Advanced Studies in Pure Mathematics 24, 1996
Progress in Algebraic Combinatorics
pp. 1–80

Combinatorial Cell Complexes

Michael Aschbacher

We define and discuss a category of combinatorial objects we call *combinatorial cell complexes* and a functor T from this category to the category of topological spaces with cell structure, whose image is closely related to the category of CW-complexes. This formalism was developed to study finite group actions on topological spaces. In order to make effective use of our detailed knowledge of the finite simple groups, it seems necessary to make such a translation from a purely topological setting to the language of geometric combinatorics.

Our functor T assigns to each combinatorial cell complex X its geometric realization $T(X)$. We show the functor T defines an equivalence of categories between the category of combinatorial cell complexes whose cell boundaries are spheres, and a certain subcategory of CW-complexes we call normal CW-complexes.

We often concentrate on a subcategory of combinatorial cell complexes we call restricted combinatorial cell complexes; the restricted CW-complexes are the CW-complexes corresponding to the restricted combinatorial cell complexes under our equivalence of categories. Restricted CW-complexes include regular CW-complexes but also many other classical examples like the torus, the Klein bottle, and the Poincaré dodecahedron, which are discussed here as illustrations.

We associate to each restricted combinatorial cell complex X, a simplicial complex $K(X)$ and a canonical triangulation of $T(X)$ by $K(X)$. The geometric realization of a general combinatorial cell complex can also be canonically triangulated, but by a more complicated simplicial complex than $K(X)$. However we do not supply a proof of this last fact here.

We define cellular homology combinatorially, and show that if X is restricted and the boundary of each cell is homologically spherical, then

Received November 3, 1994.
Revised February 22, 1995.
This work was partially supported by NSF DMS–9101237 and BSF 92-00320

the homology of $T(X)$ is the cellular homology of X. We define a duality operator on restricted complexes which will be used in a later paper to establish a version of Poincaré duality for homology manifolds with cell structure which is more concrete than the usual version.

Now some specifics. Let $\mathcal{P}$ be the category whose objects are the posets P such that each $a \in P$ is of finite height, and whose morphisms are the maps preserving order and height. Let $\mathcal{P}^*$ consist of those members of $\mathcal{P}$ with a greatest element. A *combinatorial cell complex* consists of a poset $X \in \mathcal{P}$, a function $f : X \to \mathcal{P}^*$, a map $\zeta : V = \coprod_{x \in X} f(x) \to X$, and maps $f_v : f(x)(\leq v) \to f(\zeta(v))$ for each $v \in V$, such that

(i) For each $x \in X$, $\zeta : f(x) \to X(\leq x)$ is a map of posets preserving height.

(ii) For each $x \in X$ and $v \in f(x)$, $f_v : f(x)(\leq v) \to f(\zeta(v))$ is an isomorphism of posets.

(iii) If $u, v \in f(x)$ with $u \leq v$ then $f_u = f_{f_v(u)} \circ f_v$.

(iv) For $v \in f(x)$, $\zeta = \zeta \circ f_v$ on $f(x)(\leq v)$.

(v) For each $x \in X$, $f_{\infty_x} : f(x) \to f(x)$ is the identity map and $\zeta(\infty_x) = x$, where ∞_x is the greatest element of $f(x)$.

The posets $f(x)$, $x \in X$, are the *cells* of X and the *boundary* of the cell $f(x)$ is $\dot{f}(x) = f(x) - \{\infty_x\}$. The *faces* of the cell are its subposets $f(x)(\leq v)$, $v \in f(x)$. A combinatorial cell complex is *restricted* if ζ is injective on $f(x)(\geq v)$ for each $v \in f(x)$. Equivalently, for each $x \in X$ and $y \leq x$, the faces $f(x)(\leq v)$ with $v \in \zeta^{-1}(y) \cap f(x)$ are pairwise disjoint.

Intuitively a combinatorial cell complex consists of a collection $f(x)$, $x \in X$, of cells, with the poset structure on X corresponding to inclusions among the cells. The maps ζ and f_v, $v \in V$, keep track of identifications of cells with faces of larger cells, and glue cells together at their boundaries. Extra structure in some category $\mathcal{C}$ can be adjoined to each combinatorial cell to obtain $\mathcal{C}$-cells and a $\mathcal{C}$-cell complex.

The triangulating complex $K(X)$ of a combinatorial cell complex is defined in Section 5 and the geometric realization $T(X)$ of X is defined in Section 10. Our major results are Theorem 10.6, which provides the canonical triangulation of $T(X)$ by $K(X)$ when X is restricted, Theorem 12.16, which shows the cellular homology of X is isomorphic to the homology of $T(X)$ when X is restricted with homologically spherical cell boundaries, and Theorem 15.15, which establishes the equivalence of categories between combinatorial cell complexes whose cell boundaries are spheres and normal CW- complexes. The definitions of normal and restricted CW-complexes appear in Section 15. The reader may wish to refer to Sections 13 and 16 for various examples such as the torus, the Klein bottle, and the Poincaré dodecahedron.

§1. Posets and typed simplicial complexes

Let X be a poset. For $x \in X$ let $h(x)$ be the *height* of x in X. That is $h(x)$ is the maximum length of a chain in X with greatest element x, if the length of such chains is bounded, and ∞ otherwise. Write $X(\leq x)$ for the set of elements $y \in X$ such that $y \leq x$ and define $X(< x)$, $X(\geq x)$, etc. similarly.

Denote by $\mathcal{P}$ the category of posets X such that each $x \in X$ is of finite height. The morphisms in $\mathcal{P}$ are the maps of posets which preserve height. Let $\mathcal{P}^*$ be the subcategory of those $X \in \mathcal{P}$ such that X has a unique maximal member ∞_X.

We regard X as a category whose objects are the members of X and with $\mathrm{Mor}(x, y) = \{(x, y)\}$ if $x \leq y$ and $\mathrm{Mor}(x, y) = \varnothing$ otherwise.

Recall the *order complex* of a poset X is the simplicial complex $\mathcal{O}(X)$ whose vertices are the members of X and whose simplices are the finite chains. Often we write X for the order complex $\mathcal{O}(X)$ of X.

Example (1). If K is a simplicial complex its simplices form a poset under the inclusion relation and the *barycentric subdivision* $\mathrm{sd}(K)$ of K is the order complex of this poset. Thus the vertices of $\mathrm{sd}(K)$ are the finite simplices of K and the simplices of $\mathrm{sd}(K)$ are the chains of simplices of K.

A *typed simplicial complex* over an index set I is a simplicial complex $K = (V, \Sigma)$ together with a type function $h : V \to I$ such that h is injective on simplices. The morphisms of typed complexes over I are the simplicial maps which preserve type.

Example (2). The order complex of a poset is a typed complex where $h(x)$ is the height of x.

We will use the following notational conventions in discussing the homology of a typed simplicial complex. Let K be a typed simplicial complex with type function $h : V \to I$ and pick a total ordering of I. Given a k-simplex s in K, write

$$s = \prod_{v \in s} v = v_0 \wedge \cdots \wedge v_k \in \bigwedge^k (V)$$

for the generator of $C_k(K) \leq \bigwedge^k(V)$ corresponding to s, where $s = \{v_0, \ldots, v_k\}$ with $h(v_0) < \cdots < h(v_k)$. Then our boundary map becomes

$$\partial(s) = \sum_{i=0}^{k} (-1)^i s^i$$

where $s^i = v_0 \wedge \cdots \wedge v_{i-1} \wedge v_{i+1} \wedge \cdots \wedge v_k$.

If $s = s_1 \cup \cdots \cup s_r$ is a partition of s we write $s = \prod_{i=1}^r s_i$. More generally if J is some subset of I we can consider

$$\Sigma(J) = \{s \in \Sigma : h(s) = J\}.$$

Then if $J \subseteq L \subseteq I$ and $c = \sum_{s\in\Sigma(L)} a_s s \in C_*(K)$ with $a_s \in \mathbf{Z}$, then $s = s_J s_{L-J}$, where $s_U = \{v \in s : h(v) \in U\}$ for $U = J$ or $L - J$, and $c = \sum_{t\in\Sigma(J)} tc_t$, where $c_t = \sum_{t\subseteq s} a_s s_{L-J}$. Indeed

(1.1) *Let K be a typed simplicial complex with type function $h : V \to I$, $J \subseteq L \subseteq I$, and*

$$c = \sum_{s\in\Sigma(L)} a_s s = \sum_{t\in\Sigma(J)} tc_t \in C_*(K)$$

with $a_s \in \mathbf{Z}$. Then

(1) $\partial(c) = \sum_t (-1)^N \partial(t) c_t + (-1)^M t\partial(c_t)$

(2) *If $\partial(c) = 0$ then $\partial(c_t) = 0$ for each $t \in \Sigma(J)$ and $\sum_{t\in\Sigma(J)} \partial(t)c_t$ =0.*

(3) *If $L - J = \{l\}$ is of order 1 then*

$$c_t = \sum_{u\in V_l(t)} a_{t,u} u$$

where $V_l(t) = \{v \in \mathrm{Link}_K(t) : h(v) = l\}$, $a_{t,u} = a_{t\cup\{u\}}$, and $\partial(c_t) = \sum_{u\in V_l(t)} a_{t,u}$.

(4) *If $|L| = k+1$ and $J = \{j\}$ with j the maximal member of L then $\partial(c) = \sum_t t\partial(c_t) + (-1)^k c_t$, so $\partial(c) = 0$ if and only if $\partial(c_t) = 0$ for each $t \in \Sigma(J)$ and $\sum_t c_t = 0$.*

Proof. Take $L = \{0, \ldots, k\}$. We first prove (1). Since

$$v_0 \wedge \cdots \wedge v_k = \mathrm{sgn}(\pi)(v_{\pi(0)} \wedge \cdots \wedge v_{\pi(k)})$$

for $\pi \in \mathrm{Sym}(L)$, changing our ordering of I by a suitable permutation π, we may assume $J = \{0, \ldots, j\}$ and the ordering of J and $L - J$ are induced from L. Subject to this choice of ordering, we prove

$$(*) \qquad \partial(c) = \sum_t (\partial(t)c_t + (-1)^{j+1} t\partial(c_t)).$$

As ∂ is linear, it suffices to take $c = s$. Let $t = s_J$; then $c_t = s_{L-J}$. Now for $i \leq j$, $s^i = t^i c_t$, while for $i > j$, $s^i = tc_t^{i-j-1}$. Therefore

$$\begin{aligned}\partial(c) = \partial(s) = \sum_{i=0}^{k}(-1)^i s^i &= \sum_{i=0}^{j}(-1)^i t^i c_t + \sum_{i=0}^{k-j-1}(-1)^{i+j+1} t c_t^i \\ &= \partial(t)c_t + (-1)^{j+1} t\partial(c_t).\end{aligned}$$

So (1) is established. Further under the hypotheses of (4), the sign of a permutation π mapping k to the first member of L and preserving the order on $L - J$, is $(-1)^k$ and of course $\partial(t) = 1$, so by (*), $\partial(c) = \sum_t (-1)^k c_t + t\partial(c_t)$, establishing the first statement in (4).

Suppose $\partial(c) = 0$. Then by (1), $\partial(c) = A + (-1)^M \sum_t t\partial(c_t)$ with $A = \sum_{r \in S_A} b_r r$ and $(-1)^M t\partial(c_t) = \sum_{r \in S_t} b_r r$, where S_A and $S_t, t \in \Sigma(J)$ are suitable subsets of $\Sigma^{k-1}(K)$ and S_t is the set of simplices $s \in S_A \cup \bigcup_t S_t$ with $t \subseteq s$. In particular the sets $S_A, S_t,\ t \in \Sigma(J)$, are pairwise disjoint. As $\partial(c) = 0$, $b_r = 0$ for all r, so as our index sets are pairwise disjoint, $(-1)^M t\partial(c_t) = \sum_{r \in S_t} b_r r = 0$ and hence $\partial(c_t) = 0$. As this holds for all t, also $0 = \sum_t \partial(t)c_t$.

Thus (2) is established. Finally (2) completes the proof of (4), since under the hypotheses of (4), $\partial(t)c_t = c_t$. The proof of (3) is straightforward.

§2. Cells

Let $\mathcal{C}$ be a category. Define $Cell(\mathcal{C})$ to be the category of covariant functors $F : X \to \mathcal{C}$, where $X \in \mathcal{P}^*$ is regarded as a category as in Section 1. In addition we almost always impose extra conditions on the cells.

Regard $(X, F) \in Cell(\mathcal{C})$ as a category whose objects are the pairs $(x, F(x))$, $x \in X$, with

$$\mathrm{Mor}((x, F(x)), (y, F(y))) = \{(x, y), F(x, y)\} \text{ if } x \leq y$$

where $F(x, y) : F(x) \to F(y)$ is the $\mathcal{C}$-morphism associated to (x, y) by F and

$$\mathrm{Mor}((x, F(x)), (y, F(y))) = \varnothing \text{ otherwise.}$$

Notice that as F is a functor, $F(x,z) = F(y,z) \circ F(x,y)$ whenever $x \leq y \leq z$.

The morphisms in $Cell(\mathcal{C})$ are the covariant functors $\phi : (X,F) \to (Y,G)$ together with isomorphisms $\phi_x : F(x) \to G(\phi(x))$ such that for all $x \leq z$ in X, the following diagram commutes:

$$\begin{array}{ccc} F(z) & \xrightarrow{\phi_z} & G(\phi(z)) \\ {\scriptstyle F(x,z)}\uparrow & & \uparrow{\scriptstyle G(\phi(x),\phi(z))} \\ F(x) & \xrightarrow{\phi_x} & G(\phi(x)) \end{array}$$

Here for $x \in X$, $(x, F(x))$ is an object of (X,F), so $\phi(x, F(x))$ is an object of (Y,G) which must then be of the form $(\phi(x), G(\phi(x)))$ for some $\phi(x) \in Y$. Further as ϕ is a functor, if $x_1 \leq x_2$ then $\mathrm{Mor}((x_1, F(x_1)), (x_2, F(x_2))) \neq \varnothing$, so $\mathrm{Mor}((\phi(x_1), G(\phi(x_1))), (\phi(x_2), G(\phi(x_2)))) \neq \varnothing$ and hence $\phi(x_1) \leq \phi(x_2)$, so $\phi : X \to Y$ is a map of posets.

The category $Cell(\mathcal{C})$ is the category of *$\mathcal{C}$-cells.* Intuitively a $\mathcal{C}$-cell is an object $F(\infty_X)$ in $\mathcal{C}$ together with a distinguished family $(F(x) : x \in X)$ of subobjects, called the *faces* of $F(\infty_X)$, indexed by the poset X, with the inclusion relation on these subobjects corresponding to the partial order on X. In the case of combinatorial and topological cells this intuition is made precise in the following two examples:

Example. (1) Let $\mathcal{C} = \mathcal{P}^*$. Then $Cell(\mathcal{P}^*)$ is the category of *combinatorial cells.* We also require that a combinatorial cell (X,F) satisfy $F(x) = X(\leq x)$ and $F(x,y) : F(x) \to F(y)$ be inclusion for all $x, y \in X$ with $x \leq y$. So a combinatorial cell is nothing more than a poset X in $\mathcal{P}^*$ together with its faces $X(\leq x)$, $x \in X$.

(2) Let $\mathcal{C} = Top$ be the category of topological spaces with morphisms closed injections. In this case $Cell(Top)$ is the category of *topological cells.* We also demand that a topological cell (X,F) satisfy the requirements that if $x, y \in X$ then

(i) $F(x,\infty)(F(x)) \cap F(y,\infty)(F(y)) = \bigcup_{z \leq x,y} F(z,\infty)(F(z))$.

(ii) If $x \neq y$ then $F(x,\infty)(F(x)) \neq F(y,\infty)(F(y))$.

(iii) For each $Y \subseteq X$, $\bigcup_{y \in Y} F(y,\infty)(F(y))$ is closed in T.

Since for $x \in X$, $F(x,\infty) : F(x) \to F(\infty)$ is a closed injection, we can identify $F(x)$ with its image in $F(\infty)$. Condition (i) says that if $x \leq y$ then $F(x) \subseteq F(y)$, and then condition (ii) says the map $x \mapsto F(x)$ is an isomorphism of the poset X with the poset $\{F(x) : x \in X\}$ of distinguished closed subspaces of $F(\infty)$.

§3. Cell complexes

Again let $\mathcal{C}$ be a category. Let $Complex(\mathcal{C})$ be the category whose objects consist of

(1) Some $X \in \mathcal{P}$.

(2) A function $f : X \to Cell(\mathcal{C})$.

(3) A map $\zeta : V = \coprod_{x\in X} f(x) \to X$ such that for each $x \in X$, $\zeta : f(x) \to X(\leq x)$ is a morphism in $\mathcal{P}$.

(4) For each $x \in X$ and $v \in f(x)$, an isomorphism $f_v : f(x)(\leq v) \to f(\zeta(v))$ of $\mathcal{C}$-cells satisfying:

(5) If $u, v \in f(x)$ with $u \leq v$ then $f_u = f_{f_v(u)} \circ f_v$.

(6) For $v \in f(x)$, $\zeta = \zeta \circ f_v$ on $f(x)(\leq v)$.

(7) For each $x \in X$, $f_{\infty_x} : f(x) \to f(x)$ is the identity map and $\zeta(\infty_x) = x$, where ∞_x is the greatest element of $f(x)$.

Formally $f(x)$ is a pair (X_x, F_x) where X_x is a poset and $F_x : X_x \to \mathcal{C}$ is a functor, but we usually write $f(x)$ for the poset X_x and F for F_x. In particular this is the convention in axioms (3) and (7). However the isomorphism f_v of axiom (4) is an isomorphism of cells, so it consists of a covariant functor $f_v : f(x)(\leq v) \to f(\zeta(v))$ and isomorphisms $f_{v,u} : F_x(u) \to F_{\zeta(v)}(f_v(u))$ for each pair $u, v \in f(x)$ with $u \leq v$, and these isomorphisms satisfy $F_{\zeta(v)}(f_v(u), f_v(w)) \circ f_{v,w} = f_{v,w} \circ F_x(w, u)$ for each $w \leq u \leq v$.

The morphisms $\psi : (X, f) \to (Y, g)$ are morphisms $\psi : X \to Y$ in $\mathcal{P}$ together with morphisms $\psi_x : f(x) \to g(\psi(x))$ in $Cell(\mathcal{C})$ for each $x \in X$ such that

(a) For each $x \in X$ and $v \in f(x)$, the following diagram commutes:

$$\begin{array}{ccc} f(x)(\leq v) & \xrightarrow{\psi_x} & g(\psi(x))(\leq \psi_x(v)) \\ f_v\downarrow & & \downarrow g_{\psi_x(v)} \\ f(\zeta(v)) & \xrightarrow{\psi_{\zeta(v)}} & g(\psi(\zeta(v))) \end{array}$$

(b) For each $x \in X$ the following diagram commutes:

$$\begin{array}{ccc} f(x) & \xrightarrow{\zeta} & X(\leq x) \\ \psi_x\downarrow & & \downarrow\psi \\ g(\psi(x)) & \xrightarrow{\zeta} & Y(\leq \psi(x)) \end{array}$$

Further we define composition in our category so that if $\phi : (Y, g) \to (Z, h)$ is a morphism then $(\phi \circ \psi)_x = \phi_{\psi(x)} \circ \psi_x$ for each $x \in X$.

The members $X = (X, f)$ of *Complex*$(\mathcal{C})$ are called *$\mathcal{C}$-cell complexes.* X is a *combinatorial cell complex* if the cells are combinatorial cells. X is a *topological cell complex* if the cells of X are topological.

Intuitively a cell complex consists of cells indexed by the poset X together with identifications of the faces of cells accomplished by the maps ζ and f_v, $v \in V$.

Example (1). Let $X \in \mathcal{P}$ and for $x \in X$ and $v \in X(\leq x)$ let $f(x) = X(\leq x)$ and let ζ and f_v be the appropriate identity maps. Then X is a combinatorial cell complex. We call this cell complex the *simplicial cell complex* of the poset X.

Remarks. (1) Given any $\mathcal{C}$-cell complex (X, f, F) we can suppress the $\mathcal{C}$-structure supplied by the functor F and obtain the *combinatorial cell complex* (X, f) of (X, f, F). This gives us a forgetful functor from $\mathcal{C}$-cell complexes to combinatorial cell complexes.

(2) The combinatorial cell complexes are the simplest cell complexes, and we can give a somewhat simpler definition of this category equivalent to the specialization of the general definition above to the case of combinatorial cells: A *combinatorial cell complex* consists of a poset $X \in \mathcal{P}$, a function $f : X \to \mathcal{P}^*$, a map $\zeta : V = \coprod_{x \in X} f(x) \to X$, and maps $f_v : f(x)(\leq v) \to f(\zeta(v))$ for each $v \in V$, such that

(i) For each $x \in X$, $\zeta : f(x) \to X(\leq x)$ is a map of posets preserving height.

(ii) For each $x \in X$ and $v \in f(x)$, $f_v : f(x)(\leq v) \to f(\zeta(v))$ is an isomorphism of posets.

(iii) If $u, v \in f(x)$ with $u \leq v$ then $f_u = f_{f_v(u)} \circ f_v$.

(iv) For $v \in f(x)$, $\zeta = \zeta \circ f_v$ on $f(x)(\leq v)$.

(v) For each $x \in X$, $f_{\infty_x} : f(x) \to f(x)$ is the identity map and $\zeta(\infty_x) = x$.

Moreover a *morphism* $\psi : (X, f) \to (Y, g)$ of combinatorial cell complexes consists of a height preserving map $\psi : X \to Y$ of posets together with height preserving maps $\psi_x : f(x) \to g(\psi(x))$ of posets for each $x \in X$ such that ψ and ψ_x satisfy the commutative diagrams (a) and (b) for morphisms of cell complexes given earlier in this section.

Define a cell complex (X, f) to be *regular* if $\zeta : f(x) \to X(\leq x)$ is an isomorphism for each $x \in X$. For example the simplicial cell complex of a poset (defined in Example (1)) is a regular cell complex. We will see in a moment that, up to isomorphism, all regular combinatorial cell complexes are simplicial cell complexes. Define (X, f) to be *restricted* if ζ is injective on $f(x)(\geq v)$ for each $v \in f(x)$. For example regular cell complexes are restricted, but the converse is certainly not true. The combinatorial cell complexes of the torus and Klein bottle, discussed in Section 13, are examples of restricted cell complexes which are not regular, as is the complex of the Poincaré dodecahedron, discussed in Section 16.

The *cells* of the cell complex (X, f, F) are the $\mathcal{C}$-cells $f(x) = (X_x, F_x)$. The *boundary* of the combinatorial cell $f(x)$ is $\dot{f}(x) = f(x) - \{x\}$ and if $f(x)$ is a $\mathcal{C}$-cell with extra structure supplied by $F(x)$ then $\dot{f}(x)$ is the $\mathcal{C}$-cell complex with extra structure $\dot{F}(x) = F(x)_{|\dot{f}(x)}$. We say (X, f, F) is of *height* n if X is of height n.

