University of Nevada, Reno

Witt Rings and Algebraic Knot Concordance

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Mathematics

by

Clinton Reece

Dr. Swatee Naik/Thesis Advisor

May, 2013

University of Nevada, Reno
Statewide • Worldwide

THE GRADUATE SCHOOL

We recommend that the thesis
prepared under our supervision by

CLINTON REECE

entitled

Witt Rings And Algebraic Knot Concordance

be accepted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE

Swatee Naik, Ph.D., Advisor

Christopher Herald, Ph.D., Committee Member

Sushil Louis, Ph.D., Graduate School Representative

Marsha H. Read, Ph. D., Dean, Graduate School

May, 2013

Abstract

The (knot) concordance group was introduced by Fox and Milnor in 1966. Since then some progress has been made studying both slice knots and concordance, though even some basic questions remain unanswered. We discuss the construction of the concordance group starting from knotted circles in S^3. Of the two related theories of concordance, we focus on the smooth setting rather than the topological setting.

In 1969, Levine classified higher dimensional analogues of concordance and provided great insight into the classical version mentioned above. His scheme allows tools developed in the purely algebraic setting of the theory of symmetric bilinear forms (specifically, Witt theory) to be applied to questions of concordance through use of an algebraic concordance group. We develop this notion of algebraic concordance alongside the corresponding notions from Witt theory.

List of Figures

Contents

Chapter 1

Introduction

This work grew out of a desire to understand the algebraic underpinnings behind knot concordance. Both *bilinear algebra* and *classical knot theory* are accessible topics for first or second year graduate students, so the essential elements of these are given in the ensuing two chapters. Afterwards, *knot concordance* is introduced and shown to have the algebraic structure of a group – though this group turns out to be difficult to work with directly. Therefore, we follow J. Levine's novel program, described in [7], to make a series of connections from knot concordance to *algebraic concordance* and ending in the setting of the theory of *Witt rings*.

Chapter 2

Bilinear Algebra

For our purposes, *bilinear algebra* will be thought of as the study of bilinear mappings defined on vector spaces over a field. With some care, the results given in this chapter may be extended to the study of module homomorphisms defined on modules over a commutative ring, as beautifully presented in [12] or [14]. For modules over a principal ideal domain (PID) the extension is essentially trivial. Consequently, despite some later discussion of the ring of integers $\mathbb{Z}$, we limit our focus to fields to avoid inessential abstraction. Fields of particular interest are $\mathbb{R}$ (the real numbers), $\mathbb{Q}$ (the rational numbers), $\mathbb{Q}_p$ (the p-adic numbers, with p a prime), and $\mathbb{F}_p$ (finite fields with p elements).

As with linear algebra, the case of $\mathbb{F}_2$ (in the general theory, rings of characteristic 2) is somewhat special and results must often be tailored specifically for this situation. As such, $\mathbb{F}$ is assumed to be of characteristic not equal to 2 unless explicitly stated otherwise.

2.1 Basic Definitions

2.1.1 Bilinear Forms

Definition 2.1.1. Let V be a finite dimensional vector space over a field $\mathbb{F}$. A *bilinear form* on V is a function $\mathfrak{b} : V \times V \to \mathbb{F}$, where $\mathfrak{b}(x, \bullet) : V \to \mathbb{F}$ is linear for each fixed $x \in V$, and $\mathfrak{b}(\bullet, x) : V \to \mathbb{F}$ is linear for each fixed $y \in V$. That is,

$$\mathfrak{b}(ax + by, z) = a\mathfrak{b}(x, z) + b\mathfrak{b}(y, z)$$

and

$$\mathfrak{b}(x, by + cz) = b\mathfrak{b}(x, y) + c\mathfrak{b}(x, z)$$

for every $x, y, z \in V$ and $a, b, c \in \mathbb{F}$. A bilinear form is less commonly called a *bilinear functional.*

Remark 2.1.2. Alternatively, one can separate the two conditions into three conditions: additivity in the first slot, additivity in the second slot, and scalar multiplication. This is accomplished by letting the scalars a, b, c take on combinations of the values 0 or 1. That is, if $a = b = 1$,

$$\mathfrak{b}(x + y, z) = \mathfrak{b}(x, z) + \mathfrak{b}(y, z)$$

and if $b = c = 1$,

$$\mathfrak{b}(x, y + z) = \mathfrak{b}(x, y) + \mathfrak{b}(x, z)$$

and if $b = 0$ (left equality), or if $c = 0$ with relabeling $b = a$ (right equality),

$$\mathfrak{b}(ax, y) = a\mathfrak{b}(x, y) = \mathfrak{b}(x, ay)$$

Definition 2.1.3. Let $\mathfrak{b}$ be a bilinear form on V.

(i) If $\mathfrak{b}(x, y) = \mathfrak{b}(y, x)$ for every $x, y \in V$, then $\mathfrak{b}$ is called *symmetric.*

(ii) If $\mathfrak{b}(x, y) = -\mathfrak{b}(y, x)$ for every $x, y \in V$, then $\mathfrak{b}$ is called *skew-symmetric.*

(iii) If $\mathfrak{b}(x, x) = 0$ for every $x \in V$, then $\mathfrak{b}$ is called *alternating.*

(iv) If $\mathfrak{b}(x, x) = 0$ for a nonzero $x \in V$, then the vector x is called *isotropic.* The bilinear form $\mathfrak{b}$ is called *isotropic* if V contains an isotropic vector. If all (nonzero) vectors are isotropic, then $\mathfrak{b}$ is called *totally isotropic.*

(v) If $\mathfrak{b}(x, x) \neq 0$ for $x \in V$, then the vector x is called *anisotropic.* The bilinear form $\mathfrak{b}$ is called *anisotropic* if V does not contain an isotropic vector (that is, if all vectors are anisotropic).

Example 2.1.4. This will be a continuing example. Let $V = \mathbb{R}^n$, and let $\mathbb{F} = \mathbb{R}$. Then the usual scalar (dot) product on $\mathbb{R}^n$

$$\mathfrak{b}(\mathbf{x}, \mathbf{y}) = \mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^{n} x_i y_i$$

is a symmetric bilinear form:

$$\begin{aligned}
\mathfrak{b}(a\mathbf{x}+b\mathbf{y},\mathbf{z}) &= \sum_{i=1}^{n}(ax_i+by_i)z_i = \sum_{i=1}^{n}ax_iz_i+\sum_{i=1}^{n}by_iz_i = a\mathfrak{b}(\mathbf{x},\mathbf{z})+b\mathfrak{b}(\mathbf{y},\mathbf{z})\\
\mathfrak{b}(\mathbf{x},b\mathbf{y}+c\mathbf{z}) &= \sum_{i=1}^{n}x_i(by_i+cz_i) = \sum_{i=1}^{n}bx_iy_i+\sum_{i=1}^{n}cx_iz_i = b\mathfrak{b}(\mathbf{x},\mathbf{y})+c\mathfrak{b}(\mathbf{x},\mathbf{z})\\
\mathfrak{b}(\mathbf{x},\mathbf{y}) &= \sum_{i=1}^{n}x_iy_i = \sum_{i=1}^{n}y_ix_i = \mathfrak{b}(\mathbf{y},\mathbf{x})
\end{aligned}$$

This demonstrates in our present language the well-known result that finite sums are $\mathbb{R}$-bilinear.

Example 2.1.5. Let $V=\mathbb{R}^n$ as a vector space over $\mathbb{F}=\mathbb{R}$ and define

$$\mathfrak{b}(\mathbf{x},\mathbf{y}) = \sum_{i<j}x_iy_j - x_jy_i.$$

Then $\mathfrak{b}$ is evidently a bilinear form by following similar steps as in the previous example. However, this bilinear form $\mathfrak{b}$ is skew-symmetric, as can be seen by factoring out a negative:

$$\mathfrak{b}(\mathbf{x},\mathbf{y}) = \sum_{i<j}x_iy_j - x_jy_i = -\sum_{i<j}x_jy_i - x_iy_j = -\mathfrak{b}(\mathbf{y},\mathbf{x})$$

Since there are generally many bilinear forms that may be defined on a given finite dimensional vector space V, we will want to have a way to discuss the two together as a pair.

Definition 2.1.6. If $\mathfrak{b}$ is a bilinear form on the finite dimensional vector space V, then the pair $(V,\mathfrak{b})$ is called a *bilinear form space* (or simply, *bilinear space*) over the field $\mathbb{F}$.

Definition 2.1.7. Let $\mathfrak{b}$ be either a symmetric or alternating bilinear form on V. Two elements x and y of a bilinear space are called *orthogonal* if $\mathfrak{b}(x,y)=0=\mathfrak{b}(y,x)$.

Example 2.1.8. Orthogonal vectors in $\mathbb{R}^n$ with the scalar product symmetric bilinear form described in Example 2.1.4 are orthogonal in the above sense.

2.1.2 Inner Products

Definition 2.1.9. A bilinear form $\mathfrak{b}$ is an *inner product* if for each $\mathbb{F}$-linear map $\varphi : V \to \mathbb{F}$ the following two conditions hold:

(i) There is a unique element $x_0 \in V$ so that the homomorphism $y \mapsto \mathfrak{b}(x_0,y)$ is equal to φ.

(ii) There is a unique element $y_0 \in V$ so that the homomorphism $x \mapsto \mathfrak{b}(x, y_0)$ is equal to φ.

Said differently, the two homomorphisms $x_0 \mapsto \mathfrak{b}(x_0, \bullet)$ and $y_0 \mapsto \mathfrak{b}(\bullet, y_0)$ from V to the dual space $V^* = \mathrm{Hom}_{\mathbb{F}}(V, \mathbb{F})$ are required to be bijective.

Sometimes inner products are called *nondegenerate*, *regular*, or *nonsingular* bilinear forms. We will not use this terminology (see Lemma 2.2.8 for justification for its use elsewhere.)

Remark 2.1.10. The notions of *symmetric*, *skew-symmetric*, and *alternating* also hold for inner products. Every alternating inner product is necessarily skew-symmetric: $0 = \mathfrak{b}(x + y, x + y) = \mathfrak{b}(x, x) + \mathfrak{b}(x, y) + \mathfrak{b}(y, x) + \mathfrak{b}(y, y) = \mathfrak{b}(x, y) + \mathfrak{b}(y, x)$. If the field $\mathbb{F}$ is not of characteristic 2, the converse holds since for every $x \in V$, $\mathfrak{b}(x, x) = -\mathfrak{b}(x, x)$ implies that $\mathfrak{b}(x, x) = 0$.

Definition 2.1.11. If $\mathfrak{b}$ is an inner product on a vector space V over a field $\mathbb{F}$, then the bilinear form space $(V, \mathfrak{b})$ is called an *inner product space* over $\mathbb{F}$.

Example 2.1.12. The scalar product symmetric bilinear form (see Example 2.1.4) is actually a symmetric inner product. Thus the symmetric bilinear space $(\mathbb{R}^n, \mathfrak{b})$ is actually a symmetric inner product space.

Proposition 2.1.13. *If V is an inner product space, an element x is orthogonal to every $y \in V$ if and only if $x = 0$. If $\mathbb{F}$ is a field, the converse statement is true: If $(V, \mathfrak{b})$ is a bilinear space over a field such that only the zero vector in V is orthogonal to every $x \in V$, then $\mathfrak{b}$ is an inner product on V.*

2.2 Bilinear Space Structure

Let V be a finite dimensional vector space with dimension n so that V has a basis consisting of n elements. Recall from linear algebra that the positive integer n is called the *dimension* of V and is denoted $\dim V$.

2.2.1 Isometries of Bilinear Forms

In linear algebra, vector space isomorphisms are bijective linear transformations, and vector spaces are classified up to isomorphism by their dimension. To get a analogous notion of isomorphism between bilinear spaces, we will impose another condition on these vector space isomorphisms that in some sense preserves the behavior of the associated bilinear forms.

Definition 2.2.1. Let $(V_1, \mathfrak{b}_1)$ and $(V_2, \mathfrak{b}_2)$ be bilinear spaces.

(i) An *isometry* from $(V_1, \mathfrak{b}_1)$ to $(V_2, \mathfrak{b}_2)$ is an $\mathbb{F}$-linear isomorphism $T : V_1 \to V_2$ such that $\mathfrak{b}_1(x, y) = \mathfrak{b}_2(T(x), T(y))$. We will frequently abuse the notation and simply write $T : \mathfrak{b}_1 \to \mathfrak{b}_2$.

(ii) The bilinear space $(V_1, \mathfrak{b}_1)$ is *isometric* to $(V_2, \mathfrak{b}_2)$ if there exists an isometry $T : \mathfrak{b}_1 \to \mathfrak{b}_2$. If such an isometry exists, write $\mathfrak{b}_1 \simeq \mathfrak{b}_2$.

(iii) The triple $(V, \mathfrak{b}, T)$, where $T : V \to V$ is an isometry with respect to $\mathfrak{b}$, is called an *isometric structure* over $\mathbb{F}$ if $n = \dim(V)$ is even.

Remark 2.2.2. The set $\text{Bil}(V)$ of bilinear forms on V is a vector space over the field $\mathbb{F}$. The sets $\text{Alt}(V)$ and $\text{Sym}(V)$ of alternating and symmetric bilinear forms on V, respectively, are subspaces of $\text{Bil}(V)$.

Proposition 2.2.3. *The isometry relation, $\simeq$, is an equivalence relation on* $\text{Bil}(V)$*.*

Proof. Let $(V_1, \mathfrak{b}_1)$ and $(V_2, \mathfrak{b}_2)$ be arbitrary bilinear spaces. Then $\mathfrak{b}_1 \simeq \mathfrak{b}_1$ by way of the identity linear transformation $\text{id} : V_1 \to V_2$, so $\simeq$ is reflexive. If $\mathfrak{b}_1 \simeq \mathfrak{b}_2$, then there exists an isomorphism $T : V_1 \to V_2$ such that $\mathfrak{b}_1(x, y) = \mathfrak{b}_2(T(x), T(y))$. In particular, there also exists an inverse $T^{-1} : V_2 \to V_1$ such that $\mathfrak{b}_1(T^{-1}(x), T^{-1}(y)) = \mathfrak{b}_2(TT^{-1}(x), TT^{-1}(y)) = \mathfrak{b}_2(x, y)$, so $\mathfrak{b}_2 \simeq \mathfrak{b}_1$. Thus $\simeq$ is symmetric. Finally, details of transitivity are ommitted for brevity but readily follow from the composition of isomorphisms being an isomorphism. □

Proposition 2.2.4. *If $(V_1, \mathfrak{b}_1)$ and $(V_2, \mathfrak{b}_2)$ are isometric, then $\mathfrak{b}_1$ is symmetric (skew-symmetric, alternating, an inner product) if and only if $\mathfrak{b}_2$ is also.*

2.2.2 Matrix Representations of Bilinear Forms

Fix a basis for a vector space V over $\mathbb{F}$ of dimension n. As it turns out, there is a bijective correspondence between the set $\text{Bil}(V)$ of bilinear forms on V and the set of all $n \times n$ matrices over $\mathbb{F}$, denoted $M_n(\mathbb{F})$.

If V has a basis $e_1, e_2, \ldots, e_n$, then any bilinear form $\mathfrak{b}$ on V gives rise to an $n \times n$ matrix $B = (\mathfrak{b}_{ij})$ with entries in $\mathbb{F}$, where $\mathfrak{b}_{ij} = \mathfrak{b}(e_i, e_j)$. This matrix determines the bilinear form uniquely, since if $x = \sum a_i e_i$ and $y = \sum b_j e_j$, then $\mathfrak{b}(x, y) = \sum \mathfrak{b}_{ij} a_i b_j$.

Definition 2.2.5. Let $(V, \mathfrak{b})$ be a bilinear space over $\mathbb{F}$ with basis $e_1, e_2, \ldots, e_n$. The matrix $B = (\mathfrak{b}_{ij})$ with $\mathfrak{b}_{ij} = \mathfrak{b}(e_i, e_j)$ will be called the *matrix representation* of the bilinear form $\mathfrak{b}$.

Notation 2.2.6. When the matrix representation B for the bilinear form $\mathfrak{b}$ is known, $\langle B\rangle$ will denote the bilinear form space $(V, \mathfrak{b})$.

