1$ the theta function is a modular form for a congruence subgroup of $SL(2,\mathbb{Z})$ by considering it as a special value of a theta function on $\mathcal{H}_k$. Andrianov and Maloletkin [15] generalize this result to any n—evaluating the 8th root of unity involved when k is even. Stark [576] extends these results further by evaluating the 8th root of unity in a case that can be reached by theorems on matrix primes in progressions (see also Styer [584]). Other references for theta functions are Freitag [185] and Igusa [316].

Theta functions and zeta functions for indefinite quadratic forms have been studied by many authors (see Koecher [360], Maass [418, 423, 425], Siegel [565, Vol. I, pp. 410–443; Vol. II, pp. 41–96, 421–466; Vol. III, pp. 85–91, 105–142, 154–177], Andrianov and Maloletkin [16], and Friedberg [190]).

It is possible to obtain examples of modular forms of various types by integrating against theta functions of indefinite quadratic forms. For example, one can obtain holomorphic Hilbert modular forms for a real quadratic field by integrating ordinary holomorphic modular forms for $SL(2,\mathbb{Z})$ multiplied by an appropriate theta function. One can similarly lift Maass wave forms for $SL(2,\mathbb{Z})$. And one can replace real by imaginary quadratic fields, etc. References for such constructions, often referred to as "base change," include Friedberg [190], Goldfeld [230], Kudla [378, 379], Stark [579], Tsuyumine [618], Marie-France Vignéras [634], and Waldspurger [641].

Theta functions have many other applications. There are, in fact, entire books devoted to them (e.g., Igusa [316] and Mumford [471]). Siegel's main theorem on quadratic forms can be viewed as an equality between linear combinations of theta functions and generalized Eisenstein series (see Siegel [565, Vol. I, pp. 326–405, 410–443, 469–548]). Theta functions on the Siegel upper half space can be used to obtain an expression for the generalization of elliptic integrals known as abelian integrals (see Siegel [564]). This leads to many applications in physics. We have already mentioned the work of Sofya Kovalevskaya [368] and Dubrovin et al. [143]. Other references for applications to the Korteweg–deVries equation are McKean and Trubowitz [440] and Novikov [474]. Siegel's work on quadratic forms has been connected with quantum mechanics and representation theory via the Segal–Shale–Weil representation (see Cartier's talk in Borel and Mostow [68, pp. 361–368], Gérard Lion and Michèle Vergne [406], Shale [551], Weil [662, Vol. III, pp. 1–157], and Wallach [652]). The role of theta functions in algebraic geometry and purely algebraic constructions of these functions are discussed in Mumford [471] and Shafarevitch [549]. Connections with Jordan algebras are pursued by Resnikoff [508]. Connections with knot theory and physics are to be found in Gelca [211].

In order to discuss cusp forms, Siegel defined the **Φ-operator** taking $f \in \mathcal{M}(Sp(n,\mathbb{Z}),k)$ to $f|\Phi \in \mathcal{M}(Sp(n-1,\mathbb{Z}),k)$ by:

$$f|\Phi(W) = \lim_{t\to\infty} f\begin{pmatrix} W & 0 \\ 0 & it \end{pmatrix}, \text{ for } W \in \mathcal{H}_{n-1}. \tag{2.57}$$

A Siegel modular form $f \in \mathcal{M}(Sp(n,\mathbb{Z}),k)$ is said to be a **cusp form** if $f|\Phi$ vanishes identically. Let $\mathcal{S}(Sp(n,\mathbb{Z}),k)$ denote the **space of Siegel cusp forms of weight k**. The Fourier expansion (2.55) of $f \in \mathcal{S}(Sp(n,\mathbb{Z}),k)$ can have no terms corresponding to singular symmetric semi-integral matrices T. Here a singular T is one with determinant equal to zero.

Exercise 2.2.10. Prove the last statement about cusp forms.
Hint. See Klingen [355].

There is an opposite concept to that of cusp form—the concept of **singular form** which is a form in $\mathcal{M}(Sp(n,\mathbb{Z}),k)$ whose Fourier coefficients $a(T)$ in (2.55) vanish unless the T are singular matrices. Certain theta functions give examples. An integer $k \geq 0$ is said to be a **singular weight** for $Sp(n,\mathbb{Z})$ if $k < n/2$. Maass [426] showed singular forms have singular weights. Resnikoff [507] and Freitag [184] proved the converse. See Klingen [355, Chapter 8].

Let us say a bit about dimensions of spaces of Siegel modular forms for $Sp(n,\mathbb{Z})$. Christian showed that for nonvanishing forms to exist, the weights must be integers. See Klingen [355, p. 43]. Moreover the weights must be nonnegative. See Klingen [355, p. 47]. Finally any form of weight 0 is constant. See Klingen [355, p. 49].

It can be proved that there is a positive constant c_n depending only on n such that

$$\dim \mathcal{M}(Sp(n,\mathbb{Z}),k) \leq c_n k^{n(n+1)/2}.$$

See Eicher [157, 158], Freitag [185, p. 52], Klingen [355, p. 51], Maass [426, p. 194], and Siegel [565, Vol. II, pp. 97–137]. The main principle needed to prove such an inequality is that which gave us Theorem 3.5.4 in Volume I. This principle says: The vanishing of sufficiently many terms in the Fourier series of $f \in \mathcal{M}\,(Sp(n,\mathbb{Z}),k)$ implies the vanishing of f itself. Freitag [185, p. 50] shows that $\mathcal{S}\,(Sp(n,\mathbb{Z}),k)$ vanishes in the following situations:

$n=1$	$k<12$
$n=2$	$k<9$
$n=3$	$k<8$
$n=4$	$k<5$

See the very last pages of this volume for more information.

It is possible to use the Selberg trace formula or the Hirzebruch–Riemann–Roch theorem to give formulas for dimensions of spaces of Siegel cusp forms for congruence subgroups of $Sp(n,\mathbb{Z})$ acting without elliptic fixed points. See Arakawa [19], Christian [109], Hirzebruch [297, Appendix], Langlands [389], Morita [463], Petra Ploch [489], and Yamazaki [673]. See also Eie [160] and Hashimoto [264] for the case of $Sp(n,\mathbb{Z})$ for small n. Let us examine one such computation—that of Arakawa [19] using the Selberg trace formula. One begins with the dimension formula of Godement which writes the dimension of the space of cusp forms as an integral of a sum over Γ. The identity and parabolic elements of Γ produce the only nonzero contributions to the trace formula. Work of Shintani [559] is needed to compute special values of zeta functions arising in the calculations. See the last section of this book for a few more details and references.

The spaces $\mathcal{M}\,(Sp(2,\mathbb{Z}),k)$ were completely determined by Igusa [315]. See also Freitag [183].

The **Petersson inner product** for weight k Siegel modular forms f,g for $\Gamma = Sp(n,\mathbb{Z})$, at least one of which is a cusp form, is defined by:

$$\langle f,g\rangle = \int\limits_{\Gamma\backslash\mathcal{H}_n} f(Z)\overline{g(Z)}\det\left(\operatorname{Im}(Z)\right)^k d\mu_n^*,$$

where $d\mu_n^*$ is the invariant volume element on $\mathcal{H}_n$.

Hecke operators for $Sp(n,\mathbb{Z})$ were first systematically investigated by Maass [414]. Let m be a positive integer and J_n as in the definition of $Sp(n,\mathbb{Z})$, define

$$M_n = \left\{ g \in \mathbb{Z}^{2n\times 2n} \;\middle|\; {}^t g J_n g = mJ_n \right\}.$$

Then we will characterize the ***m*th Hecke operator** $T(m)$ by what it does to a form $f \in \mathcal{M}\,(Sp(n,\mathbb{Z}),k)$, namely:

$$T(m)f(Z) = m^{nk-n(n+1)/2} \sum_{\gamma=\begin{pmatrix} * & * \\ C & D\end{pmatrix} \in Sp(n,\mathbb{Z})\backslash M_n} |CZ+D|^{-k} f(\gamma Z).$$

It is possible to show that symplectic Hecke operators have similar properties to those of Hecke operators for $GL(n,\mathbb{Z})$ which were obtained in Theorem 1.5.2 of Section 1.5.2. The Euler products involved are more complicated though. Some other references for these Hecke operators are Andrianov [9–14], Freitag [185], and Shimura [554].