Example (2). Let $0 \leq n \in \mathbf{Z}$ and let $X(n)$ be the poset $\{0, 1, \ldots, n\}$ under the usual order. For $k \in X(n)$ define $f(k) = \{(k, i) : 0 \leq i \leq k\}$ and order $f(k)$ so that the map $\zeta : f(k) \to X(n)$ defined by $\zeta(k, i) = i$ preserves order. Define $f_{(k,i)} : f(k)(\leq (k, i)) \to X(n)(\leq i)$ by $f_{(k,i)}(k, j) = (i, j)$. Then $(X(n), f)$ is a combinatorial cell complex isomorphic to the simplicial cell complex of the poset $X(n)$. As $X(n)$ is an n-simplex, we call $X(n)$ the *simplicial cell complex of the n-simplex.*

(3.1) *Let $X = (X, f)$ be a regular combinatorial cell complex. Then (X, f) is isomorphic to the simplicial cell complex of the poset X.*

Proof. Let $\bar{X} = (\bar{X}, \bar{f})$ be the simplicial cell complex of X. Thus $\bar{X} = X$, $\bar{f}(x) = X(\leq x)$ for each $x \in X$, and $\bar{\zeta}$ and $\bar{f}_v$ are the appropriate identity maps. Define

$$\psi : (X, f) \to (\bar{X}, \bar{f})$$

and

$$\bar{\psi} : (\bar{X}, \bar{f}) \to (X, f)$$

to be the morphisms with $\psi : X \to \bar{X}$ and $\bar{\psi} : \bar{X} \to X$ the identity maps, $\psi_x : f(x) \to \bar{f}(x)$ the restriction of ζ to $f(x)$, and $\bar{\psi}_x = \psi_x^{-1}$. It is essentially immediate from the definition of ψ and $\bar{\psi}$ and from axiom 6 for cell complexes that each of these maps *is* a morphism of cell complexes. Of course $\bar{\psi} = \psi^{-1}$, so ψ is an isomorphism.

(3.2) *Let (X, f) be a combinatorial cell complex and $x \in X$. Then the poset $f(x)$ is of height $h(x)$ and ∞_x is the unique member of $f(x)$ of height $h(x)$.*

Proof. As $f(x) \in \mathcal{P}^*$, ∞_x is the unique element of $f(x)$ of maximal height. Then as $\zeta : f(x) \to X(\leq x)$ preserves height, $h(f(x)) = h(\infty_x) = h(\zeta(\infty_x)) = h(x)$ by axiom 7 for cell complexes.

§4. Topological cell complexes

Let X, f, F be a topological cell complex. That is $X \in \mathcal{P}$, for $x \in X$, $f(x)$ is a topological cell with $F(v)$ the topological space associated to $v \in f(x) \in \mathcal{P}^*$, etc.

Write x for ∞_x. As in Example (2) in Section 2, we regard each cell $f(x)$ as a topological space $F(x)$ together with a distinguished class $\{F(v) : v \in f(x)\}$ of closed subspaces. Namely, for $v \in f(x)$, we have a closed injection $F(v,x) : F(v) \to F(x)$, and we identify $F(v)$ with its image under this injection and regard it as a closed subspace of $F(x)$. Because F is a functor, these identifications are compatible with the ordering on $f(x)$; that is if $u < v < x$ then $F(u) \subseteq F(v) \subseteq F(x)$ and the identification of $F(u)$ with a subspace of $F(x)$ factors through the identification of $F(u)$ with a subspace of $F(v)$. Subject to these conventions, $F(x)$ is a topological space with a poset of distinguished closed subspaces, and that poset is isomorphic to $f(x)$.

Recall from Section 2 that topological cells are required to satisfy the property that if $u, v \in f(x)$ then $F(u,x)(F(u)) \cap F(v,x)(F(v)) = \bigcup_{w \leq u,v} F(w,x)(F(w))$, which under our new notational conventions translates into the statement that $F(u) \cap F(v) = \bigcup_{w \leq u,v} F(w)$. Also if $u \neq v$ then $F(u,x)(F(u)) \neq F(v,x)(F(v))$, which in our new language reads if $u \neq v$ then $F(u) \neq F(v)$. In particular it follows that

(4.1) *For each $x \in X$ and $a \in F(x)$ there exists a unique $v \in f(x)$ of minimal height such that $a \in F(v)$.*

Next let $v \in f(x)$. Then we have an isomorphism $f_v : f(x)(\leq v) \to f(\zeta(v))$ of topological cells. The identifications above identify $f(x)(\leq v)$ with $F(v)$ and the subspaces determined by the poset $f(x)(\leq v)$, and identify $f(\zeta(v))$ with the space $F(\zeta(v))$ and its family of subspaces. As f_v is an isomorphism of topological cells, it induces an isomorphism $F_v : F(v) \to F(\zeta(v))$ such that if $u \leq v$ then $F_v(F(u)) = F(f_v(u))$. Further as F is a functor, $F_u = F_{f_v(u)} \circ F_v$.

Let $F_n = \coprod_{\dim(x)=n} F(x)$ be the disjoint union of the spaces associated to the n-cells of X, and $\hat{A}_n = \coprod_{m \leq n} F_n$. Thus for each $a \in \hat{A}_n$ there is a unique $x(a) \in X$ with $a \in F(x(a))$ and by 4.1 there is a unique $v(a) \in f(x)$ of minimal height such that $a \in F(v(a))$. Let $y(a) = \zeta(v(a))$ and observe that $F_{v(a)}(a) \in F(y(a))$ with $x(F_{v(a)}(a)) = y(a)$.

For $x \in X$ of dimension n, $a \in F(x)$, and $b \in \hat{A}_{n-1}$, define $a \searrow b$ if $F_v(a) = b$ for some $v \in f(x)$ with $a \in F(v)$ and $b \in F(\zeta(v))$.

We construct a topological space A_n by factoring out a suitable equivalence relation $\sim_n$ from $\hat{A}_n$. The definition is recursive. Namely the equivalence relation $\sim_n$ on $\hat{A}_n$ is defined to be the equivalence relation on $\hat{A}_n$ generated by $\searrow$, regarded as a relation on $F_n \cup \hat{A}_{n-1}$, and $\sim_{n-1}$. We first observe that:

(4.2) *For $a \in \hat{A}_{n-1}$, $[a]_n \cap \hat{A}_{n-1} = [a]_{n-1}$, where $[a]_k$ is the equivalence class of a with respect to $\sim_k$.*

Proof. It suffices to show that if $a \in F_n$ and $b, c \in \hat{A}_n$ with $a \searrow b$ and $a \searrow c$ then $[b]_{n-1} = [c]_{n-1}$. Let $x = x(a)$, $r = x(b)$, and $s = x(c)$. Then $a \in F(u) \cap F(w)$ where $u \in \zeta^{-1}(r)$ and $w \in \zeta^{-1}(s)$. By an earlier remark, $F(u) \cap F(w) = \bigcup_{v \leq u,w} F(v)$, so $a \in F(v)$ for some $v \leq u, w$. Now $F_{f_u(v)}(b) = (F_{f_u}(v) \circ F_u)(a) = F_v(a) = F_{f_w(v)}(c)$, so by induction on n, $[b]_{n-1} = [c]_{n-1}$.

Let $\hat{A} = \bigcup_n \hat{A}_n$ and $\sim = \bigcup_n \sim_n$. By 4.2, $\sim$ is an equivalence relation on $\hat{A}$. Write $\tilde{a}$ for the equivalence class of $a \in \hat{A}$ and let $A = A(X) = \hat{A}/\sim$. We conclude from 4.2 that

(4.3) *For each $a \in \hat{A}$ and nonnegative integer n, $\tilde{a} \cap \hat{A}_n = [a]_n$.*

For $x \in X$, $u \in f(x)$, define $I(u) = F(u) - \bigcup_{u > v \in f(x)} F(v)$. Further define

$$\begin{aligned} \lambda_x : F(x) &\to A \\ a &\mapsto \tilde{a} \end{aligned}$$

and let $\tilde{F}(x)$, $\tilde{I}(x)$ be the image of $F(x)$, $I(x)$ in A under the map λ_x.

(4.4) (1) *The map $\lambda_x : I(x) \to A$ is an injection.*

(2) *For $\tilde{a} \in A$ there exists a unique $y(\tilde{a}) \in X$ such that $\tilde{a} \cap I(y(\tilde{a})) \neq \varnothing$.*

(3) *There exists a unique element $\xi(\tilde{a}) \in F(y(\tilde{a}))$ with $\xi(\tilde{a}) \in \tilde{a}$.*

(4) *If $x \in X$ and $b \in \tilde{a} \cap F(x)$, then there exists $v \in f(x) \cap \zeta^{-1}(y(\tilde{a}))$ with $b \in F(v)$ and $F_v(b) = \xi(\tilde{a})$. Further $\tilde{F}(y(\tilde{a})) = \widetilde{F(v)} \subseteq \tilde{F}(x)$.*

Proof. Let $n = h(x)$ and $a, b \in I(x)$. If $\tilde{a} = \tilde{b}$ then by 4.3, $[a]_n = [b]_n$. But by definition of $\sim_n$, $[a]_n = \{a\}$ for $a \in I(x)$ as $h(x) = n$. This establishes (1) and (3).

Let $x \in X$ and suppose x is minimal subject to $\tilde{a} \cap F(x) \neq \varnothing$. Let $b \in F(x) \cap \tilde{a}$; claim $b \in I(x)$. For if not then $b \in F(v)$ for some $x \neq v \in f(x)$ and then $b \sim F_v(b) \in F(y)$ with $y = \zeta(v)$. Now $\tilde{F}(y) = \widetilde{F(v)} \subseteq \tilde{F}(x)$, and the minimality of x is contradicted. In particular there exists some $y \in X$ with $\tilde{a} \cap I(y) \neq \varnothing$.

On the other hand suppose $a_i \in \tilde{a}$ with $a_i \in I(y_i)$, for $i = 1, 2$, and let $n = \max\{h(y_1), h(y_2)\}$. Then, as we saw in paragraph one, $|\tilde{a} \cap \hat{A}_n| = 1$, so $a_1 = a_2$ and hence as a_1 is in $F(y)$ for a unique $y \in X$, $y_1 = y_2$. This establishes (2).

We saw in paragraph two that if $x \in X$ and $b \in \tilde{a} \cap F(x)$, then either $b \in I(x)$ or there is $y < x$ and $v \in \zeta^{-1}(y)$ with $b \in F(v)$ and $\tilde{F}(y) = \tilde{F}(v) \subseteq \tilde{F}(x)$. In the former $x = y(a)$ by (2), so that (4) holds, and in the latter (4) holds by induction on $h(x)$.

(4.5) (1) *The sets* $\tilde{I}(x)$, $x \in X$, *partition* A.
(2) $\tilde{F}(x) \cap \tilde{F}(y) = \bigcup_{z \leq x,y} \tilde{F}(z)$.

Proof. Part (1) follows from 4.4.2. Let $\tilde{a} \in \tilde{F}(x) \cap \tilde{F}(y)$. Then by 4.4.4, $\tilde{a} \in \tilde{F}(y(\tilde{a})) \subseteq \tilde{F}(x) \cap \tilde{F}(y)$ and $y(a) \leq x, y$. Thus (2) holds.

We topologize $\tilde{F}(x)$ by defining a subset C of $\tilde{F}(x)$ to be closed if and only if $\lambda_x^{-1}(C)$ is closed in $F(x)$. Then we topologize A by decreeing that $C \subseteq A$ is closed in A if and only if $C \cap \tilde{F}(x)$ is closed in $\tilde{F}(x)$ for each $x \in X$.

(4.6) (1) $\tilde{F}(x)$ *is closed in* A *so a subset* C *of* $\tilde{F}(x)$ *is closed in* $\tilde{F}(x)$ *if and only if* C *is closed in* A.
(2) $\lambda_x : F(x) \to A$ *is continuous.*
(3) *If* $\zeta : f(x) \to X(\leq x)$ *is injective then* $\lambda_x : F(x) \to \tilde{F}(x)$ *is a homeomorphism.*
(4) *For each* $Y \subseteq X$, $\bigcup_{y \in Y} \tilde{F}(y)$ *is closed in* A.

Proof. Let C be closed in A. Then by definition of the topology on A, C is closed in $\tilde{F}(x)$. Conversely if $\tilde{F}(x)$ is closed in A and C is closed in $\tilde{F}(x)$ then C is closed in A, so to prove (1) it remains to show $\tilde{F}(x)$ is closed in A. We must show $\tilde{F}(x) \cap \tilde{F}(y)$ is closed in $\tilde{F}(y)$ for all $y \in X$. By 4.5.2,

$$\tilde{F}(x) \cap \tilde{F}(y) = \bigcup_{z \leq x,y} \tilde{F}(z)$$

Further

$$\lambda_y^{-1}(\tilde{F}(z)) = \bigcup_{v \in \zeta^{-1}(z) \cap f(y)} F(v)$$

so

$$\lambda_y^{-1}(\tilde{F}(x) \cap \tilde{F}(y)) = \bigcup_{\substack{z \leq x,y \\ v \in \zeta^{-1}(z) \cap f(y)}} F(v)$$

is closed in $F(y)$ by axiom (iii) for topological cells in Example (2) of Section 2. Now by definition of the topology on A, $\tilde{F}(x) \cap \tilde{F}(y)$ is closed in $\tilde{F}(y)$, completing the proof of (1). A similar argument establishes (4).

By definition of the topology on $\tilde{F}(x)$, $\lambda_x : F(x) \to \tilde{F}(x)$ is continuous, so (2) follows from (1). Assume $\zeta : f(x) \to X(\leq x)$ is injective. Claim $\lambda_x : F(x) \to \tilde{F}(x)$ is bijective. For if $a, b \in F(x)$ with $\tilde{a} = \tilde{b}$ then by 4.4.4, there is $u, v \in f(x) \cap \zeta^{-1}(y(\tilde{a}))$ with $a \in F(u)$, $b \in F(v)$, and $F_u(a) = F_v(b) = \xi(\tilde{a})$. As ζ is injective, $u = v$. Then as F_u is injective, $a = b$. So λ_x is bijective. Now for $D \subseteq F(x)$ closed, $D = \lambda_x^{-1}(\lambda_x(D))$ is closed, so $\lambda_x(D)$ is closed in $\tilde{F}(x)$ by definition of the topology on $\tilde{F}(x)$. This proves (3).

(4.7) *Let $\varphi : (X, f, F) \to (Y, g, G)$ be a morphism of topological cell complexes. Then*

(1) *φ induces a continuous map $A(\varphi) : A(X) \to A(Y)$ via $A(\varphi)(\tilde{a}) = \widetilde{\varphi(a)}$.*

(2) *A is a covariant functor from the category of topological cell complexes to the category of topological spaces.*

Proof. Let $\tilde{\varphi} = A(\varphi)$. Observe first that $\tilde{\varphi}$ is well defined, since if $a_i \in \tilde{a}$, $i = 1, 2$, then $a_i \in F(v_i) \subseteq F(x_i)$ with $v_i \in \zeta^{-1}(y)$, where $y = y(\tilde{a})$ and $F_{v_i}(a_i) = \xi(\tilde{a}) \in F(y)$. Then $G_{\varphi(v_i)}(\varphi(a_i)) = \varphi(\xi(\tilde{a}))$, so $\varphi(a_1) \sim \varphi(\xi(\tilde{a})) \sim \varphi(a_2)$.

Next claim $\tilde{\varphi} : \tilde{F}(x) \to \tilde{G}(\varphi(x))$ is continuous for each $x \in X$. For if C is a closed subset of $\tilde{G}(\varphi(x))$ then $\lambda_{\varphi(x)}^{-1}(C)$ is closed in $G(\varphi(x))$ and then as $\varphi : F(x) \to G(\varphi(x))$ is continuous, $\varphi^{-1}(\lambda_{\varphi(x)}^{-1}(C))$ is closed in $F(x)$. So as $\lambda_{\varphi(x)} \circ \varphi = \tilde{\varphi} \circ \lambda_x$, $\lambda_x^{-1}(\tilde{\varphi}^{-1}(C))$ is closed in $F(x)$. Therefore $\tilde{\varphi}^{-1}(C)$ is closed in $\tilde{F}(x)$, so indeed $\tilde{\varphi} : \tilde{F}(x) \to \tilde{G}(\varphi(x))$ is continuous. Therefore by 4.6 and the definition of the topology on $A(X)$ and $A(Y)$, $\tilde{\varphi} : A(X) \to A(Y)$ is continuous. Hence part (1) of the lemma is established. Part (2) is straightforward.

§5. The triangulating complex of a combinatorial cell complex

Let (X, f) be a combinatorial cell complex. Let $V = V(X) = \coprod_{x \in X} f(x)$ be the disjoint union of the posets $f(x)$, $x \in X$. So for each $v \in V$ there exists a unique $\hat{\zeta}(v) \in X$ with $v \in f(\hat{\zeta}(v))$.

For $v \in V$ define

$$L(v) = \{u \in V : f_w(v) \geq u \text{ for some } w \in f(\hat{\zeta}(v))(\geq v)\}.$$

(5.1) *Let $x \in X$, $v \in f(x)$, $u \in L(v)$, and $w \in f(x)(\geq v)$ with $f_w(v) \geq u$. Then*

(1) $\hat{\zeta}(u) = \zeta(w)$.

(2) *If (X, f) is restricted then w is the unique $z \in f(x)$ $(\geq v)$ with $\zeta(z) = \hat{\zeta}(u)$ and we denote w by $\hat{f}_u(v)$.*

Proof. As $f_w(v) \geq u$, $u \in f(\zeta(w))$, so $\hat{\zeta}(u) = \zeta(w)$. If X is restricted then for each $y \in X$ there is at most one $z \in f(x)$ $(\geq v)$ with $\zeta(z) = y$, so (2) holds.

Define the *graph* $\Delta = \Delta(X)$ of the cell complex X to be the graph with vertex set V and u adjacent to v if $u \in L(v)$ or $v \in L(u)$. The *clique complex* of a graph Γ is the simplicial complex with vertex set Γ and simplices the cliques of Γ. Denote by $K(X)$ the clique complex of $\Delta(X)$. We call $K(X)$ the *triangulating complex* of X.

(5.2) *If $u \in L(v)$ then $\zeta(u) \leq \zeta(v) \leq \hat{\zeta}(u) \leq \hat{\zeta}(v)$.*

Proof. As $u \in L(v)$, there is $w \in f(x)(\geq v)$ with $f_w(v) \geq u$. Then $\hat{\zeta}(u) = \zeta(w) \leq \hat{\zeta}(v)$, $\zeta(v) = \zeta(f_w(v)) \geq \zeta(u)$, and $\zeta(v) \leq \zeta(w) = \hat{\zeta}(u)$.

(5.3) *If $u, v \in V$ are adjacent with $\hat{\zeta}(u) = \hat{\zeta}(v)$ and $\zeta(u) = \zeta(v)$ then $u = v$.*

Proof. We may take $u \in L(v)$. Let $x = \hat{\zeta}(v)$ and $w \in f(x)(\geq v)$ with $f_w(v) \geq u$. Then $\zeta(w) = \hat{\zeta}(u) = \hat{\zeta}(v) = x$, so by 3.2, $w = \infty_x = x$. Then $v = f_x(v) \geq u$, so as $\zeta(v) = \zeta(u)$ and ζ preserves height, $v = u$.

For $s \subseteq \Delta$ define $X(s) = \{\hat{\zeta}(v) : v \in s\}$ and $\zeta(s) = \{\zeta(v) : v \in s\}$.

(5.4) *Let s be a simplex in $K(X)$. Then*

(1) *There is a unique ordering $v_0, \ldots, v_k$ of the vertices of s such that $v_i \in L(v_j)$ for $0 \leq i \leq j \leq k$.*

(2) $\zeta(v_0) \leq \cdots \leq \zeta(v_k) \leq \hat{\zeta}(v_0) \leq \cdots \leq \hat{\zeta}(v_k)$.

(3) *Assume (X, f) is restricted and let $w_i = \hat{f}_{v_i}(v_k)$. Then w_i is the unique $w \in f(\hat{\zeta}(v_k))(\geq v_k)$ with $\zeta(w) = \hat{\zeta}(v_i)$. Moreover $w_0 \leq \cdots \leq w_k = \hat{\zeta}(v_k)$ and $f_{w_j}(w_i) \geq v_j$ for $j \geq i$.*

Proof. Induct on the dimension k of s. The case $k = 0$ is trivial, so take $k > 0$. By 5.2, $X(s)$ and $\zeta(s)$ are chains, so pick $v_k \in s$ with $\hat{\zeta}(v_k)$ maximal, and subject to this constraint, with $\zeta(v_k)$ maximal. Let $t = s - \{v_k\}$. By induction on k, there is a unique ordering $v_0, \ldots, v_{k-1}$ of t satisfying the conditions of the lemma. Let $i < k$. By 5.3 and the choice of v_k, either $\hat{\zeta}(v_i) < \hat{\zeta}(v_k)$ or $\zeta(v_i) < \zeta(v_k)$. Then by 5.2, $v_i \in L(v_k)$, and $\zeta(v_i) \leq \zeta(v_k) \leq \hat{\zeta}(v_i) \leq \hat{\zeta}(v_k)$, establishing (1) and (2).

Thus it remains to prove (3), so we may assume X is restricted. By 5.1.2, w_i is the unique $w \in f(\hat{\zeta}(v_k))(\geq v_k)$ with $\zeta(w) = \hat{\zeta}(v_i)$. By induction, $z_i = \hat{f}_{v_i}(v_{k-1})$ is the unique $z \in f(\hat{\zeta}(v_{k-1}))(\geq v_{k-1})$ with $\zeta(z) = \hat{\zeta}(v_i)$ for $i < k$. Further $z_0 \leq \cdots \leq z_{k-1}$ and $f_{z_i}(z_j) \geq v_i$. Now $f^{-1}_{w_{k-1}}(z_i) \geq f^{-1}_{w_{k-1}}(v_{k-1})$ and $w_i \geq v_k \geq f^{-1}_{w_{k-1}}(v_{k-1})$, so as X is restricted, $w_i = f^{-1}_{w_{k-1}}(z_i)$. So as $z_j \geq z_i$ for $j \geq i$, $w_j \geq w_i$. Then $f_{w_j}(w_i) \geq f_{w_j}(v_k) \geq v_j$, completing the proof of (3).

For s a simplex of $K(X)$, Lemma 5.4 says that $X(s)$ has a greatest element $\hat{\zeta}(s)$.

(5.5) *If s is a simplex of $K(X)$ then* $\dim(s) \leq h(\hat{\zeta}(s))$ *and* $\dim(s) \leq |X(s)| + |\zeta(s)| - 2$.