Furthermore, if the matrix B is a block diagonal matrix with diagonal matrices B_i for $1 \le i \le n$, i.e.

$$B = \begin{pmatrix} B_1 & & 0 \\ & \ddots & \\ 0 & & B_n \end{pmatrix}$$

then we will write $\langle B_1\rangle \oplus \langle B_2\rangle \oplus \cdots \oplus \langle B_n\rangle$ instead of $\langle B\rangle$. In the special case where the matrices B_i are 1×1 matrices (b_i), respectively, we will write $\langle b_1\rangle \oplus \langle b_2\rangle \oplus \cdots \oplus \langle b_n\rangle$.

Example 2.2.7. With the scalar product on $\mathbb{R}^n$ as our bilinear form (see Example 2.1.4) and $e_1, e_2 \ldots, e_n$ as the standard basis on $\mathbb{R}^n$, the matrix $B = (\mathfrak{b}_{ij})$ is the identity matrix I_n:

$$B = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} = I_n$$

This bilinear space is written $\langle I_n\rangle$ or $\underbrace{\langle 1\rangle \oplus \cdots \oplus \langle 1\rangle}_{n \text{ times}}$.

Lemma 2.2.8. *On a bilinear space $(V, \mathfrak{b})$, the bilinear form $\mathfrak{b}$ is an inner product if and only if its matrix representation B is invertible.*

Proof. Let $\mathfrak{b}$ be an inner product, and let V have basis $e_1, e_2, \ldots, e_n$ and $V^* = \mathrm{Hom}_{\mathbb{F}}(V, \mathbb{F})$ have dual basis $\eta_1, \eta_2, \ldots, \eta_n$. Then the bijective map $x \mapsto \mathfrak{b}(x, \bullet)$ from V to V^* is determined by the map $e_i \mapsto \sum_j \mathfrak{b}_{ij}\eta_j$ (for example, by writing $x = \sum_i a_i e_i$). The bijective map $y \mapsto \mathfrak{b}(\bullet, y)$ from V to V^* is determined by the same map (for example, by writing $y = \sum_i b_i e_i$). Thus $e_i \mapsto \sum_j \mathfrak{b}_{ij}\eta_j$ is bijective, and it follows that $B = (\mathfrak{b}_{ij}) = (\mathfrak{b}(e_i, e_j))$ is invertible. The converse is similar. □

Example 2.2.9. We mentioned in Example 2.1.12 that the scalar product is an inner product over $\mathbb{R}^n$, i.e. that $\langle I_n\rangle$ is an inner product space. With the above Lemma in hand this is immediate since I_n is invertible.

Lemma 2.2.10. *A bilinear space $\langle B\rangle = (V, \mathfrak{b})$ is isometric to a bilinear space $\langle B'\rangle = (V', \mathfrak{b}')$ of the same dimension n if and only if $B' = ABA^T$ for some invertible $n \times n$ matrix A.*

Proof. If $e'_1, e'_2, \ldots, e'_n$ is a basis for $\langle B' \rangle$, then $e'_i = a_{i1}e_1 + \cdots + a_{in}e_n$ for some invertible matrix $A = (a_{ik})$. It follows that $B' = \mathfrak{b}'(e'_i, e'_j) = \sum_{k,l=1}^{n} a_{ik}\mathfrak{b}_{kl}a_{jl}$. □

The following Proposition is easy to obtain:

Proposition 2.2.11. *Let $(V, \mathfrak{b})$ be a bilinear space with a fixed basis. The following are true:*

(i) $\mathfrak{b}$ *is symmetric if and only if* $B = B^T$

(ii) $\mathfrak{b}$ *is skew-symmetric if and only if* $B = -B^T$

(iii) $\mathfrak{b}$ *is alternating if and only if* $B = -B^T$ *and the diagonal entries of* B *are all zeros.*

Definition 2.2.12. Given a basis $e_1, e_2, \ldots, e_n$ for an inner product space $(V, \mathfrak{b})$, a dual basis $\eta_1, \eta_2, \ldots, \eta_n$ for $(V, \mathfrak{b})$ is defined by the conditions $e_i \cdot \eta_k = 0$ for $i \neq k$, and $e_i \cdot \eta_i = 1$.

Lemma 2.2.13. *To each basis for an inner product space* $\langle B \rangle = (V, \mathfrak{b})$ *there corresponds a unique dual basis.*

Proof. Since $\langle B \rangle$ is an inner product space, the matrix $B = (\mathfrak{b}_{ij}) = (\mathfrak{b}(e_i, e_j))$ is invertible, with inverse matrix $B^{-1} = (\gamma_{jk})$. The equations $\eta_k = \gamma_{1k}e_1 + \cdots + \gamma_{nk}e_n$ now yield the dual basis. □

Example 2.2.14. For the inner product space from Example 2.2.7 with standard basis in $\mathbb{R}^n$, the dual basis is also the standard basis in $\mathbb{R}^n$.

2.2.3 Invariants of Bilinear Spaces

Definition 2.2.15. The *dimension* of a bilinear space $(V, \mathfrak{b})$ is the dimension of the vector space V, or equivalently, the number of rows or columns of a matrix representation of $\mathfrak{b}$.

Notation 2.2.16. Let $\mathbb{F}$ be a field. Then denote by $\mathbb{F}^\times$ the multiplicative group of units of $\mathbb{F}$, and denote by $\mathbb{F}^{\times 2}$ the subgroup of $\mathbb{F}^\times$ consisting of all squares of units.

Definition 2.2.17. Refer to the setup of Lemma 2.2.10. Since $\det B' = \det B \cdot (\det A)^2$ and $\det A \neq 0$, it follows that $\det B'$ and $\det B$ are the same up to squares of units in $\mathbb{F}$. Thus the *determinant* of an inner product $\mathfrak{b}$ is $\det \mathfrak{b} = \det B \cdot \mathbb{F}^{\times 2}$ in $\mathbb{F}^\times / \mathbb{F}^{\times 2}$, where B is a matrix reprensetation of $\mathfrak{b}$.

Example 2.2.18. Following Example 2.2.7, $V = \mathbb{R}^n$ with standard basis, $B = I_n$, and $\langle I_n \rangle_{\mathbb{R}} \cong \mathbb{R}^n$. Then $\det(I_n) = 1$.

2.3 Orthogonality

In this section we discuss orthogonality in its various incarnations (orthogonal sums, orthogonal decomposition, orthogonal basis, etc.)

2.3.1 Orthogonal Sums

The following definition of orthogonal sums of bilinear spaces leads to what might be called *Witt addition*, that is, the addition operation in a Witt ring (to be defined in Section 2.7).

Definition 2.3.1. Let $(V_1, \mathfrak{b}_1), (V_2, \mathfrak{b}_2), \ldots, (V_n, \mathfrak{b}_n)$ be bilinear spaces. The *orthogonal sum* $(V_1, \mathfrak{b}_1) \oplus (V_2, \mathfrak{b}_2) \oplus \cdots \oplus (V_n, \mathfrak{b}_n)$ is defined to be the direct sum $V = V_1 \oplus V_2 \oplus \cdots \oplus V_n$ with bilinear form $\mathfrak{b} = \mathfrak{b}_1 \perp \mathfrak{b}_2 \perp \cdots \perp \mathfrak{b}_n$ defined by the equation

$$\mathfrak{b}(x, y) = (\mathfrak{b}_1 \perp \mathfrak{b}_2 \perp \cdots \perp \mathfrak{b}_n)(x, y) = \sum_{i=1}^{n} \mathfrak{b}_i(x_i, y_i).$$

where $x = x_1 \oplus x_2 \oplus \cdots \oplus x_n$ and $y = y_1 \oplus y_2 \oplus \cdots \oplus y_n$ are elements of V. The resulting bilinear space $(V, \mathfrak{b})$ is then $(V_1 \oplus V_2 \oplus \cdots \oplus V_n, \mathfrak{b}_1 \perp \mathfrak{b}_2 \perp \cdots \perp \mathfrak{b}_n)$.

Notation 2.3.2. As before, we will abuse the notation and write $\langle V_i \rangle$ for the bilinear space $(V_i, \mathfrak{b}_i)$. The symbol V_i can be thought of as either the vector space or matrix representation of $\mathfrak{b}_i$ where appropriate.

Proposition 2.3.3. *The orthogonal sum $\langle V_1 \rangle, \langle V_2 \rangle, \ldots, \langle V_n \rangle$ is an inner product space if and only if $\langle V_i \rangle$ is an inner product space for each $i = 1, \ldots, n$.*

Proposition 2.3.4. *The dimension and determinant of orthogonal sums are additive and multiplicative, respectively. That is,*

$$\dim(\langle V_1 \rangle \oplus \langle V_2 \rangle \oplus \cdots \oplus \langle V_n \rangle) = \sum_{i=1}^{n} \dim(\langle V_i \rangle)$$

and

$$\det(\langle V_1 \rangle \oplus \langle V_2 \rangle \oplus \cdots \oplus \langle V_n \rangle) = \prod_{i=1}^{n} \det(\langle V_i \rangle).$$

Example 2.3.5. Let $V_i = \mathbb{R}$ for each $1 \leq i \leq n$, and $\mathfrak{b}_i$ be the scalar product on $\mathbb{R}$. Then the orthogonal sum $\underbrace{\langle \mathbb{R} \rangle \oplus \cdots \oplus \langle \mathbb{R} \rangle}_{n \text{ times}}$ is an inner product space. Indeed, $\langle \mathbb{R} \rangle \oplus \cdots \oplus \langle \mathbb{R} \rangle \cong \langle \mathbb{R}^n \rangle$.

Following Example 2.2.7 and writing $\langle I_1 \rangle$ as $\langle 1 \rangle$,

$$\dim\left(\langle 1 \rangle \oplus \cdots \oplus \langle 1 \rangle\right) = \sum_{i=1}^{n} \dim\left(\langle 1 \rangle\right) = n = \dim\left(\langle \mathbb{R}^n \rangle\right)$$

and

$$\det\left(\langle 1 \rangle \oplus \cdots \oplus \langle 1 \rangle\right) = \prod_{i=1}^{n} \det\left(\langle 1 \rangle\right) = 1 = \det\left(\langle \mathbb{R}^n \rangle\right).$$

2.3.2 Orthogonal Decomposition

Recall Definition 2.1.7: For $\mathfrak{b}$ either a symmetric or alternating bilinear form on V, two elements x and y of V are called *orthogonal* if $\mathfrak{b}(x, y) = 0 = \mathfrak{b}(y, x)$.

It often happens that a vector is orthogonal to every vector in a subspace of the underlying vector space, prompting the following definition.

Definition 2.3.6. Let $(V, \mathfrak{b})$ be a bilinear space, and let M be a subspace of V. Then the *orthogonal complement* of M, denoted $M^{\perp}$, consists of all $x \in V$ such that $\mathfrak{b}(x, M) = 0$. That is,

$$M^{\perp} = \{x \in V : \mathfrak{b}(x, y) = 0 \text{ for all } y \in M\}$$

However, a nonzero vector of an (either symmetric or alternating) inner product space cannot be orthogonal to every vector in the entire underlying vector space:

Proposition 2.3.7. *Let* $(V, \mathfrak{b})$ *be a bilinear form space. Then* $\mathfrak{b}$ *is an inner product if and only if* $V^{\perp} = 0$.

Proof. This can be done directly from Definition 2.1.9. If $\mathfrak{b}$ is an inner product, then the two homomorphisms $x_0 \mapsto \mathfrak{b}(x_0, \bullet)$ and $y_0 \mapsto \mathfrak{b}(\bullet, y_0)$ are bijective. Without loss generality, consider $x_0 \mapsto \mathfrak{b}(x_0, \bullet)$. The kernel of this map is zero since the map is injective, but the kernel is also the set of vectors v such that the maps $\mathfrak{b}(x_0, v) = 0 \in V^*$. But this is the same as the set $\{v \in V : \mathfrak{b}(v, x) = 0 \text{ for all } x \in V\}$. Comparing definitions, we arrive at $V^{\perp} = 0$. ☐

Remark 2.3.8.

(i) $M \subseteq (M^{\perp})^{\perp}$. This follows since M is orthogonal to $M^{\perp}$ (in the sense that every vector in M is orthogonal to $M^{\perp}$), and $(M^{\perp})^{\perp}$ is every vector in V orthogonal to $M^{\perp}$.

(ii) Since $M \cap M^{\perp} = \{v \in M : \mathfrak{b}(v, x) = 0 \text{ for all } x \in M\}$, $M \cap M^{\perp} = 0$ if and only if M is an inner product subspace of V. The proof is very similar to the proof of the Proposition 2.3.7.

Theorem 2.3.9. *Let $(V, \mathfrak{b})$ be an (either symmetric or alternating) inner product space. If M is an inner product subspace of V, then*

$$V = M \oplus M^{\perp}$$

Proof. As mentioned above, if M is an inner product subspace of V, then $M \cap M^{\perp} = 0$. From linear algebra, this means that $M \oplus M^{\perp}$ is a disjoint union of M and $M^{\perp}$. Thus by the dimension theorem, $\dim(M \oplus M^{\perp}) = \dim M + \dim M^{\perp} = \dim V$. Since $M \oplus M^{\perp}$ is a subspace of V of the same dimension, the result follows. □

Definition 2.3.10. Let $(V, \mathfrak{b})$ be a bilinear space. A basis $\{e_1, e_2, \ldots, e_n\}$ for V is called an *orthogonal basis* if the basis vectors are pairwise orthogonal, i.e. $\mathfrak{b}(e_i, e_j) = 0$ for $i \neq j$.

Said differently, a basis $\{e_1, e_2, \ldots, e_n\}$ for the bilinear space $\langle V \rangle$ is orthogonal if and only if $\langle V \rangle = \langle \mathfrak{b}(e_1, e_1) \rangle \oplus \langle \mathfrak{b}(e_2, e_2) \rangle \oplus \cdots \oplus \langle \mathfrak{b}(e_n, e_n) \rangle$.

Corollary 2.3.11. *Let $(V, \mathfrak{b})$ be a symmetric bilinear space. Then V has an orthogonal basis.*

Proof. Recall the definition of isotropic and anisotropic vectors from Definition 2.1.3. The proof is by induction on $\dim V$.

If $\dim V = 1$, then any singleton set containing a nonzero vector is the required orthogonal basis.

Next assume that $\dim V = n > 1$ and each symmetric bilinear space of smaller dimension has an orthogonal basis. Note that the two cases form a dichotomy:

(i) If every vector in V is an isotropic vector, then for all $u, v \in V$

$$2\mathfrak{b}(u, v) = \mathfrak{b}(u, u) + \mathfrak{b}(v, v) + 2\mathfrak{b}(u, v) = \mathfrak{b}(u + v, u + v) = 0$$

showing that $\mathfrak{b}(u, v) = 0$ (warning: here we are explicitly using the fact that the characteristic of the ground field is not 2). Thus in this case every basis of V is orthogonal.

(ii) If V contains an anisotropic vector v, then the subspace $M = \mathbb{F}v$ spanned by v is an inner product subspace. This is because in this case $M \cap M^{\perp} \neq 0$ if and only if $v \in M$ if and only if $\mathfrak{b}(v, v) = 0$. Thus by Theorem 2.3.9, $\dim M^{\perp} < \dim V$, so by the induction hypothesis it follows that $M^{\perp}$ has an orthogonal basis. Adding the vector v to the orthogonal basis for $M^{\perp}$ (which has $n - 1$ basis vectors) then gives the requried orthogonal basis for V.

□

2.4 Tensor Products

In this section we discuss the tensor product of bilinear spaces. This notion will lead to what might be called *Witt multiplication*, that is, the tensor product becomes the multiplication operation of a Witt ring. The distinction of Witt ring versus a Witt group is minimal for our purposes. Thus this section might be considered as optional.

Let $(V_1, \mathfrak{b}_1), (V_2, \mathfrak{b}_2), \ldots, (V_n, \mathfrak{b}_n)$ be bilinear spaces over $\mathbb{F}$. Then the content of what follows is to show how to use these bilinear spaces to form a tensor product, denoted $\langle V_1 \rangle \otimes \langle V_2 \rangle \otimes \cdots \otimes \langle V_n \rangle$.