There are many sorts of **Dirichlet series or *L*-functions associated with Siegel modular forms** $f \in \mathcal{M}(Sp(n,\mathbb{Z}),k)$. For simplicity, let us assume that f is a cusp form. Then, given a Maass form v for $GL(n,\mathbb{Z})$ in $\mathcal{A}(GL(n,\mathbb{Z}),\lambda)$ as in Section 1.5.1, we can consider the following Mellin transform:

$$M(f,v) = \int\limits_{Y\in\mathcal{P}_n/GL(n,\mathbb{Z})} f(iY)v(Y)\,d\mu_n(Y).$$

Suppose that $f(Z)$ has the Fourier expansion (2.55) above and that

$$v(Y) = |Y|^s\,u(Y^o), \quad \text{for} \quad Y^o = |Y|^{-1/n}\,Y \in \mathcal{SP}_n, \quad s\in\mathbb{C}.$$

Set

$$u^*(W) = u(W^{-1}).$$

As in the proof of part (5) of Theorem 1.5.2 of Section 1.5.2, we have:

$$\begin{aligned} M(f,v) &= \sum_{T>0}\ \int\limits_{Y\in\mathcal{P}_n/GL(n,\mathbb{Z})} a(T)\exp\{-2\pi\operatorname{Tr}(TY)\}\,v(Y)\,d\mu_n(Y) \\ &= (2\pi)^{-s}\,\Gamma_n(r(u,s)) \sum_{0<T/GL(n,\mathbb{Z})} a(T)\ |T|^{-s}\ u^*(T)\,, \end{aligned}$$

for some $r(u,s)\in\mathbb{C}^n$. Thus it is natural to associate to $f\in\mathcal{S}(Sp(n,\mathbb{Z}),k)$ with Fourier expansion (2.55) having Fourier coefficients $a(T)$, the Dirichlet series:

$$L(f,v) = \sum_{\substack{0<T/GL(n,\mathbb{Z}) \\ T \text{ symmetric, semi-integral}}} a(T)v\left(T^{-1}\right).$$

Maass [426, Section 15] obtains the analytic continuation and functional equation of $L(f,v)$, even when f is not a cusp form. See the proof of part (5) of Theorem 1.5.2 of Section 1.5.2 for a discussion of a similar analytic continuation.

Harmonic analysis on $\mathcal{P}_n/GL(n,\mathbb{Z})$ gives a converse to this Hecke correspondence between f and $L(f,v)$, as Kaori Imai (Ota) [317] shows when $n = 2$. Weissauer [663] obtains a converse result for congruence subgroups of $Sp(n,\mathbb{Z})$, for general n. It would be nice not to have to know that there are Dirichlet series with functional equations for all the v in $\mathcal{A}(GL(n,\mathbb{Z}),\lambda)$. Just how many such v are necessary is a very interesting question. Similar questions exist for Weil's theory of the Hecke correspondence for congruence subgroups of $SL(2,\mathbb{Z})$ (see Section 3.6 of Volume I).

Andrianov [10–14] investigates various sorts of Dirichlet series associated with eigenfunctions of Hecke operators. The language of adelic representation theory leads to the same sort of results. See Piatetski-Shapiro's talk in Borel and Casselman [66, Vol. 1, pp. 185–188]. See also Piatetski-Shapiro [486]. Tamara Veenstra [626] investigates L-functions corresponding to Siegel eigenforms of Hecke operators and shows that the p-factors for almost all primes determine the L-function.

Bounds on Fourier coefficients of Siegel cusp forms are investigated by Kathrin Bringmann [80]. The explicit action of the standard generators of the Hecke algebra on Fourier coefficients of Siegel modular forms of half-integral weight is studied by Lynne Walling [654].

Very general Poincaré series for $\Gamma_n = Sp(n, Z)$ are considered by Klingen [355, Ch. 6]. One example which was introduced by Maass is:

$$g_n^k(Z,t) = \sum_{\Upsilon_n\backslash\Gamma_n} |CZ+D|^{-k} \exp\left(t\,(Az+B)\,(Ca+D)^{-1}\right), \tag{2.58}$$

for $k > 2n$, $kn \equiv 0(\mathrm{mod}\,2)$, $Z \in \mathcal{H}_n$, and half integral positive t. Here Υ_n is the subgroup of matrices of the form $\begin{pmatrix} \pm I_n & * \\ 0 & \pm I_n \end{pmatrix}$. Klingen [355, p. 90] notes the main properties of the Maass Poincaré series; the most important being

(1) The Petersson inner product $\left\langle f, g_n^k(*,t)\right\rangle$ pulls out the tth Fourier coefficient of a Siegel cusp form f of weight k multiplied by a constant and a power of t.
(2) The Maass Poincaré series (2.58) span the space $\mathcal{S}\,(Sp(n,\mathbb{Z}),k)$ for a finite set of values of t.

Example 2.2.7 (Eisenstein Series for $GL(2,\mathfrak{O}_K)$).

References for this subject include: Asai [27], Efrat and Sarnak [154], Elstrodt et al. [168], Fueter [191], Gelbart [208], Grosswald [250], Hecke [268], Hoffstein [303], Hunter [311], Jacquet and Langlands [324], Mennicke [442], Mordell [460], Ramanathan [497, 498], Sarnak [527], Siegel [565, Vol. I, pp. 173–179], Stark [579], Tamagawa [587], Terras [597–600], and Weil [660].

Let K be an algebraic number field of degree m over $\mathbb{Q}$. We use the notation set up earlier during the discussion of the fundamental domain for $GL(n,\mathfrak{O}_K)$. The **Epstein zeta function** for $GL(n,\mathfrak{O}_K)$ is defined for $Y \in \mathcal{P}_2^K$, $\mathfrak{a}$ an ideal of K, and s a complex number with $\mathrm{Re}\, s > 1$, by:

$$Z(\mathfrak{a}, Y, s) = \sum_{0 \neq b \in \mathfrak{a}^2/U_K} N\left(Y\{b\}\right)^{-s}. \tag{2.59}$$

Here the **norm** $N\left(Y\{b\}\right)$ is defined by

$$N(Y\{b\}) = \prod_{j=1}^{r_1+r_2} \left(\overline{{}^t b^{(j)}} Y^{(j)} b^{(j)}\right)^{e_j}, \text{ for } e_j = \begin{cases} 1, j = 1, \ldots, r_1, \\ 2, j = r_1+1, \ldots, r_1+r_2. \end{cases} \tag{2.60}$$

The sum in (2.59) is over a complete system of nonzero column vectors in $\mathfrak{a}^2$ inequivalent under the equivalence relation

$${}^t b = (b_1, b_2) \sim (b_1 u, b_2 u), \text{ for } u \in U_K = \text{the group of units of } \mathfrak{O}_K.$$

In order to prove the convergence of Epstein's zeta function (2.59), one could devise an integral test similar to that used in the case that K is the field of rational numbers (see Corollary 1.4.4 in Section 1.4.4). Related methods are used by Siegel [563, p. 290] and Godement in Borel and Mostow [68, p. 207]. It is also possible to deduce the convergence from bounds on theta functions as in Ramanathan [498, p. 54].

Exercise 2.2.11. Obtain the analytic continuation and functional equation of Epstein's zeta function for $GL(2, \mathfrak{O}_K)$ in (2.59) by imitating Riemann's proof of the analytic continuation of $\zeta(s)$ given in Section 1.4 of Volume I. See also Hecke's proof of the analytic continuation of the Dedekind zeta function in Lang [386, pp. 255–258]. You will find that $Z(\mathfrak{a}, Y, s)$ has a simple pole at $s = 1$ and a functional equation relating it to $Z(\mathfrak{a}', Y^{-1}, 1-s)$, where $\mathfrak{a}' = (\mathfrak{a}\mathfrak{d}_K)^{-1}$, if $\mathfrak{d}_K$ is the different of K.

We have a **simultaneous Iwasawa decomposition** of the vector $Y \in \mathcal{P}_2^K$

$$Y = \begin{pmatrix} v & 0 \\ 0 & w \end{pmatrix} \begin{bmatrix} 1 & q \\ 0 & 1 \end{bmatrix}, \text{ with } v, w \in Y \in \mathcal{P}_1^K, \ q \in \mathbb{R} \otimes_{\mathbb{Q}} K. \tag{2.61}$$

This means that for $1 \leq j \leq r_1 + r_2$,

$$Y^{(j)} = \begin{pmatrix} v^{(j)} & 0 \\ 0 & w^{(j)} \end{pmatrix} \begin{bmatrix} 1 & q^{(j)} \\ 0 & 1 \end{bmatrix}, \text{ with } v^{(j)}, w^{(j)} > 0,$$
$$q^{(j)} \in E_j, \quad E_j = \begin{cases} \mathbb{R}, j = 1, \ldots r_1, \\ \mathbb{C}, j = r_1+1, \ldots, r_1+r_2. \end{cases}$$

Clearly the Epstein zeta function $Z(\mathfrak{a}, Y, s)$ has the **invariance property**:

$$Z(\mathfrak{a}, Y, s) = Z(\mathfrak{a}, Y\{\gamma\}, s) \text{ for all } \gamma \in GL(2, \mathfrak{O}_K), Y \in \mathcal{P}_2^K.$$

It follows that if we view $Z(\mathfrak{a}, Y, s)$ as a function of the q-variable in the Iwasawa decomposition (2.61) of Y, we are looking at a function that is periodic modulo $\mathfrak{O}_K$. Thus we can obtain a Fourier expansion of $Z(\mathfrak{a}, Y, s)$ in the q-variable.

Let us eliminate the dependence on the ideal $\mathfrak{a}$ by defining a new zeta summed over classes C in the ideal class group I_K:

$$Z^*(Y,s) = \sum_{C\in I_k} N\mathfrak{b}^{2s}\, Z(\mathfrak{b},Y,s), \quad \text{for} \quad \mathfrak{b}\in C. \tag{2.62}$$

Here $N\mathfrak{b}$ denotes the norm of the ideal $\mathfrak{b}$, which is $\#(\mathfrak{O}_K/\mathfrak{b})$. Note that it does not matter what ideal $\mathfrak{b}\in C$ is chosen. Because the ideal class group I_K is finite, so is the sum in (2.62) provided that $s\neq 1$.