Proof. Let $s = \{v_0, \ldots, v_k\}$ be ordered as in 5.4. By 5.3, for each $1 \leq i \leq k$, $\hat{\zeta}(v_{i-1}) < \hat{\zeta}(v_i)$ or $\zeta(v_{i-1}) < \zeta(v_i)$. Let $\phi(i) = \hat{\zeta}(v_{i-1})$ or $\zeta(v_{i-1})$ in the respective case. Then the map $\phi : \{i : 1 \leq i \leq k\} \to X(s) - \{\hat{\zeta}(s)\} \cup \zeta(s) - \{y\}$ is an injection, where $y = \zeta(v_k)$. Therefore the second remark in the lemma holds. Also by 5.4, $\{\phi(i) : i\}$ is a chain of length $k-1$ in $X(< \hat{\zeta}(s))$, so the first remark holds.

(5.6) *Let (X, f) be of height n and $V(n) = \{(m,k) : 0 \leq k \leq m \leq n\}$. Define $\tau : V \to V(n)$ by $\tau(v) = (h(\hat{\zeta}(v)), h(\zeta(v)))$. Then $K(X)$ is a typed simplicial complex over $V(n)$ with type function τ.*

Proof. This follows from 5.3 and 5.4.

Remark 5.7. Observe we have a covariant functor K from the category of combinatorial cell complexes to the category of typed simplicial complexes. We have already associated a typed simplicial complex $K(X)$ to X. Suppose $\alpha : X \to \bar{X}$ is a morphism of combinatorial cell complexes. Then for $x \in X$, we have a map $\alpha_x : f(x) \to \bar{f}(\alpha(x))$ of posets which induces a map $K(\alpha) : V(X) \to V(\bar{X})$ defined by $K(\alpha)(v) = \alpha_{\hat{\zeta}(v)}(v)$. If $u \in L(v)$ there is $w \in f(x)(\geq v)$ with $f_w(v) \geq u$. Then as

$$\begin{array}{ccc} f(x)(\leq w) & \xrightarrow{\alpha_x} & \bar{f}(\alpha(x))(\leq \alpha_x(w)) \\ f_w \downarrow & & \downarrow \bar{f}_{\alpha_x(w)} \\ f(\zeta(w)) & \xrightarrow{\alpha_{\zeta(w)}} & \bar{f}(\alpha(\zeta(w))) \end{array}$$

commutes,

$$K(\alpha)(u) = \alpha_{\hat{\zeta}(u)}(u) \le \alpha_{\hat{\zeta}(u)}(f_w(v)) \\ = \bar{f}_{\alpha_x(w)}(\alpha_x(v)) = \bar{f}_{K(\alpha)(w)}(K(\alpha)(v)),$$

with $K(\alpha)(v) = \alpha_x(v) \le \alpha_x(w) = K(\alpha)(w)$, so $K(\alpha)(u) \in L(K(\alpha)(v))$. Thus $K(\alpha) : \Delta(X) \to \Delta(\bar{X})$ is a map of graphs, and thus induces a simplicial map from $K(X)$ to $K(\bar{X})$. As α preserves height, $K(\alpha)$ also preserves the type function τ of Lemma 5.6. It is easy to check that $K(\alpha \circ \beta) = K(\alpha) \circ K(\beta)$, so K is indeed a functor.

Example 5.8. Consider the simplicial cell complex $(X(n), f)$ of the n-simplex defined in Example (2) in Section 3. The set $V(n) = \{(k,i) : 0 \le i \le k \le n\}$ is $V(X(n))$. Denote by $K(n)$ the triangulating complex $K(X(n))$ of $X(n)$. Then for $(a,b), (\alpha,\beta) \in V(n)$, $(a,b) \in L(\alpha,\beta)$ if and only if $b \le \beta \le a \le \alpha$. The complex $X(n)$ is regular and hence restricted. Observe that if $(a,b) \in L(\alpha,\beta)$ then $\hat{f}_{(a,b)}((\alpha,\beta)) = (\alpha, a)$.

In the remainder of this section we discuss the triangulating complex $K(n)$ of the simplicial cell complex $X(n)$ of the n-simplex. Observe first that we may regard $V(n)$ as the lower diagonal elements in an $n+1$ by $n+1$ square array. From this point of view, for $(\alpha,\beta) \in V(n)$, $L(\alpha,\beta)$ is the set of entries in $V(n)$ living in the rectangle with corners $(\beta,0)$, (β,β), $(\alpha,0)$, (α,β) sitting directly above and to the left of (α,β). Similarly those (a,b) with $(\alpha,\beta) \in L(a,b)$ form the rectangle with corners (α,β), (α,α), (n,β), (n,α), sitting directly below and to the right of (α,β).

Next 5.4 translates into the statement:

(5.9) *A subset s of $V(n)$ is in the set $\Sigma(n)$ of simplices of $K(n)$ if and only if we can order s so that $s = \{(\alpha_i,\beta_i) : 0 \le i \le k\}$ with $\beta_0 \le \cdots \le \beta_k \le \alpha_0 \le \cdots \le \alpha_k$.*

For $(\alpha,\beta) \in V(n)$ define

$$l(\alpha,\beta) = \{(a,b) : \alpha = a \text{ and } b = \beta - 1 \text{ or } \beta = b \text{ and } a = \alpha - 1\}$$

and a directed graph structure on $V(n)$ by $e \to f$ if $e \in l(f)$. Notice if $e \to f$ then e is adjacent to f in the graph $\Delta(n)$ of $X(n)$. Also $e \to f$ if e and f are adjacent lattice points in the array $V(n)$.

For $s \in \Sigma(n)$, let $\alpha^*(s) = \max\{\alpha : (\alpha,\beta) \in s\}$ and $\alpha_*(s) = \min\{\alpha : (\alpha,\beta) \in s\}$. Define $\beta^*(s)$ and $\beta_*(s)$ similarly.

(5.10) *The maximal simplices $\Sigma^*(n)$ of $K(n)$ are precisely the directed paths $p = p_0 \cdots p_n$ of length n in the directed graph $(V(n), \to)$ such that $p_i \to p_{i-1}$, $p_0 = (\alpha_*(p), 0)$, $p_n = (n, \beta^*(p))$, and $\alpha_*(p) = \beta^*(p)$. In particular $\alpha^*(p) = n$ and $\beta_*(p) = 0$.*

Proof. Let p be a maximal path and order p as in 5.9. Notice $\beta_0 = \beta_*(p)$, $\alpha_0 = \alpha_*(p)$, $\beta_k = \beta^*(p)$, and $\alpha_k = \alpha^*(p)$. Now $p \cup \{(n, \beta_k), (\alpha_0, 0)\} \in \Sigma(n)$, so by maximality of p, $n = \alpha_k = \alpha^*(p)$ and $0 = \beta_0 = \beta_*(p)$. If $p_i \notin l(p_{i+1})$ then $\alpha_{i+1} - \alpha_i > 1$ or $\beta_{i+1} - \beta_i > 1$, or $\alpha_{i+1} - \alpha_i = \beta_{i+1} - \beta_i = 1$, and we adjoint $(\alpha_i + 1, \beta_i)$, $(\alpha_i, \beta_i + 1)$, or $(\alpha_i + 1, \beta_i)$ to p in the respective case to contradict the maximality of p. Thus p is a directed path in $(V(n), \to)$. Notice the length of the path p is the number N of changes down and to the right as the path proceeds from p_0 to p_k, since at each step there is exactly one such change. Finally $p \cup \{(n, \alpha_0)\} \in \Sigma(n)$, so $(n, \alpha_0) = p_k$ by maximality of p. Thus $\beta^*(p) = \alpha_0 = \alpha_*(p)$. This implies that $n = N$, so p is of length n, completing the proof.

Remark 5.11. Lemma 5.10 says the maximal simplices p of $\Sigma(n)$ are all of dimension n and are the paths in the directed graph $(V(n), \to)$ within rectangles $R(k)$ with corners $(k, 0)$, (k, k), $(n, 0)$, (n, k) running from the upper left hand corner $(k, 0)$ to the lower right hand corner (n, k), where $k = \beta^*(p) = \alpha_*(p)$.

§6. Affine space, convex sets, and triangulations

Let $\mathbf{R}^n$ be n-dimensional Eulidean space. An *affine subspace* of $\mathbf{R}^n$ is a coset $U + x$ of a linear subspace U of $\mathbf{R}^n$. The *dimension* of the affine subspace $U + x$ is $\dim(U)$, with the empty set of dimension -1.

A subset C of $\mathbf{R}^n$ is *convex* if for each $x, y \in C$ and each real number t with $0 \le t \le 1$, $tx + (1 - t)y \in C$. The intersection of any family of convex sets is convex, so for each subset S if $\mathbf{R}^n$ there is a smallest convex subset $[S]$ of $\mathbf{R}^n$ containing S. We call $[S]$ the *convex closure* of S.

Define the *affine dimension* of a subset S of $\mathbf{R}^n$ to be the smallest dimension of an affine subspace containing S. Thus the affine dimension of S is $\dim(U(S))$, where $U(S) = \langle x - y : x, y \in S \rangle$, since $U(S) + s$ is the smallest affine subspace containing S for any $s \in S$. In particular $\dim(S) \le |S| - 1$ and we say S is *affine independent* if $\dim(S) = |S| - 1$ achieves this bound.

The next lemma is well known and easy to prove:

(6.1) (1) *For* $S \subseteq \mathbf{R}^n$,

$$[S] = \{ \sum_{x \in X} a_x x : 0 \leq a_x \in \mathbf{R},\ \sum_x a_x = 1,\ \textit{and}$$
$$X \textit{ is a finite subset of } S\}.$$

(2) *Let* $S = \{x_0, \ldots, x_k\}$ *be an affine independent subset of* $\mathbf{R}^n$. *Then each* $x \in [S]$ *can be written uniquely as* $x = \sum_i a_i x_i$ *with* $0 \leq a_i$ *and* $\sum_i a_i = 1$.

(6.2) *Let* $X, Y \subseteq \mathbf{R}^n$ *be convex,* $X = [x, X \cap Y]$, $Y = [y, X \cap Y]$, *and* $[x, y] \subseteq X \cup Y$. *Then* $X \cup Y$ *is convex.*

Proof. $X \cup Y \subseteq Z = [X, Y] = [X \cap Y, x, y]$. Let $z \in Z - [x, y]$. Then $z = ax + by + (1 - a - b)v$ for some $v \in X \cap Y$ and $0 \leq a, b \in \mathbf{R}$ with $a + b < 1$. Then $w = (ax + by)/(a + b) \in [x, y] \subseteq X \cup Y$, so we may take $w \in X$. Hence $z = (1 - a - b)v + (a + b)w \in X$, so that $Z \subseteq X \cup Y$ as desired.

(6.3) *Let* $x, y, z, w \in \mathbf{R}^n$, $0 < \varepsilon < 1$, $p = \varepsilon z + (1 - \varepsilon)y$, $q = \varepsilon x + (1 - \varepsilon)w$, $X = [p, e, f]$, *and* $Y = [q, e, f]$, *where either*
(1) $e = \varepsilon x + (1 - \varepsilon)y$ *and* $f = \varepsilon z + (1 - \varepsilon)w$, *or*
(2) $e = \varepsilon z + (1 - \varepsilon)x$ *and* $f = \varepsilon y + (1 - \varepsilon)w$.
Then $[p, q] \subseteq X \cup Y$ *and* $X \cup Y$ *is convex.*

Proof. We assume (1) holds; the proof of when (2) holds is essentially the same. Notice if $[p, q] \subseteq X \cup Y$ then $X \cup Y$ is convex by 6.4, so it remains to show $[p, q] \subseteq X \cup Y$.

Let $0 \leq t \leq 1$ and $v = tp + (1 - t)q$. Suppose first $t \geq 1/2$ and let $a = b = 1 - t$. Then $1 - a - b = 2t - 1$ and $0 \leq 2t - 1 \leq 1$ as $1/2 \leq t \leq 1$. Then by 6.1, $v = ae + bf + (1 - a - b)p \in X$. So let $t \leq 1/2$ and this time take $a = b = t$, so that $1 - a - b = 1 - 2t$ and $0 \leq 1 - 2t \leq 1$ because $0 \leq t \leq 1/2$. Now $v = ae + bf + (1 - a - b)q \in Y$.

In (2) take $a = (1 - t)\varepsilon/(1 - \varepsilon)$ and $b = 1 - t$ if $t \geq \varepsilon$, while if $t \leq \varepsilon$ take $a = t$ and $b = t(1 - \varepsilon)/\varepsilon$.

Let $K = (V, \Sigma)$ be a finite dimensional simplicial complex with vertex set V and simplices Σ. A *triangulation* of a topological space T by K is a map φ of Σ into the set of closed subspaces of T together with homeomorphisms

$$\varphi_s : \varphi(s) \to \hat{\varphi}(s) = [u(s, v) : v \in s] \subset \mathbf{R}^k$$

for each k-simplex s of K such that

(T1) For $s, t \in \Sigma$, $\varphi(s) \cap \varphi(t) = \varphi(s \cap t)$, where $\varphi(\varnothing) = \varnothing$.

(T2) $T = \bigcup_{s\in\Sigma} \varphi(s)$ and $C \subseteq T$ is closed in T if and only if $C \cap \varphi(s)$ is closed in $\varphi(s)$ for all $s \in \Sigma$.

(T3) For each k-simplex s of K and $t \subseteq s$, $\hat{\varphi}(s) = [u(s,v) : v \in s]$ is of affine dimension k and $\varphi_{t,s} = \varphi_s \circ \varphi_t^{-1}$ acts on $\hat{\varphi}(t) = [u(t,v) : v \in t]$ via $\varphi_{t,s} : \sum_{v\in t} a_v u(t,v) \mapsto \sum_{v\in t} a_v u(s,v)$.

A *morphism of topological spaces* $\varphi^i : K^i \to T^i$, $i = 1, 2$, *with triangulation* is a pair (α, β) where $\alpha : K^1 \to K^2$ is a simplicial map, $\beta : T^1 \to T^2$ is continuous, and for each $s \in \Sigma^1$,

(T_1) $\beta(\varphi^1(s)) \subseteq \varphi^2(\alpha(s))$, and

(T_2) $\alpha_s \circ \varphi^1_s = \varphi^2_{\alpha(s)} \circ \beta$, where $\alpha_s : \hat{\varphi}^1(s) \to \hat{\varphi}^2(\alpha(s))$ is defined by

$$\alpha_s : \sum_{v\in s} a_v u(s,v) \mapsto \sum_{v\in s} a_v u(\alpha(s), \alpha(v)).$$

Example 6.4. Let I be an index set and for $J \subseteq I$ let $T_J = [u_j : j \in J]$ be a convex subset of $\mathbf{R}^k$ of affine dimension $k = |J| - 1$. Let $K = (X_0, \Sigma)$ be a typed simplicial complex with type function $h : X_0 \to I$ (cf. Section 1) and let $X = \mathrm{sd}(K)$ be the barycentric subdivision of K.

We now construct a topological cell complex $\chi(K)$. We begin by defining X to be the poset of $\chi(K)$. (Notice $X \in \mathcal{P}$.) Then for $x \in X$ we form a topological cell $f(x) = (f(x), F(x))$ by letting $f(x) = X(\leq x)$ and for $u \leq v \leq x$, defining $F(v) = T_{h(v)}$, (where $h(v) = \{h(z) : z \in v\} \subseteq I$, keeping in mind that v is a simplex of K) and defining $F(u,v) : F(u) \to F(v)$ to be the inclusion map. The map ζ is defined to be the identity map on each $f(x)$, and for $v \in f(x)$, $f_v : f(x)(\leq v) \to f(\zeta(v))$ is also the identity map. It is straightforward to check that $\chi(K)$ *is* a topological cell complex. Notice that the combinatorial cell complex of the topological cell complex $\chi(K)$ is the simplicial cell complex of $\mathrm{sd}(K)$. In particular $\chi(K)$ is a regular complex.

We next extend χ to a covariant functor from the category of typed simplicial complexes over I to the category of topological cell complexes. Namely if $\alpha : K \to \bar{K}$ is a morphism of typed complexes over I, then α extends to a map $\chi(\alpha) : \mathrm{sd}(K) \to \mathrm{sd}(\bar{K})$ of posets via $\chi(\alpha)(x) = \{\alpha(v) : v \in x\}$. Next for $x \in X$, define $\chi(\alpha)_x : F(x) \to F(\chi(\alpha)(x))$ to be the identity map. This makes sense, since as α preserves the type function h, $h(x) = h(\chi(\alpha)(x))$, so $F(x) = F(\chi(\alpha)(x))$. It is easy to check that $\chi(\alpha)$ is a morphism of topological cell complexes and that $\chi(\alpha \circ \beta) = \chi(\alpha) \circ \chi(\beta)$, so that χ is indeed a covariant functor.

Form the topological space $A = A(\chi(K))$ as in Section 4. We next construct a triangulation $\varphi : K \to A$. Namely for $x \in X$ define $\varphi(x) = \tilde{F}(x) \subseteq A$. Then for $v \in x$, define $u(x,v) = u_{h(v)} \in T_{h(x)}$ and define $\varphi_x : \varphi(x) \to T_{h(x)}$ by $\varphi_x : \tilde{a} \mapsto a$. The map φ_x is just the inverse of $\lambda_x : F(x) \to \tilde{F}(x)$ defined in Section 4 by $\lambda_x(a) = \tilde{a}$. As ζ is injective, 4.6.3 says the map λ_x is a homeomorphism, so φ_x is a well defined homeomorphism.

By definition of the space A, $\varphi(x) = \tilde{F}(x)$ is a closed subspace of A for each $x \in X$ and axiom (T2) for triangulation holds. By 4.5, for $x, y \in X$, $\varphi(x) \cap \varphi(y) = \tilde{F}(x) \cap \tilde{F}(y) = \bigcup_{z \leq x,y} \tilde{F}(z) = \tilde{F}(x \cap y) = \varphi(x \cap y)$, so axiom (T1) holds. Finally if $y \leq x$ then

$$\varphi_{y,x} = \varphi_x \circ \varphi_y^{-1} = \lambda_x^{-1} \circ \lambda_y : \sum_i a_i u_i \to \lambda_x^{-1}(\widetilde{\sum_i a_i u_i}) = \sum_i a_i u_i,$$

so axiom (T3) is satisfied.

Let $T(K) = A(\chi(K))$ and $\varphi^K : K \to T(K)$ the triangulation just constructed. The space $T(K)$ is the *geometric realization* of the simplicial complex K.

To complete our discussion in this example, we extend T to a covariant functor from the category of typed simplicial complexes over I to the category of topological spaces with triangulation by essentially viewing T as the composition $T_0 = A \circ \chi$. We have just seen that χ is a functor and by 4.7, A is a functor, so $T_0 = A \circ \chi$ is a functor from typed complexes to topological spaces. Suppose $\alpha : K^1 \to K^2$ is a morphism of typed simplicial complexes over I. Define $T(\alpha) = (\alpha, T_0(\alpha))$. As T_0 is a covariant functor, $T(\alpha \circ \beta) = T(\alpha) \circ T(\beta)$, so it remains to check that $T(\alpha)$ is a morphism of triangulated spaces. We leave that as an exercise for the reader.

§7. Polyhedral cell complexes

A *polyhedral cell complex* is a $\mathcal{C}$-cell complex (X, f) where $\mathcal{C}$ is the category of triangulated topological spaces. Thus for $x \in X$ and $v \in f(x)$, $F(v)$ is a topological space together with a triangulation $B^v : f(x)(\leq v) \to F(v)$, where $f(x)(\leq v)$ is regarded as an order complex. Moreover for each simplex s of $f(x)(\leq v)$, $\hat{B}^v(s) = \hat{B}^{\zeta(v)}(f_v(s))$ and if $u \leq v$ and $s \subseteq f(x)(\leq u)$ then $\hat{B}^u(s) = \hat{B}^v(s)$ and $F(u,v) : F(u) \to F(v)$ satisfies $F(u,v)(B^u(s)) = B^v(s)$ and $B^v_s \circ F(u,v) = B^u_s$ on $B^u(s)$. Finally $f_v : f(x)(\leq v) \to f(\zeta(v))$ as an isomorphism of polyhedral cells satisfies $F_v(B^v(s)) = B^{\zeta(v)}(f_v(s))$ and $B^{\zeta(v)}_{f_v(s)} \circ F_v = B^v_s$ on $B^v(s)$.

A *morphism of polyhedral cell complexes* $\alpha : (X, f, F, B) \to (\bar{X}, \bar{f}, \bar{F}, \bar{B})$ is a morphism $\alpha : (X, f, F) \to (\bar{X}, \bar{f}, \bar{F})$ of topological cell complexes such that for each $x \in X$ and each simplex s of $f(x)$, $\alpha_x(B^x(s)) \subseteq B^{\alpha(x)}(\alpha(s))$ and $\alpha_s \circ B^x_s = B^{\alpha(x)}_{\alpha(s)} \circ \alpha_x$ on $B^x(s)$.

Example 7.1. We proceed as in Example 6.4. In particular let $I = \{0, 1, 2, \ldots\}$ and for $J \subseteq I$ let $T_J = [u_j : j \in J]$ be a convex subset of $\mathbf{R}^k$ of affine dimension $k = |J| - 1$.

Let $X = (X, f)$ be a combinatorial cell complex. We associate a polyhedral cell complex $\mathcal{P}(X) = (X, f, F, B)$ to X. For $x \in X$ the polyhedral cell associated to x is obtained using the construction of Example 6.4. Namely if $h(x) = n$ then $f(x)$ is a typed complex over I with respect to the height function $h : f(x) \to I$, so we can apply the geometric realization functor T of Example 6.4 to $f(x)$ and obtain a topological space $F(x) = T(f(x))$ (the geometric realization of the order complex of $f(x)$) and a triangulation $B^x : f(x) \to F(x)$. Then for $v \in f(x)$, $F(v)$ is the subspace $\bigcup_{u \leq v} B^x(u) \cong T(f(x)(\leq v))$ and $B^v : f(x)(\leq v) \to F(v)$ is the restriction of B^x to $f(x)(\leq v)$. For $u \leq v$, $F(u, v) : F(u) \to F(v)$ is the inclusion map. Notice as B^v and B^u are restrictions of B^x, $B^v_s = B^u_s$ on $B^u(s)$ for each simplex s of $f(x)(\leq v)$.

Axioms (i) and (iii) for topological cells (given in Example (2) of Section 2) are satisfied by 4.5.2 and 4.6.4.

Next as T is a functor, the isomorphism f_v of posets induces an isomorphism $T(f_v) = (f_v, F_v)$ of spaces with triangulation from $B^v : f(x)(\leq v) \to F(v)$ to $B^{\zeta(v)} : f(\zeta(v)) \to F(\zeta(v))$. In particular $F_v(B^v(s)) = B^{\zeta(v)}(f_v(s))$ for each simplex s of $f(x)(\leq v)$ and $B^{\zeta(v)}_{f_v(s)} \circ F_v = B^v_s$ on $B^v(s)$.

