The Semihull Theorem for Siegel Modular Forms

A Thesis
Presented to
The Division of Mathematics and Natural Sciences
Reed College

In Partial Fulfillment
of the Requirements for the Degree
Bachelor of Arts

Nadir Hajouji

December 2012

Approved for the Division
(Mathematics)

_____

Jerry Shurman

Acknowledgements

I'd like to thank my family for all of their support. I would certainly not be here without it. I'd like to thank all of my professors at Reed for helping me grow intellectually. Most of all, I'd like to thank Jerry, my advisor, for working tirelessly to make the thesis process as enjoyable and productive as possible.

Table of Contents

Abstract

The first half of the thesis focuses on the geometry of numbers, specifically the theory of height functions and kernels. The main idea introduced in the first half of the thesis is the dyadic trace, a new and useful height function. The second half of the thesis focuses on Siegel modular forms. The main result is the Semihull Theorem for Siegel modular forms. The Problem of Witt is solved using theory developed in this thesis.

Dedication

This thesis is dedicated to my dad.

Preface

Siegel modular forms are a generalization of classical modular forms, sharing many properties with them. Siegel modular forms are defined on Siegel upper half-space, a space of matrices remarkably similar to the upper half-plane. Furthermore, Siegel modular forms have unique Fourier expansions, and the space of Siegel modular forms of a given dimension and weight is a finite-dimensional vector space. See Geer (2007) and Klingen (1990) for surveys of results on Siegel modular forms, and see Hulek and Sankaran (1998) for applications of Siegel modular forms.

Siegel proved that one needs only finitely many Fourier coefficients to completely determine a Siegel modular form, so one naturally wonders how many Fourier coefficients are necessary. One could also phrase this question as, *given two Siegel modular forms, how many of their Fourier coefficients do we need to compare before we can say the two forms are equal?*.

For most of the 20th century, the best lower bounds for the number of necessary Fourier coefficients were very large. In 1990, for example, Schiemann showed that two forms were equal by showing that 375 of their Fourier coefficients were equal. The large number of Fourier coefficients that need to be tracked obviously makes computations very difficult. In 2000, Cris Poor and David S. Yuen proved the Semihull Theorem, and that theorem provided much more manageable lower bounds. For dimension 4 Siegel cusp forms (an important subclass of Siegel modular forms), for example, we need only 1 Fourier coefficient in weight 6, 2 Fourier coefficients in weight 8 and 10 Fourier coefficients in weight 10.

The Semihull Theorem is the central result of this thesis. Poor and Yuen came to realize that with a better understanding of the geometry of numbers, one can reason about Siegel modular forms much more clearly. The theorem is so named because it is phrased entirely in terms of semihulls, geometric objects. The Fourier coefficients of a Siegel modular are indexed by symmetric matrices that lie discretely in a Euclidean space. The indices of nonzero Fourier coefficients determine a semihull with some special properties. In particular, the semihull of a Siegel cusp form is a kernel, meaning that it does not approach the origin, and that in some sense it is wide. The Semihull Theorem says that if the semihull of a function is sufficiently far from the origin then the function vanishes.

The first chapter of the thesis defines the mathematical objects that will be used throughout the thesis. After the first chapter, there are two primary sections. The first section focuses on the geometry of numbers, with no mention of Siegel modular forms. The second section defines Siegel modular forms and then proves the Semihull

Theorem and two corollaries. The final chapter applies the theorems to address a mathematical question known as the Problem of Witt.

Chapter 1

Preliminaries

This chapter introduces the background/notation one needs to understand the thesis.

1.1 Lattices

Lattices will be present throughout this thesis.

Definition 1.1.1. *Lattice*
Let $\beta = \{v_1, \ldots, v_r\} \subset \mathbb{R}^n$ be any set of linearly independent column vectors. The *lattice generated by* β is the set:

$$\Lambda_\beta = \text{span}_{\mathbb{Z}}\{v_1, \ldots, v_r\}$$

Definition 1.1.2. *Gram Matrix*
Let M be the $n \times r$ matrix whose columns are $v_1, \ldots, v_r$, let M' denote its transpose and let $S = M'M$, an $r \times r$ matrix regardless of the dimension of the ambient space. We call S a Gram matrix of λ_β. The Gram matrix of M takes the form:

$$S = \begin{bmatrix} \langle v_1, v_1 \rangle & \langle v_1, v_2 \rangle & \ldots & \langle v_1, v_r \rangle \\ \langle v_2, v_1 \rangle & \langle v_2, v_2 \rangle & \ldots & \langle v_2, v_r \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle v_r, v_1 \rangle & \langle v_r, v_2 \rangle & \ldots & \langle v_r, v_r \rangle \end{bmatrix}$$

where $\langle \cdot, \cdot \rangle$ denotes inner product. The Gram matrix of a lattice is not unique, because lattices can have many ordered bases. Furthermore, a Gram matrix can correspond to 'different' (but necessarily isometric) lattices.

Definition 1.1.3. *Integral Lattice*
Let Λ be a lattice. We say that Λ is *integral* if $\langle v, \tilde{v} \rangle \in \mathbb{Z}$ for all $v \in \Lambda$, or equivalently, if all Gram matrices of Λ have integer entries.

Definition 1.1.4. *Even Lattice*
Let Λ be an integral lattice. We say that Λ is even if $\langle v, v \rangle = |v|^2 \in 2\mathbb{Z}$ for all $v, \tilde{v} \in \Lambda$. The relevant condition in terms of Gram matrices is that the Gram matrices of Λ have even numbers down the diagonal, and integers off the diagonal.

Definition 1.1.5. *Unimodular Lattice*
Let Λ be an integral lattice, and let S be a Gram matrix for Λ. We say that Λ is *unimodular* if $\det S = 1$.

1.2 Positive Definite and Semidefinite Matrices

The $n \times n$ symmetric real matrices form a Euclidean space with inner product

$$\langle S, T \rangle = \operatorname{tr}(ST)$$

that we denote $\mathcal{V}_n$. For any $S \in \mathcal{V}_n$ and any column vector $v \in \mathbb{R}^n$, introduce the notation:

$$S[v] = v'Sv$$

where v' is the transpose of v.

Definition 1.2.1. *Positive Definite and Semidefinite Matrices*
Let $S \in \mathcal{V}_n$. We say S is *positive definite* if $S[v] > 0$ for all nonzero $v \in \mathbb{R}^n$, *positive semidefinite* if $S[v] \geq 0$ for all $v \in \mathbb{R}^n$. The space of $n \times n$ positive definite matrices is denoted $\mathcal{P}_n$, and the space of $n \times n$ positive semidefinite matrices is denoted $\overline{\mathcal{P}}_n$. Here $\overline{\mathcal{P}}_n$ is so denoted because it is the topological closure of $\mathcal{P}_n$.

There are plenty of ways to characterize positive definite and semidefinite matrices.

Proposition 1.2.1. *Characterizing Properties for Positive Definite Matrices*
Let S be an $n \times n$ symmetric real matrix. The following are equivalent:

1. S is positive definite.
2. $S[v] > 0$ for all nonzero $v \in V$.
3. Every eigenvalue of S is positive.
4. The determinant of the upper left $k \times k$ corner of S is positive for $k = 1, \ldots, n$.
5. There exists a matrix M, with $\det M > 0$, such that $S = M'M$.
6. There exists a unique, positive definite matrix R such that $S = R^2$. We can refer to R as $S^{1/2}$, since R is a square root.

   Replacing every $>$ by a $\geq$ in items 2 through 5 produces the relevant characterizing conditions for semidefinite matrices.

The equivalence of these conditions is well-known. See Bhatia (2007) for more on positive definite matrices.

Let $V \in \mathrm{GL}_n(\mathbb{Z})$. Extend the notation from earlier to allow $S[V] = V'SV$. If S is a Gram matrix for some matrix M then:

$$S[V] = V'SV = V'M'MV = (MV)'(MV)$$

Since the columns for M form a basis of a lattice, and the columns of MV are integral linear combinations of the columns of M having the same total rank (since V is invertible), it follows that MV is also a basis of the same lattice, and so $S[V]$ is also positive definite. Thus, given a lattice and a Gram matrix S, we can write $[S]$ to denote $S[\mathrm{GL}_n(\mathbb{Z})]$, with the understanding that $[S]$ denotes the set of all possible Gram matrices for the lattice's isometry class.

The set $\overline{\mathcal{P}}_n$ doesn't carry a total ordering, but one can impose a partial ordering on its elements by setting $S > T$ if $S - T \in \mathcal{P}_n$.

1.3 Dyadic Matrices

Definition 1.3.1. *Dyadic Square*
Let $v \in \mathbb{R}^n$. The *dyadic square of* v is the $n \times n$ matrix vv'.

Lemma 1.3.1. *Dyadic Squares are Positive Semidefinite*

Proof. Let $z, v \in \mathbb{R}^n$. Then:

$$(zz')[v] = v'(zz')v = (v'z)(z'v) = \langle v, z\rangle^2 \geq 0$$

□

An identity that relates dyadic squares to the operation $S[\cdot]$ is:

$$S[v] = v'Sv = \mathrm{tr}\,(v'Sv) = \mathrm{tr}\,(Svv') = \langle S, vv'\rangle$$

Definition 1.3.2. *Dyadic Matrix*
Let $S \in \overline{\mathcal{P}}_n$ be a matrix. We say that S is a *dyadic matrix* if there exist finitely many $\alpha_i \in \mathbb{R}_{>0}$ and $z_i \in \mathbb{Z}^n$ such that:

$$S = \sum_i \alpha_i z_i z_i'$$

It is easy to see that every diagonal matrix D with diagonal entries $d_1, \ldots, d_n \in \mathbb{R}_{\geq 0}$ is dyadic using the formula:

$$D = \sum_{i=1}^{n} d_i e_i e_i'$$

We say S is *diagonally dominant* if the entries of S, denoted s_{ij}, satisfy:

$$s_{ii} \geq \sum_{j \neq i} |s_{ij}| \qquad for\ all\ i.$$

Every diagonally dominant S has the following dyadic representation:

$$S = \sum_i \left(s_{ii} - \sum_{j \neq i} |s_{ij}| \right) e_i e_i' + \sum_{i<j} |s_{ij}| (e_i + \mathrm{sgn}(s_{ij}) e_j)(e_i + \mathrm{sgn}(s_{ij}) e_j)'$$

For matrices in $\mathcal{P}_n(\mathbb{Q})$, we can use the *Completing the Square* algorithm (see Appendix A). The algorithm produces, for any $S \in \mathcal{P}_n(\mathbb{Q})$, an expression:

$$S = \sum_{i=1}^{n} \alpha_i z_i z_i'$$

where the $\alpha_i \in \mathbb{Q}_{\geq 0}$ and $z_i \in \mathbb{Q}^n$. By absorbing the denominators of the z_i into the α_i, we can transform the expression into a dyadic representation of S. Thus, positive matrices with rational entries are dyadic. Finally, the property of being dyadic is linear, i.e. if S, T are dyadic and $c \in \mathbb{R}_{>0}$, then $S + T$ and cS are also dyadic.

Proposition 1.3.1. *Positive Matrices are Dyadic*
Let $S \in \mathcal{P}_n$ be any real-valued, positive matrix. Then S is dyadic.

Proof. We show that any $S \in \mathcal{P}_n$ can be broken down as $S = R+T$, where $R \in \mathcal{P}_n(\mathbb{Q})$ and T is a diagonally dominant matrix.

Let $\lambda \in \mathbb{R}_{>0}$ be such that the matrix $S - \lambda I \in \mathcal{P}_n$. Let $R \in \mathcal{P}_n(\mathbb{Q})$ be a matrix whose entries differ from those of $S-\lambda I$ by at most λ/n. Such a matrix exists because $\mathcal{P}_n(\mathbb{Q})$ is a dense subset of $\mathcal{P}_n$.

Let $T = S - R = (S - \lambda I - R) + \lambda I$. Then the-off diagonal entries of T and the off-diagonal entries of $S - \lambda I - R$ are equal, and the entries of the second matrix are smaller than λ/n, so the sum of the off-diagonal entries along any row of T is at most $\lambda(n-1)/n$. Furthermore, the diagonal entries of T are at least $\lambda - \lambda/n = \lambda(n-1)/n$, so T is diagonally dominant, and thus dyadic. Thus, S is dyadic. □

1.4 Geometry

Throughout this section, $\mathcal{V}$ denotes a Euclidean space.

Definition 1.4.1. *Cone*
A subset $\mathcal{C} \subset \mathcal{V}$ is called a *cone* if $\mathcal{C}$ is closed under addition and $\mathbb{R}_{>0}$-dilations. Equivalently, $\mathcal{C}$ is a cone if it is convex and closed under $\mathbb{R}_{>0}$-dilations.

Definition 1.4.2. *Cone of a Set*
Let $\mathcal{S} \subset \mathcal{V}$. The *cone generated by* $\mathcal{S}$ is:

$$\langle \mathbb{R}_{>0}\mathcal{S} \rangle = \text{span}_{\mathbb{R}_{>0}}\mathcal{S} = \left\{ \sum \lambda_i v_i \ : \ v_i \in \mathcal{S},\ \lambda_i \in \mathbb{R}_{>0} \right\}$$

Definition 1.4.3. *Dual Cone of a Set*
Let $\mathcal{S} \subset \mathcal{V}$. The *dual cone of* $\mathcal{S}$ is:

$$\mathcal{S}^\vee = \{v \in \mathcal{V} : \langle v, \mathcal{S} \rangle \subset \mathbb{R}_{\geq 0}\}$$

It is clear that the dual cone of any set is a cone, because for any $v, w \in \mathcal{S}^\vee$, $\lambda \in \mathbb{R}_{>0}$, $s \in \mathcal{S}$:

$$\langle \lambda v, s \rangle = \lambda \langle v, s \rangle \geq 0$$
$$\langle v + w, s \rangle = \langle v, s \rangle + \langle w, s \rangle \geq 0$$

Proposition 1.4.1. *Dual of $\mathcal{P}_n$*
Let $T \in \mathcal{P}_n^\vee$. Then $T \in \overline{\mathcal{P}}_n$.

Proof. Let $v \in \mathbb{R}^n$. For any $S \in \mathcal{P}_n$ and $\epsilon \in \mathbb{R}_{>0}$, $vv' + \epsilon S \in \mathcal{P}_n$, so:

$$\langle T, vv' + \epsilon S \rangle = \langle T, vv' \rangle + \epsilon \langle T, S \rangle = T[v] + \epsilon \langle T, S \rangle \geq 0$$

Let $\epsilon \to 0$ to see that $T[v] \geq 0$.

□

Definition 1.4.4. *Semihull*
A subset $\mathcal{M} \in \mathcal{V}$ is called a *semihull* if $\mathcal{M}$ is convex and closed under $\mathbb{R}_{\geq 1}$-dilations. Equivalently, we can say that $\mathcal{M}$ is closed under superconvex combinations.

Definition 1.4.5. *Semihull of a Set*
The semihull of a set $\mathcal{S} \subset \mathcal{V}$ is:

$$\langle \mathbb{R}_{\geq 1} \mathcal{S} \rangle = \left\{ \sum_i \lambda_i v_i \ : \ v_i \in \mathcal{S}, \ \lambda_i \in \mathbb{R}_{\geq 0}, \sum_i \lambda_i \geq 1 \right\}$$

Definition 1.4.6. *Dual Semihull of a Set*
The dual semihull of a set $\mathcal{S} \subset \mathcal{V}$ is:

$$\mathcal{S}^\sqcup = \{ v \in \mathcal{V} : \langle v, \mathcal{S} \rangle \subset \mathbb{R}_{\geq 1} \}$$

We can show that the dual semihull of a set is a semihull: if $v \in \mathcal{S}^\sqcup$, then for any $\lambda \in \mathbb{R}_{\geq 1}$ and any $s \in \mathcal{S}$, $\langle \lambda v, s \rangle = \lambda \langle v, s \rangle \geq \lambda \geq 1$. Furthermore, for any $w \in \mathcal{S}^\sqcup$ and for any $\lambda \in (0, 1)$:

$$\langle \lambda v + (1 - \lambda) w, s \rangle = \lambda \langle v, s \rangle + (1 - \lambda) \langle w, s \rangle \geq \lambda + 1 - \lambda = 1$$

which shows that $\lambda v + (1 - \lambda) w \in \mathcal{S}^\sqcup$, so $\mathcal{S}^\sqcup$ is convex, and thus a semihull.

The dual cone and the dual semihull of any set are necessarily closed by continuity of the inner product. Also, for any $A, B \subset \mathcal{V}$:

$$A \subset B \iff B^\vee \subset A^\vee$$

and

$$A \subset B \iff B^\sqcup \subset A^\sqcup$$

Proposition 1.4.2. *Semihull Decomposition*
Let $\mathcal{M} \subset \mathcal{V}$ be a nonempty closed semihull. Let $v \in \mathcal{V}$ be arbitrary, and let $c \in \mathcal{M}$ be the closest point in $\mathcal{M}$ to v. Such a point necessarily exists because $\mathcal{M}$ is nonempty and closed, and that point is unique because $\mathcal{M}$ is convex. Then:

$$c - v \in \langle \mathbb{R}_{>0} (\mathcal{M} - c) \rangle^\vee \qquad \text{and} \qquad \langle c, c - v \rangle \geq 0$$

Proof. Let c be the closest point in $\mathcal{M}$ to v. For all other $\tilde{m} \in \mathcal{M}$:

$$|\tilde{m} - v| \geq |c - v|$$

Let $d = v - c$. We need to show $d \in \langle \mathbb{R}_{>0}(\mathcal{M} - c)\rangle^\vee$, i.e.:

$$\langle v - c, \langle \mathbb{R}_{>0}(\mathcal{M} - c)\rangle\rangle \subset \mathbb{R}_{>0}.$$

By linearity of the inner product, it suffices to show:

$$\langle v - c, (\mathcal{M} - c)\rangle \subset \mathbb{R}_{>0}$$

Let $m \in \mathcal{M}$ be arbitrary. For any $r \in [0, 1]$, we have:

$$rm + (1 - r)c = r(m - c) + c \in \mathcal{M}$$

because $\mathcal{M}$ is convex. Thus, letting $\tilde{m} = r(m - c) + c$ gives us the following display:

$$0 \leq |r(m - c) + c - v|^2 - |c - v|^2 = |r(m - c) + d|^2 - |d|^2$$

We can rewrite the display using inner products:

$$0 \leq \langle r(m - c) + d, r(m - c) + d\rangle - \langle d, d\rangle = r^2 \langle m - c, m - c\rangle + 2r \langle m - c, d\rangle + \langle d, d\rangle - \langle d, d\rangle$$

$$\therefore 0 \leq r^2|m - c|^2 + 2r \langle m - c, d\rangle$$

If $\langle m - c, d\rangle < 0$, then for $r < \langle c - m, d\rangle / |m - c|^2$, the inequality in the display above doesn't hold. Thus, $\langle m - c, d\rangle \geq 0$, so $d \in \langle \mathbb{R}_{>0}\mathcal{M} - c\rangle^\vee$ as claimed.

Next, we show that $\langle c, d\rangle \geq 0$. By the same reasoning as earlier, we know that for any $r > 0$, $(r + 1)c \in \mathcal{M}$ by closure under $\mathbb{R}_{\geq 1}$-dilations. Thus, setting $\tilde{m} = (r + 1)c$ yields the following display:

$$\langle (r + 1)c - v, (r + 1)c - v\rangle - \langle c - v, c - v\rangle \geq 0$$

Expanding out the inner products and then simplifying the expression gives us:

$$0 \leq r^2|c|^2 + 2r \langle c, d\rangle$$

Once again, letting $r \to 0^+$ shows that $\langle c, d\rangle \geq 0$, because for very small r, we have $\operatorname{sgn}(r^2|c|^2 + 2r \langle c, d\rangle) = \operatorname{sgn}(\langle c, d\rangle)$. □

Proposition 1.4.3. *Empty Dual Semihull*
Let $\mathcal{M}$ be a closed semihull. Then:

$$\mathcal{M}^\sqcup \neq \emptyset \iff 0 \notin \mathcal{M}$$

Proof. First, I'll take care of some easy cases. Suppose $0 \in \mathcal{M}$. For all $v \in \mathcal{V}$, $\langle v, 0\rangle = 0 < 1$, so $v \notin \mathcal{M}^\sqcup$. Thus, $\mathcal{M}^\sqcup = \emptyset$.