Set

$$A = 2^{-r_2}\pi^{-m/2}D_K^{1/2}, \quad D_K = \text{ the absolute value of the discriminant of } K,$$

and define **Dedekind's zeta function** by

$$\zeta_K(s) = \sum_{\mathfrak{b}\subset\mathfrak{O}_K} N\mathfrak{b}^{-s}, \quad \text{for} \quad \operatorname{Re} s > 1,$$

where the sum is over ideals $\mathfrak{b}$ of $\mathfrak{O}_K$. As proved in Section 1.4 of Volume I, the functional equation of Dedekind's zeta function is:

$$\Lambda_K(s) = A^s\Gamma\left(\frac{s}{2}\right)^{r_1}\Gamma(s)^{r_2}\zeta_K(s) = \Lambda_K(1-s). \tag{2.63}$$

Motivated by this functional equation, we define:

$$\Lambda^*(Y,s) = A^{2s}\Gamma(s)^{r_1}\,\Gamma(2s)^{r_2}Z^*(Y,s).$$

Theorem 2.2.1 (Fourier Expansion of Epstein's Zeta Function for *K*). *Using the notation of formulas (2.59)–(2.63), we have the Fourier expansion:*

$$\begin{aligned}\Lambda^*(Y,s) = {} & Nv^{-s}\Lambda_K(2s) + Nv^{-\frac{1}{2}}Nw^{\frac{1}{2}-s}\Lambda_K(2s-1)\\ & + \frac{2^{r_1+r_2}D_K^{s-\frac{1}{2}}}{Nv^{\frac{1}{2}}N|Y|^{-\frac{1}{4}+s/2}} \sum_{0\neq u\in\mathfrak{d}_K^{-1}} |Nu|^{s-\frac{1}{2}}\sigma_{1-2s}(u\mathfrak{d}_K)\, e^{2\pi i \operatorname{Tr}(qu)}\\ & \times \prod_{j=1}^{r_1+r_2} K_{e_j(s-\frac{1}{2})}\left(2\pi e_j\sqrt{\left(\frac{wu^2}{v}\right)^{(j)}}\right),\end{aligned}$$

where $K_s(y)$ is the ordinary K-Bessel function and for any ideal $\mathfrak{b}\subset\mathfrak{O}_K$ the function $\sigma_s(\mathfrak{b})$ is the ***divisor function****:*

$$\sigma_s(\mathfrak{b}) = \sum_{\mathfrak{c}|\mathfrak{b}} N\mathfrak{c}^s.$$

Here $\mathfrak{c}|\mathfrak{b}$ is equivalent to $\mathfrak{c}\supset\mathfrak{b}$.

Proof. See Terras [598, 599]. The idea is to generalize Exercise 3.5.4 in Volume I. Set

$$\Lambda(\mathfrak{a}, Y, s) = A^{2s}\Gamma(s)^{r_1}\,\Gamma(2s)^{r_2} Z(\mathfrak{a}, Y, s).$$

Then, if Y has the Iwasawa decomposition (2.61), it follows that:

$$Y\left\{\begin{matrix} a \\ b \end{matrix}\right\} = v\,\{a + qb\} + w\,\{b\}\,.$$

Thus

$$\Lambda(\mathfrak{a}, Y, s){=}Nv^{-s}\Lambda_K(\mathfrak{a}, 2s){+}A^{2s}\Gamma(s)^{r_1}\,\Gamma(2s)^{r_2}\sum_{\substack{0\neq b\in\mathfrak{a}/U_K\\ a\in\mathfrak{a}}} N\left(v\,\{a{+}qb\} + w\,\{b\}\right)^{-s},$$

if $\mathrm{Re}\, s > 1$ and we define

$$\Lambda_K(\mathfrak{a}, 2s) = A^{2s}\Gamma(s)^{r_1}\,\Gamma(2s)^{r_2}\sum_{0\neq b\in\mathfrak{a}/U_K} |Nb|^{-2s}\,.$$

The Poisson summation formula from Section 1.3 of Volume I or Weil [661, p. 106] shows that the sum over $a \in \mathfrak{a}$ equals the sum of Fourier transforms over $c \in \mathfrak{a}' = (\mathfrak{a}\mathfrak{d}_K)^{-1}$ which is the dual ideal to $\mathfrak{a}$. The Fourier transforms here are:

$$\widehat{f}(b, c) = \int\limits_{x\in K\otimes_{\mathbb{Q}}\mathbb{R}} N\left(v\,\{a + qb\} + w\,\{b\}\right)^{-s} \exp\left(-2\pi i \mathrm{Tr}\left({}^t cx\right)\right)\, d\mu(x),$$

where the measure $d\mu(x)$ is chosen so that

$$\int\limits_{x\in K\otimes_{\mathbb{Q}}\mathbb{R}/\mathfrak{a}} d\mu(x) = 1.$$

Now the ideal $\mathfrak{a}$ has an **integral basis**, i.e.,

$$\mathfrak{a} = \sum_{j=1}^{m}{}^{\oplus}\,\mathbb{Z}w_j \quad \text{and} \quad K\otimes_{\mathbb{Q}}\mathbb{R} = \sum_{j=1}^{m}{}^{\oplus}\,\mathbb{R}w_j.$$

So if

$$x = \sum_{j=1}^{m} x_j w_j, \quad x_j \in \mathbb{R},$$

we can take

$$d\mu(x) = \prod_{j=1}^{m} dx_j \text{ with } dx_j = \text{Lebesgue measure on } \mathbb{R}.$$

We can also see that

$$K \otimes_{\mathbb{Q}} \mathbb{R} \cong \sum_{j=1}^{r_1+r_2} {}^{\oplus} E_j, \quad \text{where} \quad E_j = \begin{cases} \mathbb{R}, j = 1, \ldots r_1, \\ \mathbb{C}, j = r_1 + 1, \ldots, r_1 + r_2. \end{cases}$$

Therefore we can define the mapping:

$$T : K \otimes_{\mathbb{Q}} \mathbb{R} \to \sum_{j=1}^{r_1+r_2} {}^{\oplus} E_j \text{ by } \quad T\left(\sum_{j=1}^{m} x_j w_j\right) = y = \left(y^{(1)}, \ldots, y^{(r_1+r_2)}\right),$$
$$\text{where} \quad y^{(i)} = \sum_{j=1}^{m} x_j w_j^{(i)}. \tag{2.64}$$

The Jacobian of the map T is (**Exercise**)

$$\left|\frac{\partial y}{\partial x}\right| = \left|\frac{\partial\left(y_1, \ldots, y_{r_1}, \operatorname{Re} y_{r_1+1}, \operatorname{Im} y_{r_1+1}, \ldots, \operatorname{Re} y_{r_1+r_2}, \operatorname{Im} y_{r_1+r_2}\right)}{\partial\left(x_1, \ldots, x_m\right)}\right| = 2^{-r_2} D_K^{\frac{1}{2}} N\mathfrak{a}.$$

Therefore

$$\begin{aligned}\widehat{f}(b, c) = &D_K^{-\frac{1}{2}} N\mathfrak{a}^{-1} 2^{r_2} \prod_{j=1}^{r_1+r_2} \int_{E_j} N_{E_j/\mathbb{R}}\left(\left(v\{a + qb\} + w\{b\}\right)^{(j)}\right)^{-s} \\ &\times \exp\left(-2\pi i \operatorname{Tr}_{E_j/\mathbb{R}}\left(\left(cy\right)^{(j)}\right)\right) dy^{(j)}.\end{aligned}$$

Make the change of variables $x_j = \left(\left(w\{b\}\right)^{-\frac{1}{2}} t\left(y + qb\right)\right)^{(j)}$, where $v = t^2, t^{(j)} > 0$, to see that:

$$\begin{aligned}\widehat{f}(b, c) = &D_K^{-\frac{1}{2}} N\mathfrak{a}^{-1} 2^{r_2}\, Nv^{-\frac{1}{2}} N\left(w\{b\}\right)^{\frac{1}{2}-s} \exp\left(2\pi i \operatorname{Tr}\left(cqb\right)\right) \\ &\times \prod_{j=1}^{r_1+r_2} \int_{E_j} \left(1 + \overline{x_j} x_j\right)^{-se_j} \exp\left(-2\pi i \operatorname{Tr}_{E_j/\mathbb{R}}\left(\left(\frac{c}{t}\left(w\{b\}\right)^{\frac{1}{2}}\right)^{(j)} x_j\right)\right) dx_j.\end{aligned}$$

Define

$$I_j(a, s) = \int_{E_j} \left(1 + \bar{x}x\right)^{-se_j} \exp\left(-2\pi i \operatorname{Tr}_{E_j/\mathbb{R}}\left(ax\right)\right) dx.$$

By part (a) of Exercise 3.2.1 in Volume I, we find that for $j = 1, \ldots, r_1$:

$$I_j(a,s) = \begin{cases} 2\pi^{\frac{1}{2}}\Gamma(s)^{-1}\,|\pi a|^{s-\frac{1}{2}}\,K_{s-\frac{1}{2}}\,(2\pi\,|a|)\,, & a \neq 0, \\ \Gamma\left(\frac{1}{2}\right)\Gamma\left(s-\frac{1}{2}\right)\Gamma\left(s\right)^{-1}, & a = 0. \end{cases}$$