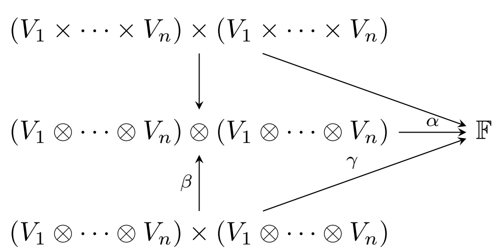
We will use heavily the notion of tensor product for vector spaces, which we will denote $V_1 \otimes V_2 \otimes \cdots \otimes V_n$. See [15], Chapter 14 for a detailed reference.

Theorem 2.4.1. *There is a unique bilinear form* $\mathfrak{b}$ *on* $V_1 \otimes V_2 \otimes \cdots \otimes V_n$ *which satisfies the identity*

$$\mathfrak{b}(x_1 \otimes \cdots \otimes x_n, y_1 \otimes \cdots \otimes y_n) = \prod_{i=1}^{n} \mathfrak{b}_i(x_i, y_i)$$

for all x_i *and* y_i *in* V_i*,* $1 \leq i \leq n$*.*

Proof. Consider the following diagram:



The function $(V_1 \times \cdots \times V_n) \times (V_1 \times \cdots \times V_n) \to \mathbb{F}$ is defined by

$$(x_1, \ldots, x_n, y_1, \ldots, y_n) \mapsto \prod_{i=1}^{n} \mathfrak{b}_i(x_i, y_i),$$

and is linear in its $2n$ variables. The universal property of the tensor product ([15], Theorem 14.1.3, page 179) makes the top triangle commute, where the function $\alpha : (V_1 \otimes \cdots \otimes V_n) \otimes (V_1 \otimes \cdots \otimes V_n) \to \mathbb{F}$ is defined by

$$(x_1 \otimes \cdots \otimes x_n \otimes y_1 \otimes \cdots \otimes y_n) \mapsto \prod_{i=1}^{n} \mathfrak{b}_i(x_i, y_i).$$

In the bottom triangle, there is the ([15], Proposition 14.3.3, page 180) function β assigning

$$(x_1 \otimes \cdots \otimes x_n, y_1 \otimes \cdots \otimes y_n) \mapsto (x_1 \otimes \cdots \otimes x_n \otimes y_1 \otimes \cdots \otimes y_n).$$

To get the required bilinear form, simply precompose α with β to get γ:

$$\begin{aligned}\mathfrak{b}(x_1 \otimes \cdots \otimes x_n, y_1 \otimes \cdots \otimes y_n) &\coloneqq \gamma(x_1 \otimes \cdots \otimes x_n, y_1 \otimes \cdots \otimes y_n)\\ &= \alpha(\beta(x_1 \otimes \cdots \otimes x_n, y_1 \otimes \cdots \otimes y_n))\\ &= \alpha(x_1 \otimes \cdots \otimes x_n \otimes y_1 \otimes \cdots \otimes y_n)\\ &= \prod_{i=1}^{n} \mathfrak{b}_i(x_i, y_i)\end{aligned}$$

The uniqueness of the form depends on the uniqueness of the linear functions $\mathfrak{b}(x_1 \otimes \cdots \otimes x_n, \bullet)$ and $\mathfrak{b}(\bullet, y_1 \otimes \cdots \otimes y_n)$, taken as understood from the references given. ☐

Definition 2.4.2. The *tensor product of bilinear spaces* $(V_1, \mathfrak{b}_1), (V_2, \mathfrak{b}_2), \ldots, (V_n, \mathfrak{b}_n)$, denoted $\langle V_1\rangle \otimes \langle V_2\rangle \otimes \cdots \otimes \langle V_n\rangle$, is the bilinear space with underlying vector space $V_1 \otimes V_2 \otimes \cdots \otimes V_n$ and bilinear form as described in Theorem 2.4.1.

Remark 2.4.3. If each of the $\langle V_i\rangle$ are symmetric, then so is $\langle V_1\rangle \otimes \langle V_2\rangle \otimes \cdots \otimes \langle V_n\rangle$. This is immediate from the definition of $\mathfrak{b}$ in Theorem 2.4.1.

2.5 Witt Cancellation

Theorem 2.5.1 (Witt Cancellation). *Let $\langle V\rangle$, $\langle U_1\rangle$, and $\langle U_2\rangle$ be symmetric inner product spaces. If $\langle V\rangle \oplus \langle U_1\rangle \simeq \langle V\rangle \oplus \langle U_2\rangle$, then $\langle U_1\rangle \simeq \langle U_2\rangle$.*

The idea of the proof is to show that $\langle U_1\rangle^\perp$ and $\langle U_2\rangle^\perp$ are the same up to isometry. To do this, we will explicitly construct the isometry. A key piece of that construction is given in the following definition.

Definition 2.5.2. Let $\langle V\rangle = \langle V_1\rangle \oplus \langle V_2\rangle$ be a symmetric inner product space. The linear map $r : V \to V$ that is constant on V_1 and sends every element in V_2 to its negative is called the *reflection* of V with respect to V_1 and V_2.

Lemma 2.5.3. *If $x, y \in V$ are anisotropic vectors such that $\mathfrak{b}(x, x) = \mathfrak{b}(y, y)$, then there is a reflection of V that maps x to y.*

Proof. If $x = y$, then the identity id_V is the required reflection. So assume $x \neq y$. Next write $u = \frac{1}{2}(x+y)$ and $v = \frac{1}{2}(x-y)$, so that $\mathfrak{b}(x,x) = \mathfrak{b}(u,u)+\mathfrak{b}(v,v)$. Since x is anisotropic by hypothesis, this means that u, v, or both u and v are anisotropic. If u is anisotropic, then $\langle V\rangle = \langle \mathbb{F}u\rangle \oplus \langle \mathbb{F}u^{\perp}\rangle$. In this case, the reflection of V with respect to $\mathbb{F}u$ and $\mathbb{F}u^{\perp}$ takes $x = u + v$ to $y = u - v$. If v is anisotropic, the reflection of V with respect to $\mathbb{F}v^{\perp}$ and $\mathbb{F}v$ maps x to y.

□

Now we are ready to show Witt cancellation.

Proof of Theorem 2.5.1. The first step in the proof is to make a reduction. By Corollary 2.3.11, $\langle V\rangle$ has an orthogonal basis, and we can write $\langle V\rangle = \langle \mathfrak{b}(e_1,e_1)\rangle \oplus \langle \mathfrak{b}(e_2,e_2)\rangle \oplus \cdots \oplus \langle \mathfrak{b}(e_n,e_n)\rangle$. Fix an i from $0,\ldots,n$, and define $e \coloneqq \mathfrak{b}(e_i,e_i)$. The result obtained below can easily be extended to the remaining summands.

Let 0_V, 0_{U_1}, and 0_{U_2} be the zero vectors of V, U_1, and U_2, respectively. Now let $T : V \oplus U_1 \to V \oplus U_2$ be an isometry. Then

$$\mathfrak{b}_{V\oplus U_2}(T(\langle e\rangle \oplus \langle 0_{U_1}\rangle)) = \mathfrak{b}_{V\oplus U_2}(\langle e\rangle \oplus \langle 0_{U_2}\rangle),$$

so by Lemma 2.5.3, there is a reflection r of $V \oplus U_2$ that maps $T(\langle e\rangle \oplus \langle 0_{U_1}\rangle)$ to $\langle e\rangle \oplus \langle 0_{U_2}\rangle$. But then

$$rT : V \oplus U_1 \to V \oplus U_2$$

takes $\langle e\rangle \oplus \langle 0_{U_1}\rangle$ to $\langle e\rangle \oplus \langle 0_{U_2}\rangle$. Hence their orthogonal complements $\langle 0_V\rangle \oplus \langle U_1\rangle$ and $\langle 0_V\rangle \oplus \langle U_2\rangle$ are isometric, so $\langle U_1\rangle \simeq \langle U_2\rangle$. □

2.6 Hyperbolic and Metabolic Bilinear Forms

2.6.1 Hyperbolic Bilinear Forms

Let $(V, \mathfrak{b})$ be a bilinear space over $\mathbb{F}$ with $\mathfrak{b}$ an (either symmetric or alternating) bilinear form. A *hyperbolic plane*, denoted $\mathbb{H}$, is a 2-dimensional subspace on which $\mathfrak{b}$ is also an inner product.

If $\mathbb{H}$ is a hyperbolic plane and $x \in \mathbb{H}$ is nonzero, then there is a $y \in \mathbb{H}$ such that $\mathfrak{b}(x,y) \neq 0$. Thus, if $\mathfrak{b}(x,y) = c$ for $c \in \mathbb{F}$, we may consider instead $y' = c^{-1}y$, so that $\mathfrak{b}(x,y') = 1$. If $\mathfrak{b}$ is symmetric, then $\mathfrak{b}(y',x) = 1$. If $\mathfrak{b}$ is alternating, it is also skew-symmetric. Thus $\mathfrak{b}(y',x) = -1$ in this case. Hence the matrix representation H of $\mathfrak{b}$ with respect to the basis $\{x, y'\}$ is

$$H = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

in the symmetric case and

$$H = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

in the alternating case.

Definition 2.6.1. If the $(V, \mathfrak{b})$ is an orthogonal sum of hyperbolic planes, then we will call $\mathfrak{b}$ a *hyperbolic bilinear form.*

2.6.2 Metabolic Bilinear Forms

Definition 2.6.2. A symmetric inner product space S over a field $\mathbb{F}$ is *metabolic* (or *split*) if there exists a subspace N of S such that N is a direct summand of S, and such that N is precisely equal to its orthogonal complement $N^{\perp}$. Such a direct summand is sometimes called a *metabolizer* (or *Lagrangian*).

Remark 2.6.3. A subspace N of S is a direct summand if there is a subspace M of S such that $N \oplus M = S$. That is, if $N + M = S$ and $N \cap M = 0$. If additionally $N = N^{\perp}$ (as is the case when S is metabolic), then N is a metabolizer.

Lemma 2.6.4. *An inner product space over a field $\mathbb{F}$ is metabolic if and only if it possesses a basis so that the associated inner product matrix has the form $\begin{pmatrix} 0 & I \\ I & A \end{pmatrix}$. If we also assume that 2 is a unit, then every metabolic inner product space is a direct sum of hyperbolic planes.*

Proof. Given any direct summand $N \subseteq V$, choose a basis $e_1, \ldots, e_n$ for N, and extend to a basis $e_1, \ldots, e_k$ for V. Let $\eta_1, \ldots, \eta_k$ be the dual basis. Then clearly $\eta_{n+1}, \ldots, \eta_k$ form a basis for the orthogonal complement $N^{\perp}$.

Suppose that $N = N^{\perp}$. Then substituting $e_1, \ldots, e_n$ for $\eta_{n+1}, \ldots, \eta_k$, we see that the elements $e_1, \ldots, e_n, \eta_1, \ldots \eta_n$ form a basis for V. (In particular the dimension k of V must be equal to $2n$.) The inner product matrix of V with respect to this new basis takes the form $\begin{pmatrix} 0 & I \\ I & A \end{pmatrix}$ for some symmetric matrix A. The converse is clear.

Now suppose that 2 is a unit in R. Setting $B = -\frac{1}{2}A$, computation shows that

$$\begin{pmatrix} I & 0 \\ B & I \end{pmatrix} \begin{pmatrix} 0 & I \\ I & A \end{pmatrix} \begin{pmatrix} I & 0 \\ B & I \end{pmatrix}^T = \mathbb{H} \oplus \mathbb{H} \oplus \cdots \oplus \mathbb{H}$$

This completes the proof. □

Since hyperbolic planes are 2-dimensional, we have the immediate consequence:

Corollary 2.6.5. *Every metabolic space over a field $\mathbb{F}$ has even dimension.*

Theorem 2.6.6. *Let S and S' be metabolic inner product spaces, and let $(V, \mathfrak{b})$ be an arbitrary symmetric inner product space. Then the following are true:*

(i) The orthogonal sum $(V, \mathfrak{b}) \oplus (V, -\mathfrak{b})$ is metabolic.

(ii) The orthogonal sum $S \oplus S'$ is metabolic.

Proof. To prove (i), we use the metabolic inner product space $\langle 1 \rangle \oplus \langle -1 \rangle$ to conclude that the tensor product

$$\begin{aligned} (\langle 1 \rangle \oplus \langle -1 \rangle) \otimes (V, \mathfrak{b}) &\cong (\langle 1 \rangle \otimes (V, \mathfrak{b})) \oplus (\langle -1 \rangle \otimes (V, \mathfrak{b})) \\ &\cong (V, \mathfrak{b}) \oplus (V, -\mathfrak{b}) \end{aligned}$$

is also metabolic.

To prove (ii), let N and N' be metabolizers for S and S', respectively. Then $S \oplus S'$ has metabolizer $N \oplus N'$. □

2.7 The Witt Ring $W(\mathbb{F})$

We restrict our attention to symmetric inner products. In Proposition 2.2.3, we saw that isometry is an equivalence relation on the vector space of bilinear forms, $\mathrm{Bil}(V)$, and it follows that it is also an equivalence relation on the subspace of symmetric bilinear forms $\mathrm{Sym}(V)$.

Isometry classes of symmetric inner product spaces form a semiring under orthogonal sum and tensor product. In particular, isometry classes are associative under orthogonal sum, and the inner product space $(0, 0)$ (i.e. $V = 0$ and $\mathfrak{b} = 0$, the inner product space of dimension 0) acts as the identity element, but there is no notion of inverses under orthogonal sum.

In situations such as these, there is a technique to form a ring from a semiring - the so-called *Grothendeick construction* (see [14], page 29). The resulting ring is called the Witt-Grothendeick ring and is denoted $\widehat{W}(\mathbb{F})$. The Witt ring can then be defined to be the quotient ring $W(\mathbb{F}) = \widehat{W}(\mathbb{F})/\mathbb{H}$. Under this definition, Witt equivalence could be shown to imply the existence of metabolic symmetric inner product spaces with which to form orthogonal sums.

Though this is a natural way to approach Witt rings, we will *not* use the Grothendeick construction and ring $\widehat{W}(\mathbb{F})$. Instead of defining the Witt ring as a quotient and investigating its structure, we will start from a explicit description of the elements of $W(\mathbb{F})$.