We next extend $\mathcal{P}$ to a functor from combinatorial cell complexes to polyhedral cell complexes. Let $\alpha : X \to \bar{X}$ be a morphism of combinatorial cell complexes. Our morphism $\mathcal{P}(\alpha) : \mathcal{P}(X) \to \mathcal{P}(\bar{X})$ is defined so that its image under the forgetful functor is α. Further for $x \in X$, $\alpha_x : f(x) \to \bar{f}(\alpha(x))$ as a map of posets is a morphism of typed complexes, so applying our functor T we get a morphism $T(\alpha_x) = (\alpha_x, T_x)$ of spaces with triangulation from $T(f(x)) = F(x)$ to $T(f(\alpha(x))) = F(\alpha(x))$, where $T_x = A(\chi(\alpha_x))$. We let $\mathcal{P}(\alpha)_x = T(\alpha_x)$.

Check that $\mathcal{P}(\alpha)$ is a morphism of polyhedral cell complexes. Moreover as $T(\alpha \circ \beta) = T(\alpha) \circ T(\beta)$ we have $\mathcal{P}(\alpha \circ \beta) = \mathcal{P}(\alpha) \circ \mathcal{P}(\beta)$.

In the remainder of this section assume (X, f, F, B) is a polyhedral cell complex. We adopt the notational conventions of Section 4. In particular for $x \in X$ and $v \in f(x)$ we regard $F(v)$ as a subspace of

$F(x)$, so that $F(u,v)$ becomes inclusion for each $u \leq v$. Similarly for s a simplex of $f(x)(\leq v)$ we can regard $B^v(s) = B^x(s)$ and write both as $B(s)$. Already by the definition of polyhedral cell complex we have $\hat{B}^x(s) = \hat{B}^v(s)$, and as $F(v,x)$ is inclusion, $B_s^v = B_s^x \circ F(v,x) = B_s^x$, and we denote both by B_s. That is $B : f(x) \to F(x)$ is a triangulation and for each $v \in f(x)$, B restricts to a triangulation $B : f(x)(\leq v) \to F(v)$.

Let S be a simplex in $f(x)$. We have a homeomorphism $B_S : B(S) \to \hat{B}(S) = [u_j : j \in J]$, where J is the set of heights of vertices in S and by axiom (T3), $B_S(v) = u_{h(v)}$. Now if $U \subseteq S$ and $a_v \geq 0$ with $\sum_{v\in U} a_v = 1$, $\sum_v a_v u_{h(v)}$ is a well defined element of $\hat{B}(S)$ and we define $\sum_{v\in U} a_v B(v) = B_S^{-1}(\sum_v a_v u_{h(v)})$. Notice by axiom (T3) that this definition is independent of the choice of S containing U. Further

(7.2) *If $w \in f(x)$ and U is a simplex in $f(x)(\leq w)$ then $Z = f_w(U)$ is a simplex in $f(\zeta(w))$, $F_w(B(U)) = B(Z)$, $\sum_{v\in U} a_v B(v) \in B(U) \subseteq F(w)$, and*

$$F_w\big(\sum_{v\in U} a_v B(v)\big) = \sum_{z\in Z} a_z B(z).$$

Proof. Let $q = \sum_{v\in U} a_v B(v)$. As $U \subseteq f(x)(\leq w)$, $S = U \cup \{w\}$ is a simplex in $f(x)$ and by the discussion above $\sum_v a_v u_{h(v)} \in \hat{B}(w)$ so $q \in B(w) \subseteq F(w)$. As f_w is a map of posets, $Z = f_w(U)$ is a simplex in $f(\zeta(w))$. Finally by definition of polyhedral cell complex, $B_Z \circ F_w = B_U$, so $F_w(q) = F_w(B_U^{-1}(\sum_v a_v u_{h(v)})) = B_Z^{-1}(\sum_v a_v u_{h(v)}) = \sum_z a_z u_{h(z)}$.

Next we associate to each $x \in X$ a graph $\Gamma = \Gamma(x)$ called the *residual graph of X at x*. Let

$$V(x) = \{(w,v) : w \in f(x) \text{ and } v \in f(\zeta(w))\}.$$

For $(w,v) \in V(x)$ define

$$L(w,v) = \{(w',v') \in V(x) : w' \leq w,\ f_w(w') \geq v, \text{ and } f_{f_w(w')}(v) \geq v'\}$$

Finally let $\Gamma(x)$ be the graph with vertex set $V(x)$ and (w,v) adjacent to (w',v') if $(w',v') \in L(w,v)$ or $(w,v) \in L(w',v')$.

Let $K(x)$ be the clique complex of $\Gamma(x)$. We call $K(x)$ the *residual complex of X at x*. Observe

(7.3) *Let $x \in X$, $(f(x), g)$ the simplicial cell complex of the poset $f(x)$, and for $w \in f(x)$ and $u \leq w$, write (w,u) for u regarded as an* *element of $f(x)(\leq w)$ in $V(f(x), g)$. Then the map $(w, v) \mapsto (w, f_w^{-1}(v))$ is an isomorphism of $K(x)$ with $K(f(x), g)$.*

Next pick a real number $0 < \varepsilon < 1$. For $v \in f(x)$ define

$$P(v) = \varepsilon B(x) + (1 - \varepsilon) B(v) \in F(x).$$

As x is the greatest element of $f(x)$, $\{x, v\}$ is a simplex of $f(x)$ and so the notation is well defined.

Next for $(w, v) \in V(x)$ define $P(w, v) \in F(x)$ by

$$P(w, v) = F_w^{-1}(P(v)) = \varepsilon B(w) + (1 - \varepsilon) B(f_w^{-1}(v)).$$

Thus $P(w, v) \in F(w)$. Indeed

(7.4) *For $w \in f(x)$, $P(w) \in I(x)$ and for $(w, v) \in V(x)$, $P(w, v) \in I(w)$.*

Proof. Recall from Section 4 that $I(w) = F(w) - \bigcup_{u \in f(x)(<w)} F(u)$. Now B is a triangulation of $F(x)$ with $F(w)$ the union of the topological simplices $B(S)$, $S \subseteq f(x)(\leq w)$, so

$$\dot{F}(x) = F(x) - I(x) = \bigcup_{x \neq w \in f(x)} F(w) = \bigcup_{x \notin S} B(S).$$

So as $P(w) = \varepsilon B(x) + (1 - \varepsilon) B(w)$ with $0 < \varepsilon < 1$, $P(w)$ is contained only in $B(S)$ for simplices S containing x, and hence $P(w) \in I(x)$. Similarly $P(v) \in I(\zeta(w))$, so $P(w, v) = F_w^{-1}(P(v)) \in F_w^{-1}(I(\zeta(w))) = I(w)$.

(7.5) *If $s = \{(w_0, v_0), \ldots, (w_k, v_k)\}$ is a simplex in $K(x) = K(\Gamma(x))$ then*

(1) We can order s so that $\bar{v}_0 \leq \bar{v}_1 \leq \cdots \leq \bar{v}_k \leq w_0 \leq \cdots \leq w_k$, where $\bar{v}_i = f_{w_i}^{-1}(v_i)$.

(2) $(w_j, v_j) \in L(w_i, v_i)$ for $j \leq i$.

(3) $S(s) = \{\bar{v}_i, w_i : 0 \leq i \leq k\}$ is a simplex of $f(x)$ with $P(w_i, v_i) \in B(S(s))$ for each i.

(4) $P(w, v) \in B(S)$ for some simplex S of $f(x)$ if and only if w and $f_w^{-1}(v) \in S$.

Proof. Parts (1) and (2) follow from 7.3 and 5.4. Notice (1) implies $S(s)$ is a simplex of $f(x)$. Then (4) completes the proof of (3), so it remains to prove (4). But (4) holds as $P(w, v) = \varepsilon B(w) + (1 - \varepsilon) B(f_w^{-1}(v))$ and $B : f(x) \to F(x)$ is a triangulation.

Let s be a simplex in $K(x)$. By 7.5, there is a smallest simplex $S(s)$ of $F(x)$ such that $P(e) \in B(S(s))$ for each vertex $e \in s$. For S a simplex of $f(x)$ containing $S(s)$, identifying $B(S)$ with $\hat{B}(S)$ we can consider the convex closure $\varphi(s) = [P(e) : e \in s]$ of the points $P(e)$ in $B(S)$. Because B is a triangulation, these identifications are independent of S. We let $u(s,e) = P(e)$, $\hat{\varphi}(s) = \varphi(s)$, and $\varphi_s : \varphi(s) \to \hat{\varphi}(s)$ be the identity map. The remainder of this section and all of the next section are devoted to showing:

Theorem 7.6. *$\varphi : K(x) \to F(x)$ is a triangulation of $F(x)$.*

As $\varphi(s)$ is a convex subset of $B(S(s))$, $\varphi(s)$ is closed in $B(S(s))$, and then as $B(S(s))$ is closed in $F(x)$, we conclude $\varphi(s)$ is closed in $F(x)$. If $t \subseteq s$ then by 7.5, $S(t) \subseteq S(s)$, so $\varphi(t) \subseteq \varphi(s)$ and then as φ_s and φ_t are identity maps, $\varphi_{t,s} = \varphi_s \circ \varphi_t^{-1} : \varphi(t) \to \varphi(s)$ is the inclusion map. Therefore axiom (T3) in the definition of triangulation is satisfied by φ *if* $\{P(e) : e \in s\}$ is affine independent for each s. Hence

(7.7) *To establish Theorem* 7.6 *is suffices to verify:*

(1) *For $s, t \in \Sigma(x)$ the set of simplices of $K(x)$, $\varphi(s) \cap \varphi(t) = \varphi(s \cap t)$.*

(2) $F(x) = \bigcup_{s \in \Sigma(x)} \varphi(s)$.

(3) *$\{P(e) : e \in s\}$ is affine independent of order* $\dim(s) + 1$ *for each* $s \in \Sigma(s)$.

(4) *$C \subseteq F(x)$ is closed in $F(x)$ if and only if $C \cap \varphi(s)$ is closed in $\varphi(s)$ for each $s \in \Sigma(x)$.*

For 7.7.1 is axiom (T1), while 7.7.2 and 7.7.4 are axiom (T2). Finally we have seen that 7.7.3 implies axiom (T3).

So it remains to verify 7.7.1 through 7.7.4.

(7.8) (1) *We may assume $X = \{x_0 < \cdots < x_n = x\}$ is a chain and $\zeta : f(x_i) \to X(\le x_i)$ is an isomorphism for each i.*

(2) *Under this assumption on X,* 7.7.2 *implies* 7.7.4.

Proof. Consider the polyhedral cell complex $\bar{X} = (\bar{X}, \bar{f}, \bar{F}, \bar{B})$, where $\bar{X} = f(x)$, $\bar{f}(v) = f(v)$, $\bar{F}(v) = F(v)$, $\zeta : f(v) \to f(x)$ is inclusion, $\bar{f}_v$ is the identity map, and $\bar{B}(v) = B(v)$ for each $v \in f(x)$. Notice that the combinatorial cell complex $(\bar{X}, \bar{f})$ of this polyhedral complex is just the simplicial cell complex of $\mathcal{O}(f(x))$; in particular by 7.3, there is a natural isomorphism of $K(x)$ with $K(\bar{X}, \bar{f})$. Then if Theorem 7.6 holds for $\bar{X}$ it also holds for X, so replacing X by $\bar{X}$, it suffices to take $X = f(x)$ and $\zeta : f(x) \to X$ an isomorphism.

Let S be a simplex of X, $\Sigma_S = \{s \in \Sigma(x) : B(s) \subseteq S\}$, $s \in \Sigma(x)$, and $s_S = \{(w,v) \in s : w, f_w^{-1}(v) \in S\}$. Then $\varphi(s) \cap B(S) \subseteq B(S(s)) \cap$

$B(S) = B(S(s) \cap S)$ as B is a triangulation. Also $\varphi(s) \cap B(S(s) \cap S) = \varphi(s_S)$ as the elements of $\varphi(s)$ are of the form

$$\sum_i a_i(\varepsilon B(w_i) + (1-\varepsilon)B(f_{w_i}^{-1}(v_i)))$$

while the elements $B(w_i), B(f_{w_i}^{-1}(v_i))$, are affine independent in $B(S(s))$. Therefore $\varphi(s) \cap B(S) = \varphi(s_S)$. In particular for $s, t \in \Sigma(x)$, $\varphi(s) \cap \varphi(t) = \varphi(s) \cap B(S(t)) \cap \varphi(t) \cap B(S(s)) = \varphi(s_{S(t)}) \cap \varphi(t_{S(s)}) = \varphi(s_T) \cap \varphi(t_T)$, where $T = S(s) \cap S(t)$, since $s \subseteq S(s)$ so $S_{s(t)} = s_T$ and similarly $t_{S(s)} = t_T$. Therefore if for each simplex T and each $e, f \in \Sigma_T$, we have $\varphi(e) \cap \varphi(f) = \varphi(e \cap f)$, then 7.7.1 holds. Similarly if $B(T) = \bigcup_{e \in \Sigma_T} \varphi(e)$, for each T, then 7.7.2 holds as does 7.7.4. The latter holds because B is a triangulation so C is closed in $F(x)$ if and only if $C \cap B(T)$ is closed in $B(T)$ for each T, and because 7.7.4 holds when $T = X$ and $B(T) = \bigcup_{e \in \Sigma_T} \varphi(e)$, since in that case $\Sigma(x)$ is finite. So it suffices to show for each simplex S of X that $\varphi : \Sigma_S \to B(S)$ satisfies 7.7.1 through 7.7.3. Hence replacing X by S, we may assume X is a chain. That is (1) holds, and we have already observed that (2) holds.

In the remainder of the proof we assume X is as described in Lemma 7.8. Therefore X is an n-simplex with greatest element x and ζ is injective, so by 3.1, (X, f) is the simplicial cell complex of the n- simplex. That is (X, f) is isomorphic to the complex $(X(n), f)$ of Example (2) in Section 3, so $K = K(X)$ is isomorphic to the complex $K(n) = K(X(n))$ discussed in Example 5.8 and subsequent lemmas in Section 5. Further by 7.3, the residual complex $K(x)$ is isomorphic to K. Thus without loss we may take $X = X(n)$, $V(x) = V(X) = V(n)$, etc. We adopt the notational conventions of Section 5 used to discuss $X(n)$.

Next $F(x) = [B_i : 0 \leq i \leq n]$ is the convex affine subspace of $\mathbf{R}^n$ generated by the affine independent set of vectors $B_i = B(x_i)$, $0 \leq i \leq n$. Further for $(\alpha, \beta) \in V(n)$, $P(\alpha, \beta) = \varepsilon B_\alpha + (1-\varepsilon)B_\beta$.

(7.9) *For each* $s \in \Sigma(x)$, $\{P(e) : e \in s\}$ *is affine independent of order* $\dim(s) + 1$.

Proof. Without loss s is a maximal simplex, so $s = \{s_0, \ldots, s_n\}$ is described in 5.10. Translating, we may take $B_0 = 0$, so it remains to show $Y = \{P(s_i) : 0 \leq i \leq n\}$ contains a basis for $\mathbf{R}^n$ as a linear space. As $\{B_0, \ldots, B_n\}$ is affine independent and generates $F(x)$ of affine dimension n, $\{B_1, \ldots, B_n\}$ contains a basis for $\mathbf{R}^n$ so it suffices to show $B_i \in \langle Y \rangle = U$. Assume not; then as $i = \alpha$ or β for some $(\alpha, \beta) \in s$, we can pick k minimal subject to B_{α_k} or $B_{\beta_k} \notin U$, where $s_k = (\alpha_k, \beta_k)$.

As s is a path in $(V(n), \rightarrow)$, $s_k \in l(s_{k-1})$ so without loss $\alpha_k = \alpha_{k-1}$ and $\beta_k = \beta_{k-1} + 1$. Now $P(s_k) = \varepsilon B_{\alpha_k} + (1-\varepsilon)B_{\beta_k} \in U$. By minimality of k, $B_{\alpha_{k-1}} \in U$, so as $\alpha_k = \alpha_{k-1}$, $B_{\beta_k} \in U$, contrary to the choice of k.

Remark 7.10. Observe that if (α, β) and $(\gamma, \delta) \in V(n)$ with $\alpha \geq \gamma$ then one of the following holds:

(0) $\gamma \geq \beta \geq \delta$ and $y \in x^{\perp}$.
(I) $\beta > \gamma$ and $y \notin x^{\perp}$.
(II) $\beta < \delta$ and $y \notin x^{\perp}$.

Define a subset θ of $V(n)$ to be *convex* if whenever $u = (\alpha, \beta)$ and $v = (\gamma, \delta)$ are in θ with $\alpha \geq \gamma$ and $u \notin v^{\perp}$ then

(i) (α, γ) and (β, δ) are in θ if $\beta > \gamma$, and
(ii) (γ, β) and (α, δ) are in θ if $\beta < \delta$.

Theorem 7.11. *If $\theta \subseteq V(n)$ is convex then $D(\theta) = C(\theta)$ is convex, where $D(\theta) = \bigcup_{s \subseteq \theta, s \in \Sigma} \varphi$ and $C(\theta) = [P(x) : x \in \theta]$.*

Proof. First if θ is a clique then $D(\theta) = \varphi(\theta) = C(\theta)$ by definition of $\varphi(\theta)$. So we may assume θ is not a clique. In particular as we prove Theorem 7.11 by induction on $|\theta|$, the induction is anchored.

Let $N = \min\{\alpha : (\alpha, \beta) \in \theta\}$ and $M = \max\{\beta : (\alpha, \beta) \in \theta\}$. Assume first that $N \geq M$. Then whenever $(\alpha, \beta), (\gamma, \delta) \in \theta$ with $\alpha \geq \gamma$, we have $\beta \leq M \leq N \leq \gamma$, so by Remark 7.10, either $(\gamma, \delta) \in (\alpha, \beta)^{\perp}$ or $\beta < \delta$, $\gamma < \alpha$, and as θ is convex, (γ, β) and (α, δ) are in θ.

Let $\theta^* = \{(\alpha, \beta) \in \theta : \theta \nsubseteq (\alpha, \beta)^{\perp}\}$. As θ is not a clique, $\theta^* \neq \varnothing$. Define

$$\beta_0 = \min\{\beta : (\alpha, \beta) \in \theta^*\},$$
$$\alpha_0 = \max\{\alpha : (\alpha, \beta_0) \in \theta^*\},$$
$$\alpha_1 = \min\{\alpha : (\alpha, \beta) \in \theta^*\},$$
$$\beta_1 = \max\{\beta : (\alpha_1, \beta)\theta^*\},$$

$v_i = (\alpha_i, \beta_i)$, and $\theta_i = \theta - \{v_i\}$.

By definition of v_0, there exists $(\gamma, \delta) \in \theta - v_0^{\perp}$, so by definition of β_0, we have $\delta \geq \beta_0$. Thus by an earlier remark, $\beta_0 < \delta$ and $\gamma < \alpha_0$. Then by definition of v_1, $\alpha_1 \leq \gamma < \alpha_0$. By symmetry, $\beta_0 < \beta_1$.

Claim θ_0 is convex. For if $(\alpha, \beta), (\gamma, \delta) \in \theta$ with $\alpha \geq \gamma$ and $(\gamma, \delta) \notin (\alpha, \beta)^{\perp}$, then by an earlier remark, $\beta < \delta$, $\gamma < \alpha$, and $(\gamma, \beta), (\alpha, \delta) \in \theta$. As $\beta_0 \leq \beta < \delta$, $(\alpha, \delta) \neq v_0$, while if $\beta = \beta_0$ then $\alpha_0 \geq \alpha > \gamma$, so $(\gamma, \beta) \neq v_0$. Hence $(\gamma, \beta), (\alpha, \delta) \in \theta_0$ and θ_0 is indeed convex. Similarly θ_1 is convex.

Next let $X = C(\theta_0)$, $Y = C(\theta_1)$, $q = P(v_0)$, and $p = P(v_1)$. Then $X = [p, C(\theta_0 \cap \theta_1)]$ and $C(\theta_0 \cap \theta_1) \subseteq X \cap Y$, so $X = [p, X \cap Y]$ and similarly $Y = [q, X \cap Y]$. By induction on the order of θ, $X = D(\theta_0)$ and $Y = D(\theta_1)$.

Also $p = P(v_1) = \varepsilon B_{\alpha_1} + (1-\varepsilon)B_{\beta_1}$ and $q = P(v_0) = \varepsilon B_{\alpha_0} + (1-\varepsilon)B_{\beta_0}$. Further $\alpha_0 > \alpha_1$ and $\beta_1 > \beta_0$, so $v_0 \notin v_1^\perp$ and hence by an earlier remark, $(\alpha_0, \beta_1), (\alpha_1, \beta_0) \in \theta$, and indeed as neither is v_0 nor v_1, each is even in $\theta_0 \cap \theta_1$. Therefore

$$e = P(\alpha_0, \beta_1) = \varepsilon B_{\alpha_0} + (1-\varepsilon)B_{\beta_1} \in X \cap Y$$

and

$$f = \varepsilon B_{\alpha_1} + (1-\varepsilon)B_{\beta_0} \in X \cap Y.$$

Thus $X \cup Y$ is convex by 6.3.1 and 6.2. Finally observe that $X \cup Y = D(\theta)$, so that Theorem 7.11 holds in this case. For as $v_0 \notin v_1^\perp$, each $s \in \Sigma$ with $s \subseteq \theta$ is contained in θ_0 or θ_1.

This leaves the case $N < M$. This time let $\alpha_0 = N$, $\beta_0 = \max\{\beta : (\alpha_0, \beta) \in \theta\}$, $\beta_1 = M$, $\alpha_1 = \min\{\alpha : (\alpha, \beta_1) \in \theta\}$, $v_i = (\alpha_i, \beta_i)$, and $\theta_i = \theta - \{v_i\}$. This time $\beta_0 \leq \alpha_0 = N < M = \beta_1 \leq \alpha_1$, so $v_0 \notin v_1^\perp$. Again θ_i is convex. For if $(\alpha, \beta), (\gamma, \delta) \in \theta$ with $\alpha \geq \gamma$ and $(\gamma, \delta) \notin (\alpha, \beta)^\perp$ then by Remark 7.10, either (i) $\beta > \gamma$ and $(\alpha, \gamma), (\beta, \delta) \in \theta$, or (ii) $\beta < \delta$ and $(\gamma, \beta), (\alpha, \delta) \in \theta$. In case (i), $\alpha \geq \beta > \gamma \geq N = \alpha_0$, so $v_0 \neq (\alpha, \gamma)$ or (β, δ). Similarly $\beta_1 = M \geq \beta > \gamma \geq \delta$, so $v_1 \neq (\alpha, \gamma)$ or (β, δ). Thus $(\alpha, \gamma), (\beta, \delta) \in \theta_0 \cap \theta_1$ in case (i). On the other hand in case (ii), $\beta < \delta$, so $\alpha > \gamma \geq \alpha_0$ and if $\gamma = \alpha_0$ then $\beta_0 \geq \delta > \beta$, so $v_0 \neq (\gamma, \beta)$ or (α, δ). Also $\beta_1 \geq \delta > \beta$ and if $\beta_1 = \delta$ then $\alpha_1 \leq \gamma < \alpha$, so $v_1 \neq (\gamma, \beta)$ or (α, δ). So again $(\alpha, \gamma), (\beta, \delta) \in \theta_0 \cap \theta_1$.