Next, suppose $0 \notin \mathcal{M}$. If $\mathcal{M} = \emptyset$, then $\mathcal{M}^\sqcup = \mathcal{V}$, so again the proposition holds. Otherwise, $\mathcal{M} \neq \emptyset$. By semihull decomposition, we know that we can write $0 = c - d$,

where $c \in \mathcal{M}$ is the closest point to 0, and $d \in \langle \mathbb{R}_{>0}\mathcal{M} - c\rangle^\vee$. Obviously, $c = d$ for this particular choice of v, so we also have $c \in \langle \mathbb{R}_{>0}\mathcal{M} - c\rangle^\vee$. Let $\tilde{c} = c/|c|$; note that $\tilde{c} \in \langle \mathbb{R}_{>0}\mathcal{M} - c\rangle^\vee$. For any $m \in \mathcal{M}$:

$$\langle m, \tilde{c}\rangle = \langle m - \tilde{c}, \tilde{c}\rangle + \langle \tilde{c}, \tilde{c}\rangle = \langle m - \tilde{c}, \tilde{c}\rangle + |\tilde{c}|^2$$

We know $|\tilde{c}| = 1$, and because $\tilde{c} \in \langle \mathbb{R}_{>0}\mathcal{M} - c\rangle^\vee$, $\langle m - \tilde{c}, \tilde{c}\rangle > 0$. Thus:

$$\langle m, \tilde{c}\rangle > |\tilde{c}| = 1 \Longrightarrow \tilde{c} \in \mathcal{M}^\sqcup$$

so $\mathcal{M}^\sqcup \neq \emptyset$. □

Proposition 1.4.4. *Duality*
Let $\mathcal{M} \subset \mathcal{V}$ be a nonempty semihull whose closure doesn't contain 0. Then:

$$\mathcal{M}^{\sqcup\sqcup} = \overline{\mathcal{M}}$$

Proof. First, observe that $\overline{\mathcal{M}} \subset \mathcal{M}^{\sqcup\sqcup}$ because $\langle \mathcal{M}, \mathcal{M}^\sqcup\rangle \subset \mathbb{R}_{\geq 1}$ and because the inner product is continuous. Thus we only need to show $\mathcal{M}^{\sqcup\sqcup} \subset \overline{\mathcal{M}}$.

Let $z \in \mathcal{M}^\sqcup$. Such a z exists because $0 \notin \overline{\mathcal{M}}$. Fix $v \in \mathcal{M}^{\sqcup\sqcup}$. By Prop. 1.4.1., there exists $c \in \overline{\mathcal{M}}$ and $d \in \left\langle \mathbb{R}_{>0}\overline{\mathcal{M}} - c\right\rangle^\vee$ such that $v = c - d$, $\langle c, d\rangle \geq 0$. To complete the argument, we can show that $d = 0$.

For all $m \in \mathcal{M}$, we know that $\langle m - c, d\rangle \in \mathbb{R}_{\geq 0}$ because $d \in \langle \mathbb{R}_{>0}\mathcal{M} - c\rangle^\vee$, and we know $\langle m, z\rangle \geq 1$ because $z \in \mathcal{M}^\sqcup$, so we have:

$$\langle m - c, d\rangle \geq 0 \Longrightarrow \langle m, d\rangle \geq \langle c, d\rangle \geq 0.$$

Thus, for all $r \in \mathbb{R}_{>0}$,

$$\langle m, z\rangle \geq 1 \Longrightarrow r\langle m, z\rangle = \langle m, rz\rangle \geq r \Longrightarrow \langle m, rz\rangle + \langle c, d\rangle \geq \langle c, d\rangle + r$$

which allows us to deduce that:

$$\langle m, rz + d\rangle = \langle m, rz\rangle + \langle m, d\rangle \geq \langle m, rz\rangle + \langle c, d\rangle \geq \langle c, d\rangle + r.$$

The results of the display hold for all $m \in \mathcal{M}$, $r \in \mathbb{R}_{>0}$ so we can rewrite it:

$$(\forall r \in \mathbb{R}_{>0}) \quad \langle \mathcal{M}, rz + d\rangle \geq \langle c, d\rangle + r.$$

This is equivalent to:

$$\frac{rz + d}{\langle c, d\rangle + r} \in \mathcal{M}^\sqcup$$

Now, since $v \in \mathcal{M}^{\sqcup\sqcup}$, we have:

$$(\forall r \in \mathbb{R}_{>0}) \quad \left\langle v, \frac{rz + d}{\langle c, d\rangle + r}\right\rangle \geq 1 \quad so \quad \langle v, rz + d\rangle \geq \langle c, d\rangle + r$$

If we let $r \to 0$, we can use the continuity of the inner product to deduce that:

$$\langle v, d\rangle \geq \langle c, d\rangle \Longrightarrow 0 \geq \langle c, d\rangle - \langle v, d\rangle = \langle c - v, d\rangle = \langle d, d\rangle = |d|^2$$

Thus, $d = 0$, so $v \in \overline{\mathcal{M}}$ as claimed. □

Chapter 2

Height Functions and Kernels

2.1 Definitions and Basic Properties

2.1.1 Height Functions

We want to be able to order the positive matrices, so we introduce an auxiliary function called a *height function* that assigns a "height" to positive matrices.

Definition 2.1.1. *Height Function*
Let $\mathcal{C}$ be a cone of positive semidefinite matrices that contains the cone of positive matrices, i.e. let $\mathcal{P}_n \subseteq \mathcal{C} \subseteq \overline{\mathcal{P}}_n$.

A function $\phi : \mathcal{C} \longrightarrow \mathbb{R}_{\geq 0}$ is called a *height function* if it has the following properties:

1. For any $S \in \mathcal{P}_n$, $\phi(S) \in \mathbb{R}_{>0}$.

2. For any $\lambda \in \mathbb{R}_{>0}$ and $S \in \mathcal{C}$, $\phi(\lambda S) = \lambda\phi(S)$.

3. For all $S, T \in \mathcal{C}$, $\phi(S) + \phi(T) \leq \phi(S + T)$.

If a height-function has the property $\phi([S]) = \phi(S)$ for all $S \in \text{dom } \phi$, we say that ϕ is a class function. For reasons that will be clear later, we will not distinguish between height functions that agree on $\mathcal{P}_n$.

Lemma 2.1.1. *Height functions preserve order*
Let $\phi : \mathcal{C} \to \mathbb{R}_{\geq 0}$ be a height function, and let $S, T \in \mathcal{C}$ be such that $S < T$. Then $\phi(S) < \phi(T)$.

Proof. Since $S < T$, there exists $U \in \mathcal{P}_n$ such that $S + U = T$. By superadditivity:

$$\phi(T) = \phi(S + U) \geq \phi(S) + \phi(U) > \phi(S)$$

□

Proposition 2.1.1. *Height functions are continuous on $\mathcal{P}_n$.*

Proof. Let $\epsilon > 0$ be given, and let $S \in \mathcal{P}_n$ be arbitrary. Take $q \in (0,1)$ such that $q\phi(S) < \epsilon$. Define a neighborhood of S as follows:

$$N_S = \{T \in \mathcal{P}_n : (1-q)S < T < (1+q)S\}.$$

Now, we have that:

$$\epsilon > q\phi(S) \Longrightarrow -\epsilon < -q\phi(S) \Longrightarrow \phi(S) - \epsilon < (1-q)\phi(S)$$

and for all $T \in N_S$, we have:

$$\phi(S) - \epsilon < (1-q)\phi(S) + \phi(T - (1-q)S)$$

since $T > (1-q)S$ means $T - (1-q)S \in \mathcal{P}_n$. By superadditivity and homogeneity, we have:

$$(1-q)\phi(S) + \phi(T - (1-q)S \leq \phi(T)$$

so we've shown $\phi(S) - \epsilon < \phi(T)$.

Since $(1+q)S > T$, $(1+q)S - T \in \mathcal{P}_n$, so $\phi((1+q)S - T) > 0$, which gives us:

$$\phi(T) < \phi(T) + \phi((1+q)S - T).$$

By superadditivity,

$$\phi(T) + \phi((1+q)S - T) \leq \phi((1+q)S) = (1+q)\phi(S) < \phi(S) + \epsilon$$

In sum, we've shown that $\phi(S) - \epsilon < \phi(T) < \phi(S) + \epsilon$, so $|\phi(T) - \phi(S)| < \epsilon$ for all $T \in N_S$, so ϕ is continuous on $\mathcal{P}_n$. □

2.1.2 Examples of Height Functions

There are many famous height functions that arise in linear algebra.

For example, one can easily verify that the trace is a height function. To see that the trace is nonnegative on $\overline{\mathcal{P}}_n$, it suffices to observe that $s_{ii} = S[e_i] \geq 0$ by positive semidefiniteness. Thus, the trace is a sum of nonnegative real numbers. Homogeneity and superadditivity follow from the linearity of the trace.

The trace isn't a class function, but it has other interesting properties. Specifically, a matrix $S \in \overline{\mathcal{P}}_n$ is 0 if and only if $\operatorname{tr}(S) = 0$. Furthermore, for any $c \in \mathbb{R}_{>0}$, the set:

$$\{S \in \overline{\mathcal{P}}_n : \operatorname{tr}(S) \leq c\}$$

is compact.

Proposition 2.1.2. *Dilational Dominance of the Trace*
Let $\phi : \mathcal{C} \to \mathbb{R}_{\geq 0}$ be a height function. There exists a ϕ-dependent constant c with the property that:

$$\phi(S) \leq c\operatorname{tr}(S) \quad \textit{for all } S \in \mathcal{C}$$

Proof. Start by defining a compact set:

$$K = \{T \in \overline{\mathcal{P}}_n : \mathrm{tr}\,(T) \leq 1\}$$

Let N be the translation of K by I_n. It is clear that N is also compact. Furthermore, one can easily very that $N \subset \mathcal{P}_n$, because for any $T \in K$ and any $v \in \mathbb{Z}^n$:

$$(T + I_n)[v] = T[v] + I_n[v] \geq I_n[v] > 0$$

Thus, N is a compact subset of $\mathcal{P}_n$. Since ϕ is continuous on $\mathcal{P}_n$, it must be bounded on $\mathcal{N}$ by some constant c, so $\phi(N) \subset [0, c]$.

Let $S \in \mathcal{C}$ be nonzero, and let $\tilde{S} = S/\mathrm{tr}\,(S)$. Clearly, $\tilde{S} \in K$ because $\mathrm{tr}\,(\tilde{S}) = 1$. Thus,

$$\phi(\tilde{S}) < \phi(\tilde{S}) + \phi(I_n) \leq \phi(\tilde{S} + I_n) \leq c$$

and by homogeneity,

$$\phi(S) = \mathrm{tr}\,(S)\phi(\tilde{S}) \leq c\mathrm{tr}\,(S)$$

as desired. □

Three other height functions that will be useful later on are the *least eigenvalue*, denoted λ_1, and the *reduced determinant*, denoted δ:

$$\lambda_1(S) = \min\{\text{eigenvalues of } S\} \quad \delta(S) = \sqrt[n]{\det S} \quad \rho(S) = \min_{V \in \mathrm{GL}_n(\mathbb{Z})} \mathrm{tr}\,(S[V])$$

Both of these functions are defined on all of $\overline{\mathcal{P}}_n$, but vanish off $\mathcal{P}_n$. Furthermore, both of them are class functions.

2.1.3 Kernels

Let $\mathcal{K} \subset \overline{\mathcal{P}}_n$ be a semihull. We say that $\mathcal{K}$ is a *kernel* if $\mathcal{K}$ satisfies the following:

1. $\mathcal{P}_n \subset \mathbb{R}_{>0}\mathcal{K}$
2. $0 \notin \overline{\mathcal{K}}$

It should be obvious that a semihull $\mathcal{K}$ is a kernel if and only if $\overline{\mathcal{K}}$ is a kernel.

Lemma 2.1.2. *Kernels Expand in all Directions*
Let $\mathcal{K}$ be a kernel, $S \in \mathcal{K}$. Then $S + \mathcal{P}_n \subset \mathcal{K}$.

Proof. Let $U \in \mathcal{P}_n$, and set $T = S + U$. Since $\mathcal{P}_n \subset \mathbb{R}_{>0}\mathcal{K}$, there exists $r \in \mathbb{R}_{>0}$ for which $rU \in \mathcal{K}$. A quick computation shows:

$$\frac{r}{r+1}T = \frac{r}{r+1}S + \frac{r}{r+1}U = \frac{r}{r+1}S + \frac{1}{r+1}rU$$

Thus, $r/(r+1)T \in \mathcal{K}$ because $\mathcal{K}$ is convex, so $T \in \mathcal{K}$ because $\mathcal{K}$ is closed under $\mathbb{R}_{\geq 1}$-dilations. □

Proposition 2.1.3. *Persistence of Closure*
For any kernel $\mathcal{K}$, we have:

$$\overline{\mathcal{K}} = \overline{\mathcal{K} \cap \mathcal{P}_n} = \overline{\mathcal{K} \cap \mathcal{P}_n(\mathbb{Q})}$$

Proof. It is obvious that $\overline{\mathcal{K}} \supset \overline{\mathcal{K} \cap \mathcal{P}_n} \supset \overline{\mathcal{K} \cap \mathcal{P}_n(\mathbb{Q})}$, so it suffices to show that $\overline{K} \subset \overline{K \cap \mathcal{P}_n(\mathbb{Q})}$.

Recall that $\mathcal{P}_n$ is an open space in $\mathcal{V}_n$, and $I_n \in \mathcal{P}_n$, so there exists a neighborhood N of I that is entirely contained in $\mathcal{P}_n$.

Now, let $S \in \overline{K}$. Let $\{S_i\} \subset \mathcal{K}$ be a sequence that approaches S. By the preceding lemma, the set of neighborhoods $\{S_i + \frac{1}{i}N\}$ must also lie in $\mathcal{K}$ because the neighborhoods contain only positive matrices. Furthermore, each of those neighborhoods contains matrices in $\mathcal{P}_n(\mathbb{Q})$. Thus, we can define a sequence $\{\tilde{S}_i\}$, with each $\tilde{S}_i \in ((S_i + \frac{1}{i}N) \cap \mathcal{P}_n(\mathbb{Q}))$, that lies in $\mathcal{K}$ and approaches S. Thus, $S \in \overline{\mathcal{K} \cap \mathcal{P}_n(\mathbb{Q})}$. □

In general, we will not make a distinction between kernels that have the same closure.

Lemma 2.1.3. *Containment*
Let $\mathcal{K}$ be a closed kernel, and let $\mathcal{S} \subset \mathcal{P}_n$. Suppose $\mathcal{K}^{\sqcup} \cap \mathcal{P}_n(\mathbb{Q}) \subset \mathcal{S}^{\sqcup}$. Then $\mathcal{S} \subset \mathcal{K}$.

Proof. We have:

$$\mathcal{K}^{\sqcup} \cap \mathcal{P}_n(\mathbb{Q}) \subset \mathcal{S}^{\sqcup}$$

Since $\mathcal{S}^{\sqcup}$ is closed, we can take the closure of the left hand side of the display and preserve the containment:

$$\overline{\mathcal{K}^{\sqcup} \cap \mathcal{P}_n(\mathbb{Q})} \subset \mathcal{S}^{\sqcup}$$

By Prop. 2.1.3, and the fact that duals are always closed, we have $\overline{\mathcal{K}^{\sqcup} \cap \mathcal{P}_n(\mathbb{Q})} = \mathcal{K}^{\sqcup}$, so:

$$\mathcal{K}^{\sqcup} \subset \mathcal{S}^{\sqcup} \Longrightarrow \mathcal{S}^{\sqcup\sqcup} \subset \mathcal{K}^{\sqcup\sqcup}$$

But $\mathcal{K}^{\sqcup\sqcup} = \overline{\mathcal{K}} = \mathcal{K}$, and $\mathcal{S} \subset \mathcal{S}^{\sqcup\sqcup} = \overline{\langle \mathbb{R}_{\geq 1}\mathcal{S} \rangle}$, so:

$$\mathcal{S} \subset \mathcal{K}.$$

□

2.2 Height Function-Kernel Correspondence

Definition 2.2.1. *Kernel of a Height Function*
Let $\phi : \mathcal{C} \to \mathbb{R}_{\geq 0}$ be a height function. The *kernel of with* ϕ, denoted $\mathcal{K}_\phi$, is $\phi^{-1}(\mathbb{R}_{\geq 1})$.

Proposition 2.2.1. *Kernel of a Height Function*
Let $\phi : \mathcal{C} \to \mathbb{R}_{\geq 0}$ be a height function. Then $\mathcal{K}_\phi$ is a kernel.

Proof. It is clear that $\mathcal{K}_\phi \subset \overline{\mathcal{P}}_n$, because $\mathcal{C} \subset \overline{\mathcal{P}}_n$. If $S \in \mathcal{K}_\phi$, and $r \in \mathbb{R}_{\geq 1}$, then $\phi(rS) = r\phi(S) \geq r \geq 1$, so $\mathcal{K}_\phi$ is closed under $\mathbb{R}_{\geq 1}$-dilations. Furthermore, if $S, T \in \mathcal{K}_\phi$, and $0 < \lambda < 1$, then by superadditivity and then homogeneity:

$$\phi(\lambda S + (1-\lambda)T) \geq \phi(\lambda S) + \phi((1-\lambda)T) = \lambda\phi(S) + (1-\lambda)\phi(T)$$

Since $\phi(S), \phi(T) \geq 1$, we get:

$$\phi(\lambda S + (1-\lambda)T) \geq \lambda + (1-\lambda) = 1$$

which shows that $\mathcal{K}_\phi$ is convex. Thus, $\mathcal{K}_\phi$ is a semihull in $\overline{\mathcal{P}}_n$.

Now we can check the conditions in the definition of a kernels. For any $S \in \mathcal{P}_n$, $\phi(S) = c > 0$, so $\phi(c^{-1}S) = 1$ by homogeneity, which shows that $c^{-1}S \in \mathcal{K}_\phi$. Thus, $\mathcal{P}_n \subset \mathbb{R}_{>0}\mathcal{K}_\phi$.

The final thing to check is that $0 \notin \overline{K}$. If $\{S_i\} \subset \mathcal{C}$ approaches 0, then $\{\mathrm{tr}\,(S_i)\}$ must also approach 0. By the dilational dominance of the trace, that means $\{\phi(S_i)\}$ must approach 0, so the sequence is not entirely contained in $\mathcal{K}_\phi$. Thus, 0 is not a limit point of $\mathcal{K}_\phi$, so $\mathcal{K}_\phi$ is a kernel. □

Definition 2.2.2. *Height Function of a Kernel*
Let $\mathcal{K}$ be a kernel. The height function associated with $\mathcal{K}$, which depends only on the closure of $\mathcal{K}$, is:

$$\phi_{\mathcal{K}}(S) = \frac{1}{\inf\{r \in \mathbb{R}_{>0} : rS \in \mathcal{K}\}} \qquad (S \in \mathcal{P}_n)$$

Since $\mathcal{K}$ is a kernel, $0 \notin \overline{\mathcal{K}}$, and since $\mathcal{P}_n \subset \mathbb{R}_{>0}\mathcal{K}$, for any $S \in \mathcal{P}_n$, the set

$$\{r \in \mathbb{R}_{>0} : rS \in \mathcal{K}\}$$

must have positive lower bound. Thus, there is no risk of division by zero.

We can also characterize the height function of a kernel using one of the following equivalent properties:

$$\mathbb{R}_{>0}S \cap \phi_{\mathcal{K}}(S)\overline{\mathcal{K}} = \mathbb{R}_{\geq 1}S \quad \textit{or} \quad \mathbb{R}_{>0}S \cap \overline{\mathcal{K}} = \mathbb{R}_{\geq 1/\phi_{\mathcal{K}}(S)}S$$

Proposition 2.2.2. *Height Function of a Kernel.*
Let $\mathcal{K}$ be a kernel. Then $\phi_{\mathcal{K}}$ is a height function.

Proof. It is clear that the domain of $\phi_{\mathcal{K}}$ is simply $\mathbb{R}_{>0}\mathcal{K}$, a cone of positive semidefinite matrices that contains the cone of positive matrices. Furthermore, the definition of $\phi_{\mathcal{K}}$ makes it clear that $\phi_{\mathcal{K}}$ is positive on $\mathcal{P}_n$.

Homogeneity follows from the definition:

$$\phi_{\mathcal{K}}(\lambda S) = \frac{1}{\inf\{r \in \mathbb{R}_{>0} : r(\lambda S) \in \overline{\mathcal{K}}\}} = \frac{1}{\lambda^{-1}\inf\{r \in \mathbb{R}_{>0} : rS \in \overline{\mathcal{K}}\}}$$

$$\implies \phi_{\mathcal{K}}(\lambda S) = \lambda \cdot \frac{1}{\inf\{r \in \mathbb{R}_{>0} : rS \in \overline{\mathcal{K}}\}} = \lambda\phi_{\mathcal{K}}(S)$$

In order to show superadditivity holds, first observe that for any $S \in \mathbb{R}_{>0}\mathcal{K}$:

$$\frac{S}{\phi_{\mathcal{K}}(S)} = S \cdot \inf\{r \in \mathbb{R}_{>0} : rS \in \overline{\mathcal{K}}\} \Longrightarrow \frac{S}{\phi_{\mathcal{K}}(S)} \in \overline{\mathcal{K}}$$

Now let $S, T \in \mathbb{R}_{>0}\overline{\mathcal{K}}$, and check that:

$$S + T = (\phi_{\mathcal{K}}(S) + \phi_{\mathcal{K}}(T)) \cdot \left(\frac{\phi_{\mathcal{K}}(S)}{\phi_{\mathcal{K}}(S) + \phi_{\mathcal{K}}(T)} \frac{S}{\phi(S)} + \frac{\phi_{\mathcal{K}}(T)}{\phi_{\mathcal{K}}(S) + \phi_{\mathcal{K}}(T)} \frac{T}{\phi(T)} \right)$$

Since $\frac{S}{\phi_{\mathcal{K}}(S)}, \frac{T}{\phi_{\mathcal{K}}(T)} \in \overline{\mathcal{K}}$ and $\overline{\mathcal{K}}$ is convex, it follows that:

$$\left(\frac{\phi_{\mathcal{K}}(S)}{\phi_{\mathcal{K}}(S) + \phi_{\mathcal{K}}(T)} \frac{S}{\phi(S)} + \frac{\phi_{\mathcal{K}}(T)}{\phi_{\mathcal{K}}(S) + \phi_{\mathcal{K}}(T)} \frac{T}{\phi(T)} \right) \in \overline{\mathcal{K}}$$

so

$$S + T \in (\phi_{\mathcal{K}}(S) + \phi_{\mathcal{K}}(T))\overline{\mathcal{K}}.$$

□

Proposition 2.2.3. *Correspondence of Height Functions and Kernels*
Recall that kernels are equivalent if they have the same closure, and height functions are equivalent if they agree on $\mathcal{P}_n$.