Definition 2.7.1. Two symmetric inner product spaces $\langle V_1 \rangle$ and $\langle V_2 \rangle$ belong to the same *Witt class*, written $\langle V_1 \rangle \sim \langle V_2 \rangle$, if there exist metabolic inner product spaces $\langle M_1 \rangle$ and $\langle M_2 \rangle$ so that $\langle V_1 \rangle \oplus \langle M_1 \rangle$ is isometric to $\langle V_2 \rangle \oplus \langle M_2 \rangle$. That is,

$$\langle V_1 \rangle \sim \langle V_2 \rangle \quad \Rightarrow \quad \exists \text{ metabolic } M_1, M_2, \quad \langle V_1 \rangle \oplus \langle M_1 \rangle \simeq \langle V_2 \rangle \oplus \langle M_2 \rangle$$

Lemma 2.7.2. *The relation $\sim$ defined above is an equivalence relation on the set of all symmetric inner product spaces.*

Proof. The reflexivity and symmetry properties of $\sim$ follow immediately from the corresponding properties for $\simeq$ (see Proposition 2.2.3.) To see that $\sim$ is transitive, let $\langle V_1 \rangle \sim \langle V_2 \rangle$ and $\langle V_2 \rangle \sim \langle V_3 \rangle$. Then there are metabolic symmetric inner product spaces $\langle M_1 \rangle$, $\langle M_2 \rangle$, $\langle M_3 \rangle$, and $\langle M_4 \rangle$ such that

$$\langle V_1 \rangle \oplus \langle M_1 \rangle \simeq \langle V_2 \rangle \oplus \langle M_2 \rangle$$

and

$$\langle V_2 \rangle \oplus \langle M_3 \rangle \simeq \langle V_3 \rangle \oplus \langle M_4 \rangle.$$

This implies that

$$\langle V_1 \rangle \oplus \langle M_3 \rangle \oplus \langle M_1 \rangle \simeq \langle V_2 \rangle \oplus \langle M_3 \rangle \oplus \langle M_2 \rangle$$

and

$$\langle V_2 \rangle \oplus \langle M_3 \rangle \oplus \langle M_2 \rangle \simeq \langle V_3 \rangle \oplus \langle M_4 \rangle \oplus \langle M_2 \rangle.$$

By transitivity of $\simeq$,

$$\langle V_1 \rangle \oplus \langle M_3 \rangle \oplus \langle M_1 \rangle \simeq \langle V_3 \rangle \oplus \langle M_4 \rangle \oplus \langle M_2 \rangle.$$

Since the orthogonal sum of metabolic symmetric inner product spaces is metabolic, it follows that $\langle V_1 \rangle \sim \langle V_3 \rangle$. □

Lemma 2.7.3. *If* $\langle V_1 \rangle \sim \langle V_3 \rangle$ *and* $\langle V_2 \rangle \sim \langle V_4 \rangle$*, then* $\langle V_1 \rangle \oplus \langle V_2 \rangle \sim \langle V_3 \rangle \oplus \langle V_4 \rangle$*.*

Proof. Refer to Theorem 2.6.6 and Proposition 2.2.3. Since $\langle V_1 \rangle \sim \langle V_3 \rangle$ and $\langle V_2 \rangle \sim \langle V_4 \rangle$, there are metabolic symmetric inner product spaces $\langle M_1 \rangle$, $\langle M_2 \rangle$, $\langle M_3 \rangle$, and $\langle M_4 \rangle$ such that

$$\langle V_1 \rangle \oplus \langle M_1 \rangle \simeq \langle V_3 \rangle \oplus \langle M_3 \rangle$$

and

$$\langle V_2 \rangle \oplus \langle M_2 \rangle \simeq \langle V_4 \rangle \oplus \langle M_4 \rangle.$$

Combining these and using the commutativity of $\simeq$ gives

$$(\langle V_1 \rangle \oplus \langle V_2 \rangle) \oplus (\langle M_1 \rangle \oplus \langle M_2 \rangle) \simeq (\langle V_3 \rangle \oplus \langle V_4 \rangle) \oplus (\langle M_3 \rangle \oplus \langle M_4 \rangle),$$

and the result follows. □

Lemma 2.7.4. *If* $\mathbb{F}$ *is a field of characteristic not equal to* 2*, then two symmetric inner product spaces over* $\mathbb{F}$ *are isometric if and only if they belong to the same Witt class and have the same dimension.*

Proof. This follows easily from Corollary 2.5.1 (Witt Cancellation) and Lemma 2.6.4. □

Recall that $(V, \mathfrak{b}) \oplus (V, -\mathfrak{b}) \sim 0$ and $\langle 1 \rangle \otimes (V, \mathfrak{b}) \simeq (V, \mathfrak{b})$.

Definition 2.7.5. The collection $W(\mathbb{F})$ of all Witt classes of symmetric inner product spaces over a field $\mathbb{F}$ forms an abelian group with orthogonal sum of Section 2.3 as the addition operation, called the *Witt group* over $\mathbb{F}$. This group can be made into the *Witt ring* over $\mathbb{F}$ by utilizing the tensor product of Section 2.4 as the multiplication operation.

In the Chapter 3, we will develop some topological ideas before returning to the sorts of algebraic structures we have just constructed.

Chapter 3

Knot Theory

We expect the reader to be familiar with many ideas from point-set topology. Beyond that, some homology and cohomology theory is used to provide the rigor needed in a couple definitions. In practice, however, the use of these definitions is intuitive and the reader without an extensive algebraic topology toolkit will not be at much of a disadvantage. Nevertheless, a good background reference is [5].

For the remainder of the thesis, unless explicitly stated otherwise, we will be confined to the setting of smooth oriented manifolds. This convention is especially crucial in 3.5, since the distinction between the smooth and topological categories is paramount in dimension four (the latter not being discussed here).

3.1 Knots in S^3

Recall that a mapping $f : X \to Y$ is an *embedding* if $f : X \to f(X)$ is a homeomorphism (i.e. if the map f is a homeomorphism onto its image).

Definition 3.1.1. A *knot* is an embedding of the circle S^1 into the 3-sphere S^3.

Definition 3.1.2. A *link* is a disjoint union of circles embedded into S^3.

We will mostly concern ourselves with knots and avoid the complications links bring to the table. To keep the notation under control, we will abuse the terminology *knot* by allowing it to mean any of the following: the embedding itself, a class of embeddings, the image of S^1 under the embedding, or a class of these images.

Definition 3.1.3. A *(regular) knot diagram* is a projection of a knot into a plane in such a way that the resulting immersed planar curve

- has finitely many double points (also called *crossings*), and
- contains the additional information of whether a crossing was an overcrossing or undercrossing.

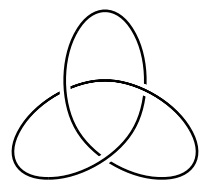


Figure 3.1: A knot with three crossings: the (left-handed) trefoil.

We will need a way to distinguish between embeddings.

Definition 3.1.4. Two embeddings (knots) $K_0, K_1 : S^1 \to S^3$ are *isotopic* if there is an embedding $F : S^1 \times I \to S^3 \times I$ where $I = [0, 1]$ such that $F(x, t) = (f(x, t), t)$, $x \in S^1$, $t \in I$, and with $f(x, 0) = K_0(x)$ and $f(x, 1) = K_1(x)$. We will call such an F an *isotopy* between K_0 and K_1.

The last two conditions can be subsumed by use of the notation $f_t(x) = K_t(x) = f(x, t)$.

Unfortunately, the problem with an isotopy is that while it changes the image of K_0 into K_1 continuously it ignores the points in S^3 that are outside of the image of $S^1 \times I$. So this notion of isotopy turn out to be not quite what we are looking for.

To remedy this, we will require that the ambient space S^3 change continuously along with the image of f_t, as in the following definition.

Definition 3.1.5. Two embeddings $K_0, K_1 : S^1 \to S^3$ are *ambient isotopic* if there is an isotopy $G : S^3 \times I \to S^3 \times I$ where $I = [0, 1]$ such that $G(y, t) = (g_t(y), t)$, $y \in S^3$, $t \in I$, and with $K_1 = g_1 K_0$ and $g_0 = \mathrm{id}_{S^3}$. We will call G an *ambient isotopy* between K_0 and K_1.

A key fact is that the homeomorphism $g_1 : S^3 \to S^3$ restricts to a homeomorphism $\widetilde{g}_1 : S^3 \setminus K_0(X) \to S^3 \setminus K_1(X)$ if K_0 and K_1 are ambient isotopic. This allows one to consider whether the complements of knots are homeomorphic in order to distinguish between the knots.

Definition 3.1.6. Two knots are equivalent if they are ambient isotopic.

In the sequel, we will say that $K_1 = K_2$ if K_1 and K_2 are ambient isotopic, and we will abuse the terminology by calling *ambient isotopies* simply *isotopies* as there will be no confusion.

Our definition of equivalent knots only requires that an isotopy exists, but it says nothing about how to construct isotopies in general. More to the point, we do not have an algorithm to determine whether two knots are equivalent. In order to tell when two knots are *not* equivalent, we will develop knot invariants by bringing in more mathematics.

A knot invariant ([13] page 47) is a function $K \rightsquigarrow f(K)$ which assigns to each knot K an object $f(K)$ in such a way that isotopic knots are assigned equivalent objects. As with invariants in other parts of mathematics (e.g. invariants of topological spaces), there is unfortunately no perfect invariant. The best invariant often depends on the particular knots in question. Desirable qualities are that they can distinguish between the knots in question, are easily computable, and preserve properties of knots.

Example 3.1.7. For knot diagrams, the number of crossings is not a knot invariant.

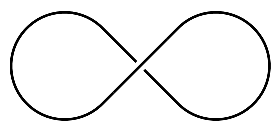
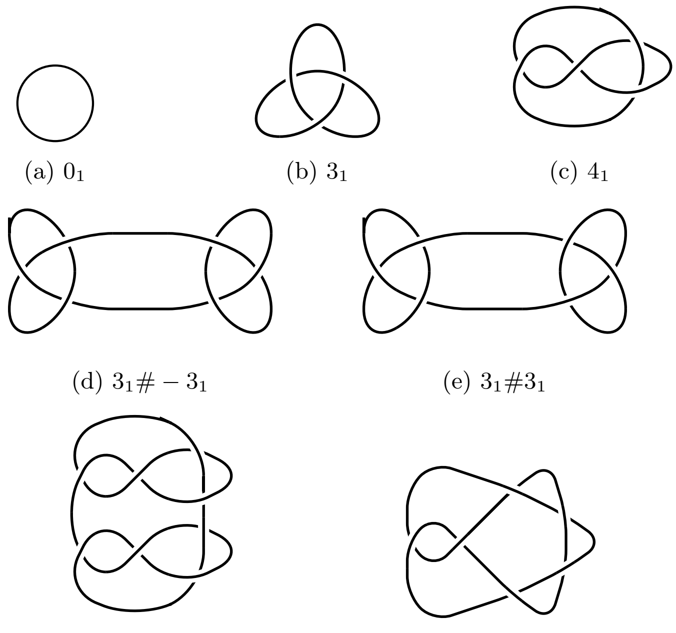


Figure 3.2: An unknot with a kink has one crossing that can be undone.

The *minimum number of crossings* on the set of all knot diagrams, however, is a knot invariant.

Notation 3.1.8. For knots with less than ten crossings the most prevalent classification is by the minimum number of crossings described above. For knots with the same minimum crossing number, there is an arbitrarily chosen but fixed order called the *Alexander-Briggs* notation.

In this notation, the unknot is 0_1 (the first knot with zero crossings), the trefoil is 3_1 (the first knot with three crossings), 5_2 (the second knot with five crossings), and so on. A table of these low crossing number knots is available in Rolfsen's book [13] and on the Internet at the Knot Atlas [16] or Knot Info [2].



(a) 0_1 (b) 3_1 (c) 4_1

(d) $3_1\# - 3_1$ (e) $3_1\#3_1$

(f) $4_1\#4_1$ (g) 6_1

Figure 3.3: Some of the knots used in this thesis.

3.2 Connected Sums of Knots

Next, we define an operation on the set of all equivalence classes of knots, the connected sum of knots. First we need to address something we have neglected thus far: orientations.

Definition 3.2.1. A knot K is oriented if the circle S^1, embedded in S^3, is oriented.

Remark 3.2.2. In light of orientations on K, we revise Definition 3.1.6 to say that two oriented knots are equivalent if they are ambient isotopic and any ambient isotopy respects the orientation of the knots.

Note that the ambient space S^3 can also be oriented (and ambient isotopy respects this orientation). Combinations of these two types of orientations give rise to three more variants of a knot.

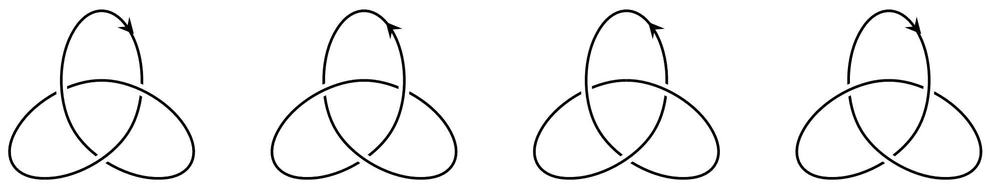
Definition 3.2.3. Given an oriented knot K embedded in the oriented 3-sphere S^3,

- the *mirror image* of K, denoted by mK, is K with the opposite orientation on S^3,
- the *reverse* of K, denoted by rK, is the opposite orientation for K in S^3,

- and the *inverse* of K, denoted by rmK, is the opposite orientation for K with the opposite orientation on S^3.

We will use the notation $-K$ for the inverse of the knot (note the potential for conflict with the standard notation $-M$ for opposite orientation for an oriented manifold M.)

Example 3.2.4. The left-handed trefoil knot, $K = 3_1$.



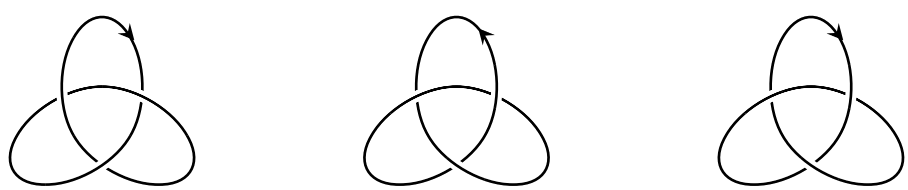
(a) K (b) mK (c) rK (d) $-K$

Figure 3.4: Possible variants of the trefoil, 3_1.**??**

It so happens that for $K = 3_1$, K is equivalent to rK but not to mK. According to Appendix C: Table I in [1], the first knot where K, mK, rK, and $-K$ are distinct is the knot 9_{32}.

A natural question to ask is if there is a way to attach two knots together to get another knot. In the trefoil knot example above, it is easy to see that cutting two knots and connecting them together can sometimes work. For instance, if you cut 3_1 and -3_1, you can attach them together in a way that respects orientation.

Example 3.2.5. Consider the trefoil 3_1, its mirror image $m3_1$, and its inverse -3_1 from Example 3.2.4.



(a) 3_1 (b) $m3_1$ (c) -3_1

If one of the three corner loops of 3_1 is connected with any of the three corner loops of $m3_1$, the orientations do not agree. If instead one of the three corner loops of 3_1 is connected with any of the three corner loops of -3_1, then the orientations do agree. The result is a square knot:

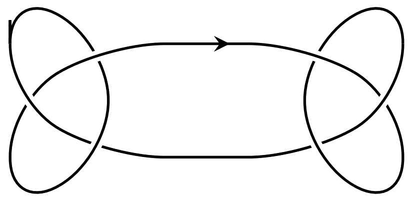


Figure 3.5: The square knot, $3_1\# - 3_1$.

If one of the three corner loops of 3_1 is connected with any of the three corner loops of another copy of 3_1, then the orientations also agree. The result is a granny knot:

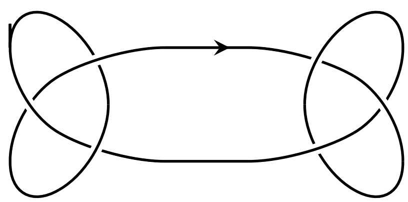


Figure 3.6: The granny knot, $3_1\#3_1$.

Remark 3.2.6. In fact, any two oriented knots K_1 and K_2 can be connected together by cutting K_1 and K_2, then connecting K_1 to K_2 in a way that respects orientation:

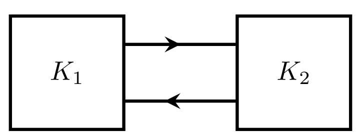


Figure 3.7: The connected sum operation.

Proposition 3.2.7. *If the knots K_1 and K_2 are oriented, the connecting operation shown in the previous remark is well-defined. The resulting knot is denoted $K_1\#K_2$ and the connecting operation is called the connected sum.*

Proposition 3.2.8. *This connected sum commutes and is associative, i.e. the knots $K_1\#K_2$ and $K_2\#K_1$ are isotopic, and the knots $(K_1\#K_2)\#K_3$ and $K_1\#(K_2\#K_3)$ are isotopic.*

Example 3.2.9. The connected sum of the unknot with itself is the unknot:

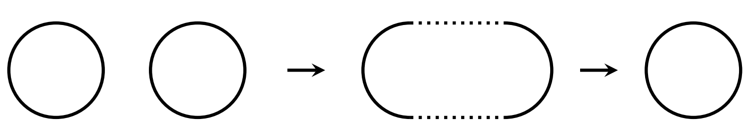


Figure 3.8: The unknot connected sum with itself is the unknot.

Example 3.2.10. In fact, connected sums involving the unknot do nothing:

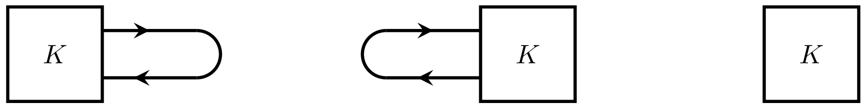


Figure 3.9: The (ambient) isotopic knots $K\#0_1 = 0_1\#K = K$.