For $j = r_1 + 1, \ldots, r_1 + r_2$, we must compute:

$$\begin{aligned} I_j(a_1 + ia_2, s) &= \int\limits_{x_1 + ix_2 \in \mathbb{C}} \left(1 + x_1^2 + x_2^2\right)^{-2s} \exp\left(-4\pi i\,(a_1 x_1 - a_2 x_2)\right) dx \\ &= k_{1,2}\,(2s|I, 2\pi\,(a_1, a_2))\,, \end{aligned}$$

where $k_{1,2}$ is the function defined in formula (1.60) in Section 1.2.2. We can use part (2) of Theorem 1.2.2 in Section 1.2.2 to see that in terms of the K-Bessel function defined by (1.61) in Section 1.2.2, we have, for $j = r_1 + 1, \ldots, r_1 + r_2$,

$$I_j(a_1 + ia_2, s) = \pi\,\Gamma(2s)^{-1} K_1\left(1 - 2s|4\pi^2\left(a_1^2 + a_2^2\right), 1\right).$$

It follows then that for $j = r_1 + 1, \ldots, r_1 + r_2$, we have the following formula when $a = a_1 + ia_2 \in \mathbb{C} = E_j$,

$$I_j(a,s) = \begin{cases} 2^{2s}\Gamma(2s)^{-1}\left(\pi^2\left(a_1^2 + a_2^2\right)\right)^{s-\frac{1}{2}} K_{2s-1}\left(4\pi\sqrt{a_1^2 + a_2^2}\right), & a = a_1 + ia_2 \neq 0, \\ \pi\,\Gamma\left(2s-1\right)\Gamma\left(2s\right)^{-1}, & a = 0. \end{cases}$$

Substituting these results into the original Poisson sum leads to:

$$\begin{aligned} \Lambda(\mathfrak{a}, Y, s) = {} & Nv^{-s}\Lambda_K(\mathfrak{a}, 2s) + Nv^{-\frac{1}{2}} Nw^{\frac{1}{2}-s} N\mathfrak{a}^{-1}\Lambda_K(\mathfrak{a}, 2s-1) \\ & + N\,|Y|^{\frac{1}{4}-\frac{s}{2}}\, Nv^{-\frac{1}{2}} A^{2s-1} N\mathfrak{a}^{-1} 2^{r_1 + 2sr_2} \\ & \times \sum_{\substack{0 \neq b \in \mathfrak{a}/U_K \\ 0 \neq c \in (\mathfrak{a}\mathfrak{d}_K)^{-1}}} \left|\frac{Nc}{Nb}\right|^{s-\frac{1}{2}} e^{2\pi i \mathrm{Tr}(qbc)} \prod_{j=1}^{r_1 + r_2} K_{e_j\left(s-\frac{1}{2}\right)}\left(2\pi e_j\sqrt{\tfrac{w^{(j)}}{v^{(j)}}}\left|b^{(j)}c^{(j)}\right|\right). \end{aligned}$$

To complete the proof of Theorem 2.2.1, suppose that C is an ideal class of K and $\mathfrak{b} \in C$ is as in formula (2.62). Note that the equation

$$\mathfrak{a}\mathfrak{b} = b\mathfrak{O}_K$$

defines a one-to-one mapping from ideals $\mathfrak{a} \in C^{-1}$ onto elements $b \in \mathfrak{b} \bmod U_K$. Set $u = bc$, for $c \in (\mathfrak{b}\mathfrak{d}_K)^{-1}$. Then define the map

$$L : (\mathfrak{b}/U_K) \times (\mathfrak{b}\mathfrak{d}_K)^{-1} \to \left(C^{-1}\right) \times (\mathfrak{d}_K)^{-1} \quad \text{by} \quad L\,(b,c) = (\mathfrak{a}, u = bc)\,.$$

The map L is easily seen to be one-to-one. It is not onto, since the image consists of $(\mathfrak{a}, u)$ such that $\mathfrak{a}$ divides $\mathfrak{d}_K u$.

Finally observe that

$$N\mathfrak{d}_K^{2s-1}\left|\frac{Nc}{Nb}\right|^{s-\frac{1}{2}} = N\mathfrak{a}^{1-2s}\ |Nu|^{s-\frac{1}{2}}\,.$$

This completes the proof of Theorem 2.2.1. ■

We have corrected the following Corollaries after reading the observation of Elstrodt et al. [168, p. 395]. We had made the mistake of thinking that $e_1 = 1$. But recall that Nv involves $v_1^{e_1}$ Of course in the case considered by Elstrodt, Grunewald, and Mennicke $e_1 = 2$ as the field K is imaginary quadratic.

Corollary 2.2.1 (Relations Between $\zeta_K(s)$ and $\zeta_K(s-1)$). *Set*

$$M_s(z) = K_s(z) + \frac{2}{e_1} z \frac{d}{dz} K_s(z)$$

and

$$T(s,u) = M_{e_1 s}\left(2\pi e_1\left|u^{(1)}\right|\right)\prod_{j=2}^{r_1+r_2} K_{e_j s}\left(2\pi e_j\left|u^{(j)}\right|\right).$$

Then

$$\begin{aligned}&(1-s)\,\Lambda_K\,(2s-1) + s\Lambda_K(2s)\\ &\quad = -e_1 2^{r_1+r_2-1} D_K^{s-\frac{1}{2}} \sum_{0\neq u\in(\mathfrak{d}_K)^{-1}} |Nu|^{s-\frac{1}{2}}\,\sigma_{1-2s}\,(u\mathfrak{d}_K)\,T\left(s-\frac{1}{2}, u\right).\end{aligned}$$

Proof. This is the analogue of a generalization of part (b) of Exercise 3.5.7 of Volume I.

Substitute

$$Y = \begin{pmatrix} v & 0\\ 0 & 1\end{pmatrix} \quad \text{in} \quad Z^*\,(Y, s)$$

and use the following functional equation:

$$Z^*\left(\begin{pmatrix} v & 0\\ 0 & 1\end{pmatrix}, s\right) = Z^*\left(\begin{pmatrix} 1 & 0\\ 0 & v\end{pmatrix}, s\right)$$

plus Theorem 2.2.1 to deduce that

$$\begin{aligned}
&Nv^{-s}\Lambda_K(2s) + Nv^{-\frac{1}{2}}\Lambda_K(2s-1)\\
&\qquad + \frac{2^{r_1+r_2}D_K^{s-\frac{1}{2}}}{Nv^{\frac{1}{4}+s/2}} \sum_{0\neq u\in\mathfrak{d}_K^{-1}} |Nu|^{s-\frac{1}{2}}\sigma_{1-2s}\left(u\mathfrak{d}_K\right) \prod_{j=1}^{r_1+r_2} K_{e_j(s-\frac{1}{2})}\left(2\pi e_j \frac{|u^{(j)}|}{\sqrt{v^{(j)}}}\right)\\
&= \Lambda_K(2s) + Nv^{\frac{1}{2}-s}\Lambda_K(2s-1)\\
&\qquad + \frac{2^{r_1+r_2}D_K^{s-\frac{1}{2}}}{Nv^{-\frac{1}{4}+s/2}} \sum_{0\neq u\in\mathfrak{d}_K^{-1}} |Nu|^{s-\frac{1}{2}}\sigma_{1-2s}\left(u\mathfrak{d}_K\right) \prod_{j=1}^{r_1+r_2} K_{e_j(s-\frac{1}{2})}\left(2\pi e_j \sqrt{v^{(j)}}\left|u^{(j)}\right|\right).
\end{aligned}$$

Differentiate this equation with respect to v_1 and set all $v_j = 1,\ j = 1, \ldots, r_1 + r_2$ to finish the proof of Corollary 2.2.1. ■

The following corollary gives upper bounds for the product of the class number and the regulator, which should be compared with that obtained by Lang [386, p. 261]. A lower bound is more difficult to obtain. See p. 74 of Volume I for the definition of the regulator.

Corollary 2.2.2 (A Formula for the Product of the Class Number and the Regulator).
Let K be any algebraic number field of degree m, with w_K = the number of roots of unity in K, R_K = the regulator of K, h_K = the class number of K, D_K = the absolute value of the discriminant, $e_1 = 1$ if the field K has any real conjugate fields and 2 otherwise, $\mathfrak{d}_K$ = the different, r_2 = the number of complex conjugate fields, ζ_K = the Dedekind zeta function, $T(s,u)$ as defined in Corollary 2.2.1. Then

$$h_K R_K/w_K = 2\,(2\pi)^{-m}\, D_K\zeta_K(2) + e_1 2^{r_2} D_K^{\frac{1}{2}} \sum_{0\neq u\in\mathfrak{d}_K^{-1}} |Nu|^{\frac{1}{2}}\,\sigma_{-1}\left(u\mathfrak{d}_K\right) T\left(\frac{1}{2}, u\right).$$

Proof. Let s approach 1 in Corollary 2.2.1. ■

Exercise 2.2.12. Compute the Jacobian of the mapping T in formula (2.64).

Exercise 2.2.13. Complete the proof of Corollary 2.2.2.