So θ_0 and θ_1 are convex. Again let $X = C(\theta_0)$, $Y = C(\theta_1)$, $q = P(v_0)$, and $p = P(v_1)$. As before, $X = [p, X \cap Y]$ and $Y = [q, X \cap Y]$, and by induction, $X = D(\theta_0)$ and $Y = D(\theta_1)$. As $v_0 \notin v_1^\perp$, as $\alpha_1 > \alpha_0$, and as $\beta_1 = M < N = \alpha_0$, it follows from Remark 7.10 and the convexity of θ that $(\alpha_1, \alpha_0), (\beta_1, \beta_0) \in \theta$, and then even are in $\theta_0 \cap \theta_1$. So

$$e = P(\alpha_1, \alpha_0) = \varepsilon B_{\alpha_1} + (1-\varepsilon)B_{\alpha_0}$$

and

$$f = P(\beta_1, \beta_0) = \varepsilon B_{\beta_1} + (1-\varepsilon)B_{\beta_0}$$

are in $D(\theta_0 \cap \theta_1) \subseteq X \cap Y$. Hence by 6.3.2 and 6.2, $X \cup Y$ is convex. Finally as above, $X \cup Y = D(\theta)$, completing the proof.

Corollary 7.12. $A = \bigcup_{s \in \Sigma} \varphi(s)$.

Proof. Notice $V(n)$ is convex with $\Sigma = \{s \in \Sigma : s \subseteq V(n)\}$, so $D(V(n)) = \bigcup_{s \in \Sigma} \theta(s)$ is convex by Thoerem 7.11. But $B_\alpha = P(\alpha, \alpha) \in C(V(n)) = D(V(n))$ for each α, $0 \leq \alpha \leq n$. Therefore $A = [B_\alpha : 0 \leq \alpha \leq n] \subseteq D(V(n))$, completing the proof.

§8. The proof of Theorem 7.6 is completed

In this section we complete the proof of Theorem 7.6. By 7.8 and the discussion following the proof of that lemma, we have reduced to the case where $X = X(n)$ is an n-simplex with maximal member $x = n$, $A = F(x)$, $\Sigma = \Sigma(x)$, and $K = K(X) = K(x) = K(n)$. By 7.7, 7.8, 7.9, and 7.12, it suffices to prove:

(8.1) *For $s, t \in \Sigma$, $\varphi(s) \cap \varphi(t) = \varphi(s \cap t)$.*

Recall we have

$$V(n) = \{(\alpha, \beta) : 0 \le \beta \le \alpha \le n\}$$

with $(\alpha', \beta') \in L(\alpha, \beta)$ if and only if $\beta' \le \beta \le \alpha' \le \alpha$.

Let $y \in A$. Recall from Section 7 that $A = [B_i : 0 \le i \le n]$ is the polytope in $\mathbf{R}^n$ generated by the affine independent set of vectors B_i and for $(\alpha, \beta) \in V(n)$, $P(\alpha, \beta) = \varepsilon B_\alpha + (1 - \varepsilon) B_\beta$. In particular $y = \sum_{i=0}^{n} a_i B_i$ with $a_i \ge 0$ and $\sum_i a_i = 1$, and that this expression is unique as the B_i are affine independent. Define

$$\sup(y) = \{i \in X : a_i \ne 0\}$$

and for $s \in \Sigma$ let

$$\sup(s) = \{\alpha, \beta : (\alpha, \beta) \in s\}$$

(8.2) *If $s, t \in \Sigma$ with $\sup(s) \cup \sup(t) \ne X$ then $\varphi(s) \cap \varphi(t) = \varphi(s \cap t)$.*

Proof. Let $m \notin \sup(s) \cup \sup(t)$. Replace X by $\bar{X} = X - \{m\} \cong X(n-1)$, $V(n)$ by

$$\bar{V}(n) = \{(\alpha, \beta) \in V(n) : \alpha, \beta \in \bar{X}\} \cong V(n-1),$$

and A by $\bar{A} = [B_i : i \in \bar{X}]$. Then by induction on n, $\bar{\varphi} : K(\bar{V}(n)) \to \bar{A}$ is a triangulation, where $\bar{\varphi}$ is the restriction of φ to $K(\bar{V}(n))$. As $\sup(s) \cup \sup(t) \subseteq \bar{X}$, s and t are simplices of $K(\bar{V}(n))$, so $\varphi(s) \cap \varphi(t) = \bar{\varphi}(s) \cap \bar{\varphi}(t) = \bar{\varphi}(s \cap t) = \varphi(s \cap t)$.

From now on pick $s, t \in \Sigma$ such that $\varphi(s) \cap \varphi(t) \ne \varphi(s \cap t)$. If $r \subseteq s$ then $\varphi(r) = [P(v) : v \in r] \subseteq [P(v) : v \in s] = \varphi(s)$, so $\varphi(s \cap t) \subseteq \varphi(s) \cap \varphi(t)$. Thus we can pick $y \in \varphi(s) \cap \varphi(t) - \varphi(s \cap t)$. By a remark above there is a unique expression, $y = \sum_i a_i B_i$ with $0 \le a_i \in \mathbf{R}$ and $\sum_i a_i = 1$. Also by 7.9, $\{P(v) : v \in s\}$ is affine independent, so there is a unique expression $y = \sum_{v \in s} b_v P(v)$ and similarly a unique expression $y = \sum_{v \in t} c_v P(v)$. For $r \in \Sigma$ and $\gamma \in X$, let

$$I_r(\gamma) = \{\alpha : (\alpha, \gamma) \in r\} \text{ and } J_r(\gamma) = \{\beta : (\gamma, \beta) \in r\}$$

(8.3) $\sup(y) = \sup(s) = \sup(t) = X$.

Proof. If $\sup(y) = X$ then as $\sup(y) \subseteq \sup(s) \cap \sup(t)$, the lemma holds. So it remains to prove $\sup(y) = X$. So assume not. Let $X_s = \{\alpha, \beta : b_{(\alpha,\beta)} \neq 0\}$ and $s_0 = \{v \in s : b_v \neq 0\}$. Define X_t and t_0 similarly. Then $X \neq \sup(y) = X_s = X_t$, $X_s = \sup(s_0)$, and $X_t = \sup(t_0)$. By 8.2, $y \in \varphi(s_0) \cap \varphi(t_0) = \varphi(s_0 \cap t_0) \subseteq \varphi(s \cap t)$, a contradiction.

(8.4) (1) *For* $\alpha > \beta^*(s)$, $a_\alpha/\varepsilon = \sum_{\beta \in J_s(\alpha)} b_{\alpha,\beta}$.
(2) *For* $\beta < \alpha_*(s)$, $a_\beta/(1-\varepsilon) = \sum_{\alpha \in I_s(\beta)} b_{\alpha,\beta}$.
(3) *If* $\kappa = \alpha_*(s) = \beta^*(s)$ *then*

$$a_\kappa = \varepsilon\Big(\sum_{\beta \in J_s(\kappa)} b_{\kappa,\beta}\Big) + (1-\varepsilon) \sum_{\alpha \in I_s(\kappa)} b_{\alpha,\kappa}.$$

Proof. This follows as $P(\alpha, \beta) = \varepsilon B_\alpha + (1-\varepsilon) B_\beta$ and $\beta^*(s) \leq \alpha_*(s)$.

Recall from Section 5 that we can think of the members of $V(n)$ as the lower diagonal elements in an $n+1$ by $n+1$ square array. Notice we have a duality on statements concerning $V(n)$ and K corresponding to the involution on X interchanging i and $n-i$ for each $i \in X$, and this corresponds to reflecting $V(n)$ about the "diagonal" $\{(n,0), (n-1,1), \ldots, (0,n)\}$. In applying this duality, one must also interchange the roles of ε and $1-\varepsilon$. We use this duality frequently from now on. In particular for $k \in X$ define

$$R^*(k) = \Big\{z = \sum_i z_i B_i \in A : \sum_{i \geq k} z_i \geq \varepsilon\Big\}$$

and define $R_*(k)$ dually. That is

$$R_*(k) = \Big\{z = \sum_i z_i B_i : \sum_{i \leq k} z_i \geq 1-\varepsilon\Big\}$$

Define $\alpha_*(s,y)$ to be the minimum α such that $b_{\alpha,\beta} \neq 0$ for some $(\alpha,\beta) \in s$ and define $\beta^*(s,y)$ dually.

(8.5) *Let* $k \in X$. *Then*
(1) $y \in R^*(k)$ *if and only if* $k \leq \alpha_*(s,y)$,
(2) $y \in R_*(k)$ *if and only if* $k \geq \beta^*(s,y)$.

Proof. We prove (1); then (2) follows by duality. Suppose first that $k \leq \alpha_*(s,y)$. Then by 8.3, $\sum_{j \geq k} a_j = \varepsilon(\sum_{v \in s} b_v) + (1-\varepsilon) \sum_{\alpha \in I_s(k)} b_{\alpha,k}$

$= \varepsilon + (1-\varepsilon)b$, where $b \geq 0$, since $\sum_{v\in s} b_v = 1$ and $b_v \geq 0$. Thus if $k \leq \alpha_*(s,y)$, $y \in R^*(k)$.

On the other hand if $k > \beta^*(s,y)$ then by 8.4, $\sum_{j\geq k} a_j = \varepsilon(\sum_{\alpha(v)\geq k} b_v) \leq \varepsilon$ with equality if and only if $k \leq \alpha_*(s,y)$, where $v = (\alpha(v), \beta(v))$. So as $\alpha_*(s,y) \geq \beta^*(s,y)$, we conclude that if $y \in R^*(k)$ then $k \geq \alpha_*(s,y)$.

(8.6) (1) $\varphi(s) \cap R^*(k) = \{z \in \varphi(s) : \alpha_*(s,z) \geq k\} = \varphi(s^*(k))$, *where* $s^*(k) = \{(\alpha,\beta) \in s : \alpha \geq k\}$.

(2) $\varphi(s) \cap R_*(k) = \{z \in \varphi(s) : \beta^*(s,z) \leq k\} = \varphi(s_*(k))$, *where* $s_*(k) = \{(\alpha,\beta) \in s : \beta \leq k\}$.

Proof. As usual (2) is the dual of (1) so it suffices to prove (1). But (1) follows from 8.5.1.

(8.7) $\alpha_*(s) \leq \beta^*(t) + 1$.

Proof. Let $\alpha_*(s) = k$. By 8.6, $\varphi(s) \subseteq R^*(k)$ and $\varphi(t) \cap R^*(k) = \varphi(t^*(k))$. Similarly setting $j = \beta^*(t)$, we have $\varphi(t) \subseteq R_*(j)$ and $\varphi(s) \cap R_*(j) = \varphi(s_*(j))$. Therefore

$$\varphi(s) \cap \varphi(t) = \varphi(s) \cap R_*(j) \cap \varphi(t) \cap R^*(k) = \varphi(s_*(j)) \cap \varphi(t^*(k))$$

Assume $j < k-1$. Then $k-1 \notin \sup(t^*(k))$ as $\alpha_*(t^*(k)) \geq k > k-1$ and $\beta^*(t^*(k)) \leq j < k-1$. Therefore by 8.3, $\varphi(s) \cap \varphi(t) = \varphi(s_*(j)) \cap \varphi(t^*(k)) = \varphi(s_*(j) \cap t^*(k)) \subseteq \varphi(s \cap t)$, a contradiction.

(8.8) *Suppose* $v \in s \cap t$. *Then*
(1) *If* $\sup(s - \{v\}) \neq X$ *then* $b_v \neq c_v$.
(2) *If* $J_s(\alpha) = \{\beta\}$ *then* $J_t(\alpha) \neq \{\beta\}$.
(3) *If* $I_s(\beta) = \{\alpha\}$ *then* $I_t(\beta) \neq \{\alpha\}$.

Proof. Assume $\sup(s - \{v\}) \neq X$ but $b_v = c_v = b$ and let $y' = (y - bP(v))/(1-b)$. As $P(v) \in \varphi(s \cap t)$, $y' \in \varphi(s) \cap \varphi(t) - \varphi(s \cap t)$. Then as $\sup(y') \subseteq \sup(s - \{v\}) \neq X$, 8.3 supplies a contradiction.

Thus (1) is established. Notice (3) is the dual of (2) so it remains to prove (2). Assume $J_s(\alpha) = \{\beta\}$. Then by 8.4, $a_\alpha/\varepsilon = b_v$ and $\alpha \notin \sup(s - \{v\})$. Thus by (1), $b_v \neq c_v$, so by symmetry, $J_t(\alpha) \neq \{\beta\}$.

We next observe that we can choose s,t to be maximal simplices of dimension n. For $s \subseteq s_0$ and $t \subseteq t_0$ where s_0 and t_0 are maximal simplices and hence of dimension n by Remark 5.11. Notice if $r \subseteq s_0$ then by definition of φ, $\varphi(r) \cap \varphi(s) = \varphi(r \cap s)$. Hence if $\varphi(s_0) \cap \varphi(t_0) = \varphi(s_0 \cap t_0)$ then $\varphi(s) \cap \varphi(t) = \varphi(s) \cap \varphi(t) \cap \varphi(s_0 \cap t_0) = \varphi(s) \cap \varphi(s_0 \cap t) = \varphi(s \cap t)$. Therefore replacing s,t by s_0,t_0 if necessary, we may assume s

and t are of dimension n. Hence s and t are described in 5.10 and 5.11. In particular $\alpha_*(s) = \beta^*(s)$ and $\alpha_*(t) = \beta^*(t)$. Let $\kappa = \alpha_*(s)$.

(8.9) $\alpha_*(t) = \alpha_*(s) = \kappa$.

Proof. Without loss $\kappa > \alpha_*(t)$. Then as $\alpha_*(t) = \beta^*(t)$, $\beta^*(t) = \kappa - 1$ by 8.7.

Now $\alpha_*(s,y) \geq \alpha_*(s) = \kappa$, so by 8.5, $y \in R^*(\kappa)$ and then by another application of 8.5, $\alpha_*(t,y) \geq \kappa$. Then as $X = \sup(y)$, $v = (0,\kappa)$ and $u = (n, \kappa-1)$ are in t. By duality, $\beta^*(s,y) \leq \kappa - 1$ and $u, v \in s$.

Next by 8.4, either $I_s(0) = \{\kappa\}$ and $b_v = a_0/(1-\varepsilon)$ or $J_s(\kappa) = \{0\}$ and $b_v = a_\kappa/\varepsilon$. Similarly either $I_t(0) = \{\kappa-1, \kappa\}$ and $c_v = a_0/(1-\varepsilon)$ or $J_t(\kappa) = \{0\}$ and $c_v = a_\kappa/\varepsilon$. Then by 8.6, we may asssume $I_s(0) = \{\kappa\}$, $b_v = a_0/(1-\varepsilon)$, $J_t(\kappa) = \{0\}$ and $c_v = a_\kappa/\varepsilon$. Now by 8.4, $a_0/(1-\varepsilon) = c_v + C$ and $a_\kappa/\varepsilon = b_v + B$, where $C = \sum_{\alpha \in I_t(0)} c_{\alpha,0} - c_v$ and $B = \sum_{\beta \in J_s(\kappa)} b_{\kappa,\beta} - b_v$. But then

$$b_v = a_0/(1-\varepsilon) = c_v + C = a_\kappa/\varepsilon + C = b_v + B + C.$$

We conclude $B = C = 0$ and hence $b_v = c_v$. Now 8.6 supplies a contradiction.

Let S be the set of maximal simplices m of K with $\alpha_*(m) = \kappa$. Thus $s, t \in S$. We partition S into four classes S_i, $1 \leq i \leq 4$, where

$$\begin{aligned} S_1 &= \{m \in S : J_m(\kappa) = \{0\} \text{ and } I_m(\kappa) = \{n\}\}, \\ S_2 &= \{m \in S : J_m(\kappa) = \{0\} \text{ and } J_m(n) = \{\kappa\}\}, \\ S_3 &= \{m \in S : I_m(0) = \{\kappa\} \text{ and } I_m(\kappa) = \{n\}\}, \\ S_4 &= \{m \in S : I_m(0) = \{\kappa\} \text{ and } J_m(n) = \{\kappa\}\}. \end{aligned}$$

Notice that the duality map interchanges S_2 and S_3 and fixes S_1 and S_4. So anything we prove about S_2 establishes the dual statement for S_3 at the same time.

(8.10) *If* $s \in S_3 \cup S_4$ *then* $t \in S_1 \cup S_2$.

Proof. This follows from 8.8.3.

(8.11) *Let* $s \in S_3 \cup S_4$ *and define*

$$b = \sum_{\alpha \in I_s(\kappa)} b_{\alpha,\kappa}, \quad c = \sum_{\alpha \in I_t(\kappa)} c_{\alpha,\kappa}$$

$$\bar{b} = \sum_{0<\beta \in J_s(\kappa)} b_{\kappa,\beta}, \quad c^* = \sum_{\kappa<\alpha \in I_t(0)} c_{\alpha,0}$$

Then

(1) $b < c$ *and* $\bar{b} + c^* \neq 0$,

(2) $s \in S_3$.

Proof. By 8.10, $t \in S_1 \cup S_2$. Thus $J_t(\kappa) = \{0\}$, so $c_{\kappa,0} = a_\kappa/\varepsilon - (1-\varepsilon)c/\varepsilon$ by 8.4.3. Similarly as $s \in S_3 \cup S_4$, $I_s(0) = \{\kappa\}$, so $b^* = \sum_{\kappa<\alpha\in I_s(0)} b_{\alpha,0} = 0$. Now by 8.4.2,

$$(*) \qquad a_0/(1-\varepsilon) = b_{\kappa,0} + b^* = c_{\kappa,0} + c^*$$

so $b_{\kappa,0} = c_{\kappa,0} + c^* = a_\kappa/\varepsilon - (1-\varepsilon)c/\varepsilon + c^*$. Next by 8.4.3, $a_\kappa/\varepsilon = b_{\kappa,0} + \bar{b} + (1-\varepsilon)b/\varepsilon$, so

$$b_{\kappa,0} = a_\kappa/\varepsilon - (1-\varepsilon)c/\varepsilon + c^* = b_{\kappa,0} + \bar{b} + (1-\varepsilon)(b-c)/\varepsilon + c^*$$

so that

$$(**) \qquad 0 = \bar{b} + c^* + (1-\varepsilon)(b-c)/\varepsilon.$$

Therefore to prove (1) it suffices to show $b < c$. Assume otherwise. Then $\bar{b}, c^*, (b-c) \geq 0$, so by (**), $\bar{b} = c^* = b - c = 0$. But then $b^* = c^* = 0$, so by (*), $b_{\kappa,0} = c_{\kappa,0}$ and applying 8.8.1 to $v = (\kappa, 0)$, we have a contradiction.

Therefore (1) is established and it remains to prove (2), so we may take $s \in S_4$. Therefore $J_s(n) = \{\kappa\}$, so by 8.4.1, $a_n/\varepsilon = b_{n,\kappa}$. Hence $b = b_{n,\kappa} + \tilde{b} = a_n/\varepsilon + \tilde{b}$, where $\tilde{b} = \sum_{n>\alpha\in I_s(\kappa)} b_{\alpha,\kappa}$. Again by 8.4.1, $\hat{c} + c_{n,\kappa} = a_n/\varepsilon$, where $\hat{c} = \sum_{n>\beta\in J_t(n)} c_{n,\beta}$. Therefore

$$b = a_n/\varepsilon + \tilde{b} = c_{n,\kappa} + \tilde{b} + \hat{c}.$$

Finally as $s \in S_4$, $t \notin S_2$ by the dual of 8.10, so $t \in S_1$. Thus $I_t(\kappa) = \{n\}$, so $c = c_{n,\kappa}$. Therefore $b = c_{n,\kappa} + \tilde{b} + \hat{c} = c + \tilde{b} + \hat{c} \geq c$, as $\tilde{b}, \hat{c} \geq 0$. This contradicts (1).

(8.12) *Up to a permutation of* $\{s,t\}$ *and duality,* $(s,t) \in S_1 \times S_1$, $S_3 \times S_1$, *or* $S_3 \times S_2$.

Proof. This follows from 8.10 and 8.11.

Define $v = (\alpha, \beta) \in s$ to be an *inflection point* of s if $v = (\kappa, 0)$ or (n, κ), or $|J_s(\alpha)| > 1 < |I_s(\beta)|$. By 5.10, $s = \{v_0, \ldots, v_n\}$ is a directed path in $(V(n), \rightarrow)$ with $v_0 = (\kappa, 0)$ and $v_n = (n, \kappa)$. Let $v_{i_0}, \ldots, v_{i_l}$ be the inflection points with $i_j < i_{j+1}$. Then $v_{i_0} = (\alpha_0, \beta_0) = (\kappa, 0)$ and $v_{i_l} = (n, \kappa)$. If $s \in S_1 \cup S_2$ then $J_s(\kappa) = \{0\}$, so $v_{i_1} = (\alpha_1, \beta_0)$

with $\kappa = \alpha_0 < \alpha_1$. Then proceeding recursively, $v_{i_{2r}} = (\alpha_r, \beta_r)$ and $v_{i_{2r+1}} = (\alpha_{r+1}, \beta_r)$ with $\alpha_i < \alpha_{i+1}$ and $\beta_i < \beta_{i+1}$. Further if $s \in S_1$ then $I_s(\kappa) = \{n\}$, so $l = 2N$ is even and $(n, \kappa) = (\alpha_N, \beta_N)$, while if $s \in S_2$ then $J_s(n) = \{\kappa\}$ so $l = 2N + 1$ is odd and $(n, \kappa) = (\alpha_{N+1}, \beta_N)$. Then dualizing the case $s \in S_2$ to get the answer when $s \in S_3$, we conclude:

(8.13) *The inflection points for s are:*
(1) (α_i, β_i), (α_{i+1}, β_i), $0 \le i < N$, *and* $(\alpha_N, \beta_N) = (n, \kappa)$, *if* $s \in S_1$.
(2) (α_i, β_i), (α_{i+1}, β_i), $0 \le i \le N$, *if* $s \in S_2$.
(3) (α_i, β_i), (α_i, β_{i+1}), $0 \le i \le N$, *if* $s \in S_3$.