3.3 Seifert Surfaces, Forms, and Matrices

Algorithm 3.3.1 (Seifert). *On an oriented knot diagram, choose any arc and follow along in the direction of the orientation. At each crossing, switch to the opposite strand in a way that respects orientation (see Figure 3.10). Eventually this procedure will form a closed circuit, called a Seifert circle. Next choose an arc on the diagram that is not on the previous Seifert circle. Repeating the procedure gives a new Seifert circle. Continue to choose arcs not on the previous Seifert circles until there are no such remaining arcs. Now let each of the Seifert circles be the boundary of a disk, and join together the disks at the crossings by half-twisted bands that respect the crossing (again, see Figure 3.10). What remains is a Seifert surface.*

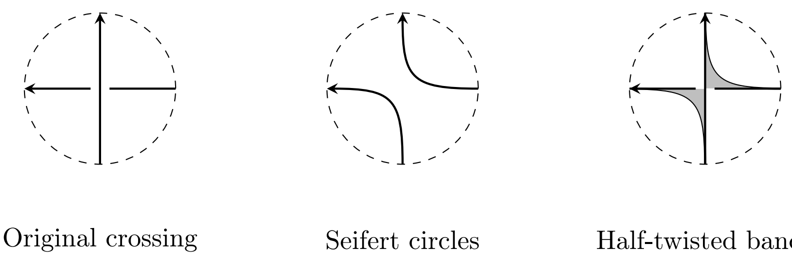


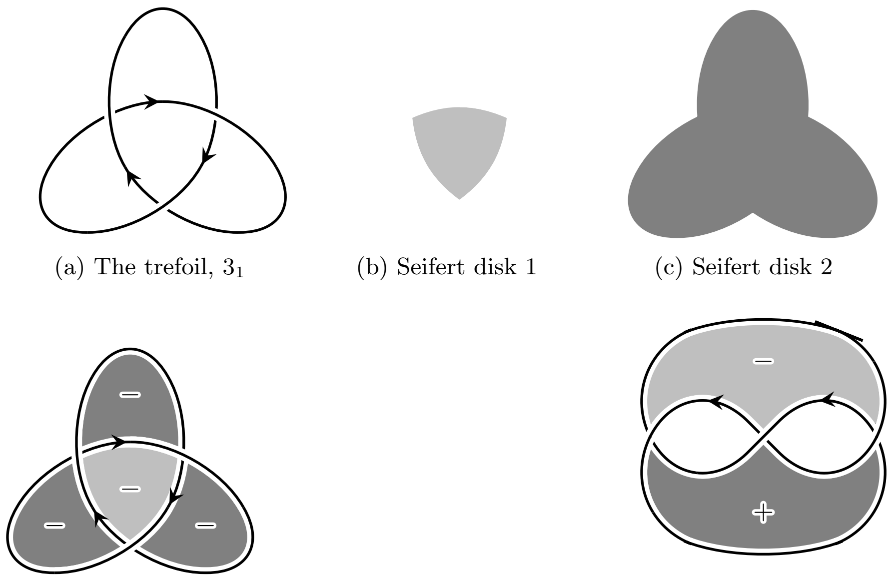
Figure 3.10: The various stages of Seifert's algorithm in the neighborhood of a crossing.

This procedure results in an orientable surface that has a consistent orientation due to the orientation on the boundary (the knot's orientation). The convention is that the oriented knot goes counterclockwise around a positive normal to the surface.

Note that Algorithm 3.3.1 constructs *a* Seifert surface, implicitly meaning that there is more than one possible surface. Indeed, the result of the algorithm depends on which oriented knot diagram one starts with.

Example 3.3.2. In Seifert's algorithm, if m is the number of Seifert circles and k is the number of crossings, the Seifert surface F produced has genus $g_F = \frac{1}{2}(k - m + 1)$. This can be seen by contracting F to a graph with m vertices and k edges. The Euler characteristic χ_F is then $m - k$, so the *genus of the Seifert surface* F is $g_F = \frac{1}{2}(2 - \chi_F - b) = \frac{1}{2}(2 - (m-k) - 1)$ where the boundary of the Seifert surface is the knot, so $b = 1$.

Example 3.3.3. What follows is a step-by-step construction of our preferred Seifert surface for 3_1.



(a) The trefoil, 3_1 (b) Seifert disk 1 (c) Seifert disk 2

(d) Seifert surface (e) Modified Seifert surface

Figure 3.11: The use of Seifert's algorithm to construct a Seifert surface for 3_1.

In the figure above, (a) is the left-handed trefoil. The layered diagram (d) is a top-down view of the Seifert surface obtained by connecting the two disks shown in (b) and (c) with half-twisted bands. Picking an arc on the *outer edge* of the trefoil gives the light gray Seifert disk (a filled in Seifert circle), while picking an arc on the *inner edge* of the trefoil gives the dark gray Seifert disk. In (e) a modified version of the same surface is shown, where the three crossings have been slid around to the bottom and the dark gray disk flipped upside-down.

Since the Seifert surface(s) has 2 Seifert circles and 3 crossings, the genus of Seifert surface(s) is $g_F = \frac{1}{2}(3 - 2 + 1) = 1$.

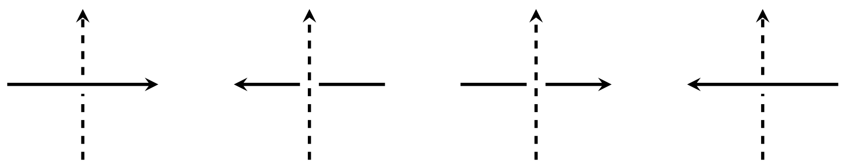
We will take for granted that if K is a knot in S^3, then K has a tubular neighborhood, denoted $N(K)$, that is a solid torus. Details surrounding this fact can be subtle, so we merely state the definition here and point to 2E9 of [13] as a reference.

Definition 3.3.4. The *exterior* of a knot K is defined to be the complement of the interior of its tubular neighborhood, i.e. $X(K) = S^3 \setminus \mathring{N}(K)$.

Definition 3.3.5. The *linking number* $lk(\alpha,\beta)$ of disjoint, oriented, simple closed curves α and β in S^3 is defined to be the class represented by β in $H_1(X(\alpha)) \cong \mathbb{Z}$.

The linking number can easily be computed by taking a signed count of the number of times that the curve β crosses over α, as described in the following algorithm.

Algorithm 3.3.6 (linking number). *On an oriented link diagram consisting of two closed curves α (solid line) and β (dashed line), crossings come in four types:*



Type 1, + Type 2, + Type 3, − Type 4, −

Figure 3.12: The four types of crossings on a link diagram of two closed curves.

To compute $lk(\alpha,\beta)$, pick a point on the curve α. Follow the curve and label each crossing with β according to the above types. The sum of the count of each signed type of the crossings is twice the linking number, i.e. $2lk(\alpha,\beta) = n_1 + n_2 - n_3 - n_4$ where n_i is the count of crossings of type i.

As a consequence of the Jordan curve theorem, $n_1 + n_3 = n_2 + n_4$, so the linking number can also be computed as either $n_1 - n_4$ or $n_2 - n_3$. Note that the former only involves overcrossings of α by β and the latter only involves undercrossings.

Example 3.3.7. Two unlinked curves α and β clearly have $lk(\alpha,\beta) = 0$, since they can be pulled apart so that there are no crossings. The next simplest situation is:

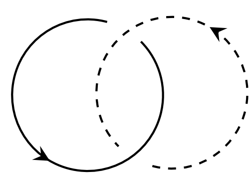


Figure 3.13: A link with linking number −1.

The top crossing is of type 3, and the bottom crossing is of type 4. Thus $2lk(\alpha,\beta) = -1-1$, where α is the dashed curve and β is the solid curve. Thus $lk(\alpha,\beta) = -1$. Notice that $lk(\alpha,\beta) = lk(\beta,\alpha)$, a fact that holds in general.

Now that we have the notions of Seifert surface and linking number, we can define a useful bilinear form that will give rise to a matrix for the surface. Recall that given any simple oriented curve i on F, we can create the *positive push-off* i^+, which is parallel to i and lies above F in the positive normal direction.

Definition 3.3.8. The *Seifert form* of a Seifert surface F is the bilinear form $\theta : H_1(F) \times H_1(F) \to \mathbb{Z}$ defined by $(x, y) \mapsto lk(x, i^+_*(y))$, where the map i^+_* is the map induced on homology by the positive push-off $i^+ : F \to (S_3 \setminus F)$.

Since $H_1(F)$ is a free abelian group of rank $2g$, where g is the genus of the Seifert surface F, we can make the following definition.

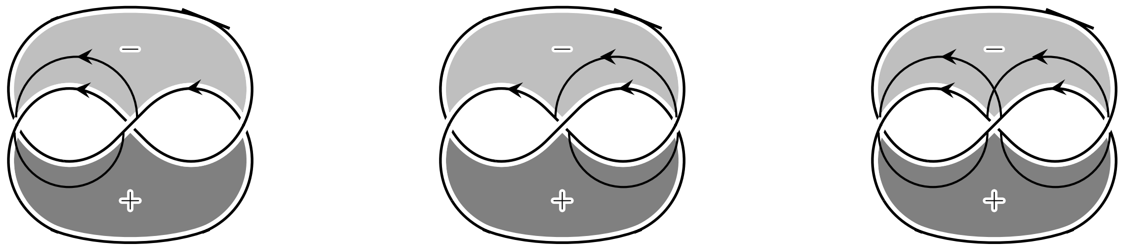
Definition 3.3.9. The *Seifert matrix* (or *linking matrix*) of a Seifert surface F is the matrix

$$A = (\theta(x_i, x_j))_{ij}$$

where $\{x_1, x_2, \ldots, x_{2g}\}$ is a fixed basis of $H_1(F)$.

Remark 3.3.10. A Seifert surface for a nontrivial knot has genus $g \geq 1$. By the previous definition a Seifert matrix for such a Seifert surface has size at least $2g \times 2g$.

Example 3.3.11. In Example 3.3.3 we showed a preferred Seifert surface for 3_1 (of genus 1).



(a) A choice of α. (b) A choice of β. (c) Both α and β.

(d) The push-off α^+. (e) The push-off β^+.

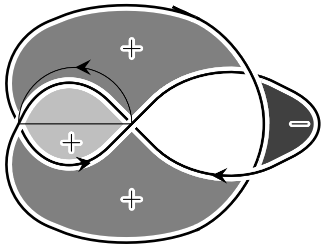
Figure 3.14: A choice of generators and their push-offs for our Seifert surface for 3_1.

Recall that a choice of generators on a genus 1 surface gives two simple closed curves α and β that intersect one another at a single point as in Figure 3.14(c). Since the light gray Seifert disk is negatively oriented coming out of the page, the positive push-off α^+ is below the page in that region (shown as a dotted arc). On the other hand, since the dark gray Seifert disk is positively oriented coming out of the page, α^+ is above the page in that region (shown as a dashed arc).

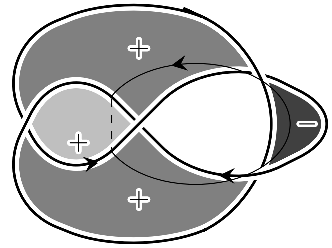
The curves α and α^+ are positively linked (see Figure 3.12 for the convention), so $lk(\alpha, \alpha^+) = 1$. Similarly, β and β^+ are positively linked. For $lk(\beta, \alpha^+)$, notice that this is the same situation as Example 3.3.7. Thus $lk(\beta, \alpha^+) = -1$. Consequently, α and β^+ are unlinked. Compiling this data gives the following Seifert matrix for 3_1:

$$A = \begin{pmatrix} lk(\alpha, \alpha_+) & lk(\alpha, \beta_+) \\ lk(\beta, \alpha+) & lk(\beta, \beta_+) \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$$

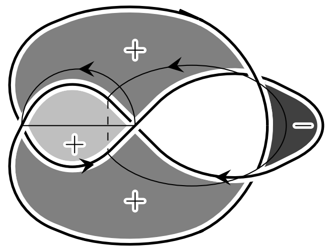
Example 3.3.12. A Seifert surface for 4_1 is shown below in Figure 3.15 and has genus 1. Thus there are 2 generators, say α and β, and the corresponding Seifert matrix will be 2×2.



(a) A choice of generator α.



(b) A choice of generator β.



(c) Both α and β.

Figure 3.15: A choice of generators for our Seifert surface for 4_1.

Note that in Figure 3.15(b), the dashed part of β indicates that β lies on the gray Seifert disk below. For the choice of generators shown, the associated Seifert matrix is:

$$A = \begin{pmatrix} lk(\alpha, \alpha_+) & lk(\alpha, \beta_+) \\ lk(\beta, \alpha+) & lk(\beta, \beta_+) \end{pmatrix} = \begin{pmatrix} -1 & -1 \\ 0 & 1 \end{pmatrix}$$

Now consider $4_1 \# 4_1$.

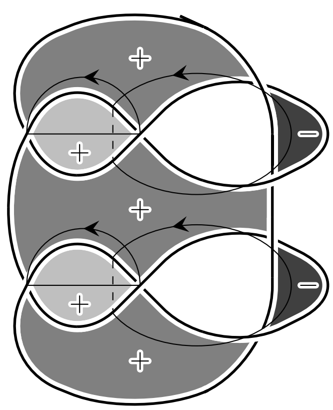


Figure 3.16: A choice of generators for our Seifert surface for $4_1 \# 4_1$.

A Seifert surface for this knot can be constructed by gluing two copies of the above Seifert surface for 4_1 along the boundary of the larger disk. The generators, α, β (top copy of 4_1) and γ, δ (bottom copy of 4_1) for the two copies clearly do not link, so:

$$A' = \begin{pmatrix} lk(\alpha,\alpha_+) & lk(\alpha,\beta_+) & 0 & 0 \\ lk(\beta,\alpha+) & lk(\beta,\beta_+) & 0 & 0 \\ 0 & 0 & lk(\gamma,\gamma_+) & lk(\gamma,\delta_+) \\ 0 & 0 & lk(\delta,\gamma+) & lk(\delta,\delta_+) \end{pmatrix} = \begin{pmatrix} -1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Notice that this is the block sum $A \oplus A$.

Remark 3.3.13. For any knot K, there is a Seifert surface called an (oriented) *disk with bands*. In practice, it is often easier to determine the linking of generators x_i by using this particular construction. For technical details of the construction and its proof, see Proposition 8.2 of [1].

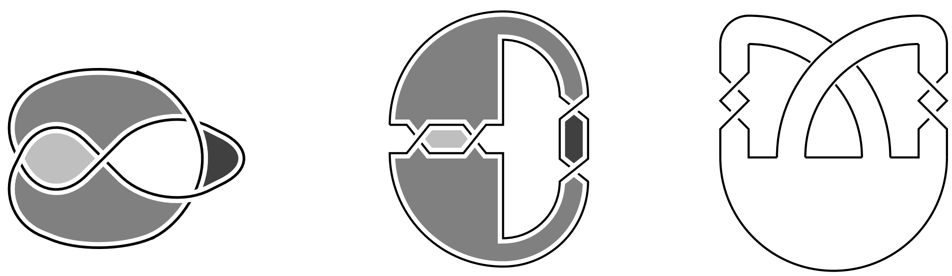


Figure 3.17: Disk with bands from our Seifert surface for 4_1.

Example 3.3.14. Figure 3.18 shows a Seifert surface for 6_1 with genus 1.

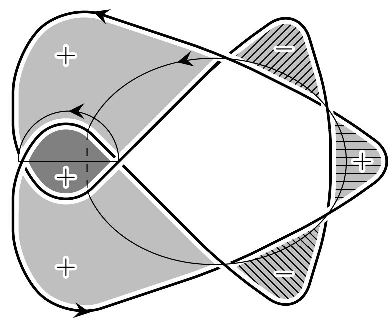


Figure 3.18: A choice of generators for our Seifert surface for 6_1.