For any height function ϕ:

$$\phi_{\mathcal{K}_\phi}|_{\mathcal{P}_n} = \phi|_{\mathcal{P}_n}$$

and for any kernel $\mathcal{K}$:

$$\mathcal{K}_{\phi_{\mathcal{K}}} = \overline{\mathcal{K}}$$

Proof. Let ϕ be given. Then for any $S \in \mathcal{P}_n$, we can use the definition of the height function of a kernel to show:

$$\phi_{\mathcal{K}_\phi}(S) = \frac{1}{\inf\{r \in \mathbb{R}_{>0} : rS \in \mathcal{K}_\phi\}}$$

Since $\mathcal{K}_\phi = \phi^{-1}(\mathbb{R}_{\geq 1})$, it is clear that $\inf\{r \in \mathbb{R}_{>0} : rS \in \mathcal{K}_\phi\} = \frac{1}{\phi(S)}$, so:

$$\phi_{\mathcal{K}_\phi}(S) = \frac{1}{\frac{1}{\phi(S)}} = \phi(S)$$

Now let a kernel $\mathcal{K}$ be given, and suppose $S \in \mathcal{K}_{\phi_{\mathcal{K}}}$. By definition of kernel of a height function, we have $\phi_{\mathcal{K}}(S) \geq 1$. By the characterizing property for kernels of height functions:

$$\mathbb{R}_{>0}S \cap \overline{\mathcal{K}} = \mathbb{R}_{\geq 1/\phi_{\mathcal{K}}(S)}S$$

Since $\phi_{\mathcal{K}}(S) \geq 1$, $\frac{1}{\phi_{\mathcal{K}}(S)} \leq 1$, so $S \in \mathbb{R}_{\geq 1/\phi_{\mathcal{K}}(S)}S \subset \overline{K}$. □

2.3 Duality

2.3.1 Dual of a Height Function

Definition 2.3.1. *Dual of a Height Function*
Let $\phi : \mathcal{C} \longrightarrow \mathbb{R}_{\geq 0}$ be a height function. The dual height function of ϕ is:

$$\hat{\phi}(S) = \inf_{T \in \mathcal{P}_n} \frac{\langle S, T \rangle}{\phi(T)}$$

For example, the dual of the trace is the least eigenvalue. Furthermore, the *scaled* reduced determinant $\tilde{\delta} = \sqrt{n}\delta$ is its own dual, i.e. for all $S \in \mathcal{P}_n$:

$$\tilde{\delta}(S) = \inf_{T \in \mathcal{P}_n} \frac{\langle S, T \rangle}{\tilde{\delta}(T)}$$

See Poor et al. for proofs of those claims, and more on height function duality.

Proposition 2.3.1. *Dual of a Height Function*
Let $\phi : \mathcal{C} \longrightarrow \mathbb{R}_{\geq 0}$. The dual of ϕ is a height function.

Proof. First, $\hat{\phi}$ is defined on all of $\overline{\mathcal{P}}_n$, so it has the type of domain specified in the definition. Thus, we can simply check that the definition holds.

Let $S \in \mathcal{P}_n$ be nonzero, arbitrary. By definition:

$$\hat{\phi}(S) = \inf_{T \in \mathcal{P}_n} \frac{\langle S, T \rangle}{\phi(T)}.$$

Since $T \in \mathcal{P}_n$, $\phi(T) > 0$ because ϕ is a height function. By duality of tr and λ_1, $0 < \operatorname{tr}(S)\lambda_1(T) \leq \langle S, T \rangle$, so it follows that $\hat{\phi}$ is positive on $\mathcal{P}_n$. Furthermore, for any $S \in \overline{\mathcal{P}}_n$, $\lambda \in \mathbb{R}_{>0}$:

$$\hat{\phi}(\lambda S) = \inf_{T \in \mathcal{P}_n} \frac{\langle \lambda S, T \rangle}{\phi(T)} = \inf_{T \in \mathcal{P}_n} \frac{\lambda \langle S, T \rangle}{\phi(T)} = \lambda \inf_{T \in \mathcal{P}_n} \frac{\langle S, T \rangle}{\phi(T)} = \lambda \hat{\phi}(S)$$

Finally, for any $S, U \in \overline{\mathcal{P}}_n$:

$$\hat{\phi}(S + U) = \inf_{T \in \mathcal{P}_n} \frac{\langle S + U, T \rangle}{\phi(T)} = \inf_{T \in \mathcal{P}_n} \left(\frac{\langle S, T \rangle}{\phi(T)} + \frac{\langle U, T \rangle}{\phi(T)} \right)$$

Taking the infimum for both fractions simultaneously is less effective than taking the infimum for each one separately, because the T that minimizes the expression for S might be different from the one that minimizes the expression for U. Thus, the infimum in the display above might not be as small as the one obtained by taking the infima separately, i.e.:

$$\implies \hat{\phi}(S + U) \geq \inf_{T \in \mathcal{P}_n} \frac{\langle S, T \rangle}{\phi(T)} + \hat{\phi}(S) \inf_{T \in \mathcal{P}_n} \frac{\langle U, T \rangle}{\phi(T)} = \hat{\phi}(S) + \hat{\phi}(U)$$

This proves superadditivity, so we're done. □

Proposition 2.3.2. *Alternative characterizations of the dual height function* Let $\phi : \mathcal{C} \longrightarrow \mathbb{R}_{\geq 0}$. The dual of ϕ is also given by the following formulas:

$$\hat{\phi(S)} = \begin{Bmatrix} \inf \langle S, \mathcal{P}_n \cap \phi^{-1}(1) \rangle \\ \inf \langle S, \mathcal{K}_\phi \rangle \end{Bmatrix}, \quad S \in \overline{\mathcal{P}}_n$$

Proof. We start by proving the equivalence of the original definition and the first formula:

$$\hat{\phi}(S) = \inf_{T \in \mathcal{P}_n} \frac{\langle S, T \rangle}{\phi(T)} = \inf_{T \in \mathcal{P}_n} \left\langle S, \frac{T}{\phi(T)} \right\rangle = \inf \left\langle S, \mathcal{P}_n \cap \phi^{-1}(T) \right\rangle$$

Next, observe that:

$$\mathcal{P}_n \cap \phi^{-1}(1) \subset \phi^{-1}(1) \subset \phi^{-1}(\mathbb{R}_{\geq 1}) = \mathcal{K}_\phi$$

so:

$$\inf \langle S, \mathcal{K}_\phi \rangle \leq \inf \left\langle S, \mathcal{P}_n \cap \phi^{-1}(T) \right\rangle = \hat{\phi}(S)$$

since the infimum of a set S is always greater than or equal to the infimum of any superset of S.

Now, rearranging the original definition of the dual height function, we get that for any $T \in \mathcal{P}_n$:

$$\phi(S)\hat{\phi}(T) \leq \langle S, T \rangle .$$

Let $T \in \mathcal{K}_\phi = \phi^{-1}(\mathbb{R}_{\geq 1})$. We know that $T \in \overline{\mathcal{P}}_n$, but we don't know $T \in \mathcal{P}_n$. However, $T + \epsilon I \in \mathcal{P}_n$ for all $\epsilon > 0$, since $\epsilon I \in \mathcal{P}_n$, so for all S:

$$\hat{\phi}(S) \leq \hat{\phi}(S)\phi(T) \leq \hat{\phi}(S)\phi(T + \epsilon I) \leq \langle S, T + \epsilon I \rangle$$

Note that we get the first inequality from the fact that $\phi(T) \geq 1$, the second inequality from the fact that height functions preserve order, and the final inequality from the original definition of $\hat{\phi}$. The inner product is continuous on $\mathcal{P}_n$, so letting $\epsilon \to 0$, we get that for all $S \in \mathcal{P}_n$, $T \in \mathcal{K}_\phi$:

$$\hat{\phi}(S) \leq \langle S, T \rangle$$

Taking the infimum over all $T \in \mathcal{K}_\phi$ yields the desired result:

$$\hat{\phi}(S) \leq \inf \langle S, \mathcal{K}_\phi \rangle \Longrightarrow \hat{\phi}(S) = \inf \langle S, \mathcal{K}_\phi \rangle = \inf \left\langle S, \mathcal{P}_n \cap \phi^{-1}(1) \right\rangle$$

□

2.3.2 The Height Function-Kernel Correspondence

The dual of a kernel $\mathcal{K} \in \overline{\mathcal{P}}_n$, denoted $\mathcal{K}^{\sqcup}$, is the dual semihull of $\mathcal{K}$:

$$\mathcal{K}^{\sqcup} = \{S \in \overline{\mathcal{P}}_n : \langle S, \mathcal{K} \rangle \subset \mathbb{R}_{\geq 1}\}$$

Lemma 2.3.1. *Dual Kernel*
The dual of a kernel $\mathcal{K}$ is contained in $\overline{\mathcal{P}}_n$.

Proof. Since $\mathcal{K}$ is a kernel, $\mathcal{P}_n \subset \mathbb{R}_{>0}\mathcal{K}$. Thus:

$$\langle \mathcal{K}^{\sqcup}, \mathcal{P}_n \rangle \subset \langle \mathcal{K}^{\sqcup}, \mathbb{R}_{>0}\mathcal{K} \rangle = \mathbb{R}_{>0} \langle \mathcal{K}^{\sqcup}, \mathcal{K} \rangle \subset \mathbb{R}_{>0} \cdot \mathbb{R}_{\geq 1} = \mathbb{R}_{>0}$$

which shows that $\mathcal{K}^{\sqcup} \subset \mathcal{P}_n^{\vee} \subset \overline{\mathcal{P}}_n$. □

Proposition 2.3.3. *Duality Passes Through Correspondence*
Let $\mathcal{K}_\phi$ be the kernel of a height function ϕ. Then:

$$\mathcal{K}_\phi^{\sqcup} = \mathcal{K}_{\hat{\phi}}$$

Proof. For any $S \in \overline{\mathcal{P}}_n$, we have

$$S \in \mathcal{K}_\phi^{\sqcup} \iff \langle S, \mathcal{K}_\phi \rangle \subset \mathbb{R}_{\geq 1}$$

by definition of dual kernel. The condition $\langle S, \mathcal{K}_\phi \rangle \subset \mathbb{R}_{\geq 1}$ is clearly equivalent to $\inf \langle S, \mathcal{K}_\phi \rangle \geq 1$, so:

$$S \in \mathcal{K}_\phi^{\sqcup} \iff \inf \langle S, \mathcal{K}_\phi \rangle \geq 1$$

But recall that $\hat{\phi}(S) = \inf \langle S, \mathcal{K}_\phi \rangle$, so:

$$S \in \mathcal{K}_\phi^{\sqcup} \iff \hat{\phi}(S) \geq 1 \iff S \in \hat{\phi}^{-1}(\mathbb{R}_{\geq 1}) = \mathcal{K}_{\hat{\phi}}$$

Thus $\mathcal{K}_\phi^{\sqcup} = \mathcal{K}_{\hat{\phi}}$. □

Since the dual of a kernel of a height function is the kernel of the dual height function, it follows that the dual of the kernel is a kernel.

Proposition 2.3.4. *Height Function Duality*
Let $\phi : \mathcal{C} \to \mathbb{R}_{\geq 0}$ be a height function. Then for all $S \in \mathcal{P}_n$, $\phi(S) = \hat{\hat{\phi}}(S)$.

Proof. Because duality passes through the correspondence:

$$\overline{\mathcal{K}_{\hat{\hat{\phi}}}} = \overline{\mathcal{K}_{\hat{\phi}}}^{\sqcup} = \overline{\mathcal{K}_\phi}^{\sqcup\sqcup}$$

By duality of kernels, $\overline{\mathcal{K}}_\phi^{\sqcup\sqcup} = \overline{\mathcal{K}}_\phi$. Since ϕ and $\hat{\hat{\phi}}$ have the same kernel, they must agree on $\mathcal{P}_n$. □

2.4 The Minimum Function and the Dyadic Trace

2.4.1 The Minimum Function

Definition 2.4.1. *Minimum Function*
Let $m : \overline{\mathcal{P}}_n \to \mathbb{R}_{\geq 0}$ be the function:

$$m(S) = \min_{v \in \mathbb{Z}^n} S[v]$$

Proposition 2.4.1. *m is a Height Function*

Proof. Homogeneity and positivity follow immediately from the definition of $S[\cdot]$, so we simply need to show superadditivity holds. For any vector v, and for any positive semidefinite matrices S, T:

$$(S+T)[v] = S[v] + T[v] \geq m(S) + m(T).$$

Taking the minimum of the left hand side shows we have superadditivity. Thus, m is a height function. □

The minimum function has some interesting properties. First, m is defined on all of $\overline{\mathcal{P}}_n$, and $m(S) = 0$ if and only if $S \notin \mathcal{P}_n$. Thus, we can use m to characterize those spaces in future sections. Also, m has a geometric interpretation. Recall that any $S \in \mathcal{P}_n$ can be written as $S = MM'$, where the columns of M determine a basis of a lattice in $\mathbb{R}^n$. The vectors in the lattice are Mz, where $z \in \mathbb{Z}^n$, and for any such vector:

$$|Mz|^2 = \langle Mz, Mz\rangle = (Mz)'(Mz) = z'M'Mz = S[z]$$

Thus, $m(S)$ is the length of the shortest nonzero vector in the lattice isometry class determined by S.

Proposition 2.4.2. *Realization of the Dual of the Minimum Function*
Let $S \in \mathcal{P}_n$. There exists a matrix $T_o \in \mathcal{P}_n$ that satisfies:

$$\hat{m}(S) = \inf_{T\in\mathcal{P}_n} \frac{\langle S, T\rangle}{m(T)} = \frac{\langle S, T_o\rangle}{m(T_o)}$$

Proof. Fix $S \in \mathcal{P}_n$. By Prop. 2.3.2., $\hat{m}(S) = \inf \langle S, \mathcal{P}_n \cap m^{-1}(1)\rangle$. Since $I \in \mathcal{P}_n$ and $m(I) = 1$, $I \in \mathcal{P}_n \cap m^{-1}(1)$, so:

$$\hat{m}(S) = \inf \left\langle S, \mathcal{P}_n \cap m^{-1}(1)\right\rangle \leq \langle S, I\rangle = \operatorname{tr}(S)$$

Furthermore, we know $\operatorname{tr}(T)\lambda_1(S) \leq \langle S, T\rangle$ for all $T \in \mathcal{P}_n$ because the trace and least eigenvalue are duals of one another, so if $\operatorname{tr}(T) > \frac{\operatorname{tr}(S)}{\lambda_1(S)}$, then:

$$\operatorname{tr}(T)\lambda_1(S) > \operatorname{tr}(S) \geq \hat{m}(S)$$

As a result, setting $K = \left\{T \in \mathcal{P}_n : \operatorname{tr}(T) \leq \frac{\operatorname{tr}(S)}{\lambda_1(S)}\right\}$ allows us to rewrite:

$$\hat{m}(S) = \inf_{T\in K} \frac{\langle S, T\rangle}{m(T)} = \inf_{T\in K\cap m^{-1}(1)} \langle S, T\rangle$$

We know K is compact, and $m^{-1}(1)$ is closed so $K \cap m^{-1}(1)$ is also compact. Because $\langle S, \cdot\rangle$ is continuous, it takes a minimum on $K \cap m^{-1}(1)$, so there exists $T_o \in K \cap m^{-1}(1)$ such that $\hat{m}(S) = \frac{\langle S, T_o\rangle}{m(T_o)}$. □

2.4.2 The Dyadic Trace

The dual of m is called the *dyadic trace* and is denoted w. Let $\mathcal{C}_w$ denote the cone of dyadic matrices. In Chapter 1, we saw that $\mathcal{P}_n \subset \mathcal{C}_w$. Furthermore, since dyadic squares are positive semidefinite, it is easy to see that $\mathcal{C}_w \subset \overline{\mathcal{P}}_n$.

Definition 2.4.2. *Dyadic Trace*
The *dyadic trace* is the function $w : \mathcal{C}_w \longrightarrow \mathbb{R}_{\geq 0}$ given by:

$$w(S) = \sup\left(\sum_i \alpha_i\right) \quad where \quad S = \sum_i \alpha_i z_i z_i', \quad \alpha_i \in \mathbb{R}_{>0}, \quad z_i \in \mathbb{Z}^n$$

where the supremum is taken over all dyadic representations of S.

Proposition 2.4.3. *The Dyadic Trace is a Height Function*

Proof. First, observe that $w(S) \in \mathbb{R}_{>0}$ for all nonzero dyadic matrices S, and in particular, $w(S) > 0$ for all $S \in \mathcal{P}_n$. Also, $w(\lambda S) = \lambda w(S)$ for all $\lambda \in \mathbb{R}_{>0}$, because every dyadic representation can be scaled by any positive real number. Finally, we need to prove superadditivity. For any $S = \sum_i \alpha_i z_i z_i'$ and $T = \sum_j \beta_j w_j w_j'$, we have:

$$w(S) \geq \sum_i \alpha_i, \quad w(T) \geq \sum_j \beta_j \quad \Longrightarrow w(S) + w(T) \geq \sum_i \alpha_i + \sum_j \beta_j$$

If we take the supremum of the right hand side over all dyadic representations of S, T, we get an equality:

$$w(S) + w(T) = \sup\left(\sum_i \alpha_i + \sum_j \beta_j\right)$$

But $S+T = \sum_i \alpha_i z_i z_i' + \sum_j \beta_j w_j w_j'$ for any dyadic representations of S, T, so $w(S+T)$ is at least that big, i.e. :

$$w(S+T) \geq \sup\left(\sum_i \alpha_i + \sum_j \beta_j\right) = w(S) + w(T).$$

Thus w is a height function. □

To show that w and m are duals of one another takes some work. Introduce the notation:

$$\mathrm{mVec}(S) = \{v \in \mathbb{Z}^n : S[v] = m(S)\}$$
$$\diamondsuit S = \langle \mathbb{R}_{>0}\{zz' : z \in \mathrm{mVec}(S)\}\rangle$$

Proposition 2.4.4. *Upper Semicontinuity of Minimal Vectors*
Let $S \in \mathcal{P}_n$. There exists a neighborhood $\mathcal{N}_S$ such that

$$\forall T \in \mathcal{N}_S, \qquad \mathrm{mVec}(T) \subset \mathrm{mVec}(S)$$

Proof. Fix $S \in \mathcal{P}_n$, and let Λ be a lattice in the isometry class of S. Let $1 < r \leq 2$ be the ratio of the two shortest nonzero vector lengths in Λ, so that $r^2 m(S)$ equals the third smallest element of $S[\mathbb{Z}^n]$ (the smaller elements being $0, m(S)$).

Define a neighborhood of S in $\mathcal{P}_n$:

$$\mathcal{N}_S = \{T \in \mathcal{P}_n : S < rT, \quad m(T) < rm(S)\}$$

Then for any $T \in \mathcal{N}_S$ and for any $z \in \mathrm{mVec}(T)$:

$$\frac{1}{r} S[z] < T[z] = m(T) < rm(S)$$

Thus, $S[z] < r^2 m(S)$, so $S[z] = m(S)$ because S is positive and z nonzero, so $z \in \mathrm{mVec}(S)$. □

Lemma 2.4.1. *Minimal Vectors*
Let $S, T \in \mathcal{P}_n$ with $T < S$, and suppose $z \in \mathrm{mVec}(S) \cap \mathrm{mVec}(T)$. Then $z \in \mathrm{mVec}(S \pm T)$.