(8.14) *For* $s \in S_1 \cup S_2$ *we have:*
(1) $b_{\alpha,\beta_i} = a_\alpha/\varepsilon$ *for* $\alpha_i < \alpha < \alpha_{i+1}$.
(2) $b_{\alpha_i,\beta} = a_\beta/(1-\varepsilon)$, *for* $\beta_{i-1} < \beta < \beta_i$.
(3) $b_{\alpha_i,\beta_i} = b_{\kappa,0} + 1/\varepsilon \sum_{\kappa<\alpha\le\alpha_i} a_\alpha - 1/(1-\varepsilon) \sum_{0\le\beta<\beta_i} a_\beta$

$$= 1 - 1/\varepsilon \textstyle\sum_{\kappa<\alpha\le n} a_\alpha - 1/(1-\varepsilon) \sum_{0\le\beta<\beta_i} a_\beta$$

(4) $b_{\alpha_i,\beta_{i-1}} = 1/(1-\varepsilon) \sum_{0\le\beta\le\beta_{i-1}} a_\beta - 1/\varepsilon \sum_{\kappa<\alpha<\alpha_i} a_\alpha - b_{\kappa,0}$

$$= 1/(1-\varepsilon) \textstyle\sum_{0\le\beta\le\beta_{i-1}} a_\beta + 1/\varepsilon \sum_{\alpha_i\le\alpha\le n} a_\alpha - 1$$

unless $s \in S_2$ *and* $i = N + 1$.
(5) $b_{\kappa,0} = 1 - 1/\varepsilon \sum_{\kappa<\alpha\le n} a_\alpha$.
(6) *If* $s \in S_2$ *then* $b_{n,\kappa} = a_n/\varepsilon$.

Proof. If $\alpha_j < \alpha < \alpha_{j+1}$ then $J_s(\alpha) = \{\beta_j\}$ because (α, β_j) is not an inflection point since $v_{i_{2j+1}} = (\alpha_{j+1}, \beta_j)$ is the next inflection point after $v_{i_{2j}} = (\alpha_j, \beta_j)$. Therefore (1) holds by 8.4.1. Similarly (2) holds. We prove the first equality in (3) and (4) by induction on i. To anchor the induction, recall $(\alpha_0, \beta_0) = (\kappa, 0)$, so (3) holds when $i = 0$. Then for $i > 0$, 8.4.1 says

$$b_{\alpha_i,\beta_i} = a_{\alpha_i}/\varepsilon - \sum_{\beta_i\ne\beta\in J_s(\alpha_i)} b_{\alpha_i,\beta} = a_{\alpha_i}/\varepsilon - \sum_{\beta_{i-1}\le\beta<\beta_i} b_{\alpha_i,\beta}$$

which by (2) and induction on i is equal to

$$a_{\alpha_i}/\varepsilon - \sum_{\beta_{i-1}<\beta<\beta_i} a_\beta/(1-\varepsilon) - 1/(1-\varepsilon) \sum_{0\le\beta\le\beta_{i-1}} a_\beta + 1/\varepsilon \sum_{\kappa<\alpha<\alpha_i} a_\alpha + b_{\kappa,0}$$

$$= b_{\kappa,0} + 1/\varepsilon \sum_{\kappa<\alpha\le\alpha_i} a_\alpha - 1/(1-\varepsilon) \sum_{0\le\beta<\beta_i} a_\beta$$

as claimed. A similar argument establishes (4), except when $s \in S_2$ and $i = N+1$, when $\beta_{i-1} = \kappa$ so that 8.4.3 must be used rather than 8.4.2, which is appropriate for smaller. In this case as $s \in S_2$, $J_s(n) = \{\kappa\}$ so $b_{n,\kappa} = a_n/\varepsilon$ by 8.4.1. But also by 8.4.3, a_κ is equal to

$$\begin{aligned}\varepsilon b_{\kappa,0} + (1-\varepsilon) \sum_{\alpha_{N-1} \leq \alpha \leq n} b_{\alpha,\kappa} =& \varepsilon b_{\kappa,0} + (1-\varepsilon)/\varepsilon \sum_{\alpha_{N-1} < \alpha < n} a_\alpha \\ &+ (1-\varepsilon)(b_{\alpha_{N-1},\kappa} + b_{n,\kappa})\end{aligned}$$

by (1), and then by induction on i, $b_{n,\kappa}$ is equal to

$$\begin{aligned}&a_\kappa/(1-\varepsilon) - \varepsilon b_{\kappa,0}/(1-\varepsilon) - 1/\varepsilon \sum_{\alpha_{N-1} < \alpha < n} a_\alpha - b_{\kappa,0}/(1-\varepsilon) \\ &- 1/\varepsilon \sum_{K < \alpha \leq \alpha_{N-1}} a_\alpha + 1/(1-\varepsilon) \sum_{0 \leq \beta < \kappa} a_\beta \\ =& 1/(1-\varepsilon) \sum_{0 \leq \beta \leq \kappa} a_\beta - 1/\varepsilon \sum_{\kappa < \alpha < n} a_\alpha - b_{\kappa,0}/(1-\varepsilon).\end{aligned}$$

Then as $b_{n,\kappa} = a_n/\varepsilon$, we have

$$b_{\kappa,0} = \sum_{0 \leq \beta \leq \kappa} a_\beta - (1-\varepsilon)/\varepsilon \sum_{\kappa < \alpha \leq n} a_\alpha = 1 - 1/\varepsilon \sum_{\kappa < \alpha \leq n} a_\alpha$$

as $\sum_{i=0}^{n} a_i = 1$. This gives (5) and (6) when $s \in S_2$.

Similarly when $s \in S_1$ we conclude from the first equality in (3) that

$$b_{n,\kappa} = b_{\kappa,0} + 1/\varepsilon \sum_{\kappa < \alpha \leq n} a_\alpha - 1/(1-\varepsilon) \sum_{0 \leq \beta < \kappa} a_\beta$$

But also by 8.4.3,

$$\begin{aligned}a_\kappa &= \varepsilon b_{\kappa,0} + (1-\varepsilon) b_{n,\kappa} \\ &= \varepsilon b_{\kappa,0} + (1-\varepsilon) b_{\kappa,0} + (1-\varepsilon)/\varepsilon \sum_{\kappa < \alpha \leq n} a_\alpha - \sum_{0 \leq \beta < \kappa} a_\beta \\ &= b_{\kappa,0} + 1/\varepsilon \sum_{\kappa < \alpha \leq n} a_\alpha - \sum_{i \neq \kappa} a_i = b_{\kappa,0} + 1/\varepsilon \sum_{\kappa < \alpha \leq n} a_\alpha - (1 - a_\kappa)\end{aligned}$$

as $\sum_i a_i = 1$. Therefore $b_{\kappa,0} = 1 - 1/\varepsilon \sum_{\kappa < \alpha \leq n} a_\alpha$, so that (5) holds in this case too. Finally substituting (5) into the first inequality in (3) and (4) gives the second inequality.

The dual of 8.14 in the case $s \in S_2$ is:

(8.15) *For* $s \in S_3$ *we have:*
(1) $b_{\alpha,\beta_i} = a_\alpha/\varepsilon$ *for* $\alpha_{i-1} < \alpha < \alpha_i$.
(2) $b_{\alpha_i,\beta} = a_\beta/(1-\varepsilon)$, *for* $\beta_i < \beta < \beta_{i+1}$.
(3) $b_{\alpha_i,\beta_{i+1}} = 1 - 1/(1-\varepsilon)\sum_{0<\beta<\beta_{i+1}} a_\beta - 1/\varepsilon\sum_{\alpha_i<\alpha\leq n} a_\alpha$.
(4) $b_{\alpha_i,\beta_i} = 1/\varepsilon\sum_{\alpha_i\leq\alpha\leq n} a_\alpha + 1/(1-\varepsilon)\sum_{0\leq\beta\leq\beta_i} a_\beta - 1$ *unless* $i = 0$.
(5) $b_{n,\kappa} = 1 - 1/(1-\varepsilon)\sum_{0\leq\beta<\kappa} a_\beta$.
(6) $b_{\kappa,0} = a_0/(1-\varepsilon)$.

(8.16) $(s,t) \notin S_1 \times S_1$.

Proof. Assume $s,t \in S_1$. Let $v = (\kappa, 0)$ and $u = (n,\kappa)$. Then by 8.14.5, $b_v = c_v$, while by 8.14.3, $b_u = c_u = 1 - 1/(1-\varepsilon)\sum_{0\leq\beta<\kappa} a_\beta$. Let

$$y' = (y - b_v P(v) - b_u P(u))/(1 - b_v - b_u).$$

Then $y' \in \varphi(s) \cap \varphi(t) - \varphi(s \cap t)$, while $\kappa \notin \sup(y')$, contradicting 8.3.

(8.17) $(s,t) \notin S_3 \times S_j$ *for* $j = 1$ *or* 2.

Proof. Assume otherwise. Recall the definition of $\bar{b}$ and c^* from 8.11. Then as $s \in S_3$, 8.15 says

$$\begin{aligned} c^* &= \sum_{\kappa<\alpha\in I_t(0)} c_{\alpha,0} = \sum_{\kappa<\alpha\leq\alpha_1} c_{\alpha,0} \\ &= 1/\varepsilon \sum_{\kappa<\alpha<\alpha_1} a_\alpha + a_0/(1-\varepsilon) + 1/\varepsilon \sum_{\alpha_1\leq\alpha\leq n} a_\alpha - 1 \\ &= 1/\varepsilon \sum_{\kappa<\alpha\leq n} a_\alpha - 1 + a_0/(1-\varepsilon). \end{aligned}$$

Similarly as $t \in S_1 \cup S_2$, 8.14 says

$$\begin{aligned} \bar{b} &= \sum_{0<\beta\in J_s(\kappa)} b_{\kappa,\beta} = \sum_{0<\beta\leq\beta_1} b_{\kappa,\beta} \\ &= 1/(1-\varepsilon) \sum_{0<\beta<\beta_1} a_\beta + 1 - 1/(1-\varepsilon) \sum_{0\leq\beta<\beta_1} a_\beta - 1/\varepsilon \sum_{\kappa<\alpha\leq n} a_\alpha \\ &= 1 - a_0/(1-\varepsilon) - 1/\varepsilon \sum_{\kappa<\alpha\leq n} a_\alpha. \end{aligned}$$

That is $\bar{b} = -c^*$, contradicting 8.11.1.

Notice 8.12, 8.16, and 8.17 constitute a contradiction. Therefore the proof of Theorem 8.1, and hence also of Theorem 7.6, is at last complete.

§9. Triangulating the space of a restricted polyhedral cell complex

In this section we continue to assume (X, f, F, B) is a polyhedral cell complex and also continue the notational conventions of Section 7. Let $K = K(X)$ be the triangulating complex of X and let Σ be the set of simplices of K.

Recall for $x \in X$ we have the residual complex $K(x)$ of X at x with vertex set

$$V(x) = \{(w, v) : w \in f(x) \text{ and } v \in f(\zeta(w))\}$$

with simplex set $\Sigma(x)$. Define $\eta : V(x) \to V$ by $\eta(w, v) = v$.

(9.1) (1) *$\eta(L(w, v)) \subseteq L(v)$ for each $(w, v) \in V(x)$.*

(2) *$\eta : V(x) \to V$ induces a morphism $\eta : \Gamma(x) \to \Delta$ of graphs and a morphism $\eta : K(x) \to K$ of simplicial complexes.*

(3) *If (X, f) is resticted then each $S \in \Sigma$ with $\hat{\zeta}(S) = x$ is the image under η of a unique $s \in \Sigma(x)$. Indeed if $S = \{v_0, \dots, v_k\}$ with $v_i \in L(v_j)$ for $i \leq j$, and $w_i = \hat{f}_{v_i}(v_k)$ then $s = s(S) = \{(w_i, v_i) : 0 \leq i \leq k\}$.*

Proof. Let $(y, u) \in L(w, v)$. Then $y \leq w$, $z = f_w(y) \geq v$, and $f_z(v) \geq u$. As $f_z(v) \geq u$, $\eta(y, u) = u \in L(v)$. Thus (1) holds. As $v = \eta(w, v)$, (1) implies (2).

Under the hypotheses and notation of (3), 5.4 says $w_0 \leq \cdots \leq w_k$ with $f_{w_j}(w_i) \geq v_j$ for $j \geq i$. Also as $v_i \in L(v_j)$, $\hat{f}_{v_i}(v_j)$ is the unique $w \in f(x_j)$ with $\zeta(w) = x_i$ and $w \geq v_j$. Thus $\hat{f}_{v_i}(v_j) = f_{w_j}(w_i)$, so $f_{f_{w_j}(w_i)}(v_j) = v_i$. Therefore $(w_i, v_i) \in L(w_j, v_j)$, so $s \in \Sigma(x)$. By construction, $\eta(s) = S$. Finally if $s' = \{(w_i', v_i') : 0 \leq i \leq k\} \in \Sigma(x)$ with $\eta(s') = S$ then $v_i' = \eta(w_i', v_i') = v_i$. Next $x = \hat{\zeta}(S) = \hat{\zeta}(v_k)$, so $v_k \in f(x)$, while as $(w_k', v_k) \in V(x)$, $v_k \in f(\zeta(w_k'))$, so $w_k' = x = w_k$. As $(w_i', v_i) \in L(x, v_k)$, $w_i' = f_x(w_i) \geq v_k$ and $f_{w_i'}(v_k) \geq v_i$, so $w_i' = \hat{f}_{v_i}(v_k) = w_i$. That is $s = s'$ is unique, completing the proof of (3).

Recall from Theorem 7.6, we have a triangulation $\varphi_x : K(x) \to F(x)$. Recall by construction that if $s \in \Sigma(x)$ then $\varphi_x(s) = \hat{\varphi}_x(s) = [P(e) : e \in s]$, where $P(w_i, v_i) = \varepsilon B_{h(w_i)} + (1 - \varepsilon) B_{h(\bar{v}_i)}$, and φ_s is the identity map.

In the remainder of this section we assume (X, f) is restricted and use the triangulations φ_x, $x \in X$, to construct a triangulation $\varphi : K \to A$, where $A = A(X)$ is the topological space of X constructed in Section 4. Recall from Section 4 that we have a map

$$\lambda : \coprod_{x\in X} F(x) \to A$$

$$a \mapsto \tilde{a}$$

For $S \in \Sigma$ define

$$\varphi(S) = \lambda(\varphi_{\hat{\zeta}(S)}(s(S)))$$

define

$$\hat{\varphi}(S) = \varphi_{\hat{\zeta}(S)}(s(S))$$

with $u(S, v_i) = u(S(s), (w_i, v_i)) = P(w_i, v_i)$ and define $\varphi_S : \varphi(S) \to \hat{\varphi}(S)$ by $\varphi_S = \lambda_S^{-1}$, where $\lambda_S : \varphi_{\hat{\zeta}(S)}(s(S)) \to \varphi(S)$ is the restriction of λ to $\varphi_{\hat{\zeta}(S)}(s(S))$.

(9.2) *For $S \in \Sigma$, $x = \hat{\zeta}(S)$, and $s = s(S)$, $\lambda_S : \varphi_x(s) \to \varphi(S)$ is a homeomorphism and $\varphi(S)$ is closed in A.*

Proof. Let $s = \{(w_i, v_i) : 0 \leq i \leq k\}$, $W = \{w_i : i\}$, $D = \varphi_x(s)$, and $D_w = I(w) \cap D$ for $w \in W$. Then $D = [P(w_i, v_i) : i]$, so by 6.3 each $d \in D$ can be written uniquely in the form $d = \sum_i d_i P(w_i, v_i)$ with $0 \leq d_i$ and $\sum_i d_i = 1$. By 7.4, $P(w_i, v_i) \in I(w_i) \cap F(w_j)$ for $j \geq i$, so

$$D_w = \{d \in D : d_j = 0 \text{ for } j > i(w) \text{ and } d_i \neq 0 \text{ for some } i \in \mathcal{I}(w)\}$$

where $\mathcal{I}(w) = \{i : w_i = w\}$ and $i(w) = \max\{i : i \in \mathcal{I}(w)\}$. In particular the D_w partition D. Next $F_w : D_w \to F(\zeta(w))$ is a homeomorphism and by 4.4, $\lambda : I(\zeta(w)) \to A$ is an injection, so as $\lambda \circ F_w = \lambda$ on $F(w)$, $\lambda : D_w \to A$ is an injection. Further by 4.5, the sets $\tilde{I}(y)$, $y \in X$, partition A, so as $w \mapsto \zeta(w)$ is an injection on W, $\tilde{I}(w) \cap \tilde{I}(u) = \varnothing$ for distinct $u, w \in W$, and hence $\lambda : D \to A$ is an injection. Thus $\lambda_S : D \to \varphi(S)$ is bijective and continuous.

Finally D is closed in $F(x)$ and $C \subseteq \tilde{F}(x)$ is closed in A if and only if $\lambda^{-1}(C)$ is closed in $F(x)$. Now as $\lambda : D \to \varphi(S)$ is a bijection, if $E \subseteq D$ is closed then $E = \lambda^{-1}(\lambda(E))$ is closed in $F(x)$, so $\lambda(E)$ is closed in A. Thus λ_S is a homeomorphism and $\varphi(S) = \lambda(S)$ is closed in A.

(9.3)*Let $T \in \Sigma$ and $S \subseteq T$. Then $\varphi(S) \subseteq \varphi(T)$ and $\varphi_{S,T} : \sum_v a_v u(S, v) \to \sum_v a_v u(T, v)$.*

Proof. Let $x = \hat{\zeta}(T)$, $t = s(T) = \{v_0, \ldots, v_k\}$, $s = s(S)$, $r = \eta^{-1}(S) \subseteq t$, $l = \max\{i : v_i \in r\}$, $w_i = \hat{f}_{v_i}(v_k)$, $w = w_l$, and $y = \zeta(w)$.

Then $u(T,v_i) = P(w_i,v_i) = \varepsilon B(w_i) + (1-\varepsilon)B(f_{w_i}^{-1}(v_i))$, so for $i \leq l$,

$$F_w(P(w_i,v_i)) = \varepsilon B(f_w(w_i)) + (1-\varepsilon)B(v_i) = P(f_w(w_i),v_i)$$

by 7.2. Then

$$F_w(\sum_i a_i u(T,v_i)) = F_w(\sum_i a_i P(w_i,v_i)) = \sum_i a_i P(f_w(w_i),v_i)$$
$$= \sum_i a_i u(S,v_i)$$

by another application of 7.2. In particular $F_w(\varphi_x(r)) = \varphi_y(s)$. Also $\lambda_S \circ F_w = \lambda_T$ on $\varphi_x(r)$ as $F_w(a) \sim a$ for each $a \in F(w)$. So $\varphi(S) = \lambda_S(\varphi_y(s)) = \lambda_S(F_w(\varphi_x(r))) = \lambda_T(\varphi_x(r)) \subseteq \varphi(T)$. Also

$$\begin{aligned}\varphi_{S,T}(\sum_i a_i u(S,v_i)) &= (\varphi_T \circ \varphi_S^{-1})(\sum_i a_i u(S,v_i)) \\ &= \varphi_T(\lambda_S(F_w(\sum_i a_i u(T,v_i)))) \\ &= \varphi_T(\lambda_T(\sum_i a_i u(T,v_i))) = \sum_i a_i u(T,v_i),\end{aligned}$$

completing the proof.

(9.4) $\varphi(S) \cap \varphi(T) = \varphi(S \cap T)$ *for all* $S,T \in \Sigma$.

Proof. By 9.3, $\varphi(S\cap T) \subseteq \varphi(S) \cap \varphi(T)$, so it remains to show that if $e \in \varphi(S)\cap\varphi(T)$ then $e \in \varphi(S\cap T)$. Then $e = \tilde{e}_S = \tilde{e}_T$ for some $e_R \in \varphi_{\hat{\zeta}(R)}(s(R))$. Then $x = y(e_T) = y(\tilde{e}_T) = y(e) = y(e_S)$ in the notation of Section 4; that is $e \in \tilde{I}(x)$ and $e_R \in I(w_R)$ with $w_R \in f(\hat{\zeta}(R))$ and $\zeta(w_R) = x$. For $R = S,T$, let $\bar{R} = \{v \in R : \hat{\zeta}(v) \leq x\}$. Then as we saw during the proof of the previous lemma, $F_{w_R}(e_R) \in \varphi_x(s(\bar{R}))$, so replacing S,T by $\bar{S},\bar{T}$, we may assume $x = \hat{\zeta}(S) = \hat{\zeta}(T)$.

Now $\varphi_x(s(R)) \subseteq F(x)$ and $e_R \in I(x)$. Then as $\lambda : I(x) \to A$ is injective, $e_S = e_T \in \varphi_x(s(T))\cap\varphi_x(s(S)) = \varphi_x(s(S)\cap s(T))$ by Theorem 7.6. Also $e_T \in I(x)$, so $s(S)\cap s(T) = s(S\cap T)$ with $\hat{\zeta}(S\cap T) = x$, so $e = \lambda(e_S) \in \lambda(\varphi_x(s(S\cap T))) = \varphi(S\cap T)$, completing the proof.

Theorem 9.5. *If* (X,f,F,B) *is a restricted polyhedral cell complex then* $\varphi : K \to A$ *is a triangulation.*

Proof. By 9.2, for each $S \in \Sigma$, $\varphi(S)$ is a closed subspace of A and

$$\lambda_S : \varphi_{\hat{\zeta}(S)}(s(S)) \to \varphi(S)$$

is a homeomorphism. Thus

$$\varphi_S = \lambda_S^{-1} : \varphi(S) \to \hat{\varphi}(S) = \varphi_{\hat{\zeta}(S)}(s(S))$$

is a homeomorphism. Axiom (T3) for triangulations holds by 9.3 and axiom (T1) holds by 9.4.

As $\varphi_x : K(x) \to F(x)$ is a triangulation, $F(x) = \bigcup_{s\in\Sigma(x)} \varphi_x(s)$ and $C \subseteq F(x)$ is closed if and only $C \cap \varphi_x(s)$ is closed in $\varphi_x(s)$ for all $s \in \Sigma(x)$. Also by 4.5, $A = \bigcup_x \tilde{F}(x)$, so as $\varphi(S) = \tilde{\varphi}_{\hat{\zeta}(S)}(s(S))$, $A = \bigcup_{S\in\Sigma} \varphi(S)$. Further by definition of the topology on A, $D \subseteq A$ is closed in A if and only if $D_x = A \cap \tilde{F}(x)$ is closed in $\tilde{F}(x)$ for all $x \in X$. Then as $\tilde{F}(x) = \bigcup_{s\in\Sigma(s)} \varphi(\eta(s))$, if $D_x \cap \varphi(\eta(s))$ is closed in $\varphi(\eta(s))$ for all $s \in \Sigma(x)$ then $\lambda_x^{-1}(D_x \cap \varphi(\eta(s)))$ is closed in $\varphi_x(s)$, so as φ_x is a triangulation, $\lambda_x^{-1}(D_x) = \bigcup_s \lambda_x^{-1}(D_x \cap \varphi(\eta(s)))$ is closed in $F(x)$, so D_x is closed in $\tilde{F}(x)$. Thus D is closed in A if and only if $D \cap \varphi(S)$ is closed in $\varphi(S)$ for all $S \in \Sigma$. That is axiom (T2) is satisfied.