With generators α and β as the left and right loops of Figure 3.18, respectively, the Seifert matrix for this Seifert surface is:

$$A = \begin{pmatrix} lk(\alpha, \alpha_+) & lk(\alpha, \beta_+) \\ lk(\beta, \alpha+) & lk(\beta, \beta_+) \end{pmatrix} = \begin{pmatrix} -1 & -1 \\ 0 & 2 \end{pmatrix}$$

Example 3.3.15. If K bounds a Seifert surface F, then rK, $-K$, and mK are bounded by natural Seifert surfaces. These surfaces correspond to matrices A^T, $-A$, and $-A^T$, respectively. To see this, recall Examples **??** and 3.3.3. We use the same basis throughout. For the reverse, the roles of α and α^+ are switched (as are β and β^+). This leaves the diagonals unchanged and swaps the off-diagonals (more precisely, it transposes the matrix A). Switching to the inverse switches crossings so introduces a negative to each linking number (i.e. $-A$). Finally, the mirror image is a combination of these two changes, so the result is $-A^T$.

The Seifert matrix of a knot is not an invariant since it is not unique. In defining the matrix we could have chosen a different Seifert surface F or a different basis for $H_1(F)$. Fortunately, the effects these choices make on the Seifert matrix are known and predictable, so there is a notion of Seifert matrix equivalence called S-equivalence. See [9] Section 1.9 for more details of this construction. An important fact is that knot invariants involving Seifert matrices must not be affected by substituting S-equivalent Seifert matrices.

3.4 Some Knot Invariants

In what follows, A is a Seifert matrix for a knot K with Seifert surface F. The notation $\doteq$ will mean equal up to multiplication by a factor of $\pm t^n$, where n is an integer.

3.4.1 The Alexander Polynomial

Definition 3.4.1. The *Alexander polynomial* of K is

$$\Delta_K(t) \doteq \det(A - tA^T)$$

Theorem 3.4.2. *The Alexander polynomial of a knot K has the following properties:*

(i) $\Delta_K(1) = 1$.

(ii) $\Delta_K(t) \doteq \Delta_K(t^{-1})$.

(iii) $\Delta_{K_1 \# K_2}(t) = \Delta_{K_1}(t) \cdot \Delta_{K_2}(t)$.

(iv) $\Delta_K(t) \doteq \Delta_{mK}(t) \doteq \Delta_{rK}(t)$.

Proof. (i) From the definition of the Alexander polynomial, $\Delta_K(1) \doteq \det(A - A^T)$. So we must show that $\det(A - A^T) = 1$. Recalling Definition 3.3.9, we can write $(A - A^T)_{ij} = lk(x_i x_j^+) - lk(x_i^+, x_j)$. Observe that this is the number of times x_i intersects x_j. By Remark 3.3.13, we then have a block diagonal matrix with blocks of the form $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, and it follows that $\det(A - A^T) = 1$.

(ii) Observe:

$$\begin{aligned} \Delta_K(t) &\doteq \det(A - tA^T) \\ &= \det(A^T - tA) \\ &= (-t)^n \det(A - t^{-1}A^T) \\ &\doteq \Delta_K(t^{-1}) \end{aligned}$$

(iii) Let A_1 and A_2 be Seifert matrices for K_1 and K_2, respectively. Then $\begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$ is a Seifert matrix for $K_1 \# K_2$, and the result follows immediately.

(iv) Recall from 3.3.15 that for a knot K with Seifert matrix A, the knots mK and rK have Seifert matrices A^T, and $-A^T$, respectively. Then the result follows since $\det(A - tA^T) \doteq \det(A^T - tA) \doteq \det(-A^T + tA)$.

□

Example 3.4.3. In Example 3.3.14, we saw that $A = \begin{pmatrix} -1 & -1 \\ 0 & 2 \end{pmatrix}$ for our chosen Seifert surface for 6_1. The Alexander polynomial in this situation is

$$\Delta_{6_1}(t) = \left|A - tA^T\right| = \begin{vmatrix} -1+t & -1 \\ t & 2-2t \end{vmatrix} = -2t^2 + 5t - 2.$$

Observe that the properties (i) and (ii) from the previous theorem hold: $\Delta_{6_1}(1) = -2 + 5 - 2 = 1$ and $\Delta_{6_1}(t^{-1}) = -2t^{-2} + 5t^{-1} - 2 = t^2 \cdot \Delta_{6_1}(t)$ so $\Delta_{6_1}(t^{-1}) \doteq \Delta_{6_1}(t)$.

Example 3.4.4. In Example 3.3.12, we saw that $A = \begin{pmatrix} -1 & -1 \\ 0 & 1 \end{pmatrix}$ for our chosen Seifert surface for 4_1. The Alexander polynomial in this situation is

$$\Delta_{4_1}(t) = \left|A - tA^T\right| = \begin{vmatrix} -1+t & -1 \\ t & 1-t \end{vmatrix} = -t^2 + 3t - 1.$$

Now consider $m4_1$ and $r4_1$. By Example 3.3.15, the new Seifert matrix for $m4_1$ is $A' = A^T$. The Alexander polynomial now becomes

$$\Delta_{m4_1}(t) = \left|A^T - tA\right| = \begin{vmatrix} -1+t & t \\ -1 & 1-t \end{vmatrix} = -t^2 + 3t - 1.$$

Similarly, for $r4_1$ the new Seifert matrix is $A'' = -A^T$, giving Alexander polynomial

$$\Delta_{r4_1}(t) = \left|-A^T + tA\right| = \begin{vmatrix} 1-t & -t \\ 1 & -1+t \end{vmatrix} = -t^2 + 3t - 1,$$

verifying property (iii) in the case of 4_1.

Example 3.4.5. By the previous theorem, the Alexander polynomial cannot distinguish between the granny knot $3_1\#3_1$ and the square knot $3_1\# - 3_1$ by properties (iii) and (iv).

3.4.2 The Signature

Definition 3.4.6. The *signature* of a knot K, denoted by $\sigma(K)$, is the number of positive entries minus the number of negative entries in a diagonalization of $A + A^T$.

Note that the matrix $A + A^T$ is symmetric.

Theorem 3.4.7. *The signature of a knot K has the following properties:*

(i) $\sigma(K)$ *is an even number for any knot K*

(ii) $\sigma(rK) = \sigma(K)$, $\sigma(mK) = -\sigma(K)$. *and* $\sigma(-K) = -\sigma(K)$

(iii) $\sigma(K_1 \# K_2) = \sigma(K_1) + \sigma(K_2)$

Proof. (i) Recall the Spectral Theorem for Symmetric Matrices (see [13], 8E1, page 216): Any symmetric matrix over a field (of characteristic not two) is congruent to a diagonal matrix over that field. Though Seifert matrices will have only integer entries, we can view these matrices as a subset of matrices with real entries. Thus it suffices to consider a diagonal matrix $A + A^T$ (which is congruent to a Seifert matrix for K).

Let P be the number of positive entries and N be the number of negative entries of $A + A^T$. Since Seifert matrices are even rank square matrices (of say, rank $2g$ where g is the genus of the corresponding Seifert surface), if we knew that $\det(A + A^T) \neq 0$ then $P + N = 2g$. But then $P = 2g - N$, so $P - N = 2g - N - N = 2(g - N)$, an even number. Hence the signature of a knot K is always even.

(ii)-(iii) Recall from 3.3.15 that for a knot K with Seifert matrix A, the knots mK and rK have Seifert matrices A^T, and $-A^T$, respectively. Changing from A to A^T leaves the diagonalization of $A + A^T$ unchanged, so $\sigma(rK) = \sigma(K)$. Similarly, changing from A to $-A^T$ only changes the sign of the diagonalization of $A + A^T$, so $\sigma(mK) = -\sigma(K)$. Combining these results gives

$$\sigma(-K) = \sigma(rmK) = \sigma(mK) = -\sigma(K)$$

(iv) This follows from the fact if A is the block sum of diagonal matrices A_1 and A_2, i.e.

$$A = A_1 \oplus A_2 = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$$

then $A + A^T$ can be diagonalized by diagonalizing $A_1 + A_1^T$ and $A_2 + A_2^T$ separately. The diagonalization of $A + A^T$ is then the block sum of these respective diagonalizations.

□

Example 3.4.8. Recall from Example 3.3.11 that a Seifert matrix for 3_1 is $\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$. Thus $\sigma(3_1) = 2$.

3.5 Slice Knots

Going back to the connected sum operation described in Section 3.2, the connected sum is commutative and associative (Proposition 3.2.8) and the unknot works as an identity element (Example 3.2.10). This gives a commutative semigroup structure to the set of all [isotopy classes of] knots. Unfortunately, there is no inverse for a nontrivial knot in this setting. In this section, we will see that we can form inverses if we expand our equivalence class of the unknot. This will underly the algebraic structure introduced in the next chapter.

Notation 3.5.1. B^4 is a 4-dimensional ball, and D^2 is a 2-dimesional disk.

It is an intuitive fact that a knot is unknotted if and only if it bounds a disk, i.e. $K = \partial D^2$, in S^3. A less intuitive fact is that if a knot is embedded into S^4 (instead of S^3), it can always be unknotted. In effect, there is too much "room to wiggle" to allow knotting. This wiggling is accomplished by requiring that the disk obtained by unknotting straddles both of the copies of D^4 that are glued along S^3 (viewing $S^4 = D^4 \cup_{S_3} D^4$), so that while unknotting one strand of a crossing goes into each copy of D^4. This begs the question: What happens when we require a knot to lie entirely in one of these copies of D^4?

With just a disk, nothing interesting happens: Any knot K in S^3 is the boundary of a disk embedded in D^4. This can be seen by taking a cone over K, which is homeomorphic to a disk.

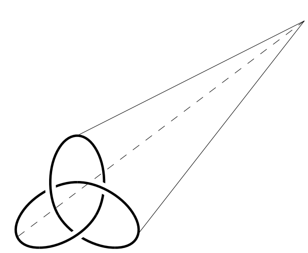


Figure 3.19: Coning the trefoil.

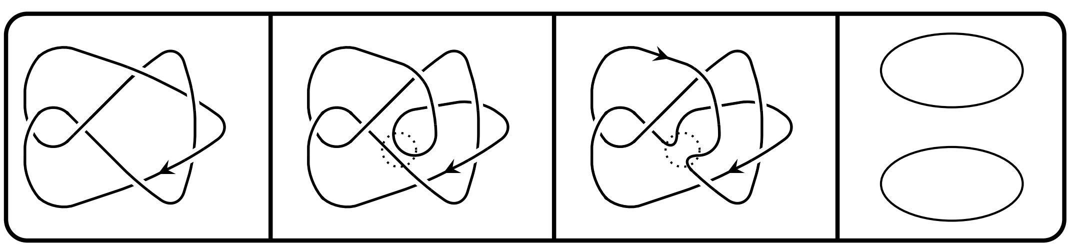
Definition 3.5.2. A knot K is *smoothly slice* (or simply *slice* since we are in the setting of smooth manifolds) if there is a pair (B^4, D^2) such that $(S^3, K) = \partial(B^4, D^2)$ where D^2 is a smooth properly embedded disk (here properly embedded means $D^2 \cap S^3 = K$). Such a disk is called a *slice disk*.

How can we tell if a knot is slice? The slice disk can be visualized by way of what is called a *slice movie*. If a knot K is slice, then it bounds a slice disk. It turns out that concentric 3-spheres intersect the slice disk in a predictable way. In particular, there are only finitely many "interesting" intersections (these are the critical points of Morse theory, see [10] for more details) each of which provides a frame in the slice movie. In short, if we are given a slice movie, we can reconstruct the slice disk by smoothing inbetween the frames (and capping off circles in the last frame). In practice, additional frames may be added for clarity (e.g., frame B in Example 3.5.5 below).

Remark 3.5.3. In general, movies can be made whether or not K is slice. If a movie does not end with unlinked circles, then the smoothing inbetween frames produces a surface that is not a disk. This more complicated surface is thus not the required *slice disk*.

Remark 3.5.4. Unfortunately, there is no algorithm to find slice movies in general. Thus it is difficult to find a slice movie even if the knot is priorly known to be slice by some other technique.

Example 3.5.5. The following figure is a slice movie starting with Stevedore's knot, 6_1, in frame A. In frame B, a kink is added and the dotted circle shows two arcs of opposite orientation. In frame C, the two arcs in the dotted circle are cut and rotated 90°, forming a saddle point. Frame D shows the resulting unlink when the two circles in the previous frame are pulled apart.



A B C D

Figure 3.20: A slice movie for Stevedore's knot 6_1.

Theorem 3.5.6. *For any knot K, the connected sum of K with its inverse, i.e. $K\# - K$, is slice.*

Proof. Let B be a 3-ball in the interior of S^3, and let $B \cap K$ be a straight arc in B. To see that $K\# - K$ is slice, observe the existence of the pair $(\overline{S^3 - B} \times I, \overline{K - B \cap K} \times I)$, where $I =$

$[0,1]$ and $\overline{K - B \cap K} \times I$ is the required slice disk. (Note: the disk is properly embedded since $(\overline{K - B \cap K} \times I) \cap \partial(\overline{S^3 - B} \times I) = K$.) □

The next series of results establish some relationships between connected sums and slice knots that will be useful in the beginning of Chapter 4.

Theorem 3.5.7. *(a) If K_1 and K_2 are slice, then $K_1\#K_2$ is also slice.*

(b) If $K_1\#K_2$ and K_1 (resp. K_2) are slice, then K_2 (resp. K_1) is also slice.

(c) If K is slice, then $-K$ is slice.

Proof. (a) To see that $K_1\#K_2$ is slice when both K_1 and K_2 are slice, observe that performing a *band move* in the following schematic for $K_1\#K_2$ separates the knots K_1 and K_2. Thus the slice disk for $K_1\#K_2$ is a saddle connecting the slice disks for K_1 and K_2 in B^4.

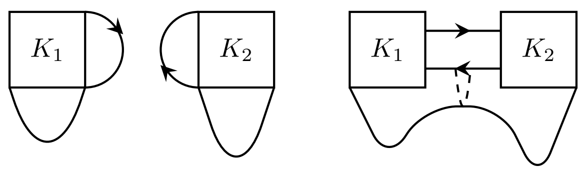


Figure 3.21: Rough schematics of the slice disks for K_1, K_2, and $K_1\#K_2$ (left to right).

Said differently, if D_1 and D_2 are the slice disks for K_1 and K_2, respectively, then the connected sum of the pairs (B^4, D_1) and (B^4, D_2) along the boundary disks is the required pair for $K_1\#K_2$ to be slice.

(b) Let $K_1\#K_2$ and K_1 be slice, with slice disks $\widehat{D}$ and D. Choose 3-balls $\widehat{B}$ and B such that $\overline{K_1 - B \cap K_1}$ is a straight arc in the 3-ball $\overline{S^3 - B}$ and $(B, B \cap K) = (\widehat{B}, \widehat{B} \cap (K_1\#K_2))$.

To see that K_2 is slice, observe the existence of the pair $(B^4, D) \sqcup_\phi (B^4, \widehat{D})$, where ϕ identifies $(B, B \cap K)$ in (B^4, D) with $(\widehat{B}, \widehat{B} \cap (K_1\#K_2))$ in $(B^4, \widehat{D})$.

(c) Using part (b) with $K = K_1 = K_2$ together with Theorem 3.5.6 immediately gives the result. That is, since $K\# - K$ is always slice, if K is known to be slice, then $-K$ must also be slice.

□

Chapter 4

Concordance

4.1 Knot Concordance

Previously, in Theorem 3.5.6, we saw that the connected sum of K with its inverse $-K$ was slice. The notation seems to suggest that we might have stumbled onto some sense of invertibility. To investigate this, we make the following definition. Here $I = [0,1]$.

Definition 4.1.1. A knot K_0 is *concordant* to K_1, denoted $K_0 \sim K_1$, if there is a 2-dimensional smooth manifold M in $S^3 \times I$ such that

(i) M is diffeomorphic to the annulus $S^1 \times I \subset S^3 \times I$,

(ii) $\partial M = M \cap (S^3 \times \{0,1\})$, and

(iii) $M \cap (S^3 \times \{0\}) = K_0$ and $M \cap (S^3 \times \{1\}) = K_1$.