Proof. Let $v \in \mathbb{Z}^n$ be any nonzero vector. Then:

$$(S \pm T)[v] = S[v] \pm T[v] \leq S[z] \pm T[z] = (S \pm T)[z]$$

Thus, z is a minimal vector of $S + T$ and $S - T$, as claimed. □

Proposition 2.4.5. *Realizing the Dyadic Trace*
Let $S \in \mathcal{C}_w$. For any $T_o \in \mathcal{P}_n$,

$$T_o \textit{ minimizes } \frac{\langle S, T\rangle}{m(T)} \textit{ over } \mathcal{P}_n \iff S \in \diamond T_o$$

Proof. Suppose $S \in \diamond T_o$. Then $S = \sum_i \alpha_i z_i z_i'$, where $z_i \in \mathrm{mVec}(T_o)$ for all i, so:

$$\frac{\langle S, T_o\rangle}{m(T_o)} = \frac{\langle \sum_i \alpha_i z_i z_i', T_o\rangle}{m(T_o)} = \frac{\sum_i \alpha_i \langle z_i z_i', T_o\rangle}{m(T_o)} = \frac{\sum_i \alpha_i T_o[z_i]}{m(T_o)} = \sum_i \alpha_i$$

Let $T \in \mathcal{P}_n$, and suppose $z_i \notin \mathrm{mVec}(T)$ for some i. We know $T[z]/m(T) \geq 1$ for all $T \in \mathcal{P}_n$, $z \in \mathbb{Z}^n$, with equality if and only if z is a minimal vector for T. Thus:

$$\frac{\langle S, T\rangle}{m(T)} = \sum_i \alpha_i \frac{T_[z_i]}{m(T_o)} > \sum_i \alpha_i = \frac{\langle S, T_o\rangle}{m(T_o)}$$

That proves that if $S \in \diamond T_o$ for some $T_o \in \mathcal{P}_n$, then T_o minimizes $\langle S, T\rangle / m(T)$ over $\mathcal{P}_n$.

Now suppose T_o minimizes $\langle S, T\rangle / m(T)$ over $\mathcal{P}_n$, so that for any $U \in \mathcal{P}_n$, we have:

$$\frac{\langle S, U\rangle}{m(U)} \leq \langle S, T_o\rangle m(T_o) \qquad (\textit{or, equivalently}) \qquad m(T_o)\langle S, U\rangle \leq m(U)\langle S, T\rangle$$

If z is a minimal vector of U, we can use the fact that $\langle U, zz' \rangle = m(U)$ to rewrite that display:

$$(\forall U \in \mathcal{P}_n)(\forall z \in \text{mVec}(U)) \qquad \langle S, T_o \rangle \langle U, zz' \rangle \leq \langle S, U \rangle m(T_o)$$

By the Upper Semicontinuity of Minimal Vectors, there exists a neighborhood $\mathcal{N}_T \subset \mathcal{P}_n$ of T_o with the property that for all $U \in \mathcal{N}_T$, $\text{mVec}(U) \subset \text{mVec}(T)$. Since $\mathcal{N}_T$ is a neighborhood of T_o, we can write its members as $U = T_o + B$, where $B \in \mathcal{V}_n$ is understood to be close to the origin; in other words, we can translate $\mathcal{N}_T$ by $-T_o$ to get a neighborhood $\mathcal{N}_o \subset \mathcal{V}_n$ of the origin with the property that for all $B \in \mathcal{N}_o$:

$$\langle S, T_o \rangle \langle T_o + B, zz' \rangle = \langle S, T_o \rangle \langle T_o, zz' \rangle + \langle S, T_o \rangle \langle B, zz' \rangle \leq \langle S, T_o \rangle m(T_o) + \langle S, B \rangle m(T_o)$$

or, after simplifying/cancelling similar terms:

$$(\forall z \in \text{mVec}(T_o + B)) \qquad \langle S, T_o \rangle \langle B, zz' \rangle \leq \langle S, B \rangle m(T_o)$$

We were trying to show that $S \in \diamondsuit T_o$. Since $\diamondsuit T_o$ is a closed cone, $\diamondsuit T_o = (\diamondsuit T_o)^{\vee\vee}$ so we can show $S \in (\diamondsuit T_o)^{\vee\vee}$, which is implied by $\langle S, (\diamondsuit T_o)^{\vee} \rangle \subset \mathbb{R}_{\geq 0}$, instead. Let $T \in \diamondsuit T_o$. We know $\mathbb{R}_{>0} \cdot T \cap \mathcal{N}_o \neq \emptyset$, so let $\widetilde{T} = \lambda T$ be a scalar multiple of T that lies in that intersection.

Then we have that:

$$\langle S, T_o \rangle \left\langle \widetilde{T}, zz' \right\rangle \leq \left\langle S, \widetilde{T} \right\rangle m(T_o) \Longrightarrow \langle S, T_o \rangle \langle T, zz' \rangle \leq \langle S, T \rangle m(T_o)$$

Since $T \in (\diamondsuit T_o)^{\vee} \subset \overline{\mathcal{P}}_n$, $\langle T, zz' \rangle \geq 0$. Furthermore, $m(T_o) > 0$ because $T_o \in \mathcal{P}_n$. Finally, $\langle S, T_o \rangle > 0$ because $S, T \in \mathcal{P}_n$, which forces $\langle S, T \rangle \geq 0$. This holds for all $T \in (\diamondsuit T_o)^{\vee}$ (though the z might change), so $\langle S, (\diamondsuit T_o)^{\vee} \rangle \subset \mathbb{R}_{\geq 0}$, so $S \in \diamondsuit T_o$. □

2.4.3 Computing the Dyadic Trace

I will show how one can compute the dyadic trace for 2×2 and 3×3 matrices. The method is similar in both cases: we use the fact that the dyadic trace is both an infimum and a supremum to obtain lower and higher bounds that agree. See the Appendix for definitions of Legendre-reduced and Minkowski-reduced.

Proposition 2.4.6. *Formula for* 2×2 *matrices*

Let $S = \begin{bmatrix} a & b \\ b & c \end{bmatrix} \in \mathcal{P}_2$ be Legendre-reduced. Then $w(S) = a + c - |b|$.

Proof. Let $T = \begin{bmatrix} 2 & \pm 1 \\ \pm 1 & 2 \end{bmatrix} \in \mathcal{P}_2$. Note that $m(T) = 2$. By duality of m and w:

$$w(S)m(T) = 2w(S) \leq \langle S, T \rangle = 2(a + c \pm b) \Longrightarrow w(S) \leq a + c - |b|$$

Furthermore, since Legendre-reduced matrices are diagonally dominant, S has the following dyadic representation

$$S = (a - |b|)e_1 e_1' + (c - |b|)e_2 e_2' + |b|(e_1 \pm e_2)(e_1 \pm e_2)'$$

so by the definition of the dyadic trace:

$$w(S) \geq (a - |b|) + (c - |b|) + |b| = a + c - |b|$$

Thus $w(S) = a + c - |b|$. □

Proposition 2.4.7. *Formula for* 3×3 *matrices*
Let:

$$S = \begin{bmatrix} a & d & e \\ d & b & f \\ e & f & c \end{bmatrix} \in \mathcal{P}_3$$

be Minkowski reduced. Then

$$w(S) = \begin{cases} a + b + c - |d| - |e| - |f| & if\ def \leq 0 \\ a + b + c - |d| - |e| - |f| + \min\{|d|, |e|, |f|\} & if\ def > 0 \end{cases}$$

Proof. Consider:

$$T = \begin{bmatrix} 2 & \delta & \epsilon \\ \delta & 2 & \phi \\ \epsilon & \phi & 2 \end{bmatrix} \qquad \delta, \epsilon, \phi \in \{0, \pm 1\} \qquad \delta\epsilon\phi \neq -1$$

One readily verifies that:

$$\det T = 8 + 2(\delta\epsilon\phi - \delta^2 - \epsilon^2 - \phi^2) \geq 2.$$

Furthermore, the upper-left 2×2 submatrix has determinant equal to 3 or 4, and the upper-left entry is 2, so $T \in \mathcal{P}_3$. In fact, T is the Gram matrix of a lattice known in the literature as A_3, so we know many things about T. In particular, we know $m(T) = 2$, so:

$$w(S) \geq \min_{\delta,\epsilon,\phi} \frac{\langle S, T \rangle}{2} = \min_{\delta,\epsilon,\phi}(a + b + c + \delta d + \epsilon e + \phi f)$$

To minimize the parenthesized expression, we want the sum to contain as many negative terms as possible. The remaining terms can should be annihilated. If $def \leq 0$, that means at least one of d, e, f is nonpositive. Pick one of the nonpositive elements, say d, and set $\delta = 1$ and $\epsilon = \phi - 1$. Then:

$$w(S) \geq a + b + c - |d| - |e| - |f|$$

Otherwise, $def > 0$. The Minkowski conditions give us $d, f \geq 0$, so all three terms are positive. However, we cannot set $\delta = \epsilon = \phi = -1$, because that would mean $\delta\epsilon\phi = -1$. Thus, we negate the two larger terms, and annihilate the smaller term by setting its coefficient equal to 0 to get:

$$w(S) \geq a + b + c - |d| - |e| - |f| + \min\{|d|, |e|, |f|\}$$

Since S is Minkowski reduced, it is diagonally dominant, so we can use the formula from earlier to get a dyadic representation for S:

$$S = (a - |d| - |e|)e_1e_1' + (b - |d| - |f|)e_2e_2' + (c - |f| - |e|)e_3e_3' +$$

$$|d|(e_1 + e_2)(e_1 + e_2)' + |e|(e_1 + \mathrm{sgn}(def)e_2)(e_1 + \mathrm{sgn}(def)e_2)' + |f|(e_2 + e_3)(e_2 + e_3)'$$

The dyadic trace of S is therefore at least:

$$w(S) \geq a + b + c - 2|d| - 2|e| - 2|f| + |d| + |e| + |f|$$

If $def \leq$, then we are done, since the upper and lower bounds agree. Otherwise, we have to keep working.

Suppose $def > 0$, and let m denote the smallest of $|d|, |e|, |f|$. Then:

$$S = m(e_1 + e_2 + e_3)(e_1 + e_2 + e_3)'$$

$$+(|d| - m)(e_1 + e_2)(e_1 + e_2)'(|e| - m)(e_1 + e_3)(e_1 + e_3)'$$

$$(|f| - m)(e_2 + e_3)' + (a - |d| - |e| + m)e_1e_1'$$

$$+(b - |d| - |f| + m)e_2e_2' + (c - |e| - |f| + m)e_3e_3'$$

Thus, a lower bound for $w(S)$ is:

$$m + |d| - m + |e| - m + |f| - m + a - |d| - |e| + m + b - |f| - |d| + m + c - |e| - |f| + m$$

$$\implies w(S) \leq m + a + b + c - |d| - |e| - |f|$$

$$\therefore w(S) = m + a + b + c - |d| - |e| - |f|$$

□

Chapter 3

Siegel Modular Forms

3.1 Definitions and Basic Properties

Fix a positive integer n. We are going to need several spaces of matrices in this section.

First, we define a few sets of matrices:

$$\begin{aligned}
\mathcal{V}_n(\mathbb{Z}) &= \{n \times n \text{ symmetric integral matrices}\} \\
V_n^*(\mathbb{Z}) &= \{V \in \mathcal{V}_n : v_{ii} \in \mathbb{Z} \text{ and } v_{ij} \in \frac{1}{2}\mathbb{Z} \text{ for all } i, j\} \\
\mathcal{X}_n^{\text{semi}} &= \mathcal{V}_n(\mathbb{Z})^* \cap \overline{\mathcal{P}}_n \\
\mathcal{X}_n &= \mathcal{V}_n(\mathbb{Z})^* \cap \mathcal{P}_n
\end{aligned}$$

Definition 3.1.1. *Siegel Upper Half Space*
The Siegel upper half space of dimension n is:

$$\mathcal{H}_n = \{\Omega = X + iY : X \in \mathcal{V}_n,\, Y \in \mathcal{P}_n\}$$

For $n = 1$, $\mathcal{H}_1$ is the familiar upper half plane from complex analysis.

Definition 3.1.2. *Symplectic Group*
The *real symplectic group of order n* is:

$$\mathrm{Sp}_n(\mathbb{R}) = \left\{\gamma \in \mathrm{GL}_{2n}(\mathbb{R}) : \begin{bmatrix} 0 & -I_n \\ I_n & 0 \end{bmatrix} [\gamma] = \begin{bmatrix} 0 & -I_n \\ I_n & 0 \end{bmatrix}\right\}$$

When $n = 1$, $\mathrm{Sp}_1(\mathbb{R}) = \mathrm{SL}_2(\mathbb{R})$.

Definition 3.1.3. *Generalized Fractional Linear Transformations*
The real symplectic group acts on Siegel upper half space by *generalized fractional linear transformations*:

$$\gamma(\Omega) = (A\Omega + B)(C\Omega + D)^{-1}$$

To show that $C\Omega + D$ is invertible takes a little bit of work. See Klingen (1990), pg. 2, for a proof of this result.

Definition 3.1.4. *Factor of Automorphy Function*
The *factor of automorphy function* is the function $j : \mathrm{Sp}_n(\mathbb{R}) \times \mathcal{H}_n \longrightarrow \mathbb{C}^\times$ given by:

$$j(\gamma, \Omega) = \det(C\Omega + D) \qquad \left(\gamma = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \in \mathrm{Sp}_n(\mathbb{R})\right)$$

The factor of automorphy function satisfies a *cocycle condition*:

$$j(\gamma\tilde{\gamma}, \Omega) = j(\gamma, \tilde{\gamma}(\Omega))j(\tilde{\gamma}, \Omega)$$

as well as the following identity:

$$\det(\mathrm{Im}(\gamma(\Omega)) = |j(\gamma, \Omega)|^{-2} \det(\mathrm{Im}(\Omega)$$

We will need those identities later, but proving them would be too much of a digression from the main ideas of this thesis.

Definition 3.1.5. *Weight-k Action*
Let k be a positive integer. Define the *weight-k action* of $\mathrm{Sp}_n(\mathbb{R})$ on the set of complex-valued functions on Siegel upper half space:

$$f[\gamma]_k(\Omega) = j(\gamma, \Omega)^{-k} f(\gamma(\Omega))$$

where $\gamma \in \mathrm{Sp}_n(\mathbb{R})$ and $f : \mathcal{H}_n \to \mathbb{C}$.

Definition 3.1.6. *Siegel Modular Forms*
Let n be a positive integer and k a nonnegative integer. We say that a function:

$$f : \mathcal{H}_n \longrightarrow \mathbb{C}$$

is a *Siegel modular form of dimension n and weight k* if f satisfies the following conditions:

1. f is holomorphic in the upper triangular entries of its argument.
2. f is invariant under the weight-k action of the discrete subgroup $\mathrm{Sp}_n(\mathbb{Z})$ of the real symplectic group of order n.
3. For any $Y_o \in \mathcal{P}_n$, f is bounded on the set $\{\Omega \in \mathcal{H}_n : \mathrm{Im}(\Omega) > Y_o\}$.

The vector space of all such functions is denoted $\mathcal{M}_k(\mathrm{Sp}_n(\mathbb{Z}))$. For $n > 1$, the third condition follows from the first two. This fact is known as the *Koecher Principle*.

The second condition tells us a lot about a Siegel modular form.

Proposition 3.1.1. *Properties of Siegel Modular Forms*
Let $f \in \mathcal{M}_k(\mathrm{Sp}_n(\mathbb{Z}))$. Then:

1. *(Translation Invariance)* For all $B \in \mathcal{V}_n(\mathbb{Z})$, $f(\Omega + B) = f(\Omega)$.
2. For all $V \in \mathrm{GL}_n(\mathbb{Z})$, $f(\Omega[V']) = (\det V)^{-k} f(\Omega)$.

Proof. Let B, V be given. Define matrices:

$$\tau = \begin{bmatrix} I_n & B \\ O & I_n \end{bmatrix} \qquad \sigma = \begin{bmatrix} V & O \\ O & V'^{-1} \end{bmatrix}$$

One can verify that $\tau, \sigma \in \mathrm{Sp}_n(\mathbb{Z})$. By definition of the group action of $\mathrm{Sp}_n(\mathbb{Z})$ on $\mathcal{H}_n$:

$$\tau(\Omega) = (I_n \cdot \Omega + B)(O \cdot \Omega + I_n)^{-1} = \Omega + B$$
$$\sigma(\Omega) = (V \cdot \Omega + O)(O \cdot \Omega + V'^{-1})^{-1} = V\Omega V' = \Omega[V'].$$

By the weight-k invariance of f with respect to $\mathrm{Sp}_n(\mathbb{Z})$:

$$f(\tau(\Omega)) = j(\tau, \Omega)^k f(\Omega) = \det(I_n) f(\Omega) = f(\Omega)$$
$$f(\sigma(\Omega)) = j(\sigma, \Omega)^k f(\Omega) = (\det V')^k f(\Omega) = (\det V)^k f(\Omega)$$

Combining results:

$$f(\tau(\Omega)) = f(\Omega + B) = f(\Omega)$$
$$f(\sigma(\Omega)) = f(\Omega[V']) = (\det V)^k f(\Omega)$$

□

A Siegel modular form is determined by its behavior on the set of orbits for $\mathrm{Sp}_n(\mathbb{Z})\backslash\mathcal{H}_n$. A set $\mathcal{S} \subset \mathcal{H}_n$ that contains exactly one representative for every orbit is called a fundamental domain. There exists a fundamental domain $\mathcal{F}_n$ satisfying:

$$m(Y) \geq \frac{\sqrt{3}}{2} \quad \textit{for all } X + iY \in \mathcal{F}_n.$$

We will use that fact later.

3.2 Fourier Series

This section will establish the Fourier expansion of Siegel modular forms.

Proposition 3.2.1. *Fourier Expansion*
Let $f \in \mathcal{M}_k(\mathrm{Sp}_n(\mathbb{Z}))$. Then f has a unique Fourier expansion:

$$f(\Omega) = \sum_{T \in \mathcal{X}_n^{\mathrm{semi}}} a(T; f) \mathrm{e}\left(\langle \Omega, T \rangle\right) \qquad (f \in \mathcal{M}_k(\mathrm{Sp}_n(\mathbb{Z})))$$

where the $a(T; f) \in \mathbb{C}$ are the Fourier coefficients, and $\mathrm{e}\left(\langle \Omega, T \rangle\right) = e^{2\pi i \cdot \mathrm{tr}\,(\Omega T)}$.

Proof. We know f is $\mathbb{Z}$-periodic and smooth in the real variables x_{jk}, $j \leq k$. The $\mathbb{Z}$-periodicity is a consequence of the weight-k invariance, and the smoothness follows from holomorphy in those variables. Thus, by real analysis, we know that f has a unique, absolutely convergent expansion as:

$$f(\Omega) = \sum_{\{t_{jk} \in \mathbb{Z}^{n(n+1)/2}\}} a(\{t_{jk}, y_{jk}\}; f) \mathrm{e}\left(\sum_{j,k} t_{jk} x_{jk}\right)$$

where the coefficients $a(\{t_{jk}, y_{jk}\}; f)$ are functions of the y_{jk}. Furthermore, f must satisfy the Cauchy-Riemann equations for each index (j, k). To see what this tells us about f, we can take the derivative term by term. The derivative of an arbitrary term with respect to some x_{mn} is:

$$\frac{\partial}{\partial x_{mn}}\left(a(\{t_{jk}, y_{jk}\}; f)\mathrm{e}\left(\sum_{j,k} t_{jk}x_{jk}\right)\right) = 2\pi i t_{mn} \cdot a(\{t_{jk}, y_{jk}\}; f)\mathrm{e}\left(\sum_{j,k} t_{jk}x_{jk}\right)$$

so we have:

$$-i\frac{\partial}{\partial y_{mn}}\left(a(\{t_{jk}, y_{jk}\}; f)\mathrm{e}\left(\sum_{j,k} t_{jk}x_{jk}\right)\right) = 2\pi i t_{mn} \cdot a(\{t_{jk}, y_{jk}\}; f)\mathrm{e}\left(\sum_{j,k} t_{jk}x_{jk}\right)$$

because Fourier series are equal if and only if they're equal term-by-term. Getting rid of all the scalars shows:

$$\frac{\partial}{\partial y_{mn}} a(\{t_{jk}, y_{jk}\}; f) = -2\pi t_{mn} a(\{t_{jk}, y_{jk}\}; f)$$

We solve the partial differential equations to get:

$$a(\{t_{jk}, y_{jk}\}; f) = a(\{t_{jk}\}; f)\mathrm{e}\left(\sum_{j,k} t_{jk} i y_{jk}\right)$$

Note that we have iy_{jk} rather than $-y_{jk}$ because $\mathrm{e}\,(ix) = e^{2\pi i^2 x} = e^{-2\pi x}$. In any case, this allows us to rewrite $f(\Omega)$ as:

$$f(\Omega) = \sum_{\{t_{jk}\in\mathbb{Z}^{n(n+1)/2}\}} a(\{t_{jk}\}; f)\mathrm{e}\left(\sum_{j,k} t_{jk} z_{jk}\right)$$

For each $n(n+1)/2$-tuple of t_{jk}'s, associate a matrix $T \in \mathcal{V}_n(\mathbb{Z})^*$ whose diagonal entries are t_{jj} and whose superdiagonal entries are $\frac{1}{2}t_{jk}$. Using this notation, it is clear that:

$$\sum_{1\le j\le k}^{n} t_{jk} z_{jk} = \langle T, \Omega\rangle$$

Thus, the previous display can be cleaned up:

$$f(\Omega) = \sum_{T\in\mathcal{V}_n(\mathbb{Z})^*} a(T; f)\mathrm{e}\left(\langle T, \Omega\rangle\right)$$

Now we just need to show that we can sum over $\mathcal{X}_n^{\mathrm{semi}}$ rather than all of $\mathcal{V}_n(\mathbb{Z})^*$. That is, we need to show that $a(T; f) = 0$ if $T \notin \overline{\mathcal{P}}_n$. Fix a nonzero vector $v \in \mathbb{R}^n$, and let $z = x + iy \in \mathcal{H}$. Consider the matrix:

$$iI + zvv' = xvv' + i(yvv' + I).$$

Since $I \in \mathcal{P}_n$, $vv' \in \overline{\mathcal{P}}_n$, $y > 0$, we have $yvv' + I \in \mathcal{P}_n$ and $vv' \in \mathcal{V}_n$ so $iI + zvv' \in \mathcal{H}_n$ for all $z \in \mathcal{H}$. As a result, we can think of

$$f(iI + zvv') = \sum_{T \in \mathcal{V}_n(\mathbb{Z})^*} a(T; f)\mathrm{e}\left(\langle iI + zvv', T\rangle\right) = \sum_{T \in \mathcal{V}_n(\mathbb{Z})^*} a(T; f)e^{2\pi i(\mathrm{tr}\,(T) + zT[v])}$$

as a holomorphic function of z in $\mathcal{H}$.