Corollary 2.2.3. *If K is a totally real algebraic number field, then, using the notation of Corollary 2.2.2, we have:*

$$h_K R_K = 4\,(2\pi)^{-m}\, D_K\zeta_K(2) - \pi 2^{3-m} D_K^{\frac{1}{2}} \sum_{0\neq u\in(\mathfrak{d}_K)^{-1}} \left|u^{(1)}\right| \sigma_{-1}\left(u\mathfrak{d}_K\right) e^{-2\pi\left(\left|u^{(1)}\right|+\cdots+\left|u^{(m)}\right|\right)}.$$

Proof. Since $K_{\frac{1}{2}}(z) = (2z/\pi)^{-\frac{1}{2}}\, e^{-z}$, we see that $M_{1/2}(z) = -(2nz)^{\frac{1}{2}} e^{-z}$. The result follows easily then from Corollary 2.2.2. ■

When $K = \mathbb{Q}$, Corollary 2.2.1 gives formulas relating $\zeta(2n)$ and $\zeta(2n+1)$ (see Exercise 3.5.7 of Vol. I). Formulas of this sort have been studied by many authors, without, however, leading to information on the rationality, irrationality, algebraicity, or transcendence of $\zeta(2n+1)$, $n = 1, 2, \ldots$. See Hunter [311].

Siegel [565, Vol. I, pp. 173–179] used the Fourier expansion of Eisenstein series for $GL(2, \mathfrak{O}_K)$ to obtain the analytic continuation and functional equation of the Dedekind zeta function. Mordell [460, pp. 518 ff.] also derives the Fourier expansion of the Eisenstein series for $GL(2, \mathfrak{O}_K)$.

Hoffstein [303] has used Fourier expansions of Eisenstein series for $GL(2, \mathfrak{O}_K)$, K a real quadratic field, to study the real zeros of these series. Asai [27] uses such Fourier expansions to generalize Kronecker's limit formula (see Exercise 3.5.6 in Vol. I). See also Zagier [674].

Grosswald [250] obtains results related to that in Corollary 2.2.1—formulas for the Dedekind zeta function involving the Meijer's G-function.

Our final goal in this section is to describe the relation between the Epstein zeta function (2.59) and the Eisenstein series for $SL(2, \mathfrak{O}_K)$. First recall that Proposition 2.2.1 gave a correspondence between the cusps of $\mathcal{SP}_2^K/SL(2, \mathfrak{O}_K)$ and ideal classes in the ideal class group I_K:

$$\begin{array}{ll} \widehat{K}/SL(2,\mathfrak{O}_K) & \leftrightarrow I_K \\ \text{represented by cusps} & \text{represented by ideals} \\ x_1, \ldots, x_h & \mathfrak{a}_1, \ldots, a_h. \end{array}$$

The map was obtained by setting $x_i = p_i/s_i$ with $p_i, s_i \in \mathfrak{O}_K$,

$$A_i = \begin{pmatrix} p_i & u_i \\ s_i & v_i \end{pmatrix} \in SL(2, K), \quad (p_i, s_i) = \begin{pmatrix} \text{the ideal generated} \\ \text{by } p_i \text{ and } s_i \end{pmatrix} = \mathfrak{a}_i,$$

$$x_i = A_i\infty, \qquad u_i,\ v_i \in \mathfrak{a}_i^{-1}.$$

We showed in Lemma 2.2.1 that

$$\begin{aligned} \Gamma_{x_i} &= \{\gamma \in SL(2, \mathfrak{O}_K) \mid \gamma x_i = x_i\} \\ &= \left\{ A_i \begin{pmatrix} w & z \\ 0 & w^{-1} \end{pmatrix} A_i^{-1} \,\middle|\, z \in \mathfrak{a}_i^{-2},\ w \in U_K = \text{units of } \mathfrak{O}_K \right\}. \end{aligned}$$

We can now define an **Eisenstein series corresponding to the cusp** x_i (cf. Kubota [377]):

$$E_i(Y, s) = N\mathfrak{a}_i^{2s} \sum_{\gamma \in SL(2,\mathfrak{O}_K)/\Gamma_{x_i}} N\left((v(Y\{\gamma A_i\})\right)^{-s}, \quad \text{for } \operatorname{Re} s > 1. \tag{2.65}$$

Here we use the notation that if $Y \in \mathcal{P}_2^K$ has Iwasawa decomposition (2.61), we write $v(Y)$ for the v-coordinate of Y. It is also the upper left entry of Y. Now taking the v-part of $Y\{\gamma A_i\}$ amounts to taking $Y\{g\}$, where g is the first column of γA_i. Since such a g must generate $\mathfrak{a}_i$, we find that

$$E_i(Y,s) = N\mathfrak{a}_i^{2s} \sum_{\substack{g\in\mathfrak{a}_i^2/U_K \\ \text{entries of } g \text{ generate } \mathfrak{a}_i}} N\left(Y\{g\}\right)^{-s}, \text{ for } \operatorname{Re} s > 1. \tag{2.66}$$

Exercise 2.2.14. Prove formula (2.66) for all $Y \in \mathcal{P}_2^K$. Then show that the Eisenstein series E_i is dependent only on the ideal class containing the ideal $\mathfrak{a}_i$, and not on the choice of $\mathfrak{a}_i$ in that ideal class. Finally, prove that, if we define for an ideal class $C \in I_K$ the **ideal class zeta function**:

$$\zeta(C,s) = \sum_{\mathfrak{c}\in C} N\mathfrak{c}^{-s}, \text{ for } \operatorname{Re} s > 1,$$

then we have a relation between Epstein's zeta function (2.59) and the Eisenstein series:

$$Z(\mathfrak{O}_K, Y, s) = \sum_{i=1}^{h} \zeta(C_i, 2s) E_i(Y,s),$$

using the notation C_i for the ideal class containing $\mathfrak{a}_i$, $i = 1, \ldots, h$.

In order to generalize the relation obtained in Exercise 2.2.14 between $Z(\mathfrak{O}_K, Y, s)$ and the vector of Eisenstein series E_i to $Z(\mathfrak{a}_i, Y, s)$, we need a **matrix of ideal class zeta functions**:

$$M_K(s) = \left(\zeta(C(i,j), 2s)\right)_{1\le i,j\le h}, \tag{2.67}$$

where $C(i,j)$ is the ideal class containing the ideal $\mathfrak{a}_j\mathfrak{a}_i^{-1}$. Define also the **column vector of Epstein zeta functions**:

$$\vec{Z}(Y,s) = {}^t\left(Z\left(\mathfrak{a}_1, Y, s\right), \ldots, Z\left(\mathfrak{a}_h, Y, s\right)\right), \tag{2.68}$$

and the **column vector of Eisenstein series**:

$$\vec{E}(Y,s) = {}^t\left(E_1(Y,s), \ldots, E_h(Y,s)\right). \tag{2.69}$$

Proposition 2.2.3 (The Relation Between Epstein's Zeta Function and the Eisenstein Series for $SL(2, \mathfrak{O}_K)$). *Using the notation (2.59), (2.65), (2.67)–(2.69), we have the following equality for* $\operatorname{Re} s > 1$*:*

$$\vec{Z}(Y,s) = M_K(s)\vec{E}(Y,s).$$

Proof. We have the following chain of equalities:

$$Z(\mathfrak{a}_i, Y, s) = N\mathfrak{a}_i^{2s} \sum_{0 \neq g \in \mathfrak{a}_i^2/U_K} N(Y\{g\})^{-s} = N\mathfrak{a}_i^{2s} \sum_{\mathfrak{a}_i|\mathfrak{b}} \sum_{\substack{g\in\mathfrak{b}^2/U_K \\ (g_1,g_2)=\mathfrak{b}}} N(Y\{g\})^{-s}$$

$$= \sum_{\mathfrak{a}_i|\mathfrak{b}} \left(\frac{N\mathfrak{b}}{N\mathfrak{a}_i}\right)^{-2s} N\mathfrak{b}^{2s} \sum_{\substack{g\in\mathfrak{b}^2/U_K \\ (g_1,g_2)=\mathfrak{b}}} N(Y\{g\})^{-s}.$$

Here the inner sum is over 2-vectors g with entries in the ideal $\mathfrak{b}$ modulo the unit group U_K such that the entries of g generate the ideal $\mathfrak{b}$. Then set $\mathfrak{b}/\mathfrak{a}_i = \mathfrak{c}$ and observe that $\mathfrak{c}$ runs through all integral ideals in the ideal class $C(i,j)$ containing the ideal $\mathfrak{a}_j\mathfrak{a}_i^{-1}$, to complete the proof. ∎

The formula in Proposition 2.2.3 raises certain questions.

Questions Arising from Proposition 2.2.3

(1) Is it possible to diagonalize the matrix $M_K(s)$ using characters of the ideal class group?
(2) What does this have to do with Hecke operators for $GL(2, \mathfrak{O}_K)$?
(3) What does this have to do with the analogue of Siegel's integral formula for $GL(n, \mathfrak{O}_K)$? See Proposition 1.4.2 of Section 1.4.4 for the integral formula when $K = \mathbb{Q}$.

Here we shall discuss only question 1. Hecke operators for these groups are treated by Herrmann [292], Shimura [554], and Styer [584]. Siegel's integral formula for $SL(n, \mathfrak{O}_K)$, K imaginary quadratic, is considered by Hunter [311]. See also Elstrodt et al. [168] as well as Efrat and Sarnak [154].