Corollary 9.6. *The homology and fundamental group of the space $A(X)$ of a restricted polyhedral cell complex X are isomorphic to the simplicial homology and fundamental group of $K(X)$.*

Remark 9.7. We can also triangulate the space A of a polyhedral cell complex (X, f, F, B) which is not restricted, but not by the triangulating complex $K(X)$. Instead consider the set $\hat{V}(X)$ of all pairs (S, ω), where $S = \{v_0, \dots, v_k\}$ is a simplex of $K(X)$ with the standard ordering, and $\omega : S \to f(\hat{\zeta}(v_k))$ with $\omega(S)$ a simplex in $f(\hat{\zeta}(v_k))$ and $w_i = \omega(v_i)$ satisfying $\zeta(w_i) = \hat{\zeta}(v_i)$, $f_{w_j}(w_i) \geq v_j$, and $f_{f_{w_i}(w_i)}(v_j) \geq v_i$ for $j \geq i$. Partially order $\hat{V}(X)$ by $(S, \omega) \geq (T, \theta)$ if $T \subseteq S$ and $\theta(v) = f_w(\omega(v))$ for each $v \in T$ and $w = \max(\omega(T))$. Finally let $\hat{K}(X)$ be the order complex of the poset $\hat{V}(X)$.

Observe we have a map of posets from $\hat{V}(X)$ into $\mathrm{sd}(K(X))$ defined by $(S, \omega) \mapsto S$, and this map is an isomorphism of $\hat{K}(X)$ with $\mathrm{sd}(K(X))$ when X is restricted.

For $x \in X$, Theorem 7.6 supplies a triangulation $\varphi_x : K(x) \to F(x)$. For s a simplex of $K(x)$, let $P(s)$ be the barycenter of $\varphi(s)$. A simplex σ of $\mathrm{sd}(K(x))$ is a chain $\{s_0 \subset \cdots \subset s_k\}$ of simplices of $K(x)$ and we have the barycentric subdivision $\psi_x : \Sigma(\mathrm{sd}(K(x))) \to F(x)$ of the triangulation φ_x, which is also a triangulation of $F(x)$, and is defined by $\psi_x(\sigma) = [P(s_i) : 0 \leq i \leq k]$.

We can use the triangulation ψ_x in place of φ_x and argue as in this section to construct a triangulation $\psi : \hat{K} \to A$. Namely suppose $\sigma =$

$\{s_0 \subset \cdots \subset s_r\}$ is a simplex in $\mathrm{sd}(K(x))$ with $s_i = \{(w_0^i, v_0^i), \ldots, (w_{k_i}^i, v_{k_i}^i)\} \in \Sigma(x)$ ordered as in 7.5 and with $w_{k_r}^r = x$. Define $\eta(\sigma) = \{(S_i, \omega_i) : 0 \leq i \leq r\}$ a simplex of $\hat{K}(X)$ by

$$S_i = \{f_{w_j^i}(v_j^i) : 0 \leq j \leq k_i\},$$

$$\omega_i(f_{w_j^i}(v_j^i)) = f_{w_{k_i}^i}(w_j^i), \quad 0 \leq j \leq k_i.$$

The map η plays the role that the map η defined at the start of this section played for restricted complexes.

Conversely given $\sigma = \{(S_i, \omega_i) : 0 \leq i \leq r\}$ a simplex in $\hat{K}(X)$ with $S_i = \{v_0^i, \ldots, v_{k_i}^i\}$ in the standard ordering and $x = \hat{\zeta}(v_{k_r}^r) = \hat{\zeta}(\sigma)$, define $\nu(\sigma) = \{s_0, \ldots, s_r\}$ by $s_i = \{(w_0^i, \hat{v}_0^i), \ldots, (w_{k_i}^i, \hat{v}_{k_i}^i)\}$, where $w_j^i = \omega_k(v_j^i)$ and $\hat{v}_j^i = f_{w_j^i}^{-1}(v_j^i)$. The simplex $\nu(\sigma)$ plays the role of the simplex $s(S)$. Thus we define

$$\psi(\sigma) = \lambda(\psi_{\hat{\zeta}(\sigma)}(\nu(\sigma))),$$

$$\hat{\psi}(\sigma) = \psi_{\hat{\zeta}(\sigma)}(\nu(\sigma)),$$

for σ a simplex of $\hat{K}(X)$. We can now repeat the proofs of Lemmas 9.1 through 9.5 with some small variation, to establish the analogous statements for general combinatorial cell complexes and the triangulation $\psi : \hat{K}(X) \to A$.

§10. The triangulation functor

By Theorem 9.5, if $X = (X, f, F, B)$ is a restricted polyhedral cell complex then there exists a triangulation $\xi^X : K(X) \to A(X)$. We seek to extend ξ to a functor from the category of restricted polyhedral cell complexes to the category of triangulated topological spaces. Our triangulation ξ^X depends on a choice of real ε, $0 < \varepsilon < 1$. Fix some choice of ε and use it to define ξ^X for all choices of X. Further given a morphism $\alpha : X \to Y$ of poyhedral cell complexes, define $\xi^\alpha : \xi^X \to \xi^Y$ by $\xi^\alpha = (K(\alpha), A(\alpha))$, where A is the functor of 4.7 and K is the functor of 5.7. We prove the following two results at the same time:

(10.1) *ξ is a covariant functor from the category of restricted polyhedral cell complexes to the category of triangulated topological spaces.*

(10.2) *Let $\alpha : X \to \bar{X}$ be a morphism of restricted polyhedral cell complexes. Then*

(1) *For $x \in X$ and σ a simplex in $f(x)$,*

$$F(x) = \{\sum_{v\in\sigma} a_v B(v) : 0 \leq a_v \in \mathbf{R} \text{ and } \sum_v a_v = 1\}$$

and

$$\alpha_x\big(\sum_{v\in\sigma} a_v B(v)\big) = \sum_{v\in\sigma} a_v \bar{B}(\alpha_x(v))$$

(2) *For $S = \{v_0, \ldots, v_k\}$ a simplex in $K(X)$ ordered as in Lemma 9.1 and $x = \hat{\zeta}(S)$,*

$$\xi^X(S) = \{\sum_i a_i \tilde{P}(w_i, v_i) : 0 \leq a_i \in \mathbf{R} \text{ and } \sum_i a_i = 1\}$$

where $w_i = \hat{f}_{v_i}(v_k) \in f(x)$, and

$$A(\alpha)\big(\sum_i a_i \tilde{P}(w_i, v_i)\big) = \sum_i a_i \tilde{P}(\bar{w}_i, \bar{v}_i)$$

with $K(\alpha)(S) = \{\bar{v}_i : 0 \leq i \leq k\}$, $\bar{v}_i = \alpha_{\hat{\zeta}(v_i)}(v_i)$, and $\bar{w}_i = \alpha_x(w_i) = \hat{f}_{\bar{v}_i}(\bar{v}_k)$.

By Theorem 9.5, $\xi^X : K(X) \to A(X)$ and $\xi^{\bar{X}} : K(\bar{X}) \to A(\bar{X})$ are triangulations. By Remark 5.7, $K(\alpha) : K(X) \to K(\bar{X})$ is a simplicial map, while by 4.7, $A(\alpha) : A(X) \to A(\bar{X})$ is a continuous map. So to prove 10.1, we must show $\xi^{\alpha\circ\beta} = \xi^\alpha \circ \xi^\beta$ and for each simplex S of $K(X)$

(T_1) $A(\alpha)(\xi^X(S)) \subseteq \xi^{\bar{X}}(K(\alpha)(S))$, and

(T_2) $\alpha_S \circ \xi^X_S = \xi^{\bar{X}}_{K(\alpha)(S)} \circ A(\alpha)$, where

$$\alpha_S : \sum_{v\in S} a_v u(S, v) \mapsto \sum_{v\in S} a_v u(K(\alpha)(S), K(\alpha)(v))$$

The first remark follows from the fact that A and K are functors. The first statements in (1) and (2) of 10.2 follow from the definition of $F(x)$ and $\xi^X(S)$, respectively. Moreover if S is as in 10.2, by definition of ξ^X_S, $u(S, v_i) = P(w_i, v_i)$ and

$$\xi^X_S : \sum_i a_i \tilde{P}(w_i, v_i) \mapsto \sum_i a_i P(w_i, v_i)$$

Therefore 10.2.2 implies (T_1) and (T_2), so it remains to prove 10.2.

As α is a morphism of polyhedral complexes, $\alpha_x(B^x(\sigma)) \subseteq \bar{B}^{\alpha(x)}(\alpha(\sigma))$ and $\alpha_\sigma \circ B^x_\sigma = \bar{B}^{\alpha(x)}_{\alpha(\sigma)} \circ \alpha_x$ on $B^x(\sigma)$ for each $x \in X$ and each simplex σ

of $f(x)$. In particular in the notation of Section 7, $\alpha_x(B(v)) = \bar{B}(\alpha_x(v))$ for each $v \in f(x)$. Also for $\sigma = \{y_0, \ldots, y_m\}$, by definition

$$\sum_i a_i B(y_i) = B_\sigma^{-1}\big(\sum_i a_i u_{h(y_i)}\big)$$

so as $\alpha_\sigma \circ B_\sigma^x = \bar{B}_{\alpha(\sigma)}^{\alpha(x)} \circ \alpha_x$ on $B^x(\sigma)$, we have

$$\alpha_x\big(\sum_i a_i B(y_i)\big) = \alpha_x(B_\sigma^{-1}\big(\sum_i a_i u_{h(y_i)}\big)) = \bar{B}_{\alpha(\sigma)}^{\alpha(x)}(\alpha_\sigma\big(\sum_i a_i u(\sigma, y_i)\big))$$
$$= \bar{B}_{\alpha(\sigma)}^{\alpha(x)}\big(\sum_i a_i u(K(\alpha)(\sigma), K(\alpha)(y_i))\big) = \sum_i a_i \bar{B}(\alpha_x(y_i))$$

so that 10.2.1 is established.

Recall the definition of $V(x)$ and $P(w,v)$ from Section 7, and observe that as α is a morphism of cell complexes, for $(w,v) \in V(x)$, $(\alpha_x(w), \alpha_{\zeta(w)}(v)) \in V(\alpha(x))$ and $\alpha_x(B(f_w^{-1}(v))) = \bar{B}(\alpha_x(f_w^{-1}(v))) = \bar{B}(\bar{f}_{\alpha_x(w)}^{-1}(\alpha_x(v)))$. Therefore

$$\alpha_x(P(w,v)) = \alpha_x(\varepsilon B(w) + (1-\varepsilon)B(f_w^{-1}(v)))$$
$$= \varepsilon \bar{B}(\alpha_x(w)) + (1-\varepsilon)\bar{B}(\bar{f}_{\alpha_x(w)}^{-1}(\alpha_x(v))) = \bar{P}(\alpha_x(w), \alpha_{x(v)}(v))$$

by an earlier remark. Therefore

$$A(\alpha)(\tilde{P}(w,v)) = \tilde{\alpha}_x(P(w,v)) = \tilde{P}(\alpha_x(w), \alpha_{x(v)}(v)).$$

Let $S = \{v_0, \ldots, v_k\}$ be a simplex of $K(X)$ ordered as in 9.1, and $x = \hat{\zeta}(S)$. Then from the definition of ξ^X in Section 9 and 9.1,

$$\xi^X(S) = \lambda(\xi_x(s(S))) = [\tilde{P}(w_i, v_i) : 0 \le i \le k]$$

where $s(S) = \{(w_i, v_i) : 0 \le i \le k\}$ and $w_i = \hat{f}_{v_i}(v_k)$. Therefore

$$\alpha_x(s(S)) = \{(\bar{w}_i, \bar{v}_i) : 0 \le i \le k\} = s(K(\alpha)(S))$$

where $\bar{v}_i = \alpha_{x(v_i)}(v_i)$ and $\bar{w}_i = \alpha_x(w_i) = \hat{f}_{\bar{v}_i}(\bar{v}_k)$. Finally

$$A(\alpha)\big(\sum_i a_i \tilde{P}(w_i, v_i)\big) = \tilde{\alpha}_x\big(\sum_i a_i P(w_i, v_i)\big) = \sum_i a_i \tilde{\alpha}_x(P(w_i, v_i))$$

by 10.2.1, and then by an earlier observation this is equal to $\sum_i a_i \tilde{P}(\bar{w}_i, v_i)$, completing our proof.

We can now use the functor ξ to construct the triangulation functor from the category of restricted combinatorial cell complexes to the category of triangulated topological spaces. Namely we define the *triangulation functor* to be the covariant functor $T = \xi \circ \mathcal{P}$, where $\mathcal{P}$ is the functor from the category of combinatorial cell complexes to the category of polyhedral cell complexes constructed in Example 7.1. As the composition of covariant functors, T is a covariant functor. Given a combinatorial cell complex $X = (X, f)$, we write $T(X)$ for the topological space $A(\mathcal{P}(X))$ and when X is restricted we write ξ^X for the triangulation $\xi^{\mathcal{P}(X)} : K(X) \to T(X)$. If $\alpha : X \to \bar{X}$ is a morphism of combinatorial cell complexes, we write $T(\alpha)$ for the morphism $\xi^{\mathcal{P}(\alpha)} = (K(\alpha), A(\mathcal{P}(\alpha)))$.

We define $T(X)$ to be the *geometric realization* of the combinatorial cell complex X.

(10.3) *Let $\gamma : K^1 \to K^2$ be a morphism of typed simplicial complexes over I and $\varphi^i : K^i \to T^i$ be triangulations. Then*

(1) *γ extends to a morphism $(\gamma, \beta^\gamma(\varphi^1, \varphi^2)) : \varphi^1 \to \varphi^2$ of triangulated topological spaces.*

(2) *If $\delta : K^2 \to K^3$ is a morphism of typed simplicial complexes over I and $\varphi^3 : K^3 \to T^3$ is a triangulation then $\beta^{\delta\circ\gamma}(\varphi^1, \varphi^3) = \beta^\delta(\varphi^2, \varphi^3) \circ \beta^\gamma(\varphi^1, \varphi^2)$.*

(3) $\beta^{id_{K^1}}(\varphi^1, \varphi^1) = id_{T^1}$.

(4) *If γ is an isomorphism then so is $(\gamma, \beta^\gamma(\gamma^1, \varphi^2))$.*

Proof. For $s \in \Sigma^1$ define

$$\gamma_s : \hat{\varphi}^1(s) \to \hat{\varphi}^2(\gamma(s))$$
$$\sum_{v\in s} a_v u^1(s, v) \mapsto \sum_{v\in s} a_v u^2(\gamma(s), \gamma(v)).$$

Observe that γ_s is continuous. Now define $\beta_s : \varphi^1(s) \to \varphi^2(\gamma(s))$ by

$$\beta_s = (\varphi^2_{\gamma(s)})^{-1} \circ \gamma_s \circ \varphi^1_s$$

and then define

$$\beta = \bigcup_{s\in\Sigma^1} \beta_s.$$

That is for $a \in \varphi^1(s)$, $\beta(a) = \beta_s(a)$.

We first check that β is well defined. To begin, if $t \subseteq s$ then by axiom (T3) for triangulations, $\varphi^i_{t,s} = \hat{\varphi}^i_{t,s}$ as maps from $\hat{\varphi}^i(t)$ to $\hat{\varphi}^i(s)$, where $\varphi^i_{t,s} = \varphi^i_s \circ (\varphi^i_t)^{-1}$ and

$$\hat{\varphi}^i_{t,s} : \sum_{v\in t} a_v u^1(t, v) \mapsto \sum_{v\in t} a_v u^1(s, v).$$

Also by definition of γ_t and γ_s,

$$\hat{\varphi}^2_{\gamma(t),\gamma(s)} \circ \gamma_t = \gamma_s \circ \hat{\varphi}^1_{t,s}$$

as maps from $\hat{\varphi}^1(t)$ to $\hat{\varphi}^2(\gamma(s))$. Therefore the diagram

$$\begin{array}{ccccccc}
\varphi^1(t) & \xrightarrow{\varphi^1_t} & \hat{\varphi}^1(t) & \xrightarrow{\gamma_t} & \hat{\varphi}^2(\gamma(t)) & \xrightarrow{(\varphi^2_{\gamma(t)})^{-1}} & \varphi^2(\gamma(t)) \\
\iota^1\downarrow & & \hat{\varphi}^1_{t,s}\downarrow & & \hat{\varphi}^2_{\gamma(t),\gamma(s)}\downarrow & & \downarrow\iota^2 \\
\varphi^1(s) & \xrightarrow{\varphi^1_s} & \hat{\varphi}^1(s) & \xrightarrow{\gamma_s} & \hat{\varphi}^2(\gamma(s)) & \xrightarrow{(\varphi^2_{\gamma(s)})^{-1}} & \varphi^2(\gamma(s))
\end{array}$$

commutes, so that $\beta_s = \beta_t$ on $\varphi^1(t)$.

Further by axiom (T1), if $r, s \in \Sigma^1$ then $\varphi^1(r) \cap \varphi^1(s) = \varphi^1(r \cap s)$. So setting $t = r \cap s$, we have $\beta_r = \beta_t = \beta_s$ on $\varphi^1(r) \cap \varphi^1(s)$, completing the proof that β is well defined.

As γ_s, φ^1_s, and $(\varphi^2_{\gamma(s)})^{-1}$ are continuous, so is β_s. Then by axiom (T2), β is continuous. That is if C is closed in T^2 then $C \cap \varphi^2(\gamma(s))$ is closed in $\varphi^2(\gamma(s))$ for each $s \in \Sigma^1$. Therefore $\beta_s^{-1}(C \cap \varphi^2(\gamma(s))) = \beta^{-1}(C) \cap \varphi^1(s)$ is closed in $\varphi^1(s)$. Hence $\beta^{-1}(C)$ is closed in T^1.

By definition of β, $\beta(\varphi^1(s)) = \beta_s(\varphi^1(s)) \subseteq \varphi^2(\gamma(s))$ and $\varphi^1_s \circ \gamma_s = \varphi^2_{\gamma(s)} \circ \beta_s = \varphi^2_{\gamma(s)} \circ \beta$. Therefore $(\gamma, \beta) : \varphi^1 \to \varphi^2$ is a morphism of triangulated topological spaces, establishing (1).

Assume the hypotheses of (2). Then

$$\beta^\gamma = \bigcup_{s \in \Sigma^1} \beta^\gamma_s, \quad \beta^\delta = \bigcup_{t \in \Sigma^2} \beta^\delta_t, \quad \beta^{\delta\circ\gamma} = \bigcup_{s \in \Sigma^1} \beta^{\delta\circ\gamma}_s.$$

Further $(\delta \circ \gamma)_s = \delta_{\gamma(s)} \circ \gamma_s$, so

$$\beta^{\delta\circ\gamma}_s = (\varphi^3_{(\delta\circ\gamma)(s)})^{-1} \circ (\delta \circ \gamma)_s \circ \varphi^1_s$$

$$= ((\varphi^3_{\delta(\gamma(s))})^{-1} \circ \delta_{\gamma(s)} \circ \varphi^2_{\gamma(s)}) \circ ((\varphi^2_{\gamma(s)})^{-1} \circ \gamma_s \circ \varphi^1_s) = \beta^\delta_{\gamma(s)} \circ \beta^\gamma_s$$

and therefore $\beta^{\delta\circ\gamma} = \beta^\delta \circ \beta^\gamma$, establishing (2).

Part (3) follows as if $\varphi^1 = \varphi^2$ and $\gamma = id_K$ then $\gamma_s = id_{\hat{\varphi}^1(s)}$, so

$$\beta_s = (\varphi^1_s)^{-1} \circ \gamma_s \circ \varphi^1_s = id_{\varphi^1(s)}.$$

Finally (2) and (3) imply that

$$\beta^{\gamma^{-1}}(\varphi^2, \varphi^1) \circ \beta^\gamma(\varphi^1, \varphi^2) = \beta^{id_K}(\varphi^1, \varphi^1) = id_{T^1}$$

so $\beta^{\gamma^{-1}}(\varphi^2, \varphi^1) = \beta^\gamma(\varphi^1, \varphi^2)^{-1}$, and hence $(\gamma, \beta^\gamma(\varphi^1, \varphi^2))$ is an isomorphism.

(10.4) *Let K be a typed simplicial complex. Then each triangulation of K is isomorphic in the category of triagulated topological spaces to the geometric realization $\varphi^K : K \to T(K)$ of K.*

Proof. Apply 10.3.4 to $K^1 = K^2 = K$, $\gamma = id_K$, $\varphi^1 = \varphi^K$, and any triangulation $\varphi^2 : K \to T^2$ of K.

Remark 10.5. It is well known that if T is the geometric realization of a simplicial complex K, then the singular homology $H_*(T)$ is isomorphic to the simplicial homology $H_*(K)$ and the fundamental group $\pi_1(T)$ is isomorphic to the fundamental group $\pi_1(K)$ of K.

As recalled in Remark 10.5, the homology and fundamental group of the geometric realization of a simplicial complex can be defined in a purely combinatorially manner in terms of the simplicial complex. We seek to do the same for combinatorial cell complexes. We know that if X is a restricted combinatorial cell complex, then $\xi^X : K(X) \to T(X)$ is a triangulation of the geometric realization $T(X)$ of X, so by Remark 10.5, $H_*(T(X)) \cong H_*(K(X))$ and $\pi_1(T(X)) \cong \pi_1(K(X))$. Thus we define $H_*(X) = H_*(K(X))$ and $\pi_1(X) = \pi_1(K(X))$. Thus we have our combinatorial definition of the homology and fundamental group of a restricted combinatorial cell complex, and from the discussion above we have:

Theorem 10.6. *The triangulation functor T is a covariant functor from the category of restricted combinatorial cell complexes to the category of triangulated topological spaces, which assigns to a restricted combinatorial cell complex (X, f) its geometric realization $T(X)$ and the triangulation $\xi^X : K(X) \to T(X)$ of the geometric realization by the triangulating complex of X. Moreover $H_*(T(X)) \cong H_*(X)$ and $\pi_1(T(X)) \cong \pi_1(X)$.*

(10.7) *Let (X, f) be the simplicial cell complex of the poset X. Then the geometric realization of the cell complex (X, f) is homeomorphic to the geometric realization of the order complex $\mathcal{O}(X)$ of the poset X.*

Proof. Let $\mathcal{P}(X, f) = (X, f, F, B)$ and $A = T(X, f)$ the topological space of this polyhedral cell complex. Thus A is the geometric realization of the simplicial cell complex (X, f). We show there exists a triangulation $\varphi : \mathcal{O}(X) \to A$. Then by 10.4, A is homeomorphic to the geometric realization of $\mathcal{O}(X)$.

The triangulation is defined by $\varphi(s) = \tilde{B}^z(s)$, $\hat{\varphi}(x) = \hat{B}^z(s)$, and $\varphi_s = B^z_s \circ \lambda_z^{-1}$, where z is the greatest element of s. As (X, f) is the simplicial cell complex for the poset X, ζ is injective, so by 4.6.3, $\lambda_x : F(x) \to \tilde{F}(x)$ is a homeomorphism and hence φ_s makes sense. Check that if $x \geq z$ then $\tilde{B}_x(s) = \tilde{B}_z(s)$, $\hat{B}^x(s) = \hat{B}^z(s)$, and $B^x_s \circ \lambda_x^{-1} = B^z_s \circ \lambda_z^{-1}$. Then use this fact to prove φ is a triangulation. We leave the details to the reader.