Now let $T \in \mathcal{V}_n(\mathbb{Z})^*$ and suppose $T[v] < 0$. By the properties of the complex exponential, we know that as $\mathrm{Im}(z) \to \infty$, the term $|e^{2\pi i(\mathrm{tr}\,(T)+zT[v])}| = |e^{-2\pi y T[v]}|$ also goes to ∞. By the boundedness condition of Siegel modular forms, that forces $a(T; f) = 0$.

Thus, if $a(T; f) \neq 0$, then there can not exist $v \in \mathbb{R}^n$ with $T[v] < 0$, so $T \in \overline{\mathcal{P}}_n$. More generally, we can write the Fourier series of f while summing over $T \in \mathcal{X}_n^{\mathrm{semi}}$ without giving up any information. Thus:

$$f(\Omega) = \sum_{T \in \mathcal{X}_n^{\mathrm{semi}}} a(T; f)\mathrm{e}\left(\langle \Omega, T\rangle\right)$$

as claimed. □

Proposition 3.2.2. *Class Invariance of Fourier Coefficients*
Let $f \in \mathcal{M}_k(\mathrm{Sp}_n(\mathbb{Z}))$. The Fourier coefficients of f have the following properties:

1. If k is even, then for all $V \in \mathrm{GL}_n(\mathbb{Z})$, $a(T; f) = a(T[V]; f)$.
2. For all positive integers k and for all $V \in \mathrm{GL}_n(\mathbb{Z})$, $a(T; f) = 0 \Leftrightarrow a(T[V]; f) = 0$.
3. For all positive integers k and for all $V \in \mathrm{SL}_n(\mathbb{Z})$, $a(T; f) = a(T[V]; f)$

Proof. Let $V \in \mathrm{GL}_n(\mathbb{Z})$. By the previous proposition:

$$f(\Omega[V'^{-1}]) = \sum_{T \in \mathcal{X}_n^{\mathrm{semi}}} a(T; f)\mathrm{e}\left(\left\langle T, \Omega[V'^{-1}]\right\rangle\right)$$

By linear algebra, we know we can rewrite the display as:

$$f(\Omega[V'^{-1}]) = \sum_{T \in \mathcal{X}_n^{\mathrm{semi}}} a(T; f)\mathrm{e}\left(\left\langle T[V^{-1}], \Omega\right\rangle\right)$$

Since $T \longmapsto T[V^{-1}]$ is an automorphism of $\mathcal{X}_n^{\mathrm{semi}}$, with inverse $T \longmapsto T[V]$, we can rewrite the previous summation as:

$$f(\Omega[V'^{-1}]) = \sum_{T \in \mathcal{X}_n^{\mathrm{semi}}} a(T[V]; f)\mathrm{e}\left(\langle T, \Omega\rangle\right)$$

By Prop. 3.1.1, we can rewrite the left hand side:

$$(\det V^{-1})^k f(\Omega) = \sum_{T \in \mathcal{X}_n^{\mathrm{semi}}} a(T[V]; f)\mathrm{e}\left(\langle T, \Omega\rangle\right).$$

Multiply both sides by $(\det V)^k$:

$$f(\Omega) = \sum_{T \in \mathcal{X}_n^{\text{semi}}} (\det V)^k a(T[V]; f) \mathrm{e}\left(\langle T, \Omega \rangle\right).$$

This is a Fourier expansion for f, and as noted earlier, Fourier expansions are unique. Thus, for all $T \in \mathcal{X}_n^{\text{semi}}$, we have:

$$a(T; f) = (\det V)^k a(T[V]; f)$$

Since $V \in \mathrm{GL}_n(\mathbb{Z})$, we know $\det V = \pm 1$. Furthermore, $\det V = 1 \Leftrightarrow V \in \mathrm{SL}_n(\mathbb{Z})$. Thus, if k is even, $V \in \mathrm{SL}_n(\mathbb{Z})$ or $a(T; f) = 0$, we have:

$$a(T; f) = a(T[V]; f)$$

□

3.3 Siegel's Map and Cusp Forms

Definition 3.3.1. *Siegel's Φ-Map*
Let n be a positive integer. Siegel's Φ-map for Siegel modular forms of dimension n, where $n \in \mathbb{Z}^+$, is the mapping:

$$\Phi : \mathcal{M}_k(\mathrm{Sp}_n(\mathbb{Z})) \longrightarrow \mathcal{M}_k(\mathrm{Sp}_{n-1}(\mathbb{Z}))$$

defined by

$$(\Phi f)(\Omega) = \lim_{y \to \infty} f\left(\begin{bmatrix} \Omega & 0 \\ 0 & iy \end{bmatrix}\right).$$

Theorem 3.1. *Siegel's Φ-map*
Siegel's Φ-map is well-defined.

Proof. First, we need to show the limit exists. Introduce the notation $\Omega_y = \begin{bmatrix} \Omega & 0 \\ 0 & iy \end{bmatrix}$, and let $\widetilde{T}$ denote the $(n-1) \times (n-1)$ upper right block of any matrix T. We can express Φf using the Fourier expansion of f:

$$(\Phi) f(\Omega) = \lim_{y \to \infty} \sum_{T \in \mathcal{X}_n^{\text{semi}}} a(T; f) \mathrm{e}\left(\langle \Omega_y, T \rangle\right)$$

Because the Fourier expansion of f is absolutely convergent, we can pull the limit into the summation:

$$(\Phi) f(\Omega) = \sum_{T \in \mathcal{X}_n^{\text{semi}}} a(T; f) \lim_{y \to \infty} \mathrm{e}\left(\langle \Omega_y, T \rangle\right)$$

For any $\Omega \in \mathcal{H}_{n-1}$, compute that:

$$\mathrm{e}\left(\langle T, \Omega_y \rangle\right) = \mathrm{e}\left(\left\langle \widetilde{T}, \Omega \right\rangle + iyt_{nn}\right) = \mathrm{e}\left(\left\langle \widetilde{T}, \Omega \right\rangle\right) e^{-2\pi y t_{nn}}$$

so

$$\lim_{y\to\infty} \mathrm{e}\left(\langle T, \Omega_y\rangle\right) = \mathrm{e}\left(\left\langle \widetilde{T}, \Omega\right\rangle\right) \lim_{y\to\infty} e^{-2\pi y t_{nn}}$$

We know $t_{nn} \geq 0$ because $T \in \mathcal{X}_n^{\text{semi}}$. If $t_{nn} > 0$, then that term is annihilated by Φ so it will certainly not cause the summation to diverge. Otherwise, $t_{nn} = 0$, and the limit is $\mathrm{e}\left(\left\langle \widetilde{T}, \Omega\right\rangle\right)$, so

$$(\Phi f)(\Omega) = \sum_{T\in\mathcal{X}_n^{\text{semi}}} a(T; f)\mathrm{e}\left(\left\langle \widetilde{T}, \Omega\right\rangle\right)$$

Now, if $t_{nn} = 0$, then $t_{nj} = t_{in} = 0$ for all i, j by the properties of definite matrices, so T is of the form $\begin{bmatrix}\widetilde{T} & 0\\ 0 & 0\end{bmatrix}$. As a result, we can simply sum over matrices $\widetilde{T} \in \mathcal{X}_{n-1}^{\text{semi}}$, because the upper left block of a $\mathcal{X}_n^{\text{semi}}$ matrix is clearly a $\mathcal{X}_{n-1}^{\text{semi}}$ matrix. Thus:

$$(\Phi f)(\Omega) = \sum_{T\in\mathcal{X}_{n-1}^{\text{semi}}} a(T; f)\mathrm{e}\left(\langle T, \Omega\rangle\right)$$

which shows that the limit exists, and $\Phi f : \mathcal{H}_{n-1} \to \mathbb{C}$. Furthermore, Φ_f inherits holomorphy and boundedness from f, so we just need to show $(\Phi f)[\gamma]_k(\Omega) = (\Phi f)(\Omega)$ for all $\gamma \in \mathrm{Sp}_{n-1}(\mathbb{Z})$.

Let $\gamma = \begin{bmatrix}A & B\\ C & D\end{bmatrix} \in \mathrm{Sp}_{n-1}(\mathbb{Z})$ be arbitrary. By definition of the weight-k action on functions on $\mathcal{H}_{n-1}$ and the definition of Siegel's Φ-map, we have:

$$(\Phi f)[\gamma]_k(\Omega) = j(\gamma, \Omega)^{-k}(\Phi f)(\gamma(\Omega)) = j(\gamma, \Omega)^{-k} \lim_{y\to\infty} f(\gamma(\Omega)_y)$$

Let δ be the $\mathrm{Sp}_n(\mathbb{Z})$ matrix:

$$\delta = \begin{bmatrix}A & 0 & B & 0\\ 0 & 1 & 0 & 0\\ C & 0 & D & 0\\ 0 & 0 & 0 & 1\end{bmatrix}$$

I claim that $\gamma(\Omega)_y = \delta(\Omega_y)$. To see this, compute:

$$\delta(\Omega_y) = \left(\begin{bmatrix}A & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}\Omega & 0\\ 0 & iy\end{bmatrix} + \begin{bmatrix}B & 0\\ 0 & 0\end{bmatrix}\right)\left(\begin{bmatrix}C & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}\Omega & 0\\ 0 & iy\end{bmatrix} + \begin{bmatrix}D & 0\\ 0 & 1\end{bmatrix}\right)^{-1}$$

$$\delta(\Omega_y) = \begin{bmatrix}A\Omega + B & 0\\ 0 & iy\end{bmatrix}\begin{bmatrix}C\Omega + D & 0\\ 0 & 1\end{bmatrix}^{-1} = \begin{bmatrix}(A\Omega + B)(C\Omega + D)^{-1} & 0\\ 0 & iy\end{bmatrix} = (\gamma(\Omega))_y$$

Thus, we can rewrite the last display we had for $(\Phi f)[\gamma]_k(\Omega)$:

$$(\Phi f)[\gamma]_k(\Omega) = j(\gamma, \Omega)^{-k}(\Phi f)(\gamma(\Omega)) = j(\gamma, \Omega)^{-k} \lim_{y\to\infty} f(\delta(\Omega_y))$$

But since $\delta \in \mathrm{Sp}_n(\mathbb{Z})$ and $f \in \mathcal{M}_k(\mathrm{Sp}_n(\mathbb{Z}))$, we know $f(\delta(\Omega_y)) = j(\delta, \Omega_y)^k f(\Omega_y)$ so:

$$(\Phi f)[\gamma]_k(\Omega) = j(\gamma, \Omega)^{-k}(\Phi f)(\gamma(\Omega)) = j(\gamma, \Omega)^{-k} \lim_{y\to\infty} j(\delta, \Omega_y)^k f(\Omega_y)$$

Finally, since $j(\delta, \Omega_y) = j(\gamma, \Omega)$ for all $y > 0$:

$$(\Phi f)[\gamma]_k(\Omega) = j(\gamma, \Omega)^{-k} \lim_{y\to\infty} j(\gamma, \Omega)^k f(\Omega_y) = \lim_{y\to\infty} f(\Omega_y) = (\Phi f)(\Omega)$$

□

3.4 Invariant Function of a Siegel Modular Form

Definition 3.4.1. *Invariant Function of a Siegel Modular Form*
Let $f \in \mathcal{M}_k(\mathrm{Sp}_n(\mathbb{Z}))$. We define the $\mathrm{Sp}_n(\mathbb{Z})$-invariant function of f as follows:

$$\phi_f : \mathcal{H}_n \longrightarrow \mathbb{R}_{\geq 0} \qquad \phi_f(\Omega) = (\det \mathrm{Im}(\Omega))^{k/2} |f(\Omega)|$$

To see that ϕ_f is, in fact, invariant with respect to $\mathrm{Sp}_n(\mathbb{Z})$, we use the following identity:

$$\det(\mathrm{Im}(\gamma(\Omega)) = |j(\gamma, \Omega)|^{-2} \det(\mathrm{Im}(\Omega)$$

For any $\gamma \in \mathrm{Sp}_n(\mathbb{Z})$, we have:

$$\phi_f(\gamma(\Omega)) = \det \mathrm{Im}(\gamma(\Omega)))^{k/2} |f(\gamma(\Omega))| = |j(\gamma, \Omega)|^{-k} \det(\mathrm{Im}(\Omega)^{k/2} |f(\gamma(\Omega))|$$

Since f is a weight k Siegel modular form:

$$\phi_f(\gamma(\Omega)) = |j(\gamma, \Omega)|^{-k} \det(\mathrm{Im}(\Omega)^{k/2} |j(\gamma, \Omega)|^k |f(\Omega)| = \det(\mathrm{Im}(\Omega)^{k/2} |f(\Omega)| = \phi_f(\Omega)$$

Lemma 3.4.1. *Bounds*
Let $f \in \mathcal{S}_k(\mathrm{Sp}_n(\mathbb{Z}))$ be a nonzero Siegel cusp form. There exist $a, b \in \mathbb{R}_{>0}$ such that for all $\Omega = X + iY \in \mathcal{H}_n$, if $Y \geq I$, then:

$$|f(\Omega)| \leq a e^{-b\delta(Y)}$$

Proof. For any $\Omega = X + iY \in \mathcal{H}_n$, we have:

$$|f(\Omega)| = \left| \sum_{T \in \mathcal{X}_n} a(T; f) e^{2\pi i \mathrm{tr}\,(\Omega T)} \right| \leq \sum_{T \in \mathcal{X}_n} \left| a(T; f) e^{2\pi i \mathrm{tr}\,(\Omega T)} \right| = \sum_{T \in \mathcal{X}_n} |a(T; f)|\, e^{-2\pi \mathrm{tr}\,(YT)}$$

Furthermore, for any $T \in \mathcal{X}_n$, and any $Y \geq I$, we have:

$$\langle T, Y \rangle \geq \mathrm{tr}\,(T)\lambda_1(Y) \qquad \langle T, Y \rangle \geq n\delta(T)\delta(Y)$$

so consequently:

$$2\langle T, Y \rangle \geq \mathrm{tr}\,(T)\lambda_1(Y) + n\delta(T)\delta(Y)$$

It is easily verified that δ is bounded below by 2^{-n} on $\mathcal{X}_n$. Furthermore, since height functions preserve order, $\lambda_1(Y) \geq \lambda_1(I) = 1$. Thus:

$$2\left\langle T, Y\right\rangle \geq \mathrm{tr}\,(T) + n2^{-n}\delta(Y)$$

Now, let $a = \sum_{T\in\mathcal{X}_n} |a(T;f)|e^{-\pi\mathrm{tr}\,(T)}$. Note a is finite because

$$f(\frac{i}{2}I) = \sum_{T\in\mathcal{X}_n} a(T;f)e^{2\pi i\mathrm{tr}\,(\frac{i}{2}IT)} = \sum_{T\in\mathcal{X}_n} a(T;f)e^{-\pi\mathrm{tr}\,(T)},$$

and Fourier expansions are absolutely convergent. Furthermore, $a \in \mathbb{R}_{>0}$ because all of the terms in the summation are positive. Let $b = \pi n2^{-n}$. Then:

$$|f(\Omega)| \leq \sum_{T\in\mathcal{X}_n} |a(T;f)|\, e^{-2\pi\langle Y,T\rangle} \leq \sum_{T\in\mathcal{X}_n} |a(T;f)|\, e^{-\pi\mathrm{tr}\,(T)}e^{-\pi n2^{-n}\delta(Y)} = ae^{-b\delta(Y)}$$

□

It is known that the $\mathrm{Sp}_n(\mathbb{Z})$-invariant function attains a global maximum(Poor and Yuen (2000)), but a proof of that claim is beyond the scope of this thesis. I will take it for granted that such a maximum exists in the next chapter, though.

Chapter 4

The Semihull Theorem and its Corollaries

4.1 Definitions and General Lemmas

Definition 4.1.1. *Accounts*
Let $\Omega \in \mathcal{H}_n$. We say that a orbit in $\mathrm{Sp}_n(\mathbb{Z})\backslash\mathcal{H}_n$ is *accounted for by* Ω if the orbit contains $(\mathrm{Im}(\Omega))^{-1}$.

Definition 4.1.2. *Support and Semihull of a Siegel Cusp Form*
Let $f \in \mathcal{S}_k(\mathrm{Sp}_n(\mathbb{Z}))$. The *support of* f is:

$$\mathrm{supp}(f) = \{T \in \mathcal{X}_n : a(T; f) \neq 0\}$$

The *semihull of* f is:

$$\nu(f) = \overline{\langle \mathbb{R}_{\geq 1} \mathrm{supp}(f) \rangle}$$

Definition 4.1.3. *Kernel of a Matrix*
For $S \in \mathcal{P}_n$, define:

$$\mathcal{K}(S) = \overline{\langle \mathbb{R}_{\geq 1} [S] \rangle} \quad \textit{where} \quad [S] = S[\mathrm{GL}_n(\mathbb{Z})]$$

Lemma 4.1.1. *Kernel of a Positive Matrix*
Let $S \in \mathcal{P}_n$. Then $\mathcal{K}(S)$ is a kernel.

Proof. First, $\mathcal{K}(S)$ is closed under superconvex combinations by construction. Furthermore, $\mathcal{K}(S) \subset \overline{\mathcal{P}}_n$: $S[V] \in \mathcal{P}_n$ for all $V \in \mathrm{GL}_n(\mathbb{Z})$, so the set of their superconvex linear combinations lies in $\mathcal{P}_n$ since $\mathcal{P}_n$ is a cone. The closure of that set is $\mathcal{K}(S)$, so $\mathcal{K}(S) \subset \overline{\mathcal{P}}_n$.

For all $T \in \mathcal{K}(S)$, we have $0 < \delta(S) \leq \delta(T)$ because:

$$T = \sum_i \lambda_i S[V_i] \qquad \sum_i \lambda_i \geq 1 \qquad V_i \in \mathrm{GL}_n(\mathbb{Z})$$

so by superadditivity, and the fact that δ is a class function:

$$\delta(T) = \delta\left(\sum_i \lambda_i S[V_i]\right) \geq \sum_i \lambda_i \delta(S[V_i]) = \sum_i \lambda_i \delta(S) \geq \delta(S).$$

Thus, for any sequence $T_i \in \mathcal{K}(S)$, if $T_i \to T$, then $\delta(T) > 0$ by continuity of height functions, so $T \neq 0$, so $0 \notin \mathcal{K}(S)$.

We need to show that $\mathcal{P}_n \subset \mathbb{R}_{>0}\mathcal{K}(S)$. Since $\mathcal{K}(S)$ is invariant under the group action of $\mathrm{GL}_n(\mathbb{Z})$, we can replace S by $S[V]$, where the second diagonal entry of $S[V]$ is positive.

For every positive integer m, define a matrix $\gamma_m = me_{21} + I_n$. The entries of γ_m are integers by construction, and since γ_m is upper triangular, one can easily check that its determinant is 1. Thus, $\gamma_m \in \mathrm{GL}_n(\mathbb{Z})$, so $S[\gamma_m] \in \mathcal{K}(S)$. Note that $S[\gamma_m] = m^2 s_{22} e_{11} + \mathcal{O}(m)$.