The ideal class group I_K is a finite abelian group of order h. Let $\widehat{I}_K$ denote the **dual group** of characters $\chi : I_K \to \mathbb{T}$, where $\mathbb{T}$ is the circle group of complex numbers of norm 1; i.e.,

$$\mathbb{T} = \{z \in \mathbb{C} \mid |z| = 1\}.$$

That is, χ is a homomorphism of multiplicative groups. Then the dual group is:

$$\widehat{I}_K = \{\chi_1, \ldots, \chi_h\}.$$

We can diagonalize the matrix M_K using Fourier transforms on I_K. Suppose that C_i is the ideal class containing the ideal $\mathfrak{a}_i$, for $i = 1, \ldots, h$. Define

$$Z(C_i) = Z(\mathfrak{a}_i, Y, s), \quad E(C_i) = E(\mathfrak{a}_i, Y, s), \quad \zeta(C) = \zeta(C, 2s). \tag{2.70}$$

Our proof of Proposition 2.2.3 rested on the equation:

$$Z(C_i) = \sum_{j=1}^{h} \zeta(C_j/C_i)E(C_j). \tag{2.71}$$

Now define **convolution** of functions $f : I_K \to \mathbb{C}$ by:

$$(f * g)(C_i) = h^{-1} \sum_{j=1}^{h} f\left(C_i/C_j\right) g(C_j). \tag{2.72}$$

Thus formula (2.71) says that

$$Z(C) = h\zeta(C^{-1}) * E(C). \tag{2.73}$$

As for the group of real numbers (see part (4) of Theorem 1.2.1 in Volume I), the Fourier transform can be used to simplify this convolution equation. We define the **Fourier transform** of a function

$$f : I_K \to \mathbb{C}$$

at the character $\chi \in \widehat{I_K}$ by:

$$\widehat{f}(\chi) = h^{-1} \sum_{y \in I_K} f(y)\overline{\chi(y)}. \tag{2.74}$$

Since I_K is a finite abelian group, there are no convergence problems. In fact, the theory of Fourier transforms on finite abelian groups has many applications, since it is just what is needed for the fast Fourier transform, an idea which has speeded computation of such transforms immensely. See Terras [608, 609] for more information on finite and fast Fourier transforms.

Proposition 2.2.4 (Some Properties of the Fourier Transform on I_K).

(1) Convolution.

$$\widehat{f * g}(\chi) = \widehat{f}(\chi) \cdot \widehat{g}(\chi).$$

(2) Inversion.

$$f(x) = \sum_{\chi \in \widehat{I_K}} \widehat{f}(\chi)\chi(x), \quad \text{for all } x \in I_K.$$

Proof. (1) Note that

$$\widehat{f*g}(\chi) = h^{-2}\sum_{z\in I_K}\sum_{y\in I_K} f(zy^{-1})g(y)\overline{\chi(z)}$$
$$= h^{-2}\sum_{y\in I_K} g(y) \sum_{w=zy^{-1}\in I_K} f(w)\overline{\chi(wy)} = \widehat{f}(\chi)\cdot\widehat{g}(\chi).$$

(2) Observe that

$$h^{-1}\sum_{\chi\in\widehat{I_K}}\chi(x)\sum_{y\in I_K} f(y)\overline{\chi(y)} = \sum_{y\in I_K} f(y)h^{-1}\sum_{\chi\in\widehat{I_K}}\chi(xy^{-1}) = f(x),$$

since we have:

$$h^{-1}\sum_{\chi\in\widehat{I_K}}\chi(xy^{-1}) = \begin{cases} 0, & x\neq y,\\ 1, & x=y;\end{cases} \text{ and } \overline{\chi(y)} = \chi(y)^{-1}. \tag{2.75}$$

■

Exercise 2.2.15. Prove formula (2.75) above.

For

$$x\in\widehat{I}_K,\quad Y\in\mathcal{P}_2^K,\quad s\in\mathbb{C} \text{ with } \operatorname{Re} s>1,$$

define the **zeta function**:

$$Z(\chi,Y,s) = \sum_{\mathfrak{a}} N\mathfrak{a}^{2s}\,\chi(\mathfrak{a}) \sum_{0\neq g\in\mathfrak{a}^2/U_K} N\left(Y\{g\}\right)^{-s}, \tag{2.76}$$

where the outer sum is over all ideals $\mathfrak{a}$ of $\mathfrak{O}_K$ and the character χ of the ideal class group is regarded in the obvious way as a function of ideals. Then $Z(\chi,Y,s)$ is the Fourier transform of $Z(C)$ (times h), where $Z(C)$ is defined in formula (2.70).

Similarly, define the **Eisenstein series** associated with $\chi\in\widehat{I}_K$, $Y\in\mathcal{P}_2^K$, and $s\in\mathbb{C}$ with $\operatorname{Re} s>1$ by:

$$E(\chi,Y,s) = \sum_{\mathfrak{a}} N\mathfrak{a}^{2s}\,\chi(\mathfrak{a}) \sum_{\substack{g\in\mathfrak{a}^2/U_K\\ (g_1,g_2)=\mathfrak{a}}} N\left(Y\{g\}\right)^{-s}, \tag{2.77}$$

where the outer sum is over all ideals $\mathfrak{a}$ of $\mathfrak{O}_K$ and the inner sum is over column vectors $g = {}^t(g_1,g_2)$ such that the ideal $\mathfrak{a}$ is generated by g_1 and g_2 and the vectors g form a complete set of representatives for the equivalence relation obtained from multiplication by units. Then $E(\chi,Y,s)$ is h times the Fourier transform of $E(C)$ defined by (2.70).

Proposition 2.2.5 (The Diagonalization of the Relation Between Epstein's Zeta Function and the Eisenstein Series for $SL(2, \mathfrak{O}_K)$). *Using the definitions (2.76), (2.77) and setting*

$$L(\chi, s) = \sum_{\mathfrak{a}} \chi(\mathfrak{a}) N\mathfrak{a}^{-s}, \quad \text{for } \operatorname{Re} s > 1,$$

with the sum running over all ideals $\mathfrak{a}$ *of* $\mathfrak{O}_K$*, we have*

$$Z(\overline{\chi}, Y, s) = L(\chi, s) E(\overline{\chi}, Y, s).$$

Proof. This is just the convolution property in Proposition 2.2.4 for the special case of the functions from Proposition 2.2.3. ∎

Our discussion of nonholomorphic automorphic forms for $GL(2, \mathfrak{O}_K)$ is now at an end, although there still remains much to do if we wish to extend all of Chapter 3 of Volume I and Chapter 1 of this volume to $GL(n, \mathfrak{O}_K)$. For we have not even begun the theory of Hecke operators, the Hecke correspondence, the Selberg trace formula. See Arthur [21–25], Bernstein and Gelbart [47], Frenkel [187], Gelbart [208], Goldfeld and Hundley [232], Jacquet and Langlands [324], and Weil [660] for a general adelic version of the subject. Many papers on automorphic forms are reviewed in *Math. Reviews*. See also the collections of math. reviews in LeVeque [401] and Guy [255] as well as a third volume compiled by the *Math. Reviews* staff [436]. In the next section we seek to address some examples of higher rank trace formulas.

2.2.3 *Trace Formulas*

In this final section we discuss special cases of the Selberg trace formula which can be applied to the special cases of automorphic forms just discussed.

Trace Formula for Discrete Γ Acting on the Quaternionic Upper Half Plane

Let us give a brief sketch of the Selberg trace formula for cocompact discrete subgroups Γ of $G = SL(2, \mathbb{C})$ or $PSL(2, \mathbb{C})$, following Elstrodt et al. [168], Define $\gamma \in \Gamma$ to be

parabolic	if $\operatorname{Tr}(\gamma) \in \mathbb{R}$ and $\lvert\operatorname{Tr}(\gamma)\rvert = 2$;
hyperbolic	if $\operatorname{Tr}(\gamma) \in \mathbb{R}$ and $\lvert\operatorname{Tr}(\gamma)\rvert > 2$;
elliptic	if $\operatorname{Tr}(\gamma) \in \mathbb{R}$ and $\lvert\operatorname{Tr}(\gamma)\rvert < 2$;
loxodromic	otherwise.

If $\gamma \in \Gamma$ is hyperbolic or loxodromic, it is conjugate in G to a diagonal matrix with diagonal entries $a(\gamma)$ and its reciprocal. We may assume $|a(\gamma)| > 1$. Define the **norm** of hyperbolic or loxodromic γ to be $N(\gamma) = |a(\gamma)|^2$. Define $N(\gamma) = 1$ is γ is elliptic.

Recall that **the quaternionic upper half plane** is

$$\mathcal{H}^c = \left\{ z + kt = x + iy + kt \,|x, y \in \mathbb{R},\ t > 0 \right\},$$

with Laplacian $\Delta = t^2(\partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial t^2) - t\partial/\partial t$. Note that $\Delta t^{1+s} = \left(s^2 - 1\right) t^{1+s}$. Thus it is natural to write the eigenvalue of $-\Delta$ acting on t^{1+s} in the form $\lambda = 1 - s^2$. Similarly the eigenvalue of $-\Delta$ acting on t^{1+r} can be written in the form $\mu = 1 - r^2$. Let $R_\lambda = (-\Delta - \lambda I)^{-1}$ denote the **resolvent** operator. This is not trace class but $R_\lambda R_\mu$ is trace class. Recall the **resolvent equation** $(\lambda - \mu) R_\lambda R_\mu = R_\lambda - R_\mu$.

Suppose that $\{e_n\}_{n \geq 0}$ is a complete orthonormal set of eigenfunctions of Δ on the (compact) fundamental domain of Γ. Write $-\Delta e_n = \lambda_n e_n$ with $\lambda_n = 1 - s_n^2$. Here we may assume that $s_n = it_n$, with $t_n \geq 0$, except for a finite number of n with $s_n \in [-1, +1]$. Why? Recall the fact that $\lambda_n \geq 0$.