Remark 4.1.2. The above definition is equivalent to the following definition: Two knots K_0 and K_1 are *concordant* if $K_0 \# -K_1$ is slice. For the proof, see [9], Theorem 2.3.2. We will call this characterization *concordance by disks*, as opposed to Definition 4.1.1, which will be called *concordance by annuli*.

There are a few things to check, including that $\sim$ is an equivalence relation and that connected sum of equivalence classes is well-defined. Then we can define a group under connected sum on the set of all isotopy classes of knots modulo $\sim$.

Theorem 4.1.3. *The following are true:*

(i) Knot concordance is an equivalence relation on the set of all knots.

(ii) Isotopic knots are concordant.

(iii) The connected sum of knots induces a well-defined binary operation on the knot concordance equivalence relation.

(iv) Under connected sum, the knot concordance equivalence classes form an abelian group $\mathcal{C}$.

Proof. We will get a lot of mileage out of Theorem 3.5.6, which states that $K\# -K$ is always slice.

(i) Reflexivity. From Theorem 3.5.6, $K\# -K$ is always slice. Using the notion of concordance by disks, we have at once $K \sim K$.

Symmetry. Let $K_1 \sim K_2$. Then $K_1\# -K_2$ is slice. By Theorem 3.5.7(c) and commutativity of connected sums, $-(K_1\# -K_2) = -K_1\#K_2 = K_2\# -K_1$. Hence $K_2 \sim K_1$.

Transitivity. Let $K_1 \sim K_2$ and $K_2 \sim K_3$. Then $K_1\# -K_2$ and $K_2\# -K_3$ are slice. By Theorem 3.5.7(a), $(K_1\# -K_2)\#(K_2\# -K_3)$ is slice. By associativity and commutativity of connected sums, this means that $(K_1\# -K_3)\#(K_2\# -K_2)$ is slice. Now, since $K_2\# -K_2$ is always slice, by Theorem 3.5.7(b) we have that $(K_1\# -K_3)$ is slice. Hence $K_1 \sim K_3$.

(ii)-(iv) Let K_1 and K_2 be isotopic (this implies that $-K_1$ and $-K_2$ are also). By Theorem 3.5.6, $K_1\# -K_1$ is slice. Since the connected sum is a well-defined operation on isotopy classes of knots, this implies that $K_2\# -K_1$ and $K_1\# -K_2$ are slice. Thus K_1 and K_2 are concordant to each other.

(iii) This can be seen from the following Lemma:

Lemma 4.1.4. *If K_1 and K_2 are concordant and J_1 and J_2 are concordant, then $K_1\#J_1$ and $K_2\#J_2$ are concordant.*

Proof of the Lemma. By assumption, $K_1\#-K_2$ and $J_1\#-J_2$ are slice. From Theorem 3.5.7(a) we know that $(K_1\#-K_2)\#(J_1\#-J_2)$ is slice. By associativity of the connected sum of knots, we have

$$(K_1\# -K_2)\#(J_1\# -J_2) = (K_1\#J_1)\# -(K_2\#J_2)$$

so $(K_1\#J_1)\# -(K_2\#J_2)$ is slice, i.e. $K_1\#J_1$ and $K_2\#J_2$ are concordant. □

(iv) Associativity and commutativity follow directly from the connected sum on isotopy classes of knots. Recall that $K\# - 0_1 = K\#0_1 = K$ by Example 3.2.10. So K is concordant to the unknot if and only if K is slice. Thus the concordance class of the unknot acts as an identity concordance. Finally, since $K\# - K$ is always slice, all knots are invertible.

□

Now that we have this knot concordance group $\mathcal{C}$, the next step is to understand its structure. As it turns out, this is a tall order: in over half a century since its original definition by Fox and Milnor, it is merely known that $\mathbb{Z}^\infty \oplus (\mathbb{Z}/2)^\infty \subset \mathcal{C}$ and that there is an epimorphism from $\mathcal{C}$ onto $\mathbb{Z}^\infty \oplus (\mathbb{Z}/2)^\infty \oplus (\mathbb{Z}/4)^\infty$. As is standard in the literature, the notation $\mathbb{Z}^\infty$ means a direct sum of countably many copies of $\mathbb{Z}$.

4.2 Algebraic Concordance

4.2.1 The Algebraic Concordance Group $\mathcal{G}^{\mathbb{Z}}$

The reader may have noticed that while it is difficult to determine whether or not a knot is slice using topological methods, it is at least easy to find a Seifert matrix. Finding these matrices allows one to perform calculations of knot invariants using relatively simple techniques. If the two notions of sliceness and Seifert matrices could be related in some sense, then we could hope to transfer techniques from the latter to results about the former. Before we establish this connection, we suss out the algebraic structure of Seifert matrices.

Recall the notion of metabolic symmetric inner product spaces from Section 2.6 (in particular, see Definition 2.6.2 and Lemma 2.6.4). Note that typically a Seifert matrix A will *not* be a symmetric matrix. For instance, the Seifert matrix $\begin{pmatrix} -1 & -1 \\ 0 & 1 \end{pmatrix}$ from Example 3.3.12 is clearly not symmetric. Moreover, the discussion of bilinear spaces in Chapter 2 handles the case of a vector space V over a field $\mathbb{F}$, as opposed to free modules over the ring $\mathbb{Z}$. Nevertheless, the constructions follow analogously.

Definition 4.2.1. A $(2n \times 2n)$ Seifert matrix A is called *algebraically slice* (or *metabolic*) if it is congruent to a matrix of the form

$$\begin{pmatrix} 0 & A_1 \\ A_2 & A_3 \end{pmatrix}$$

where each entry of the above matrix is $(n \times n)$.

Equivalently, the Seifert form associated to A is algebraically slice if there is an n-dimensional summand of the underlying free $\mathbb{Z}$-module where the form vanishes. Such summands are called *metabolizers*.

In Theorem 2.6.6 we saw that for a bilinear space (V, β), the orthogonal sum $(V, \beta) \oplus (V, -\beta)$ is metabolic. This motivates the construction of an equivalence relation using the above notion of algebraically sliceness.

Definition 4.2.2. Two Seifert matrices A_1 and A_2 are *algebraically concordant* if the orthogonal sum $A_1 \oplus -A_2 = \begin{pmatrix} A_1 & 0 \\ 0 & -A_2 \end{pmatrix}$ is algebraically slice.

Now we have Witt equivalence (Definition 2.7.1 and Lemma 2.7.2).

Lemma 4.2.3. *Algebraic concordance is an equivalence relation.*

Next we have Witt Cancellation (Corollary 2.5.1) for Seifert matrices.

Lemma 4.2.4. *Let A and N be Seifert matrices, and suppose that A and $N \oplus A$ are algebraically slice. Then N is algebraically slice.*

Finally, we form a Witt group (Definition 2.7.5).

Theorem 4.2.5. *The Seifert matrices and orthogonal sum of matrices, denoted $\mathcal{G}^{\mathbb{Z}}$. This is called the* integral algebraic concordance group*.*

Now we are ready to provide the connection between sliceness and algebraic sliceness alluded to in the beginning of this section.

Theorem 4.2.6 (Half lives, half dies). *If K is a slice knot, F a Seifert surface for K, and θ a Seifert form for F, then there exists a summand (called a* metabolizer*) M of $H_1(F)$ such that $2rk(M) = rk(H_1(F))$ and $\theta_{M \times M} = 0$.*

Proof of Theorem 4.2.6. We give the following two Lemmas, which imply the theorem.

Lemma 4.2.7. *If W is a compact connected orientable 3-manifold such that ∂W is a connected 2-manifold of genus g, then there is a summand M of $H_1(\partial W)$ such that $2 \cdot \mathrm{rk}(M) = \mathrm{rk}(H_1(W))$.*

Proof of Lemma 4.2.7. First consider the long exact sequence on homology of the pair $(W, \partial W)$:

$$\cdots 0 \longrightarrow H_3(W,\partial W) \xrightarrow{\cong} H_2(\partial W) \xrightarrow{i_{2*}} H_2(W) \xrightarrow{j_{2*}}$$
$$H_2(W,\partial W) \xrightarrow{\partial_1} H_1(\partial W) \xrightarrow{i_{1*}} H_1(W) \xrightarrow{j_{1*}} H_1(W,\partial W) \xrightarrow{\partial_0}$$
$$H_0(\partial W) \xrightarrow{\cong} H_0(W) \longrightarrow 0 \cdots$$

We will show that $M = \ker(i_{1*})$ is such a summand, so it suffices to show that $\mathrm{rk}(\ker i_{1*}) = g$ and that $\ker(i_{1*})$ is actually a summand of $H_1(\partial W)$.

By Lefschetz duality, $H_1(W,\partial W) \cong H^2(W)$ and $H_2(W,\partial W) \cong H^1(W)$. Note that $\mathrm{rk}(H_2(W)) = \mathrm{rk}(H^2(W))$ and $\mathrm{rk}(H_1(W)) = \mathrm{rk}(H^1(W))$. Thus

$$\mathrm{rk}(H_1(W,\partial W)) = \mathrm{rk}(H_2(W)) =: s$$

and

$$\mathrm{rk}(H_2(W,\partial W)) = \mathrm{rk}(H_1(W)) =: r$$

Since all homologies are finitely generated,

$$\begin{aligned} H_2(W) &\cong \mathbb{Z}^s \oplus T_1 \\ H_2(W,\partial W) &\cong \mathbb{Z}^r \oplus T_2 \\ H_1(W) &\cong \mathbb{Z}^r \oplus T_3 \\ H_1(W,\partial W) &\cong \mathbb{Z}^s \oplus T_4 \end{aligned}$$

for some torsion subgroups T_i, $i = 1, \ldots, 4$.

Exactness of the sequence together with the (First) Isomorphism Theorem for groups (and the surjectivity of ∂_2) gives,

$$\begin{aligned}
\mathrm{rk}(\ker i_{1*}) &= \mathrm{rk}(\mathrm{im}\partial_1) \\
&= \mathrm{rk}(H_2(W, \partial W)/\ker\partial_1) \\
&= r - \mathrm{rk}(\mathrm{im} j_{2*}) \\
&= r - \mathrm{rk}(H_2(W)/\ker j_{2*} \\
&= r - (s - \mathrm{rk}(\mathrm{im} i_{2*}) \\
&= r - (s - \mathrm{rk}(\mathbb{Z}/\ker i_{2*}) \\
&= r - (s - (1 - \mathrm{rk}(\mathrm{im}\partial_2)) \\
&= r - (s - (1 - \mathrm{rk}(\mathbb{Z}))) \\
&= r - s
\end{aligned}$$

On the other hand, since $H_1(\partial W) \cong \mathbb{Z}^{2g}$ by hypothesis, it follows that

$$2g - (r - s) = \mathrm{rk}(H_1(\partial W)/\ker i_{1*}) = \mathrm{rk}(\mathrm{im} i_{1*}).$$

Thus, again by exactness of the sequence and the Isomorphism Theorem (and the injectivity of i_{0*}),

$$\begin{aligned}
0 &= \mathrm{rk}(\ker i_{0*}) \\
&= \mathrm{rk}(\mathrm{im}\partial_0) \\
&= \mathrm{rk}(H_1(W, \partial W)/\ker\partial_0) \\
&= s - \mathrm{rk}(\ker\partial_0) \\
&= s - \mathrm{rk}(\mathrm{im} j_{1*}) \\
&= s - \mathrm{rk}(H_1(W)/\ker j_{1*}) \\
&= s - (r - \mathrm{rk}(\ker j_{1*})) \\
&= s - (r - \mathrm{rk}(\mathrm{im} i_{1*})) \\
&= s - (r - (2g - (r - s)) \\
&= 2g - 2(r - s)
\end{aligned}$$

hence $g = r - s$ so $\mathrm{rk}(\ker i_{1*}) = g$, $M = \ker i_{1*}$, and $2\mathrm{rk}(M) = \mathrm{rk}(H_1(\partial W))$.

Next, to see that M is actually a summand of $H_1(\partial W)$, observe that by the Universal Coefficient Theorem,

$$H^1(W) = \mathrm{Hom}(H_1(W), \mathbb{Z}) \oplus \mathrm{Ext}(H_0(W), \mathbb{Z})$$

so $H^1(W)$ is torsion-free. By Lefschetz duality, this means that $T_2 = 0$. By the Isomorphism Theorem and exactness,

$$\mathrm{im}\partial_1 \cong H_2(W, \partial W)/\mathrm{ker}\partial_1 \cong \mathbb{Z}^r/\mathrm{im}j_{2*} \cong \mathbb{Z}^r/(H_2(W)/\mathrm{ker}j_{2*})$$

so T_1 is also zero. Hence M is a direct summand. □

The following Lemma goes too far beyond the scope of this thesis to include its proof. On page 222 of Rolfsen [13], the proof is given as an exercise with a hint.

Lemma 4.2.8. *If a slice knot K (in S^3) bounds a disk Δ in D^4 and K also bounds a Seifert surface F, then there is an orientable 3-manifold W in D^4 such that $W \cap S^3 = F$ and $\partial W = F \cup \Delta$.*

We now finish the proof of the Theorem 4.2.6. From Lemma 4.2.8, we have that $\partial W = F \cup \Delta$, so $H_1(\partial W) = H_1(F)$ (note that $H_1(\Delta) = 0$ since Δ is a disk). If α, $\beta \in M$, then there are surfaces X and Y such that $\alpha = \partial X$ and $\beta = \partial Y$. Pushing off α allows X to be pushed off (write $\partial X^+ = \alpha^+$) to make the intersection of X^+ and Y empty, i.e. $lk(\alpha^+, \beta) = 0$. Hence a Seifert form of F vanishes when restricted to M. □

Recalling the definition of algebraically slice (and $lk(\alpha^+, \beta) = 0$ in the Theorem) gives the following corollary immediately:

Corollary 4.2.9. *Slice knots are algebraically slice. In particular, there is a group epimorphism from $\mathcal{C}$ onto $\mathcal{G}^{\mathbb{Z}}$.*

Corollary 4.2.10. *If K is an algebraically slice knot, then $\Delta_K(t)$ is of the form $f(t)f(t^{-1})$.*

Proof. Let A be a Seifert matrix for K. By Theorem 4.2.6, it may be assumed that

$$A = \begin{pmatrix} 0 & A_1 \\ A_2 & A_3 \end{pmatrix}$$

So that

$$\begin{aligned}\Delta_K(t) &\doteq \det(A - tA^T)\\ &= \begin{pmatrix} 0 & A_1 - tA_2^T \\ A_2 - tA_1^T & A_3 - tA_3^T \end{pmatrix}\\ &= \det(A_1 - tA_2^T)\cdot\det(A_2 - tA_1^T)\\ &\doteq \det(A_1 - tA_2^T)\cdot\det(A_1 - t^{-1}A_2^T)\end{aligned}$$

Now define $f(t) = \det(A_1 - tA_2^T)$ and the result follows. □

Corollary 4.2.11. *If K is an algebraically slice knot, then $\sigma(K) = 0$.*

Proof. Let A be a Seifert matrix for K. By Theorem 4.2.6, it may be assumed that

$$A = \begin{pmatrix} 0 & A_1 \\ A_2 & A_3 \end{pmatrix}$$

So that

$$\begin{aligned}A + A^T &= \begin{pmatrix} 0 & A_1 + A_2^T \\ A_2 + A_1^T & A_3 + A_3^T \end{pmatrix}\\ &= \begin{pmatrix} 0 & A_1 + A_2^T \\ (A_1 + A_2^T)^T & A_3 + A_3^T \end{pmatrix}\end{aligned}$$

Since $\sigma_K(t)$ is nonzero, $A + A^T$ is nonsingular. Thus $A_1 + A_2^T$ is nonsingular, and

$$P = \begin{pmatrix} (A_1 + A_2^T)^{-1} & 0 \\ (A_3 + A_3^T)\cdot(A_1 + A_2^T)^{-1} & -I \end{pmatrix}$$

is also nonsingular (where I is the identity matrix). Hence

$$P(A + A^T)P^T = \begin{pmatrix} 0 & -I \\ -I & 0 \end{pmatrix}$$

and, consequently, $\sigma_K(t) = \sigma(A + A^T) = \sigma(P(A + A^T)P^T) = 0$. □

The next three examples show how to use Corollaries 4.2.9 and 4.2.11 to determine order of elements in the groups $\mathcal{C}$ and $\mathcal{G}^{\mathbb{Z}}$.