Let $\alpha \in \mathbb{R}_{>0}$, $z \in \mathbb{Z}^n_{\text{prim}}$ be arbitrary. We know there exists $\gamma \in \mathrm{GL}_n(\mathbb{Z})$ such that $\gamma e_1 = z$, and that in turn means $e_{11}[\gamma'] = \gamma e_1 e_1' \gamma' = (\gamma e_1)(\gamma e_1)' = zz'$. Thus:

$$S[\gamma_m \gamma'] = S[\gamma_m][\gamma'] = (m^2 s_{22} e_{11} + \mathcal{O}(m))[\gamma'] = m^2 s_{22} zz' + \mathcal{O}(m)$$

We know that $S[\gamma_m \gamma'] \in \mathcal{K}(S)$ because $\gamma_m \gamma' \in \mathrm{GL}_n(\mathbb{Z})$. Furthermore, since $\mathcal{K}(S)$ is closed under superconvex combinations, it must contain the matrices:

$$S + \frac{\alpha}{m^2 s_{22}} S[\gamma_m \gamma'] = S + \alpha zz' + \mathcal{O}(1/m)$$

and since $\mathcal{K}(S)$ is a closed in the topological sense, it contains their limit point:

$$\lim_{m \to \infty} \left(S + \frac{\alpha}{m^2 s_{22}} S[\gamma_m \gamma'] \right) = S + \alpha zz'$$

Thus, we see that for $S + \alpha zz' \in \mathcal{K}(S)$ for all positive, real α and for all integral vectors z.

Now, let T be any positive matrix. From the theory of the dyadic trace, we know that every positive matrix T can be written as:

$$T = \sum_{i=1}^{k} \alpha_i z_i z_i' \qquad (z_i \in \mathbb{Z}^n_{\text{prim}})$$

so every matrix in $S + \mathcal{P}_n$ can be written as

$$S + T = S + \sum_{i=1}^{k} \alpha_i z_i z_i' = S + \frac{1}{k} \sum_{i=1}^{k} k\alpha_i z_i z_i' = \sum_{i=1}^{k} \frac{1}{k} (S + k\alpha_i z_i z_i')$$

Since $S + k\alpha_i z_i z_i \in \mathcal{K}(S)$ and $\mathcal{K}(S)$ is closed under convex combinations, it follows that $S + T \in \mathcal{K}(S)$. Thus, we have $S + \mathcal{P}_n \subset \mathcal{K}(S)$.

Finally, let $P \in \mathcal{P}_n$ be arbitrary. The theory of positive matrices tells us that for sufficiently large r, $rP - S \in \mathcal{P}_n$. Thus, $S + (rP - S) = rP \in \mathcal{K}(S)$, and the result follows. □

Lemma 4.1.2. *Kernel Lemma*:
The semihull of a nonzero cusp form $f \in \mathcal{S}_k(\mathrm{Sp}_n(\mathbb{Z}))$ is a kernel.

Proof. It is clear that $\nu(f)$ is a semihull by construction, so we just need to show that $\mathcal{P}_n \subset \mathbb{R}_{>0} \cdot \nu(f) \subset \overline{\mathcal{P}}_n$, and that the closure of $\nu(f)$ doesn't contain 0.

Let $T \in \operatorname{supp} f$. By definition, $\mathcal{K}(T) \subset \nu(f)$, so by the previous lemma, we have $\mathcal{P}_n \subset \mathbb{R}_{>0}K(T) \subset \mathbb{R}_{>0}\nu(f)$. Furthermore, f is a cusp form, so $\operatorname{supp}(f) \subset \mathcal{P}_n$, so $\nu(f) \subset \overline{\mathcal{P}}_n$. Finally, we know that for all $T \in \operatorname{supp}(f)$, $\operatorname{tr}(T) \geq 1$. Thus, $\operatorname{tr}(T) \geq 1$ for all superconvex combinations of those matrices by linearity of the trace, so $\operatorname{tr}(U) \geq 1$ for all $U \in \nu(F)$. □

Lemma 4.1.3. *Semihull Valuation Property*
For any nonzero Siegel cusp forms f, g:

$$\nu(fg) = \nu(f) + \nu(g)$$

Proof. First, let $T \in \operatorname{supp}(fg)$. By definition of support, that means $a(T; fg) \neq 0$, so

$$0 \neq a(T; fg) = \sum_{T_1+T_2=T} a(T_1; f)a(T_2; g)$$

Thus, there must exist at least one pair of matrices $T_1 \in \operatorname{supp}(f)$, $T_2 \in \operatorname{supp}(g)$ such that $T_1 + T_2 = T$. This holds for all T, so:

$$\operatorname{supp}(fg) \subset \operatorname{supp}(f) + \operatorname{supp}(g) \Longrightarrow \nu(fg) \subset \nu(f) + \nu(g)$$

Thus, we only need to show that:

$$\nu(f) + \nu(g) \subset \nu(fg)$$

By the containment lemma, we can instead show:

$$\nu(fg)^{\sqcup} \cap \mathcal{P}_n(\mathbb{Q}) \subset (\nu(f) + \nu(g))^{\sqcup}.$$

Let R be any integral domain, let $R[[x]]$ be the ring of formal power series over R, and define $\operatorname{ord}_x : (R[[x]]\backslash\{0\}) \longrightarrow \mathbb{Z}_{\geq 0}$ by

$$\operatorname{ord}_x \left(\sum_{n\in\mathbb{Z}_{\geq 0}} a_n x^n \right) = \min\{n : a_n \neq 0\}.$$

Since

$$\left(\sum_{n\in\mathbb{Z}_{\geq 0}} a_n x^n \right) \left(\sum_{n\in\mathbb{Z}_{\geq 0}} b_n x^n \right) = \sum_{n\in\mathbb{Z}_{\geq 0}} \left(\sum_{i+j=n} a_i b_j \right) x^n,$$

it follows that for any nonzero $h, k \in R[[x]]$:

$$\operatorname{ord}_x(hk) = \operatorname{ord}_x(h) + \operatorname{ord}_x(k)$$

Now, let R be the ring of holomorphic functions on $\mathcal{H}_n$, and let $U \in \overline{\mathcal{P}}_n(\mathbb{Z})$. Any nonzero holomorphic Fourier series h with support in $\mathcal{X}_n^{\text{semi}}$,

$$h(\Omega) = \sum_{T\in\mathcal{X}_n^{\text{semi}}} a(T; h)\mathrm{e}\left(\langle T, \Omega\rangle\right),$$

can be transformed into a formal power series $h^U(x) \in R[[x]]$ like so:

$$h^U(x) = \sum_{T \in \mathcal{X}_n^{\text{semi}}} a(T;h)\mathrm{e}\left(\langle T, \Omega\rangle\right) x^{\langle T,U\rangle} = \sum_{n \in \mathbb{Z}_{\geq 0}} \left(\sum_{T:\langle T,U\rangle = n} a(T;h)\mathrm{e}\left(\langle T, \Omega\rangle\right) \right) x^n$$

The terms $\mathrm{e}\left(\langle T, \Omega\rangle\right)$ are linearly independent over $\mathbb{C}$, so any inner summation that has a nonzero term cannot cancel to 0. Thus, we have:

$$\mathrm{ord}_x(h^U) = \min \langle \mathrm{supp}(f), U\rangle = \inf \langle \nu(h), U\rangle = \min \langle \mu(h), U\rangle$$

for any holomorphic Fourier series h on $\mathcal{H}_n$. Now, let h be as above, let $k(\Omega)$ be a nonzero holomorphic Fourier series, and let $T \in \mathcal{X}_n$, so that:

$$a\left(T; hk\right) \mathrm{e}\left(\langle T, \Omega\rangle\right) x^{\langle T,U\rangle} =$$

$$\sum_{T_1+T_2=T} a\left(T_1; h\right) \mathrm{e}\left(\langle T_1, \Omega\rangle\right) x^{\langle T_1,U\rangle} a\left(T_2; k\right) \mathrm{e}\left(\langle T_2, \Omega\rangle\right) x^{\langle T_1,U\rangle}$$

This shows $h^U k^U = (hk)^U$.

We now go back to the original f, g. We were trying to show that:

$$\nu(fg)^{\sqcup} \cap \mathcal{P}_n(\mathbb{Q}) \subset (\nu(f) + \nu(g))^{\sqcup},$$

so let V be an arbitrary element of $\nu(fg) \cap \mathcal{P}_n(\mathbb{Q})$. Since V has rational entries, we can express V as $V = U/q$, where $U \in \mathcal{P}_n(\mathbb{Z})$ and q is a positive integer.

Furthermore, we know that:

$$\langle \nu(fg), V\rangle = q^{-1} \langle \nu(fg), U\rangle \geq 1 \Longrightarrow \langle \nu(fg), U\rangle \geq q$$

and if we specialize to the minimal value, we get:

$$q \leq \min \langle \nu(fg), U\rangle = \mathrm{ord}_x((fg)^U) = \mathrm{ord}_x(f^U g^U)$$

The first equality follows from the definition of the order function on augmented Fourier series, and the second follows from the argument about h, k above.

The valuation property of power series gives

$$\mathrm{ord}_x(f^U g^U) = \mathrm{ord}_x(f^U) + \mathrm{ord}_x(g^U).$$

The summands on the right can be rewritten as:

$$\mathrm{ord}_x(f^U) = \min \langle \nu(f), U\rangle \qquad \mathrm{ord}_x(g^U) = \min \langle \nu(g), U\rangle$$

The sum of those two minima is the same as the minima of the sum of the sets, so:

$$q \leq \min \langle \nu(f) + \nu(g), U\rangle \Longrightarrow 1 \leq \min \langle \nu(f) + \nu(g), V\rangle$$

$$\Longrightarrow V \in (\nu(f) + \nu(g))^{\sqcup}$$

□

4.2 Main Lemmas and the Semihull Theorem

Lemma 4.2.1. *Main Lemma*
Let $f \in \mathcal{S}_k(\mathrm{Sp}_n(\mathbb{Z}))$ be a nonzero cusp form. Let $\Omega_o = X_o + iY_o$ maximize the $\mathrm{Sp}_n(\mathbb{Z})$ invariant function of f. Take matrices $A \in \mathcal{X}_n$ and $U \in \overline{\mathcal{P}}_n(\mathbb{Z})$ for which $\langle A, U\rangle \leq \inf \langle \mathrm{supp}(f), U\rangle$. Then $\langle A, U\rangle \leq \left\langle \frac{k}{4\pi} Y_o^{-1}, U\right\rangle$.

Proof. There exists a complex upper half plane

$$N = \{z \in \mathbb{C} : \mathrm{Im}(z) > -2\epsilon\}$$

such that as z varies in N, the resulting matrices $\Omega_z = \Omega_o + Uz$ vary in a neighborhood of Ω_o in $\mathcal{H}_n$. Here $\epsilon > 0$ depends on Ω_o. The function $f(\Omega_z)$ can be viewed as a holomorphic function of a single complex variable z in the neighborhood N.

The Fourier series:

$$f(\Omega_z) = \sum_{T \in \mathrm{supp}(f)} a(T; f) \mathrm{e}\left(\langle T, \Omega_o\rangle\right) \mathrm{e}\left(\langle T, U\rangle z\right)$$

converges absolutely on $\mathcal{H}_n$, and we use this to express the Fourier series as a power series of $q = e^{2\pi i z}$. The assumption that $\langle A, U\rangle \leq \langle T, U\rangle$ for all $T \in \mathrm{supp}(f)$ shows that the lowest power of q in the power series is at least $\langle A, U\rangle$, so:

$$f(\Omega_z) = \sum_{m \geq \langle A, U\rangle} \left(\sum_{T : \langle T, U\rangle = m} a(T; f) \mathrm{e}\left(\langle T, \Omega_o\rangle\right) \right) q^m$$

Since each power m of q is at least $\langle A, U\rangle$, divide by $\mathrm{e}\left(\langle A, \Omega_z\rangle\right) = \mathrm{e}\left(\langle A, \Omega_o\rangle\right) q^{\langle A, U\rangle}$ to get a new power series in q, denoted $g(q)$:

$$g(q) = \frac{f(\Omega_z)}{\mathrm{e}\left(\langle A, \Omega_z\rangle\right)} = \frac{1}{\mathrm{e}\left(\langle A, \Omega_o\rangle\right)} \sum_{m \geq 0} \left(\sum_{T : \langle T - A, U\rangle = m} a(T; f) \mathrm{e}\left(\langle T, \Omega_o\rangle\right) \right) q^m$$

Even after dividing, we have not introduced any poles. We can think of $g(q)$ as a holomorphic function of q on the open punctured disk of radius $e^{2\pi 2\epsilon}$ (since N is equivalent to the punctured disk in local q-coordinates). Furthermore, since there is no pole at the origin, the power series representation of $g(q)$ allows us to extend g holomorphically from the punctured disk to the full disk of radius $e^{2\pi 2\epsilon}$. Consequently, g is holomorphic on the closed disk of radius $e^{2\pi\epsilon}$.

By the Maximum Principle, g on the closed disk achieves its absolute maximum on the boundary. Thus, there must exist $q_o = e^{2\pi i z_o}$ for which $|g(1)| \leq |g(q_o)|$. Note that $z_o = x_o - i\epsilon$ for some x_o.

In terms of the original variables, we have shown that:

$$\left| \frac{f(\Omega_o)}{\mathrm{e}\left(\langle A, \Omega_o\rangle\right)} \right| \leq \left| \frac{f(\Omega_o + Uz_o)}{\mathrm{e}\left(\langle A, \Omega_o + Uz_o\rangle\right)} \right|$$

Let M be the maximal value of $(\det Y)^{k/2}|f(\Omega)|$ over $\mathcal{H}_n$. Note that $M > 0$ because f is nonzero. Also, let $\Omega_1 = \Omega_o + Uz_o$. We can rewrite the display in Idea 7 as:

$$(\det Y_o)^{-k/2}\frac{M}{|\mathrm{e}\left(\langle A, \Omega_o\rangle\right)|} \leq (\det(Y_o - U\epsilon))^{-k/2}\left|\frac{\det(Y_1)^{k/2}f(\Omega_1)}{\mathrm{e}\left(\langle A, \Omega_o + Uz_o\rangle\right)}\right|$$

or

$$\det(Y_o)^{-k/2}\frac{M}{|\mathrm{e}\left(\langle A, \Omega_o\rangle\right)|} \leq (\det(Y_o - U\epsilon))^{-k/2}\left|\frac{\det(Y_1)^{k/2}f(\Omega_1)}{\mathrm{e}\left(\langle A, \Omega_1\rangle\right)}\right|.$$

By using the maximality of M, we can rewrite that display as:

$$\det(Y_o)^{-k/2}\frac{M}{|\mathrm{e}\left(\langle A, \Omega_o\rangle\right)|} \leq (\det(Y_o - U\epsilon))^{-k/2}\frac{M}{|\mathrm{e}\left(\langle A, \Omega_1\rangle\right)|}$$

Since $\mathrm{e}\left(\langle A, \Omega_1\rangle\right) = \mathrm{e}\left(\langle A, \Omega_o + Uz_o\rangle\right) = \mathrm{e}\left(\langle A, \Omega_o\rangle\right)\mathrm{e}\left(\langle A, Uz_o\rangle\right)$, we can write:

$$|\mathrm{e}\left(\langle A, Uz_o\rangle\right)|(\det(Y_o))^{-k/2}\frac{M}{|\mathrm{e}\left(\langle A, \Omega_o\rangle\right)|} \leq (\det(Y_o - U\epsilon))^{-k/2}\frac{M}{|\mathrm{e}\left(\langle A, \Omega_o\rangle\right)|}$$

$$\implies |\mathrm{e}\left(\langle A, Uz_o\rangle\right)|(\det(Y_o))^{-k/2} \leq (\det(Y_o - U\epsilon))^{-k/2}$$

Finally, we can move $(\det(Y_o))^{-k/2}$ to the right hand side. Compute that:

$$(\det(Y_o))^{k/2}\cdot(\det(Y_o - U\epsilon))^{-k/2} = (\det(Y_o^{-1}\cdot(Y_o - U\epsilon)))^{-k/2} = (\det(I_n - Y_o^{-1}U\epsilon))^{-k/2}$$

so we have:

$$|\mathrm{e}\left(\langle A, Uz_o\rangle\right)| \leq (\det(I_n - Y_o^{-1}U\epsilon))^{-k/2}$$

We know that $|\mathrm{e}\left(\langle A, Uz_o\rangle\right)| = e^{-2\pi\mathrm{Im}(z_o)}$, and $\mathrm{Im}(z_o) = -\epsilon$, so we can write:

$$e^{2\pi\epsilon} \leq (\det(I_n - Y_o^{-1}U\epsilon))^{-k/2}$$

Taking logarithms preserves order, so:

$$2\pi\left\langle A, U\right\rangle\epsilon \leq \frac{k}{2}\cdot\ln(\det(I_n - Y_o^{-1}U\epsilon))^{-1} \implies \langle A, U\rangle \leq \frac{k}{4\pi\epsilon}\cdot\ln(\det(I_n - Y_o^{-1}U\epsilon))^{-1}$$

We can expand the expression in terms of powers of ϵ. First, compute that:

$$\langle A, U\rangle \leq \frac{k}{4\pi\epsilon}\cdot\ln(1 + \epsilon\mathrm{tr}(Y_o^{-1}U) + \mathcal{O}(\epsilon^2))^{-1} = \frac{k}{4\pi\epsilon}\cdot\ln(1 + \epsilon\left\langle Y_o^{-1}, U\right\rangle + \mathcal{O}(\epsilon^2))^{-1}$$

Next, by the Taylor expansion of the logarithm:

$$\langle A, U\rangle \leq \frac{k}{4\pi\epsilon}\cdot\left(\epsilon\left\langle Y_o^{-1}, U\right\rangle + \mathcal{O}(\epsilon^2)\right) = \frac{k}{4\pi}\cdot\left\langle Y_o^{-1}, U\right\rangle + \mathcal{O}(\epsilon)$$

Finally, letting $\epsilon \to 0$, we can invoke continuity of the inner product to conclude that:

$$\langle A, U\rangle \leq \left\langle \frac{k}{4\pi}Y_o^{-1}, U\right\rangle$$

□

Theorem 4.1. *Semihull Theorem*
Let $f \in \mathcal{S}_k(\mathrm{Sp}_n(\mathbb{Z}))$ be a nonzero Siegel cusp form. Let $\Omega_o = X_o + iY_o$ maximize the $\mathrm{Sp}_n(\mathbb{Z})$-invariant function $\phi_f(\Omega)$ over $\mathcal{H}_n$. Then

$$\frac{k}{4\pi} Y_o^{-1} \in \nu(f)$$

Proof. By the containment lemma, we know that it suffices to show:

$$\nu(f)^{\sqcup} \cap \mathcal{P}_n(\mathbb{Q}) \subset (\frac{k}{4\pi} Y_o^{-1})^{\sqcup}$$

Let $V \in \nu(f)^{\sqcup} \cap \mathcal{P}_n(\mathbb{Q})$ be arbitrary. Since V is an element of $\mathcal{P}_n(\mathbb{Q})$, we can express V as $V = U/q$, where $U \in \mathcal{P}_n(\mathbb{Z})$ and $q \in \mathbb{Z}_{\geq 1}$.

Since $U = qV$ and $V \in \nu(f)^{\sqcup}$,

$$q \leq \min \langle \mathrm{supp}(f), U \rangle$$

Thus, there exists a matrix $A \in \mathrm{supp}(f)$ such that:

$$q \leq \langle A, U \rangle \leq \langle T, U \rangle \qquad \text{for all } T \in \mathrm{supp}(f)$$

By the main lemma, we know:

$$q \leq \langle A, U \rangle \leq \left\langle \frac{k}{4\pi} Y_o^{-1}, U \right\rangle,$$

and dividing by q yields

$$1 \leq \left\langle \frac{k}{4\pi} Y_o^{-1}, U/q \right\rangle = \left\langle \frac{k}{4\pi} Y_o^{-1}, V \right\rangle \Longrightarrow V \in (\frac{k}{4\pi} Y_o^{-1})^{\sqcup}.$$

□

Figure 4.1, below, is a caricature of the Semihull Theorem. The ambient space is $\mathcal{V}_n$ for some n. In reality, $\mathcal{V}_n$ is a $n(n+1)/2$-dimensional space, but that is obviously much harder to depict. The shaded area is $\nu(f)$ for some nonzero f. The dots and red stars form a lattice intended to represent $\mathcal{X}_n$, with the red stars obviously representing $\nu(f) \cap \mathcal{X}_n$. The green star is Y_o^{-1} for some $\Omega_o = X_o + iY_o$ that maximizes the invariant function of f.

4.3 Corollaries

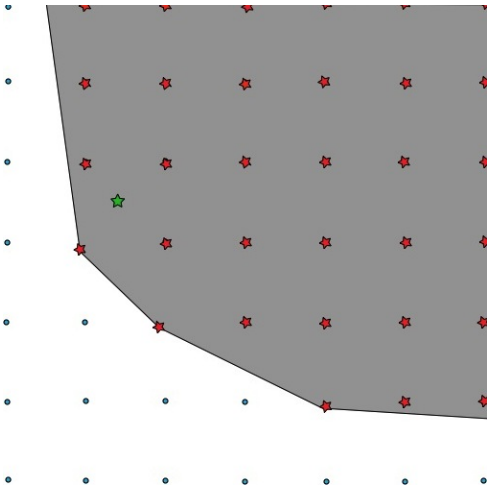
Corollary 4.1.1. *Vanishing Theorem*

Let $f \in \mathcal{S}_k(\mathrm{Sp}_n(\mathbb{Z}))$ be a Siegel cusp form. Let ν be a kernel that contains $\nu(f)$. If the set

$$\left\{ \Omega = X + iY \in \mathcal{H}_n : \frac{k}{4\pi} Y^{-1} \notin \nu \right\}$$

contains a fundamental domain for $\mathrm{Sp}_n(\mathbb{Z}) \backslash \mathcal{H}_n$, then $f = 0$.