Elstrodt et al. [168] obtain a special case of the **Selberg trace formula** says (assuming $\Gamma\backslash\mathcal{H}^c$ is compact):

$$(\lambda - \mu) \operatorname{Tr}\left(R_\lambda R_\mu\right) = \sum_{n \geq 0} \left(\tfrac{1}{s^2 - s_n^2} - \tfrac{1}{r^2 - r_n^2} \right) = \tfrac{-\mathrm{vol}(\Gamma\backslash\mathcal{H}^c)}{4\pi} (s - r)$$
$$+ \tfrac{1}{2s} \sum_{\{\gamma\}} \tfrac{\log(N(\gamma_0))}{|\varepsilon(\gamma)|\left|\mathrm{Tr}(\gamma)^2 - 4\right|} N(\gamma)^{-s} - \tfrac{1}{2r} \sum_{\{\gamma\}} \tfrac{\log(N(\gamma_0))}{|\varepsilon(\gamma)|\left|\mathrm{Tr}(\gamma)^2 - 4\right|} N(\gamma)^{-r}.$$

The sums over $\{\gamma\}$ range over the noncentral conjugacy classes of Γ. The element γ_0 is a hyperbolic or loxodromic element of the centralizer, $Z(\gamma)$, of γ in Γ having minimal norm. The maximal finite subgroup in the centralizer of γ is denoted $\varepsilon(\gamma)$. See Elstrodt et al. [168, p. 199]. They use the result to study the Selberg zeta function for Γ and to prove the **Weyl law** for the asymptotic behavior of the counting function

$$\#\left\{ n \;\middle|\; \lambda_n = 1 - s_n^2 = 1 + t_n^2,\ \ t_n < T \right\} \sim \frac{vol\,(\Gamma\backslash\mathcal{H}^c)}{6\pi^2} T^3, \text{ as } T \to \infty.$$

Error terms are also obtained. See [168, p. 211]. Moreover Elstrodt, Grunewald, and Mennicke consider the case that $\Gamma\backslash\mathcal{H}^c$ is finite volume but not compact. They give examples of groups Γ such that $\Gamma\backslash\mathcal{H}^c$ is compact as well as noncompact finite volume examples (see [168, Chapter 10]).

Sarnak [527] applies the Selberg trace formula for $SL(2, \mathfrak{O}_K)$, K imaginary quadratic of class number one, to extend his results on the asymptotics of units in number fields.

Özlem Imamoḡlu and Nicole Raulf [319] use the Selberg trace formula for $\Gamma = SL(2, \mathfrak{O}_K)$, K imaginary quadratic of class number one, to study the distribution of Hecke eigenvalues for Γ. They assume that $\{e_n\}_{n\geq 0}$ forms a complete orthonormal set of eigenfunctions of Δ in $L^2(\Gamma\backslash\mathcal{H}^c)$ and that, in addition $T_{\mathfrak{p}} e_n = \rho_n(\mathfrak{p}) e_n$, for the Hecke operator $T_{\mathfrak{p}}$ associated with the prime ideal $\mathfrak{p}$ of $\mathfrak{O}_K$. Since the class number is one, $\mathfrak{p} = v\mathfrak{O}_K$ for some element $v \in \mathfrak{O}_K$. The element v is unique up to multiplication by a unit in $\mathfrak{O}_K$. The **Hecke operator** T_v associated with any nonzero element $v \in \mathfrak{O}_K$ is defined as a sum over $SL(2, \mathfrak{O}_K)\backslash M_v$, where

$$M_v = \left\{ A \in \mathfrak{O}_K^{2\times 2} \middle| \det A = v \right\}.$$

More explicitly

$$(T_v f)(z) = \frac{1}{Nv} \sum_{A \in M_v/SL(2,\mathfrak{O}_K)} f(Az).$$

The main result proved by Imamoḡlu and Raulf [319] is that the sequence of Hecke eigenvalues $\{\rho_n(\mathfrak{p})\}_{n\geq 1}$ is equidistributed according to the measure

$$d\mu_{\mathfrak{p}}(x) = \begin{cases} \frac{1}{2\pi}\left(1 + \frac{1}{N\mathfrak{p}}\right) \frac{\sqrt{4-x^2}}{\left(1+\frac{1}{N\mathfrak{p}}\right)^2 - \frac{x^2}{N\mathfrak{p}}}, & \text{if} \quad |x| < 2, \\ 0, & \text{otherwise.} \end{cases}$$

This measure approaches an analogue of the Sato–Tate or semi-circle measure as $N\mathfrak{p} \to \infty$. To prove the result, they first use the trace formula to obtain a Weyl law for powers of the Hecke eigenvalues. Then they use the method of moments to obtain the main result.

Trace Formula for Discrete Γ Acting on $\mathcal{H}^m$

If one wants to study groups such as the Hilbert modular group $\Gamma = SL(2, \mathfrak{O}_K)$, K a totally real number field, it helps to have a trace formula for $\Gamma\backslash\mathcal{H}^m$, where m is the degree of K over $\mathbb{Q}$. See Efrat [152], Freitag [186], Müller [469], Shimizu [553], and Zograf [677]. We consider here a special case of the trace formula which gives rise to a formula for the dimension of the space of Hilbert cusp forms of weight 2 when K is real quadratic following Freitag [186].

For $\gamma \in SL(2, \mathbb{R})^m$, write

$$\gamma = \left(\underbrace{\gamma^{(1)}, \ldots, \gamma^{(k)}}_{\text{hyperbolic}}, \underbrace{\gamma^{(k+1)}, \ldots, \gamma^{(l)}}_{\text{parabolic}}, \underbrace{\gamma^{(l+1)}, \ldots \gamma^{(m)}}_{\text{elliptic}} \right).$$

This means that after conjugation

$$\begin{aligned}
\gamma^{(i)} &= \begin{pmatrix} \varepsilon_i & 0 \\ 0 & 1/\varepsilon_i \end{pmatrix}, \quad \text{for} \quad i = 1, \ldots, k; \\
\gamma^{(i)} &= \begin{pmatrix} 1 & a_i \\ 0 & 1 \end{pmatrix}, \qquad \text{for} \quad i = k+1, \ldots, l; \\
\gamma^{(i)} &\in SO(2, \mathbb{R}), \qquad \text{for} \quad i = l+1, \ldots, m.
\end{aligned}$$

One creates a self-reproducing kernel starting with (see Freitag [184, p. 74])

$$k(z, w) = N\left(\frac{z - \overline{w}}{2i}\right) = \prod_{j=1}^{m} \left(\frac{z_j - \overline{w_j}}{2i}\right)^{-2} \quad \text{for} \quad z, w \in \mathcal{H}^m.$$

Then define for $\gamma \in SL(2, \mathbb{R})^m$

$$\begin{aligned}
&k(\gamma, z) = \left[\tfrac{k(\gamma z, z)}{k(z,z)}\right]^r j(\gamma, z), \\
&\text{where} \quad j(\gamma, z) = N(cz + d)^{-2} = \prod_{j=1}^{m} \left(c^{(j)} z_j + d^{(j)}\right)^{-2}, \\
&\text{for} \quad \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, \mathbb{R})^m.
\end{aligned}$$

Let ℓ denote the order of the kernel of the natural projection of Γ into $(SL(2, \mathbb{R}) / \{\pm I\})^m$. Define for an integer $r \geq 2$,

$$K(z, w) = \frac{1}{\ell} \sum_{\gamma \in \Gamma} k(\gamma w, z)^r \, j(\gamma, w)^r.$$

Freitag [186, p. 79] proves that

$$\dim \mathcal{S}\,(\Gamma, 2r) = \left(\frac{2r - 1}{4\pi}\right) \int_{\Gamma\backslash\mathcal{H}^m} \frac{K(z, z)}{k(z, z)^r} N y^{-2} dx dy.$$

One then writes the sum over Γ giving $K(z, z)$ as a sum over conjugacy classes and finds that the only contribution comes from the central terms and the elliptic terms, assuming $\Gamma\backslash\mathcal{H}^m$ compact. Assuming $\Gamma\backslash\mathcal{H}^m$ compact and Γ irreducible (meaning that the restriction of each of the m projections of $SL(2, \mathbb{R})^m$ into $SL(2, \mathbb{R})$ is 1–1), the result is (cf. Freitag [184, p. 89]:

$$\dim \mathcal{S}\,(\Gamma, 2r) = \frac{\operatorname{vol}\,(\Gamma\backslash\mathcal{H}^m)}{(2r - 1)^m} + \sum_a E_r(\Gamma, a),$$

where the sum over a is over a set of representatives of Γ-classes of elliptic fixed points with

$$E_r(\Gamma, a) = \frac{1}{|\Gamma_a|} \sum_{\substack{\gamma \in \Gamma_a \\ \gamma \neq \text{identity}}} N \frac{\zeta^r}{1-\zeta},$$

Here the **stabilizer** of a is $\Gamma_a = \{\gamma \in \Gamma \mid \gamma a = a\}$. An elliptic γ is conjugate in $SL(2, \mathbb{C})$ with a matrix $\begin{pmatrix} \zeta & 0 \\ 0 & \zeta^{-1} \end{pmatrix}$, $\zeta^h = 1$, for some h. If $\Gamma \backslash \mathcal{H}^m$ has cusps as is the case for the Hilbert modular group, then there will also be terms corresponding to the cusps κ of the form $L(\Gamma, \kappa)$, a Shimizu L-series. See [186, p. 110].