Example 4.2.12. The following demonstrates that there are elements of infinite order in $\mathcal{C}$ and $\mathcal{G}^{\mathbb{Z}}$.

As we saw in Example 3.4.8, $\sigma(3_1) = 2$. By Corollary 4.2.11, it follows that 3_1 is not slice. Moreover, since the signature of a knot is additive, for $m \in \mathbb{Z}$,

$$\sigma(\underbrace{3_1\#3_1\#\cdots\#3_1}_{m \text{ times}}) = 2 \cdot m$$

Hence 3_1 is not does not have finite order in $\mathcal{G}^{\mathbb{Z}}$. By Corollary 4.2.9, 3_1 does not have finite order in $\mathcal{C}$ either.

Remark 4.2.13. The argument used in the above example can be used for any knot K: if $\sigma(K) \neq 0$, then K has infinite order in $\mathcal{G}^{\mathbb{Z}}$. Thus if K has finite order in $\mathcal{G}^{\mathbb{Z}}$, it must be that $\sigma(K) = 0$.

Example 4.2.14. The following demonstrates that the concordance relation is weaker than the isotopy relation.

First note two properties of the knot 6_1:

(i) Recall from Example 3.3.14 that a Seifert matrix for 6_1 is

$$A = \begin{pmatrix} -1 & -1 \\ 0 & 2 \end{pmatrix}.$$

Since

$$\begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix}$$

is a diagonalization of A, we have that $\sigma(6_1) = 0$.

(ii) Using the same Seifert matrix as above, Example 3.4.3 showed that the Alexander polynomial of 6_1 is

$$\Delta_{6_1}(t) \doteq -2t^2 + 5t - 2.$$

Notice that this polynomial is reducible in up to units in $\mathbb{Z}[t, t^{-1}]$:

$$\begin{aligned} -2t^2 + 5t - 2 &\doteq t^{-1}(-2t^2 + 5t - 2) \\ &= (2t - 1)(2t^{-1} - 1) \\ &= f(t)f(t^{-1}) \end{aligned}$$

Though we have established these two properties, *a priori* it is not known if 6_1 is slice.

On the other hand, Example 3.5.5 exhibited a slice movie for 6_1. As was mentioned then, in general, this is the only known technique to show that a knot is slice. By Corollary 4.2.9, it follows that 6_1 is also algebraically slice. This demonstrates that (algebraic) concordance relation is weaker than the isotopy relation.

Example 4.2.15. It is a well-known fact that 4_1 is isotopic to its inverse (see [3] page 10 for a series of figures demonstrating this fact.) Thus $4_1\#4_1$ is isotopic to, say, $4_1\# - 4_1$. Since $4_1\# - 4_1$ is slice by Theorem 3.5.6, it follows that $4_1\#4_1$ is slice. By Corollary 4.2.9, $4_1\#4_1$ is also algebraically slice. This implies that 4_1 is either order 1 or order 2 in $\mathcal{C}$ and $\mathcal{G}^{\mathbb{Z}}$, respectively.

Recall also that in Example 3.4.4, we saw that the Alexander polynomial of 4_1 is

$$\Delta_{4_1}(t) \doteq t^2 - 3t + 1$$

which is irreducible over $\mathbb{Z}[t, t^{-1}]$. By Corollary 4.2.10, 4_1 is not an algebraically slice knot. By Corollary 4.2.9, 4_1 is not a slice knot either. Thus 4_1 has order 2 in $\mathcal{C}$ and $\mathcal{G}^{\mathbb{Z}}$.

Note that from Example 3.3.12, a Seifert matrix for 4_1 is

$$A = \begin{pmatrix} -1 & -1 \\ 0 & 1 \end{pmatrix}.$$

Furthermore, since

$$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

is a diagonalization of A, we have that $\sigma(4_1) = 0$. We say that the signature of a knot being zero *does not obstruct* sliceness.

Theorem 4.2.16. *Every class in $\mathcal{G}^{\mathbb{Z}}$ has a nonsingular representative.*

Proof (Levine). Let A be a $2n \times 2n$ matrix with entries in $\mathbb{Z}$. If A is nonsingular, then we are done, so assume that A is singular. Then by elementary row operations on A, we can assume that A can be written with a row of zeros, let's say the first row. Then there is a nonzero $(2n-1) \times 2n$ matrix A_1 such that A is congruent to

$$\begin{pmatrix} 0 & \cdots & 0 \\ & A_1 & \end{pmatrix}$$

Column operations then leave the first row unchanged. If we now perform row operations on A_1, we can assume its first column is zero, except at one row, let's say the first row of A_1. Then A is congruent to

$$\begin{pmatrix} 0 & \cdots & 0 \\ a_1 & & \\ 0 & & \\ \vdots & A_2 & \\ 0 & & \end{pmatrix}$$

where $a_1 \in \mathbb{Z}$. Next, write A_2 as

$$A_2 = \begin{pmatrix} a_2 & M \\ N & A_3 \end{pmatrix}$$

where M is a $1 \times (2n-2)$ matrix, N is a $(2n-2) \times 1$ matrix, A_3 is a $(2n-2) \times (2n-2)$ matrix, and $a_2 \in \mathbb{Z}$. Written in this way, it is straightforward to see that if A is a Seifert matrix (that is, if $\det(A - A^T) \neq 0$), then A_3 is also a Seifert matrix (given that a_1 is nonzero).

It remains to show that A is algebraically concordant to A_3. Actually, by considering the block sum

$$\begin{pmatrix} A & 0 \\ 0 & -A_3 \end{pmatrix}$$

it suffices to show that A is algebraically slice if A_3 is algebraically slice. To see this, consider a congruent matrix PA_3P^T of the form

$$\begin{pmatrix} 0 & P_1 \\ P_2 & P_3 \end{pmatrix}$$

Then QAQ^T is of the form

$$\begin{pmatrix} 0 & Q_1 \\ Q_2 & Q_3 \end{pmatrix}$$

where

$$Q = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & \vdots & & & \\ & & & P & \\ \vdots & 0 & & & \\ 0 & 1 & 0 & \cdots & 0 \end{pmatrix}$$

□

4.2.2 The Algebraic Concordance Group $\mathcal{G}^{\mathbb{F}}$

Unfortunately, understanding the algebraic structure of Seifert matrices directly is also difficult, so first we expand our notion of Seifert matrices from integer entries to entries from a field $\mathbb{F}$ of characteristic not 2.

Definition 4.2.17. Let (V, A) be a bilinear space, where V is a vector space over $\mathbb{F}^{2n}$. Then, by Section 2.2.2, the bilinear form A can be represented by a $2n \times 2n$ dimensional matrix (of the same name) with entries in a field $\mathbb{F}$. If the matrix A satisfies the additional condition that $\det\big((A - A^T)(A + A^T)\big)$ is nonzero, then A is called an $\mathbb{F}$-*Seifert matrix* (or an *admissible matrix* in [8]).

The notions of *algebraic sliceness* and *algebraic concordance* for $\mathbb{F}$-Seifert matrices are constructed as in the previous section.

Theorem 4.2.18. *The set of $\mathbb{F}$-Seifert matrices forms a Witt group under algebraic concordance and orthogonal sum of matrices, denoted $\mathcal{G}^{\mathbb{F}}$. This is called the algebraic concordance group of the field $\mathbb{F}$.*

Recall that in Section 2.7, we introduced the Witt ring over a field $\mathbb{F}$, denoted $W(\mathbb{F})$. This construction required our bilinear spaces to be symmetric inner product spaces. We mentioned that for $\mathcal{G}^{\mathbb{Z}}$ the Seifert matrices are not necessarily symmetric, and the same holds true for $\mathcal{G}^{\mathbb{F}}$ and $\mathbb{F}$-Seifert matrices. Also, Theorem 4.2.16 holds for $\mathcal{G}^{\mathbb{F}}$ with little modification.

Since we now have a definition for Seifert matrices over a field $\mathbb{F}$, a natural choice for $\mathbb{F}$ is the rationals $\mathbb{Q}$ by the following theorem.

Theorem 4.2.19. *The natural inclusion from $\mathcal{G}^{\mathbb{Z}}$ to $\mathcal{G}^{\mathbb{Q}}$ is a monomorphism.*

Said differently, we can view $\mathcal{G}^{\mathbb{Z}}$ as a subgroup of $\mathcal{G}^{\mathbb{Q}}$. Consequently, we shift our attention to understanding the general structure of $\mathcal{G}^{\mathbb{Q}}$ and then specializing to $\mathcal{G}^{\mathbb{Z}}$. This clever approach is due to Levine [7].

4.3 Witt Group of Isometric Structures $\mathcal{G}_{\mathbb{F}}$

The presence of non-symmetric matrices in $\mathcal{G}^{\mathbb{F}}$ hampers our ability to use the construction of the Witt group in Chapter 2. If we could somehow restrict our attention to symmetric Seifert matrices, these tools would come into play.

Let $\mathbb{F}$ be a field, and let $(V, \mathfrak{b}, T)$ an isometric structure over $\mathbb{F}$ (see Definition 2.2.1). It is worth explicitly stating that the relationship between an bilinear space $(V, \mathfrak{b})$ and an isometric structure $(V, \mathfrak{b}, T)$ is that the isometry T is fixed throughout.

Definition 4.3.1. An isometric structure $(V, \mathfrak{b}, T)$ over $\mathbb{F}$ is called *admissible* if the characteristic polynomial of T, namely $\Delta_T(t) = \det(T - tI)$, satisfies the condition $\Delta_T(1) \cdot \Delta_T(-1) \neq 0$.

Our notion of algebraic sliceness of isometric structures is from Definition 2.6.2. That is, an isometric structure $(V, \mathfrak{b}, T)$ of dimension $2n$ is called *algebraically slice* (or *metabolic*) if there is an n-dimensional T-invariant subspace of V on which $\mathfrak{b}$ vanishes.

Similarly, algebraic concordance of isometric structures is a restatement of Theorem 2.6.6: Two isometric structures $(V_1, \mathfrak{b}_1, T_1)$ and $(V_2, \mathfrak{b}_2, T_2)$ are called *algebraically concordant* if the isometric structure $(V_1, \mathfrak{b}_1, T_1) \oplus (V_2, -\mathfrak{b}_2, T_2)$ is algebraically slice.

Witt Cancellation (Corollary 2.5.1) for isometric structures is then: If the isometric structures $(V_1, \mathfrak{b}_1, T_1) \oplus (V_2, \mathfrak{b}_2, T_2)$ and $(V_2, \mathfrak{b}_2, T_2)$ are algebraically slice, then isometric structure $(V_1, \mathfrak{b}_1, T_1)$ is algebraically slice.

If we restrict to admissible isometric structures, we still get a Witt group.

Lemma 4.3.2. *Algebraic concordance of isometric structures is an equivalence relation among admissible isometric structures.*

Lemma 4.3.3. *The set $\mathcal{G}_{\mathbb{F}}$ of algebraic concordance classes of isometric structures is a group under the operation described as follows: Let $(V_1, \mathfrak{b}_1, T_1)$ and $(V_2, \mathfrak{b}_2, T_2)$ be two isometric structures over a field $\mathbb{F}$. Then $(V_1, \mathfrak{b}_1, T_1) \oplus (V_2, \mathfrak{b}_2, T_2) = (V_1 \oplus V_2, \mathfrak{b}_1 \perp \mathfrak{b}_2, T_1 \oplus T_2)$, where $V_1 \oplus V_2$ is direct sum of inner product spaces, $\mathfrak{b}_1 \perp \mathfrak{b}_2$ is orthogonal sum of inner products, and $T_1 \oplus T_2$ is the direct sum of isometries.*

Theorem 4.3.4. *The groups $\mathcal{G}_{\mathbb{F}}$ and $\mathcal{G}^{\mathbb{F}}$ are isomorphic via the homomorphism induced by sending the isometric structure $(\mathbb{F}^{2n}, A + A^T, A^{-1}A^T)$ to an $\mathbb{F}$-Seifert matrix A.*

Sketch of the Proof. It needs to be shown that the map $\mathcal{G}_{\mathbb{F}} \to \mathcal{G}^{\mathbb{F}}$ given by $A \mapsto (\mathbb{F}^{2n}, A{+}A^T, A^{-1}A^T)$ is well-defined homomorphism and has a well-defined inverse. For an isometric structure $(V, \mathfrak{b}, T)$ where $\mathfrak{b}$ has matrix representation B, the inverse map is given by $(V, \mathfrak{b}, T) \mapsto B(I + T)^{-1}$. □

4.4 Further Direction

In [8] (1969), J. Levine first presented the classification scheme we have followed here.

As of the result of the last section, we have the reduction:

$$\mathcal{C} \twoheadrightarrow \mathcal{G}^{\mathbb{Z}} \hookrightarrow \mathcal{G}^{\mathbb{Q}} \cong \mathcal{G}_{\mathbb{Q}}$$

To connect the concordance group $\mathcal{C}$ with Witt theory, we need one more map, (for details see [8])

$$\mathcal{G}_{\mathbb{Q}} \to W(\mathbb{Q})$$

This gives a connection between $\mathcal{C}$ and $W(\mathbb{Q})$. From here, one can study the Witt group $W(\mathbb{Q})$. For a treatment on this, see [12]. Once this group is understood, invariants of Witt classes can be related to concordance classes using the above classification. See [9].

Bibliography

[1] Gerhard Burde and Heiner Zieschang. *Knots*. New York, 2003.

[2] Jae Choon Cha and Charles Livingston. *KnotInfo: Table of Knot Invariants*. 2013. URL: http://www.indiana.edu/~knotinfo.

[3] Richard H. Crowell and Ralph H. Fox. *Introduction to Knot Theory*. New York: Springer-Verlag, 1977.

[4] Richard Elman, Nikita Karpenko, and Alexander Merkurjev. *The Algebraic and Geometric Theory of Quadratic Forms*. American Mathematical Society Colloquim Publications, Volume 56. Providence, Rhode Island: American Mathematical Society, 2008.

[5] Allen Hatcher. *Algebraic Topology*. Cambridge, England: Cambridge University Press, 2002.

[6] Akio Kawauchi. *Survey on Knot Theory*. Tokyo: Springer-Verlag, 1990.

[7] J. Levine. "Invariants of knot cobordism". In: *Invent. Math. 8 (1969), 98–110; addendum, ibid.* 8 (1969), pp. 98–110. ISSN: 0020-9910.

[8] J. Levine. "Knot cobordism groups in codimension two". In: *Comment. Math. Helv.* 44 (1969), pp. 229–244. ISSN: 0010-2571.

[9] Charles Livingston and Swatee Naik. *Introduction to Knot Concordance (Work in Progress)*. 2013. URL: http://wolfweb.unr.edu/homepage/naik/ConcordanceBook.pdf.

[10] J. Milnor. *Morse theory*. Based on lecture notes by M. Spivak and R. Wells. Annals of Mathematics Studies, No. 76. Princeton, N.J.: Princeton University Press, 1974, pp. vii+331.

[11] John Milnor. "On isometries of inner product spaces". In: *Invent. Math.* 8 (1969), pp. 83–97. ISSN: 0020-9910.

[12] John W. Milnor and Husemoller. *Symmetric bilinear forms*. Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 73. New York: Springer-Verlag, 1973, pp. viii+147.

[13] Dale Rolfsen. *Knots and Links.* Berkeley, California: Publish or Perish, 1976.

[14] W. Scharlau. *Quadratic and Hermitian forms.* Vol. 270. Grundlehren der Mathematischen Wissenschaften. New York, Inc.: Springer-Verlag, 1985.

[15] Kazimierz Szymiczek. *Bilinear Algebra: An Introduction to the Algebraic Theory of Quadratic Forms.* Algebra, Logic and Applications Series, Volume 7. Amsterdam, The Netherlands: Gordon and Breach Science Publishers, 1997.

[16] *The Knot Atlas.* 2013. URL: http://katlas.math.toronto.edu/wiki/Main_Page.

[17] Ernst Witt. "Theorie der quadratischen Formen in beliebigen Körpern". In: *J. Reine Angew. Math.* 176 (1936), pp. 31–44.