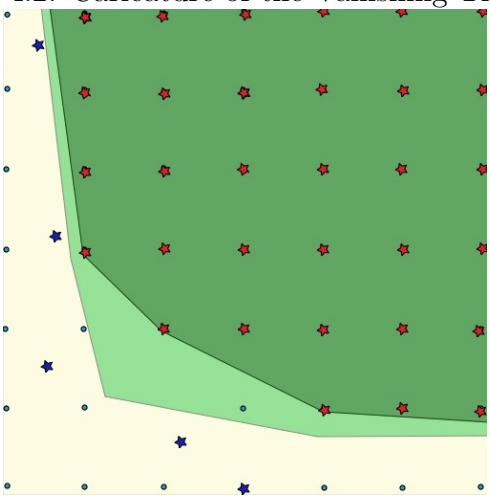
Figure 4.1: Caricature of the Semihull Theorem



Proof. Suppose $f \neq 0$. The $\mathrm{Sp}_n(\mathbb{Z})$-invariant function of f has a global maximum at some $\Omega_o = X_o + iY_o$. By the Semihull Theorem, $\frac{k}{4\pi} Y_o^{-1} \in \nu(f) \subset \nu$. Furthermore, this is true for all $\Omega \in \mathcal{H}_n$ that maximize the $\mathrm{Sp}_n(\mathbb{Z})$-invariant function of f, so it must be true for the entire $\mathrm{Sp}_n(\mathbb{Z})$-equivalence class of Ω_o. Thus, the displayed set does not contain a fundamental domain for $\mathrm{Sp}_n(\mathbb{Z})\backslash\mathcal{H}_n$ because it does not contain a representative for the class of Ω_o. □

Figure 4.2 is a caricature of the Vanishing Theorem. The dark green region represents $\nu(f)$ for some Siegel cusp form f, and the lighter green is an outer approximation ν of $\nu(f)$. The blue stars in the picture represent random elements of $\mathcal{H}_n$, although any point in the beige ambient space could have been chosen instead. The Vanishing Theorem says that if every orbit in $\mathrm{Sp}_n(\mathbb{Z})\backslash\mathcal{H}_n$ is accounted for by a blue star, then $f = 0$.

Figure 4.2: Caricature of the Vanishing Theorem



Corollary 4.1.2. *Extraction Theorem*
Let $f \in \mathcal{S}_k(\mathrm{Sp}_n(\mathbb{Z}))$, let ϕ be a height function, and let $\Omega_o = X_o + iY_o \in \mathcal{H}_n$ maximize the $\mathrm{Sp}_n(\mathbb{Z})$-invariant function of f. Define:

$$\mathcal{W}_{n,k} = \left\{ T \in \mathcal{X}_n : \phi(T) \leq \frac{k}{4\pi}\phi(Y_o^{-1}) \right\}$$

Then:

$$a(T; f) = 0 \; for \; all \; T \in \mathcal{W}_{n,k} \iff f = 0$$

Proof. The $\implies$ direction clearly holds. To prove the $\impliedby$ direction, we argue by contraposition.

Suppose f is nonzero. By the Semihull Theorem:

$$\frac{k}{4\pi}Y_o^{-1} \in \nu(f) = \overline{\langle \mathbb{R}_{\geq 1}\mathrm{supp}\,(f) \rangle}$$

The display above says that $\frac{k}{4\pi}Y_o^{-1}$ is a limit point of the set of superconvex combinations of $\mathrm{supp}\,(f)$, so we can find S of the form:

$$S = \sum_{T \in \mathrm{supp}\,(f)} a_T T \in \nu(f) \qquad \sum a_T \geq 1$$

that is arbitrarily close to $\frac{k}{4\pi}Y_o^{-1}$. Since f is a cusp form, its support contains only positive matrices, so S is positive because it is superconvex combination of positive matrices. By superadditivity, we have:

$$\phi(S) = \phi\left(\sum_{T \in \mathrm{supp}\,(f)} a_T T \right) \geq \sum_{T \in \mathrm{supp}\,(f)} a_T \phi(T).$$

Furthermore, since $T \in \mathrm{supp}\,(f)$, $\phi(T) \geq \inf \phi(\mathrm{supp}\,(f))$, so:

$$\phi(S) \geq \sum_T a_T \inf \phi(\mathrm{supp}\,(f)) = \inf \phi(\mathrm{supp}\,(f)) \sum a_T \geq \inf \phi(\mathrm{supp}\,(f))$$

Since S is arbitrarily close to $\frac{k}{4\pi}Y_o^{-1}$, S is positive and ϕ is continuous on $\mathcal{P}_n$, it follows that:

$$\inf \phi(\mathrm{supp}\,(f)) \leq \frac{k}{4\pi}\phi(Y_o^{-1})$$

Thus, if f is nonzero, there exists $T \in \mathrm{supp}\,(f)$ such that $a(T; f) \neq 0$ and $\phi(T) \leq \frac{k}{4\pi}\phi(Y_o^{-1})$, so if $\mathcal{W}_{n,k} \cap \mathrm{supp}\,(f) = \emptyset$, then $f = 0$. □

The Extraction Theorem is the tool that ultimately allows one to produce good bounds for the number of Fourier coefficients needed to study a Siegel cusp form. The Extraction Theorem says that we only need to know the Fourier coefficients T for which $\phi(T) \leq \phi(Y_o^{-1})$ to fully determine a Siegel cusp form. Since Y_o^{-1} is not unique, one can work a little harder and obtain a slightly better bound for the number

of Fourier coefficients needed by replacing $\phi(Y_o^{-1})$ with the infimum of $\phi(Y_o^{-1})$ over the $\mathrm{Sp}_n(\mathbb{Z})$-equivalence class of Ω_o. That is, one can compute the constant:

$$\phi_n = \sup_{\Omega \in \mathcal{H}_n} \inf_{\sigma \in \mathrm{Sp}_n(\mathbb{Z})} \phi(\mathrm{Im}(\sigma\Omega)^{-1})$$

and redefine:

$$\mathcal{W}_{n,k} = \left\{ T \in \mathcal{X}_n : \phi(T) \leq \frac{k}{4\pi}\phi_n \right\}$$

and the Extraction Theorem would still hold.

Siegel proved the Extraction Theorem for the special case $\phi = \mathrm{tr}$. Although useful from a theoretical standpoint, Siegel's theorem is highly impractical when $n > 1$. First, computing tr_n is very difficult, and one often has to replace tr_n with an upper bound that is not optimal. Second, the vanishing of a Fourier coefficient is a class property by Prop. 3.2.2., but the trace is not a class function. Both of those issues cause us to compute many unnecessary Fourier coefficients. Replacing the trace with the dyadic trace in the Extraction Theorem addresses both of those issues.

Recall that there exists a fundamental domain $\mathcal{F}_n$ satisfying $m(Y) \geq \frac{\sqrt{3}}{2}$ for all $X + iY \in \mathcal{H}_n$. For any such Y, we have:

$$w(Y^{-1}) \leq \frac{\langle Y^{-1}, Y \rangle}{m(Y)} = \frac{\mathrm{tr}\,(Y^{-1}Y)}{m(Y)} = \frac{n}{m(Y)} \leq n\frac{2}{\sqrt{3}}.$$

Thus, we can express a special case of the Extraction Theorem that makes no reference to the unknown matrix Ω_o or the unknown constant ϕ_n. Finally, we can save even more time by only checking $\mathrm{GL}_n(\mathbb{Z})$-equivalence classes, since w a class function.

It is important to note that before one can actually apply the following theorem, one needs to know which matrices in $\mathcal{X}_n$ have small dyadic trace. For small n, we can use Prop. 2.4.6 and Prop. 2. 4. 7, but the issue is more difficult for $n \geq 4$. See Poor and Yuen (2000), Poor and Yuen (2002) and Poor et al. for descriptions of algorithms that allow one to compute the dyadic trace of larger matrices, and tables of matrices in $\mathcal{X}_3, \mathcal{X}_4$ with small dyadic trace.

Corollary 4.1.3. *Special Extraction Theorem*
Let $f \in \mathcal{S}_k(\mathrm{Sp}_n(\mathbb{Z}))$. Let:

$$\mathcal{W}_{n,k} = \left\{ [T] : T \in \mathcal{X}_n, w(T) \leq \frac{nk}{2\sqrt{3}\pi} \right\} \qquad ([T] = T[\mathrm{GL}_n(\mathbb{Z})])$$

Then:

$$\mathcal{W}_{n,k} \cap \mathrm{supp}\,(f) = \emptyset \iff f = 0.$$

Chapter 5

Application: The Problem of Witt

5.1 Theta Series

Definition 5.1.1. *Type II Lattice*
A *Type II lattice* is an integral lattice that is even and unimodular. It is known that every Type II lattice has rank divisible by 8. The name *Type II* is borrowed from Conway and Sloane (1999).

Definition 5.1.2. *Theta Series*
Let n be a positive integer and let Λ be a lattice. The *degree n theta series of Λ* is:

$$\theta_\Lambda^{(n)}(\Omega) = \sum_{(v_1,\ldots,v_n)\in\Lambda^n} \mathrm{e}\left(\left\langle \frac{1}{2}[v_i \cdot v_j], \Omega \right\rangle\right)$$

where $[v_i \cdot v_j]$ denotes the Gram matrix of $(v_1, \ldots, v_n)$.

We restrict our attention to Type II lattices because it is known that if Λ is a Type II lattice of rank r, then:

$$\theta_\Lambda^{(n)} \in \mathcal{M}_{r/2}(\mathrm{Sp}_n(\mathbb{Z})).$$

See Freitag (1983) for a proof of that result.

Proposition 5.1.1. *Fourier Coefficients of Theta Series*
Let Λ be a Type II lattice, and let $\theta_\Lambda^{(n)}$ be the degree n theta series for Λ. Then:

$$a(T; \theta_\Lambda^{(n)}) = |\{(v_1, \ldots, v_n) \in \Lambda^n : [v_i \cdot v_j] = 2T\}|$$

Proof. As mentioned earlier, since Λ is a Type II lattice, the Gram matrices $[v_i \cdot v_j]$ are integral with even integers down the diagonal, and they're positive semidefinite by virtue of being Gram matrices, so $[v_i \cdot v_j]/2 \in \mathcal{X}_n^{\mathrm{semi}}$ for all $(v_1, \ldots, v_n) \in \Lambda^n$. Thus:

$$\theta_\Lambda^{(n)}(\Omega) = \sum_{(v_1,\ldots,v_n)\in\Lambda^n} \mathrm{e}\left(\left\langle \frac{1}{2}[v_i \cdot v_j], \Omega \right\rangle\right) = \sum_{T\in\mathcal{X}_n^{\mathrm{semi}}} \left(\sum_{(v_1,\ldots,v_n):[v_i\cdot v_j]=2T} \mathrm{e}\left(\langle T, \Omega\rangle\right) \right)$$

The inner summations are constant, so we can rewrite the double summation as:

$$\theta_\Lambda^{(n)}(\Omega) = \sum_{T \in \mathcal{X}_n^{\text{semi}}} |\{(v_1, \ldots, v_n) \in \Lambda^n : [v_i \cdot v_j] = 2T\}| \mathrm{e}\left(\langle T, \Omega \rangle\right)$$

The display above is a Fourier expansion, so by uniqueness of Fourier expansions:

$$a(T; \theta_\Lambda^{(n)}) = |\{(v_1, \ldots, v_n) \in \Lambda^n : [v_i \cdot v_j] = 2T\}|$$

□

Proposition 5.1.2. *Siegel's Φ-map and Theta Series*
Let Λ be a Type II lattice, let n be a positive integer and let $\theta_\Lambda^{(n)}$ be the degree n theta series for Λ. Then $\Phi(\theta_\Lambda^{(n)}) = \theta_\Lambda^{(n-1)}$.

Proof. Let $\Omega \in \mathcal{H}_{n-1}$. By definition of Φ, we have:

$$(\Phi\theta_\Lambda^{(n)})(\Omega) = \lim_{y \to +\infty} \sum_{(v_1, \ldots, v_n) \in \Lambda^n} \mathrm{e}\left(\frac{1}{2} \langle [v_i \cdot v_j], \Omega_y \rangle\right)$$

Recall that for any $(v_1, \ldots, v_n) \in \Lambda^n$:

$$[v_i \cdot v_j] = \begin{bmatrix} \langle v_1, v_1 \rangle & \langle v_1, v_2 \rangle & \ldots & \langle v_1, v_n \rangle \\ \langle v_2, v_1 \rangle & \langle v_2, v_2 \rangle & \ldots & \langle v_2, v_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle v_n, v_1 \rangle & \langle v_n, v_2 \rangle & \ldots & \langle v_n, v_n \rangle \end{bmatrix}$$

When we take the inner product with Ω_y, the terms $\langle v_1, v_n \rangle, \ldots, \langle v_{n-1}, v_n \rangle$ will be annihilated. Thus, we can view the inner product $\langle [v_i \cdot v_j], \Omega_y \rangle$ as the inner product of the Gram matrix for $(v_1, \ldots, v_{n-1})$ with Ω, with $\langle v_n, v_n \rangle \cdot iy$ added on. That is:

$$(\Phi\theta_\Lambda^{(n)})(\Omega) = \lim_{y \to +\infty} \sum_{(v_1, \ldots, v_{n-1}) \in \Lambda^{n-1}, v_n \in \Lambda} \mathrm{e}\left(\frac{1}{2} \langle [v_i \cdot v_j], \Omega \rangle\right) \cdot \mathrm{e}\left(\frac{1}{2} |v_n|^2 iy\right)$$

The limit doesn't depend on $(v_1, \ldots, v_{n-1})$, so we break up the summation and pull the limit inside:

$$(\Phi\theta_\Lambda^{(n)})(\Omega) = \sum_{(v_1, \ldots, v_{n-1}) \in \Lambda^{n-1}} \mathrm{e}\left(\frac{1}{2} \langle [v_i \cdot v_j], \Omega \rangle\right) \cdot \lim_{y \to +\infty} \sum_{v_n \in \Lambda} e^{-\pi |v_n|^2 y}$$

For all nonzero $v_n \in \Lambda$, the term $e^{-\pi |v_n|^2 y}$ goes to 0 as $y \to \infty$. For $v_n = 0$, the term is equal to 1 for all y. Thus, the limit of that summation is 1, so we have:

$$(\Phi\theta_\Lambda^{(n)})(\Omega) = \sum_{(v_1, \ldots, v_{n-1}) \in \Lambda^{n-1}} \mathrm{e}\left(\frac{1}{2} \langle [v_i \cdot v_j], \Omega \rangle\right) \cdot 1 = \theta_\Lambda^{(n-1)}$$

□

5.2 Problem of Witt

The theta function of a lattice Λ is a Siegel modular form if Λ is even and unimodular. The theory of lattices tells us that if Λ is unimodular and even, then Λ has rank divisible by 8. There is only one isometry class of even, unimodular lattices of rank 8, but there are two isometry classes of rank 16: $E_8 \oplus E_8$ and D_{16}^+.

The Problem of Witt asks, for which positive integers n is it true that

$$\theta^{(n)}_{E_8\oplus E_8}(\Omega) = \theta^{(n)}_{D_{16}^+}(\Omega) \qquad for\ all\ \Omega \in \mathcal{H}_n$$

Witt managed to prove that the equality holds for $n = 1, 2$, but could not prove the general result. Using the machinery developed in this thesis, one can answer the problem of Witt with relative ease.

5.2.1 The Method

We will use the Special Extraction Theorem. We know:

$$\Phi(\theta^{(n)}_{E_8\oplus E_8}(\Omega) - \theta^{(n)}_{D_{16}^+}(\Omega)) = \Phi(\theta^{(n)}_{E_8\oplus E_8}(\Omega)) - \Phi(\theta^{(n)}_{D_{16}^+}(\Omega)) = \theta^{(n-1)}_{E_8\oplus E_8} - \theta^{(n-1)}_{D_{16}^+}$$

The display tells us that $\theta^{(n)}_{E_8\oplus E_8} = \theta^{(n)}_{D_{16}^+} \implies \theta^{(n-1)}_{E_8\oplus E_8} = \theta^{(n-1)}_{D_{16}^+}$, so we can solve the Problem of Witt for all n by finding the smallest n for which the two theta series aren't equal. Also, the display tells us that the smallest nonzero difference is a cusp form. Thus, if we know that $\theta^{(n-1)}_{E_8\oplus E_8} = \theta^{(n-1)}_{D_{16}^+}$, we can determine whether $\theta^{(n)}_{E_8\oplus E_8} = \theta^{(n)}_{D_{16}^+}$ by checking:

$$a(T; \theta^{(n)}_{E_8\oplus E_8}) = a(T; \theta^{(n)}_{D_{16}^+}) \quad for\ all\ T : w(T) \le \frac{8n}{2\sqrt{3}}\pi$$

We compute the Fourier coefficients using Prop. 5.1.1.

5.2.2 The $n = 3$ Case

Witt proved that $\theta^{(2)}_{E_8\oplus E_8} = \theta^{(2)}_{D_{16}^+}$, so I will start by checking the $n = 3$ case. In order to determine whether $\theta^{(3)}_{E_8\oplus E_8} = \theta^{(3)}_{D_{16}^+}$, we need to check whether $a(T; \theta^{(3)}_{E_8\oplus E_8}) = a(T; \theta^{(3)}_{D_{16}^+})$ for all $[T] \in \mathcal{X}_n$ with $w(T) \le 4\sqrt{3}/\pi \approx 2.2$.

We can take T to be Minkowski-reduced, so T has dyadic trace:

$$w(T) = a + b + c - |d| - |e| - |f| + \min\{|d|, |e|, |f|\}$$

We know that T has positive integers on the diagonal because $T \in \mathcal{X}_3$, so:

$$w(T) \ge 1 + 1 + 1 - |d| - |e| - |f| + \min\{|d|, |e|, |f|\}$$

We need at least two of $|d|, |e|, |f|$ to be equal to $1/2$ to get the dyadic trace below 2.2. The remaining coordinate can be set equal to $1/2$ also, but we would get the

same dyadic trace either way. It turns out that it doesn't matter what we pick for $|d|, |e|, |f|$, because the resulting matrices are all Gram matrices for the well-known lattice A_3 (scaled by a factor of 1/2). That is,

$$\frac{1}{2}A_3 = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 1 & 0 \\ \frac{1}{2} & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & \frac{1}{2} & 0 \\ \frac{1}{2} & 1 & \frac{1}{2} \\ 0 & \frac{1}{2} & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 1 \end{bmatrix}$$

We will use the first Gram matrix in that display.

Furthermore, it is easy to see that there are no other matrices in $\mathcal{X}_n$ with a smaller dyadic trace. We cannot make the off-diagonal entries any bigger without making the diagonal entries bigger because of the Minkowski condition. But if we were to, say, set $c = 2$, then we would need to increase e and f for the matrix to still have a small dyadic trace. But if we make e, f bigger, then a, b would also have to increase because of the Minkowski conditions, and we will not be able to subtract enough to keep the dyadic trace below 2.2.

We want to know how many ordered triples $(v_1, v_2, v_3) \in {D_{16}^+}^3$ have Gram matrix:

$$\begin{bmatrix} \langle v_1, v_1\rangle & \langle v_1, v_2\rangle & \langle v_1, v_3\rangle \\ \langle v_1, v_2\rangle & \langle v_2, v_2\rangle & \langle v_2, v_3\rangle \\ \langle v_1, v_3\rangle & \langle v_2, v_3\rangle & \langle v_3, v_3\rangle \end{bmatrix} = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

Because D_{16}^+ is an even lattice, the smallest nonzero value that $\langle v, v\rangle$ can take is 2. As a result, we can restrict our attention to vectors $v \in D_{16}^+$ of minimal length. The minimal vectors of D_{16}^+ take the form $\pm e_i \pm e_j$, where $1 \le i < j \le 16$. There are $\binom{16}{2}$ ways to choose i, j, and 4 vectors for each choice of i, j, so D_{16}^+ has 480 minimal vectors.