Putting this together for the case of the Hilbert modular group (for which the fundamental domain does have cusps), one obtains the formula for the **arithmetic genus** $= g = 1 + (-1)^m \dim \mathcal{S}(\Gamma, 2)$, $m > 1$. Freitag [186, p. 130] gives the result for the Hilbert modular group for $K = \mathbb{Q}\left(\sqrt{p}\right)$, where p is a prime. For example, $g = 1$, if $p = 2, 3, 5$, and, if the prime p is greater than 5 and $p \equiv 1 \pmod 4$,

$$g = 1 + \dim \mathcal{S}(SL(2, \mathfrak{O}_{\mathbb{Q}(\sqrt{d})}), 2) = \frac{\zeta_K(-1)}{2} + \frac{h(-4p)}{8} + \frac{h(-3p)}{6},$$

where $h(d)$ is the class number of $\mathbb{Q}\left(\sqrt{d}\right)$. Another reference is Hirzebruch [296]. Helen Grundman and Lisa E. Lippencott [252] have computed the arithmetic genus for many examples of totally real degree 4 fields over $\mathbb{Q}$.

It is also possible to compute dimensions of spaces of holomorphic cusp forms using generalized Riemann–Roch theorems.

Trace Formula for Γ Acting on the Siegel Upper Half Space

Many people have used the trace formula to compute the dimension of the space of holomorphic cusp forms of weight k for a subgroup Γ of $Sp(n, \mathbb{R})$. We mention only a few: Arakawa [19], Christian [109], Eie [160], Hashimoto [264], Morita [463], Tsushima [617], and Wakatsuki [640]. Yamazaki [673] obtained similar results using the Riemann–Roch theorem. See also the review of Christian's paper by Resnikoff in *Math. Reviews*, 53 #2841. The method goes back to Selberg as well as Godement's Séminaire Cartan [547] lectures in which one uses a self-reproducing or Bergman kernel in the trace formula. This is the same method used for Hilbert modular forms. See Klingen [355, p. 76] for a discussion of the reproducing kernel.

Of course, one must also compute all the conjugacy classes of Γ as well as the orbital integrals corresponding to each class. For example when $n = 3$ there are 300 conjugacy classes. It is not surprising that mistakes might be made. Anyway if you read [*Math.*] *Reviews in Number Theory, Vol. 2B*, for the period 1984–1996, especially pages 576 and 586, you will find much heated discussion. Here we will definitely not delve into the details of these computations, nor take sides in this war.

What is the kernel used in the trace formula to obtain the dimension of $\mathcal{S}\left(Sp(n,\mathbb{Z}),k\right)$, the space of Siegel cusp forms of weight k? Here we follow Arakawa [19] and Hashimoto's discussion in [272, pp. 253–276]. It starts out for $Z,W \in \mathcal{H}_n$ as $k(Z,W) = \det\left(\frac{Z-\overline{W}}{2i}\right)$. See Klingen [355, p. 76] who uses the Cayley transformation to transform the integrals from $\mathcal{H}_n$ to the generalized unit disc:

$$\mathcal{D}_n = \left\{W \in \mathbb{C}^{n\times n} \mid {}^tW = W,\ I - \overline{W}W \in \mathcal{P}_n\right\}.$$

which is a bounded domain. The Cayley transform is:

$$\begin{array}{l} \mathcal{H}_n \to \mathcal{D}_n \\ Z \ \ \mapsto (Z-iI)(Z+iI)^{-1}. \end{array}$$

The self-reproducing formula on $\mathcal{D}_n$ involves the Bergman kernel and was studied by Hua [308].

Then, in the usual way of trace formulas, following Godement [547], one gets a self-reproducing kernel on the Hilbert space $L^2(\Gamma\backslash H_n, \det(\mathrm{Im}(W))^k d\mu^*)$, using the measure for the Petersson inner product. The self-reproducing kernel for the weight k cusp forms has the following form, with $j(\gamma,Z) = \det(CZ+D)$, if $\gamma = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \Gamma$,

$$K(Z,W) = a_n(k) \sum_{\gamma\in\Gamma} \det\left(\frac{Z-\overline{\gamma W}}{2i}\right)^{-k} j(\gamma,\overline{W})^{-k},$$

where $a_n(k)$ is a constant. Note that the kernel is the symplectic analogue of that for the Hilbert modular group. Moreover, the trace of the kernel should give the dimension of the space of Siegel cusp forms of weight k.

Next one must split Γ into conjugacy classes and evaluate orbital integrals for each conjugacy class. See Wakatsuki [640, pp. 203–204]. There are seven basic types: central, elliptic, hyperbolic, elliptic-hyperbolic, unipotent (or parabolic), quasi-unipotent, hyperbolic-unipotent. What happens next is similar to what happened in the Hilbert modular case. One finds that the orbital integrals vanish unless they correspond to $\gamma \in \Gamma$ which are central, elliptic, unipotent, or quasi-unipotent. Dumping factors are needed for the last two types of terms. The elliptic terms were evaluated by Langlands in [389].

Let us just mention some results on dimensions of spaces of Siegel modular forms of small weight for $Sp(n,\mathbb{Z})$, where $n = 2$. Some references are William Duke and Özlem Imamoḡlu [147], Gerard van der Geer's article in Bruinier et al. [82], Jun-ichi Igusa [315], Helmut Klingen [355], and Martin Raum et al. [500], and Tsuymine [619]. If $\Gamma = Sp(2,\mathbb{Z})$, the ring of Siegel modular forms of even weights is generated by the four Eisenstein series E_4, E_6, E_{10}, E_{12}. See Klingen [355, p. 123]. Moreover the generators are algebraically independent. Thus, for even k, the dimension of the space of Siegel modular forms $\mathcal{M}\,(Sp(2,\mathbb{Z}),k)$ is the number of nonnegative integer solutions (a,b,c,d) of $k = 4a + 6b + 10c + 12d$. The first cusp form occurs at weight 10. Breeding [78], Van der Geer in [82, p. 233] and Wakatsuki [640, p. 249] give the following list of dimensions

$$d_k(2) = \dim \mathcal{S}\,(Sp(2,\mathbb{Z}),k)$$

of the space of Siegel cusp forms of even weight for $Sp(2,\mathbb{Z})$:

$\boldsymbol{k}$	10	12	14	16	18	20	22
$\boldsymbol{d_k}(2)$	1	1	1	2	2	3	4

These authors also give more general tables for modular forms transforming according to a representation of $GL(2,\mathbb{C})$ and for more general $\Gamma \subset Sp(2,\mathbb{R})$ such as congruence subgroups and quaternion groups. Wakatsuki notes that one can use such dimension formulas to aid in understanding the Jacquet–Langlands–Ihara correspondence for $Sp(2,\mathbb{R})$.

The dimensions of the spaces of Siegel cusp forms for $Sp(3,\mathbb{Z})$ were computed by Eie and Lin [161] as well as Tsuyumine [619]. One finds for example that the first time (for even k) that $\dim \mathcal{M}\,(Sp(3,\mathbb{Z}),k) \geq 2$ is $k = 10$.

Tsuymine [619] gives a long table for

$$md_k(3) = \dim \mathcal{M}\,(Sp(3,\mathbb{Z}),k)$$

as well as 34 generators of the ring of even weight modular forms for $Sp(3,\mathbb{Z})$. He notes that the ring of Siegel modular forms for $Sp(3,\mathbb{Z})$ cannot be generated by Eisenstein series. Instead he makes use of theta series known as theta constants. We reproduce a bit of Tsuymine's table here:

$\boldsymbol{k}$	0	2	4	6	8	10	12	14	16	18	20
$\boldsymbol{md_k}(3)$	1	0	1	1	1	2	4	3	7	8	11

See Poor and Yuen [492] for dimensions of spaces of Siegel modular forms of low weight for $Sp(4,\mathbb{Z})$. They find for example that $\dim \mathcal{S}\,(Sp(4,\mathbb{Z}),8) = 1$, $\dim \mathcal{S}\,(Sp(4,\mathbb{Z}),12) = 2$, and that the dimensions in lower weights are 0. William Duke and Özlem Imamoḡlu [147] find a multitude of results of this sort for small weights and small n in $Sp(n,\mathbb{Z})$. There is a conjectural formula for dimensions of spaces $\mathcal{M}\,(Sp(n,\mathbb{Z}),k)$ in general. See T. Ibukiyama and H. Saito [314]. Wikipedia has a long table of $\dim \mathcal{M}\,(Sp(n,\mathbb{Z}),k)$ going out to $n = 9$. For more information on Siegel modular forms for $Sp(n,\mathbb{Z})$, $n \leq 4$, see the website: www.lmfdb.org.

Applications of Siegel modular forms are discussed by A. Ghitza in [217]. Duke [145] gives applications to coding theory. Applications to cryptography can be found in Kirsten Eisenträger and Kristin Lauter [162] as well as Kristin Lauter and Tonghai Yang [397].

http://www.springer.com/978-1-4939-3406-5

Harmonic Analysis on Symmetric Spaces—Higher Rank Spaces, Positive Definite Matrix Space and Generalizations
Terras, A.
2016, XV, 487 p. 41 illus., 21 illus. in color., Hardcover
ISBN: 978-1-4939-3406-5