By symmetry, we can set $v_1 = e_1 + e_2$ and multiply the total by 480. The first thing to do is count the number of v_2 we can choose that satisfy $\langle v_1, v_2\rangle = 1$. Clearly, if $\langle v_1, v_2\rangle = 1$, then $v_2 = e_1 \pm e_j$ or $v_2 = e_1 \pm e_j$ for some $j > 2$. There are 14 choices for j, and each choice of j corresponds to 4 different vectors, so we have 56 choices for v_2. Suppose we take $v_2 = e_1 + e_3$. We count the number of v_3 that satisfy $\langle v_2, v_3\rangle = 1 = \langle v_1, v_3\rangle$. By the argument from the previous paragraph, we see that there are 56 vectors whose inner product with v_2 is 1. Of those 56, 26 of those vectors take the form $e_1 \pm e_j$, $j > 3$, and those vectors are clearly not orthogonal to v_1, so we can discard them. Furthermore, 26 of the vectors take the form $e_3 \pm e_j$, $j > 3$, and those 26 *are* orthogonal to v_1. The remaining vectors are $e_1 + e_2$, $e_1 - e_2$, $e_3 + e_2$ and $e_3 - e_2$. Of these, only $e_1 - e_2$ is orthogonal to v_1, so we have a total of 27 choices for v_3. By symmetry, there are $480 \cdot 56 \cdot 27 = 725\,760$ ordered triples in ${D_{16}^+}^3$ that have the desired Gram matrix.

Next, we will compute $a(A_3; \theta^{(3)}_{E_8 \oplus E_8})$. To do so, we count the number of $(v_1, v_2, v_3) \in (E_8 \oplus E_8)^3$ with Gram matrix:

$$\begin{bmatrix} \langle v_1, v_1\rangle & \langle v_1, v_2\rangle & \langle v_1, v_3\rangle \\ \langle v_1, v_2\rangle & \langle v_2, v_2\rangle & \langle v_2, v_3\rangle \\ \langle v_1, v_3\rangle & \langle v_2, v_3\rangle & \langle v_3, v_3\rangle \end{bmatrix} = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

For any triple $(v_1, v_2, v_3) \in (E_8 \oplus E_8)^3$ with Gram matrix A_3, it is clear that no pair of vectors v_1, v_2, v_3 can be orthogonal (since $\langle v_1, v_2\rangle = \langle v_2, v_3\rangle = \langle v_1, v_3\rangle = 1$). Thus, the ordered triples we are looking for either lie in $0 \oplus E_8$ or $E_8 \oplus 0$, so we can count the triples of minimal vectors for E_8 that have Gram matrix A_3 and multiply the result by 2.

The minimal vectors of E_8 take one of two forms. There are *type one* minimal vectors that resemble the minimal vectors of D_8, i.e. vectors of the form $\pm(e_i \pm e_j)$. The number of type one minimal vectors is $\binom{8}{2} \cdot 2 \cdot 2$, because we choose two of the 8 coordinates, and then choose a sign for the first coordinate we chose, and then we choose a sign for the second coordinate we chose. Thus, there are 112 type one minimal vectors. There are also *type two* minimal vectors of the form $\left(\left(\pm\frac{1}{2}\right)^8\right)$, and have an even number of negative signs. We can pick the sign of the first 7 coordinates freely, and the sign of the final coordinate will then be determined. Thus, there are 2^7 type two minimal vectors.

First, suppose v_1, v_2, v_3 are all type one. We can set v_1 to be any of the 112 type one vectors. Say $v_1 = (1, 1, (0)^6)$. We need v_2 to have a 1 in exactly one of the coordinates v_1 has a 1. Thus, v_2 is either $e_1 \pm e_j$ or $e_2 \pm e_j$, where $j > 2$, so we have $2 \cdot 2 \cdot 6 = 24$ choices for v_2. Say we chose $v_2 = (1, 0, 1, (0)^6)$. Then v_3 is either of the form $e_1 \pm e_j$, where $j > 3$, or it is $(0, 1, 1, (0)^5)$. Overall, we have 29 568 triples of this type.

Next, we count the triples where exactly one of the three vectors is type two. Without loss of generality, suppose v_3 is type two. We will multiply this total by 3 at the end. We can choose v_1 and v_2 first; we know there are $112 \cdot 24$ ways to choose them. We need v_3's coordinates to have the same signs as the nonzero coordinates of v_1 and v_2. Thus, we know the first three coordinates of v_3 are $v_3 = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \ldots)$. The next four coordinates are free, and the sign of the final coordinate is determined by the first 7, so we have $2^4 = 16$ choices for v_3. Thus, there are 129 024 triples of this type.

Next, count the triples where exactly one of the three vectors is type one. Again, take $v_1 = e_1 + e_2$. We need v_2 to start with two positive signs. The next five coordinates are free, and the final sign of the final coordinate is determined. Thus, we have $2^5 = 32$ choices for v_2. Take $v_2 = \left(\left(\frac{1}{2}\right)^8\right)$. Finally, we choose v_3. We need the first two coordinates to be positive for $\langle v_1, v_3\rangle = 1$, and we need the remaining 6 coordinates to have exactly 2 negative signs for $\langle v_2, v_3\rangle = 1$. Thus, there are exactly $\binom{6}{2} = 15$ choices for v_3. Thus, there are $112 \cdot 32 \cdot 15 \cdot 3 = 161\,280$ triples of this type.

Finally, we count the triples where all three vectors are type 2. Set $v_1 = \left(\left(\frac{1}{2}\right)^8\right)$. We need v_2 to have exactly 2 negative signs, so we have $\binom{8}{2} = 28$ choices for v_2. Fix $v_2 = (-\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}^6)$. We need v_3 to have the same sign as v_1 and v_2 in all but 2 coordinates. Thus, the first two coordinates of v_3 have different signs, but we can choose which coordinate is positive. Exactly one of the last 6 coordinates is negative.

Thus, we have 43 008 triples of this type.

In sum, we found that the number of triples $(v_1, v_2, v_3) \in E_8^3$ with the desired Gram matrix is:

$$29\,568 + 129\,024 + 161\,280 + 43\,008 = 362\,880$$

Because we're doubling the total, and that number is exactly half of 725 760, the two theta functions are equal by the Extraction Theorem.

5.2.3 The $n = 4$ case

We repeat the process for $n = 4$. There are two Fourier coefficients that matter in this instance:

$$A_4 = \begin{bmatrix} 2 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 2 \end{bmatrix} \qquad D_4 = \begin{bmatrix} 2 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 1 & 1 & 2 & 0 \\ 1 & 1 & 0 & 2 \end{bmatrix}$$

See Poor and Yuen (2002) for a table of matrices in $\mathcal{X}_4$ with small dyadic trace. There are obviously other members of $[A_4]$ and $[D_4]$ we could have chosen, but these representatives make computations easiest. Note that for both matrices, the upper-left 3×3 submatrix is the matrix we just studied.

Now, consider $\theta^{(4)}_{D_{16}^+}$. We will compute the Fourier coefficient $a(A_4; \theta^{(4)}_{D_{16}^+})$. We can choose v_1, v_2, v_3 first, subject to the constraint $\langle v_1, v_2\rangle = \langle v_2, v_3\rangle = \langle v_1, v_3\rangle = 1$, in 725,760 different ways. However, if we chose $v_1 = (1, 1, 0^{14})$, $v_2 = (1, 0, 1, 0^{13})$, $v_3 = (0, 1, 1, 0^{13})$, there is no choice of v_4 that satisfies all 3 equalities. Thus, we have 26 choices for v_3 instead of 27. Say we chose $v_1 = (1, 1, 0^{14})$, $v_2 = (1, 0, 1, 0^{13})$, $v_3 = (1, 0, 0, 1, 0^{12})$. Then v_4 should have a 1 in the first coordinate, and 0's in the next 3 coordinates. There's a ± 1 somewhere in the last 12 coordinates, so we have 24 choices for v_4. Thus,

$$a(A_4; \theta^{(4)}_{D_{16}^+}) = 480 \cdot 56 \cdot 26 \cdot 24 = 16\,773\,120$$

Now we count $a(A_4; \theta^{(4)}_{E_8 \oplus E_8})$. First, we count the 4-tuples whose components are all type 1. We get the result almost immediately from the work we did for $a(A_4; \theta^{(4)}_{D_{16}^+})$. There are 112 choices for v_1 and 24 choices for v_2. Say we take $v_1 = (1, 1, 0^6)$, $v_2 = (1, 0, 1, 0^5)$. We cannot take $v_3 = (0, 1, 1, 0^5)$, since there will be no choice of v_4 that works given that choice, so v_3 is of the form $(1, 0, 0, \ldots)$, with a ± 1 somewhere in the remaining 5 coordinates. Thus, we have 10 choices for v_3. Say we take $v_3 = (1, 0, 0, 1, 0^4)$. It is clear that v_4 has a 1 in the first coordinate, and a ± 1 somewhere in the final four coordinates, so we have 8 choices for v_4. In sum, we have 215 040 4-tuples of this type.

Now count the 4-tuples with 3 type 1 vectors. Take v_1 to be the type 2 vector. We have 128 choices for v_1; suppose we take $v_1 = (\frac{1}{2}^8)$. We can choose any type 1 vector for v_2, as long as it has positive 1's in both nonzero coordinates. Thus,

there are $\binom{8}{2} = 28$ choices for v_2. If we take $v_2 = (1, 1, 0^6)$, then the entries of v_3 are nonnegative, and v_3 has a nonzero coordinate in exactly one of its first two coordinates, and one of its final six coordinates. Thus, there are $2 \cdot 6 = 12$ choices for v_3. Finally, v_4 can be $(0, 1, 1, 0^5)$ or it can have a 1 in the first coordinate and a 1 in one of the final 5 coordinates. Thus, there are 6 choices for v_4, so there are $128 \cdot 28 \cdot 12 \cdot 6 \cdot 4 = 1\,032\,192$ 4-tuples of this type.

Next count the 4-tuples with 2 type 1 vectors and 2 type 2 vectors. Take $v_1 = (1, 1, 0^6)$, $v_2 = (1, 0, 1, 0^5)$. Note that there are $112 \cdot 24$ ways we could have selected v_1 and v_2. We need v_3 to have positive signs in the first 3 coordinates, and an even number of minus in the remaining five coordinates. Thus, we have $2^4 = 16$ choices for v_3. Take $v_3 = (\frac{1}{2}^8)$. We need v_4 to have positive signs in its first three coordinates, and exactly 2 negative signs in the remaining 5. Thus, we have $\binom{5}{2} = 10$ choices for v_4, and $\binom{4}{2} = 6$ ways we could have assigned the types, so there are $112 \cdot 24 \cdot 16 \cdot 10 \cdot 6 = 2\,580\,480$ 4-tuples of this type.

Next count the 4-tuples with only one type 2 vector. We can choose v_1, v_2, v_3 in the same way as the final case in the $n = 3$ section. There are 43,008 ways to choose v_1, v_2, v_3. Say we chose $v_1 = (\frac{1}{2}^8)$, $v_2 = (-\frac{1}{2}^2, \frac{1}{2}^6)$, $v_3 = (-\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, \frac{1}{2}^5)$. We need the two nonzero coordinates of v_4 to match signs with the corresponding coordinates in the other three vectors. Thus, the first three coordinates of v_4 are all 0, and there are two 1's in the final 5 coordinates. Thus, we have $\binom{5}{2} = 10$ choices for v_4. Since any of the four vectors could have been the type 2 vector, there are $43\,008 \cdot 10 \cdot 4 = 1\,720\,320$ 4-tuples of this type.

Finally, we count the 4-tuples whose components are all type 2. Choose $v_1 = (\frac{1}{2}^8)$. Recall that the inner product of two type one vectors is 1 if and only if their coordinates have different signs in exactly 2 places. Thus, we have $\binom{8}{2}$ choices for v_2. Say we choose $v_2 = ((-\frac{1}{2})^2, (\frac{1}{2})^6)$. We need v_3 to have different signs from v_1 and v_2 in exactly two positions. Thus, one of the first two coordinates of v_3 must be negative, and one of the final 6 coordinates is negative. Say we choose $v_3 = (-\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, \frac{1}{2}^5)$. Finally, we need v_4 to differ in sign from the other three in exactly 2 spots each. Again, exactly one of the first two coordinates for v_4 can be positive. If we take the first coordinate to be positive, that forces $v_4 = (\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}^5)$. Otherwise, we can take the first coordinate to be negative, and the second coordinate to be positive. In that case, we have exactly one negative sign in the final six coordinates. If we place that negative sign in the third coordinate, then $v_3 = v_4$, which won't work. Thus, the final negative sign can be in any of the final five coordinates, so we have a total of 6 choices for v_4. In sum, we had 128 choices for v_1, 28 choices for v_2, 12 choices for v_3 and 6 choices for v_4, so there are 258 048 4-tuples of this type.

Overall, there are:

$$215\,040 + 1\,032\,192 + 2\,580\,480 + 1\,720\,320 + 258\,048 = 5\,806\,080$$

4-tuples of minimal vectors for E_8, so $a(A_4;\theta^{(4)}_{E_8\oplus E_8}) = 11\,612\,160$. Since:

$$a(A_4;\theta^{(4)}_{D_{16}^+}) - a(A_4;\theta^{(4)}_{E_8\oplus E_8}) = 16\,773\,120 - 11\,612\,160 = 5\,160\,960 \neq 0,$$

it follows that $\theta^{(4)}_{D_{16}^+} \neq \theta^{(4)}_{epe}$. Furthermore, this solves the Problem of Witt:

$$\theta^{(n)}_{D_{16}^+} = \theta^{(n)}_{E_8\oplus E_8} \Longleftrightarrow n \leq 3$$

Since we've solved the Problem of Witt, there's no reason to compute the Fourier coefficients for D_4. However, it might interest the reader to know that:

$$a(D_4;\theta^{(4)}_{D_{16}^+}) = 2\,096\,640 \qquad a(D_4;\theta^{(4)}_{E_8\oplus E_8}) = 7\,257\,600$$

so:

$$a(D_4;\theta^{(4)}_{D_{16}^+}) - a(D_4;\theta^{(4)}_{E_8\oplus E_8}) = -5\,160\,960 = a(A_4;\theta^{(4)}_{E_8\oplus E_8}) - a(A_4;\theta^{(4)}_{D_{16}^+})$$

The fact that those two differences are inverses of one another is not a coincidence. It is shown in Poor and Yuen (2002) that the Fourier coefficients of a Siegel cusp form $f \in \mathcal{S}_8(\mathrm{Sp}_4(\mathbb{Z}))$ satisfy $a(A_4) + a(D_4) = 0$. The difference of our two theta series is the well-known Schottky form, a Siegel cusp form.

Appendix A

Algorithms

This section gives a sketch of some of the algorithms referenced in the thesis.

A.1 Completing a Square

This algorithm allows us to find representations for positive matrices as positive linear combinations of dyadic squares of vectors in $\mathbb{R}^n$.

For $S = \begin{bmatrix} s_{11} & c' \\ c & \tilde{S} \end{bmatrix} \in \mathcal{P}_n$, define:

$$S_2 = \tilde{S} - s_{11}^{-1} cc' \qquad V_1 = \begin{bmatrix} 1 & s_{11}^{-1} c' \\ 0_{n-1} & I_{n-1} \end{bmatrix}$$

Then one readily verifies that:

$$S = \begin{bmatrix} s_{11} & 0'_{n-1} \\ 0_{n-1} & S_2 \end{bmatrix} [V_1] = s_{11} e_1 e_1' [V_1] + \begin{bmatrix} 0 & 0'_{n-1} \\ 0_{n-1} & S_2 \end{bmatrix} [V_1]$$

Furthermore, $S_2 \in \mathcal{P}_{n-1}$. To see this, compute that:

$$S[\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}] = s_{11}(x_1 + s_{11}^{-1} s_{12} x_2 + \ldots + s_{11}^{-1} s_{1n} x_n) + S_2[\begin{bmatrix} x_2 \\ \vdots \\ x_n \end{bmatrix}]$$

Thus, for all $\begin{bmatrix} x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^{n-1}$, if we set $x_1 = -\sum_{j=2}^{n} s_{11}^{-1} s_{1j}$, then by positive definiteness of S, we have:

$$0 \leq S[\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}] = s_{11}(x_1 - x_1) + S_2[\begin{bmatrix} x_2 \\ \vdots \\ x_n \end{bmatrix}] = S_2[\begin{bmatrix} x_2 \\ \vdots \\ x_n \end{bmatrix}]$$

As a result, we can repeat the process to obtain a representation of S_2 as $vv' + S_3$. The algorithm terminates when we get to S_{n-1}, and finding a dyadic representation for a matrix in $\mathcal{P}_1$ can be done by hand.

Note that if $S \in \mathcal{P}_n(\mathbb{Q})$, then the dyadic representation we get from completing the square will be defined over $\mathbb{Q}$.

A.2 Matrix Reduction

A.2.1 Legendre Reduction

The Legendre reduction algorithm takes matrices in $S \in \mathcal{P}_2$ as input, and outputs a unique, representative matrix in the same equivalence class.

Let $\begin{bmatrix} a & b \\ b & c \end{bmatrix} \in \mathcal{P}_2$. The Legendre reduction algorithm is:

1. First, check that $a \leq c$. If $a > c$, replace S by:

$$S[\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}] = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} c & -b \\ -b & a \end{bmatrix}$$

   and go to step 2.

2. Second, check that $2|b| \leq a$. If $2|b| > a$, define:

$$\lambda = \left\lfloor -\frac{b}{a} + \frac{1}{2} \right\rfloor$$

   and replace S by:

$$S[\begin{bmatrix} 1 & \lambda \\ 0 & 1 \end{bmatrix}] = \begin{bmatrix} a & b + \lambda a \\ b + \lambda a & c + 2\lambda b + \lambda^2 a \end{bmatrix}.$$

   Now $2|b| \leq a$, but we've made c smaller. If so, go back to step 1. Otherwise, go to step 3.

3. Finally, if $b < 0$, replace S by:

$$S[\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}] = \begin{bmatrix} a & -b \\ -b & c \end{bmatrix}.$$

The output is clearly in the same $\mathrm{GL}_n(\mathbb{Z})$-equivalence class as S, and satisfies:

$$0 \leq 2|b| \leq a \leq c.$$

Note that the last step is optional, but performing it ensures we have a unique class representative.

A.2.2 Minkowski Reduction

Definition A.2.1. *Minkowski Reduced Matrix*
Let $S \in \mathcal{P}_n$. We say S is *Minkowski reduced* if:

1. $S[v] \geq s_{kk}$ for all $k \in \{1, \ldots, n\}$ and all $v \in \mathbb{Z}^n$ such that the tail $(x_k, \ldots, x_n)$ is primitive.

2. $s_{k,k+1} \geq 0$ for all $k \in \{1, \ldots, n-1\}$.

A Minkowski-reduced matrix has a geometric interpretation. For a Gram matrix of some lattice to be Minkowski-reduced, it must be generated by a basis $\beta = \{v_1, \ldots, v_r\}$ with the property that $S[v_1] = m(S)$, and each successive v_i minimizes $S[\cdot]$ as much as possible. See Conway and Sloane (1999) for a more detailed description of a Minkowski-reduced basis.

There exist various reduction algorithms that take arbitrary matrices $S \in \mathcal{P}_n$ as input, and output a matrix $\tilde{S} \in [S]$ that is Minkowski-reduced. See Zhang et al. (2011) for examples. I can not present a useable Minkowski-reduction algorithm, because doing so would require an additional chapter on reduction algorithms.

For a 3×3 Minkowski-reduced matrix:

$$S = \begin{bmatrix} a & d & e \\ d & b & f \\ e & f & c \end{bmatrix}$$

the Minkowski-reduced conditions are equivalent to the following system of inequalities:

$$0 < a \leq b \leq c, \quad 2|d|, 2|e| \leq a, \quad 2|f| \leq b,$$
$$2(\pm d \pm e \pm f) \leq a + b \qquad (odd\ \#\ of\ minus\ signs)$$

Bibliography

Bhatia, Rajendra. *Positive Definite Matrices*. Princeton Series in Applied Mathematics, 2007.

Conway, John H. and Sloane, Neil J. A. *Sphere Packings, Lattices and Groups*. Springer-Verlag, 1999.

Diamond, Fred and Shurman, Jerry. *A First Course in Modular Forms*. Graduate Texts in Mathematics. Springer, 2005.

Freitag, E. *Siegelsche Modulfunktionen*. Springer Verlag, 1983.

Geer, Gerard Van Der. "Siegel Modular Forms." Nordfjordeid Summer School on Modular Forms and their Applications, 2007.

Hulek, K. and Sankaran, G.K. "The Geometry of Siegel Modular Varieties." (1998).

Igusa, Jun-ichi. *Theta Functions*. Springer-Verlag, 1972.

Klingen, Helmut. *Introductory Lectures on Siegel Modular Forms*. Cambridge Studies in Advanced Mathematics, 1990.

Poor, Cris, Shurman, Jerry, and Yuen, David S. "Computing Siegel Modular Forms.", ????

Poor, Cris and Yuen, David S. "Linear dependence among Siegel modular forms." *Math. Annalen* .318: 205–234, (2000).

———. "Restrictions of Siegel Modular Forms to Modular Curves." (2002).

Zhang, Wen, Qiao, Sanzheng, and Wei, Yimin. "Practical Algorithms for Constructing HKZ and Minkowski-reduced Bases." (2011).