(10.8) *Assume $X = (X, f)$ is a regular combinatorial cell complex. Let $\mathcal{O}(X)$ be the order complex of the poset X. Then $H_*(\mathcal{O}(X)) \cong H_*(X)$ and $\pi_1(\mathcal{O}(X)) \cong \pi_1(X)$.*

Proof. By 3.1, X is isomorphic to the simpicial cell complex Y of the poset X. Thus $H_*(X) \cong H_*(Y)$ and $\pi_1(X) \cong \pi_1(Y)$. But by 10.7, the geometric realizations A of Y and T of $\mathcal{O}(X)$ are homeomorhic, so $H_*(Y) \cong H_*(A) \cong H_*(T) \cong H_*(\mathcal{O}(X))$. Similarly the fundamental groups are isomorphic.

§11. Homology in $K(X)$

Let (X, f) be a combinatorial cell complex, $\Delta = \Delta(X)$ the graph of X, and $K = K(X)$ the triangulating complex of X. Define

$$V^n = \{v \in V : h(\hat{\zeta}(v)) \leq n\}, \quad V^n_n = \{v \in V : h(\hat{\zeta}(v)) = h(\zeta(v)) = n\}$$

and $V_n = V^n - V^n_n$. Denote by Δ^n and Δ_n the graphs on V^n and V_n induced by Δ, respectively. Let $K^n = K(\Delta^n)$ and $K_n = K(\Delta_n)$ be the clique complexes of Δ^n and Δ_n, respectively. Define the n-skeleton of X to be the combinatorial cell complex $X^n = (X^n, f_{|X^n})$, where $X^n = \{x \in X : h(x) \leq n\}$. Thus $K^n = K(X^n)$ is the triangulating complex of the n-skelton of X.

By 5.6, K is a typed simplicial complex over $I \times I$, $I = \{0, 1, \ldots\}$, with type function $\tau(v) = (h(\hat{\zeta}(v)), h(\zeta(v)))$. Order $I \times I$ lexiographically; that is $(a, b) < (i, j)$ if $a < i$ or $a = i$ and $b < j$. We use this ordering to define our boundary operator on the simplicial chain complex $C_*(K)$ of K as in Section 1. We also adopt the notional conventions established in that section.

By 5.4, if s is a k-simplex of K then there is a unique ordering $v_0, \ldots, v_k$ of s such that $v_i \in L(v_{i+1})$ for each i. By 5.2, $\tau(v_0) \leq \cdots \leq \tau(v_k)$ also, so this is the ordering of s used to define the oriented simplex $s = v_0 \cdots v_k = v_0 \wedge \cdots \wedge v_k$ via the conventions of Section 1.

Example. Consider the simplicial cell complex $X(n)$ of the n-simplex, defined in Example (2) in Section 3. Recall from Section 5 that

$$V(X(n)) = V(n) = \{(\alpha,\beta) \in X(n) \times X(n) : \alpha \geq \beta\},$$

and for $(\alpha,\beta) \in V(n)$,

$$L(\alpha,\beta) = \{(a,b) \in V(n) : b \leq \beta \leq a \leq \alpha\}$$

Recall we write $K(n)$ for the triangulating complex $K(X(n))$ of the cell complex $X(n)$. Then $K(n)$ is a typed complex over $V(n)$ with type function the identity map. As $V(n)$ is ordered lexiographically, we have $(a,b) < (\alpha,\beta)$ if $a < \alpha$ or $a = \alpha$ and $b < \beta$. Notice $K(n)^m = \{(a,b) \in V(n) : a \leq m\} \cong K(m)$, $V(n)^m_m = \{(m,m)\}$, and $K(n)_m = K(n)^m - \{(m,m)\}$. By 5.9:

(11.1) *Let $s = \{(\alpha_i,\beta_i) : 0 \leq i \leq k\}$ be a subset of $V(n)$ with $(\alpha_i,\beta_i) \leq (\alpha_{i+1},\beta_{i+1})$ for each i. Then s is in $\Sigma(n) = \Sigma(K(n))$ if and only if $\beta_0 \leq \cdots \leq \beta_k \leq \alpha_0 \leq \cdots \leq \alpha_k$.*

Observe next that

(11.2) *For $z \in V^n_n$, $\mathrm{Link}_{K^n}(z) \cong \mathcal{O}(\dot{f}(\hat{\zeta}(z)))$.*

Indeed $\mathrm{Link}_{K^n}(z) = \{u \in V_n : \hat{\zeta}(u) = \hat{\zeta}(z)\}$ and the identity map $\mathrm{Link}_{K^n}(z) \to \mathcal{O}(\dot{f}(\hat{\zeta}(z)))$ is an isomorphism. Observe also that under the notational conventions of Section 1, and by 1.1:

(11.3) *If $z \in V^n_n$ and*

$$\alpha = \sum_{s \in \Sigma^{k-1}(\mathrm{Link}_{K^n}(z))} a_s s \in C_{k-1}(\mathrm{Link}_{K^n}(z))$$

then

$$\alpha z = \sum_{s \in \Sigma^{k-1}(\mathrm{Link}_{K^n}(z))} a_s s z \in C_k(K^n)$$

where if $s = v_0 \cdots v_{k-1}$ then $sz = v_0 \cdots v_{k-1} z$. Further $\partial(\alpha z) = \partial(\alpha)z + (-1)^k \alpha$.

(11.4) *Let X be restricted of height at least $n+1$ and $\iota : K^n \to K_{n+1}$ the inclusion map. Then*

(1) *ι is a homotopy equivalence.*

(2) *$\iota_* : H_*(K^n) \to H_*(K_{n+1})$ is an isomorphism.*

Proof. We show for all simplices s of K_{n+1}, $\iota^{-1}(st_{K_{n+1}}(s))$ is contractible. Then by Theorem 1 in [3], ι is a homotopy equivalence.

Let $s = \{v_0, \ldots, v_k\}$ be a simplex in K_{n+1} ordered as in Lemma 5.4. If $v_0 \in K^n$ then $v_0 \in \iota^{-1}(st_{K_{n+1}}(s)) \subseteq st_{K^n}(v_0)$ and hence $\iota^{-1}(st_{K_{n+1}}(s))$ is contractible.

So assume $x = \hat{\zeta}(v_0)$ is of height $n+1$. Let $W = \{w \in f(x) : v_k \leq w < x\}$ and for $w \in W$ let $S(w) = f(\zeta(w))(\leq f_w(v_0))$, $T(w) = \bigcup_{u \in W(\leq w)} S(u)$, and for $U \subseteq W$ let $T(U) = \bigcup_{u \in U} T(u)$.

Observe that $\iota^{-1}(st_{K_{n+1}}(s)) = T(W)$. For if $a \in \iota^{-1}(st_{K_{n+1}}(s))$ then $f_{w_i}(v_i) \geq i$, where $w_i = \hat{f}_a(v_i)$. Then $w_i \geq v_i \geq v_0$ and $\zeta(w_i) = \hat{\zeta}(a)$, so as X is restricted, $w = w_i$ is independent of i. Thus $w = w_k \in W$ and $a \in S(w) \subseteq T(W)$.

Claim for each nonempty subset U of W, $T(U)$ is contractible. Observe $f_w(v_0) \in T(w) \subset L(f_w(v_0))$, so $T(w)$ is contractible. Further if Z is the set of maximal members of U then $C = \{T(z) : z \in Z\}$ is a cover of $T(U)$, so if $Z = \{z\}$ then $T(U) = T(z)$ is contractible. Next if $I \subseteq Z$, $T_I = \bigcap_{i \in I} T(i) = T(J)$, where $J = \bigcap_{i \in I} W(\leq i)$. So if $I = \{i\}$ is of order 1 then $T_I = T(i)$ is contractible, while if $|I| > 1$ then $M(J) < M(U) = \max\{h(u) : u \in U\}$, so by induction on $M(U)$, T_I is contractible. Here we use the fact that $J \neq \varnothing$ since v_k is the unique minimal element of J. It follows that $T(U)$ has the homotopy type of the nerve $N(C)$ of the cover C, and $N(C)$ has the homotopy type of the set of all nonempty subsets of C, so $N(C)$ is contractible. Thus the Claim is established. In particular $T(W) = \iota^{-1}(st_{K_{n+1}}(s))$ is contractible, completing the proof of (1). By (1), the inclusion $\iota : K^n \to K_{n+1}$ is a homotopy equivalence, so $\iota_* : H_*(K^n) \to H_*(K_{n+1})$ is an isomorphism.

Next some notation. If C_* and D_* are chain complexes, then write $\mathrm{Hom}(C_*, D_*)$ for the group of all maps $f = \bigcup_i f_i$, where $f_i \in \mathrm{Hom}(C_i, D_i)$ is a group homomorphism. Further if G is a group, denote by $\mathrm{Hom}(G, C_*)$ the chain complex whose ith term is $\mathrm{Hom}(G, C_i)$ and whose boundary map ∂ is defined by $\partial(x)(g) = \partial(\hat{\zeta}(g))$ for $g \in G$ and $x \in \mathrm{Hom}(G, C_i)$. Note that ∂ is a boundary map, since $\partial^2(x)(g) = (\partial(\partial(x)))(g) = \partial(\partial(x)(v)) = \partial(\partial(\hat{\zeta}(v)) = \partial^2(\hat{\zeta}(v)) = 0$, so $\partial^2 = 0$. Finally define $\hat{K}^n$ to be the full subcomplex of K^n consisting of those $v \in V$ with $h(\hat{\zeta}(v)) = n$.

Using these notions and the simplicial complex $K(n)$ discussed in the example earlier in this section, we now define

$$\phi \in \mathrm{Hom}(C_*(K(n)), \mathrm{Hom}(C_n(\hat{K}^n), C_*(K^n)))$$

For $c \in C_k(K(n))$ we write ϕ_c for the image of c under ϕ. Thus

$$\phi_c \in \mathrm{Hom}(C_n(\hat{K}^n), C_*(K^n))_k = \mathrm{Hom}(C_n(\hat{K}^n), C_k(K^n))$$

and as ϕ is a group homomorphism, it suffices to define ϕ on the generators $\sigma \in \Sigma^k(n)$ of $C_k(K(n))$. Similarly as $\phi_\sigma \in \mathrm{Hom}(C_n(\hat{K}^n), C_k(K^n))$ is a group homomorphism, it suffices to define ϕ_σ on the generators $s = \{v_0, \dots, v_n\} \in \Sigma^n(\hat{K}^n)$, ordered so that $v_0 < \cdots < v_n$. We do so by decreeing that

$$\phi_\sigma(s) = \prod_{(i,j)\in\sigma} f_{v_i}(v_j).$$

Observe

(11.5) *For $\sigma = \{(\alpha_0, \beta_0), \dots, (\alpha_k, \beta_k)\} \in \Sigma^k(n)$ and $s = \{v_0, \dots, v_n\} \in \Sigma^n(\hat{K}^n)$, with $v_0 < \cdots < v_n$ and $(\alpha_0, \beta_0) < \cdots < (\alpha_k, \beta_k)$, we have $\phi_\sigma(s) \in \Sigma^k(K^n)$ with*

$$f_{v_{\alpha_i}}(v_{\beta_i}) \in L(f_{v_{\alpha_j}}(v_{\beta_j})) \text{ for } i < j,$$

and for $(a,b) \in \sigma$, $\hat{\zeta}(f_{v_a}(v_b)) = \zeta(v_a)$, and $\zeta(f_{v_a}(v_b)) = \zeta(v_b)$.

Proof. The last two remarks in the lemma follow from the definition of $f_w(v)$. As $s \in \Sigma^n(\hat{K}^n)$, $v_n = \infty_z$, where $z = \hat{\zeta}(v_n)$, and $v_r \in f(z)$ for each r. As $\beta_r \le \alpha_r$, $v_{\beta_r} \le v_{\alpha_r}$, so $f_{v_{\alpha_r}}(v_{\beta_r})$ is defined.

Let $i < j$. Then by 11.1, $\beta_i \le \beta_j \le \alpha_i \le \alpha_j$. For $r \le \alpha_j$ let $\bar{v}_r = f_{v_{\alpha_j}}(v_r)$. Then $\bar{v}_{\beta_j} \le \bar{v}_{\alpha_i}$ with $f_{\bar{v}_{\alpha_i}}(\bar{v}_{\beta_j}) \ge f_{\bar{v}_{\alpha_i}}(\bar{v}_{\beta_i}) = f_{v_{\alpha_i}}(v_{\beta_i})$, so $f_{v_{\alpha_i}}(v_{\beta_i}) \in L(f_{v_{\alpha_j}}(v_{\beta_j}))$. In particular this shows $\phi_\sigma(s) \in \Sigma^k(K^n)$.

By 11.5, $\phi_\sigma(s) \in \Sigma^k(K^n)$, so ϕ is well defined. Thus we have proved:

(11.6) *The map ϕ defined above is in* $\mathrm{Hom}(C_*(K(n)), \mathrm{Hom}(C_n(\hat{K}^n), C_*(K^n)))$.

Next we have boundary maps ∂ on $C_*(K(n))$ and on the chain complex $\mathcal{H}(n, K) = \mathrm{Hom}(C_n(\hat{K}^n), C_*(K^n))$. We observe next that the map ϕ preserves these boundary maps:

(11.7) *$\partial(\phi_c) = \phi_{\partial(c)}$ for each $c \in C_*(K(n))$. Thus $\phi : C_*(K(n)) \to \mathcal{H}(n, K)$ preserves the boundary maps.*

Proof. Let $\sigma = \{(a_0, b_0), \dots, (a_k, b_k)\} \in \Sigma^k(K(n))$ and

$$s = \{v_0, \dots, v_n\} \in \Sigma^n(\hat{K}^n))$$

Then

$$\phi_{\sigma^i}(s) = \prod_{j\ne i} f_{v_{a_j}}(v_{b_j}) = \phi_\sigma(s)^i$$

Therefore if $c = \sum_\sigma a_\sigma \sigma \in C_k(K(n))$ then $\phi_{c^i}(s) = \sum_\sigma a_\sigma \phi_{\sigma^i}(s) = \sum_\sigma \phi_\sigma(s)^i = (\sum_\sigma a_\sigma \phi_\sigma)^i = \phi_c(s)^i$. Then

$$(\partial(\phi_c))(s) = \partial(\phi_c(s)) = \sum_i (-1)^i \phi_c(s)^i = \sum_i (-1)^i \phi_{c^i}(s) = \phi_{\partial(c)}(s)$$

Let $z \in V_n^n$ and $L = \mathrm{Link}_{K^n}(z)$. We can also regard

$$\phi \in \mathrm{Hom}(C_*(K(n)), \mathrm{Hom}(C_{n-1}(L), C_*(K^n)))$$

by composing ϕ with the map

$$C_{n-1}(L) \to C_n(K^n)$$
$$c \mapsto cz$$

using 11.3. That is $\phi_c = \phi_{cz}$ for $c \in C_{n-1}(L)$.

§12. Cellular homology

We begin by recalling a few standard facts from homological algebra and elementary algebraic topology.

(12.1) *If*

$$0 \to C \xrightarrow{\alpha} D \xrightarrow{\beta} E \to 0$$

is a short exact sequence of chain complexes, then there exists a map

$$\partial_* : H_*(E) \to H_*(C)$$

such that for $z \in H_n(E)$, $\partial_*(z) = [(\alpha^{-1} \circ \partial \circ \beta^{-1})(z)]$. *That is if* $z = e + B_n(E)$ *with* $e \in Z_n(E)$ *then* $\partial_*(z) = \alpha^{-1}(\partial(d)) + B_{n-1}(C)$ *for each choice of* $d \in D_n$ *with* $\beta(d) = e$.

Proof. See for example Lemma 4.5.3 in [5].

(12.2) *If*

$$0 \to C \xrightarrow{\alpha} D \xrightarrow{\beta} E \to 0$$

is a short exact sequence of chain complexes, then

$$\cdots \xrightarrow{\partial_*} H_n(C) \xrightarrow{\alpha_*} H_n(D) \xrightarrow{\beta_*} H_n(E) \xrightarrow{\partial_*} H_{n-1}(C) \to \cdots$$

is an exact sequence of groups.

Proof. See for example Theorem 4.5.4 in [5].

Recall if $\iota : L \to K$ is an inclusion of simplicial complexes then ι extends to a linear map $\iota : C_*(L) \to C_*(K)$ of simplicial chain groups, which induces a short exact sequence

$$(*) \qquad 0 \to C_*(L) \xrightarrow{\iota} C_*(K) \xrightarrow{j} C_*(K)/C_*(L) \to 0$$

of chain complexes. Let $H(K,L) = H_*(C_*(K)/C_*(L))$. Then by 12.1, we get a map

$$\begin{aligned} \partial_* : H(K,L) &\to H_*(L) \\ z &\mapsto [(\iota^{-1} \circ \partial \circ j^{-1})(z)] \end{aligned}$$

Let $C_n(K,L) = C_n(K)/C_n(L)$, $Z_n(K,L) = Z_n(C_*(K,L))$, and define $B_n(K,L) = B_n(C_*(K,L))$. Then $Z_n(K,L) = \partial^{-1}(C_{n-1}(L))/C_n(L)$ and $B_n(K,L) = (B_n(K) + C_n(L))/C_n(L)$. Now for $u \in Z_n(K,L)$ and $d \in C_n(K)$, $j(d) = u$ if and only if $u = d + C_n(L)$, in which case as $u \in Z_n(K,L)$, $\partial(d) \in C_{n-1}(L)$. Then by definition of ∂_*, $\partial_*([d]) = \partial(d) + B_{n-1}(L)$. Then composing ∂_* with $\iota_* : H_*(L) \to H_*(K)$ and $j_* : H_*(K) \to H_*(K,L)$, we can regard

$$\begin{aligned} \partial_* : H_n(K,L) &\to H_{n-1}(K,L) \\ [d] &\mapsto [\partial(d)] \end{aligned}$$

Notice $\partial_*^2([d]) = [\partial^2(d)] = 0$, so $\partial_*^2 = 0$. Therefore

(12.3) *If L is a subcomplex of the simplicial complex K then we have a chain complex $H_*(K,L)$ with boundary map ∂_* such that $Z_n(K,L) = \partial^{-1}(C_{n-1}(L))/C_n(L)$, $B_n(K,L) = (B_n(K)+C_n(L))/C_n(L)$, and $\partial_*([d]) = [\partial(d)]$*

(12.4) *If L is a subcomplex of the simplicial complex K containing no $(n-1)$-simplices of K then $H_n(K) \cong H_n(K,L)$.*

(12.5) *If L is a subcomplex of the simplicial complex K then*
(1) *We have an exact sequence of groups:*

$$\cdots \xrightarrow{\partial_*} H_n(L) \xrightarrow{\iota_*} H_n(K) \xrightarrow{j_*} H_n(K,L) \xrightarrow{\partial_*} H_{n-1}(L) \to \cdots$$

(2) *If $\iota_* : H_*(L) \to H_*(K)$ is an isomorphism then $H_*(K,L) = 0$.*
(3) *If $H_{n+1}(K,L) = H_n(K,L) = 0$ then $j_* : H_n(K) \to H_n(L)$ is an isomorphism.*

Proof. Applying 12.2 to the short exact sequence (*) above, we get (1). Then (1) implies (2) and (3).

(12.6) *Let $J \subseteq L \subseteq K$ be a chain of simplicial complexes. Then*

(1) *We have a short exact sequence of groups*:

$$\cdots \xrightarrow{\partial_*} H_n(L,J) \xrightarrow{\iota_*} H_n(K,J) \xrightarrow{j_*} H_n(K,L) \xrightarrow{\partial_*} H_{n-1}(L,J) \to \cdots$$

(2) *If* $H_{n-1}(L,J) = H_n(L,J) = 0$ *then* $j_* : H_n(K,J) \to H_n(K,L)$ *is an isomorphism.*

(3) *If* $H_n(L,J) = H_n(K,L) = 0$ *then* $H_n(K,J) = 0$.

(4) *Let* $\kappa : J \to L$ *be inclusion and assume* $\kappa_* : H_*(J) \to H_*(L)$ *is an isomorphism and* $H_n(K,L) = 0$. *Then* $H_n(K,J) = 0$.

Proof. We have the exact sequence

$$(^{**}) \qquad 0 \to C_*(L)/C_*(J) \xrightarrow{\iota} C_*(K)/C_*(J) \xrightarrow{j} C_*(K)/C_*(L) \to 0$$

of chain complexes. Applying 12.2 to (**), we conclude that (1) holds. Of course (1) implies (2) and (3). Assume the hypotheses of (4). As $\kappa_* : H_*(J) \to H_*(L)$ is an isomorphism, 12.5.2 says $H_*(L,J) = 0$, so as $H_n(K,L) = 0$, (3) says $H_n(K,J) = 0$.

Now let (X,f) be a restricted combinatorial cell complex, $\Delta = \Delta(X)$ the graph of X, and $K = K(X)$ the triangulating complex of X. Define the subcomplexes K^n and K_n of K as in Section 11, and adopt the notational conventions of that section. An n-dimensional simplicial complex L is *homology spherical* if $\tilde{H}_i(L) = 0$ for $i \neq n$.

(12.7) *Let* $h(X) \geq n$ *and* $V_n^n = \{z_1, \ldots, z_r\}$. *Then*

(1) $Z_i(K^n, K_n) = \bigoplus_{j=1}^r Z_{ij}$ *and* $B_i(K^n, K_n) = \bigoplus_{j=1}^r B_{ij}$, *where*

$$Z_{ij} = \{\alpha z_j + C_i(K_n) : \alpha \in \tilde{Z}_i(\mathrm{Link}_{K^n}(z_j))\} \text{ and}$$

$$B_{ij} = \{\beta z_j + C_i(K_n) : \beta \in B_i(\mathrm{Link}_{K^n}(z_j))\}.$$

(2) $H_i(K^n, K_n) \cong \bigoplus_{j=1}^r \tilde{H}_i(\mathrm{Link}_{K^n}(z_j))$.

(3) $\partial_* : H_i(K^n, K_n) \to H_{i-1}(K_n)$ *acts via*

$$\partial_* : [\sum_{j=1}^i \alpha_j z_j] \mapsto (-1)^i \sum_{j=1}^i \alpha_j + B_{i-1}(K_n).$$

(4) *If* $\mathcal{O}(\dot{f}(x))$ *is homology spherical for all* $x \in X$ *of height* n *then* $H_i(K^n, K_n) = 0$ *for* $i \neq n$.

Proof. Without loss, X is of height n, so $K = K^n$. Let

$$V_{ij} = \{\alpha z_j + C_i(K_n) : \alpha \in C_i(\mathrm{Link}_K(z_j))\}.$$

Then $C_i(K, K_n) = \bigoplus_j V_{ij}$ and by 11.3, $\partial(\alpha z_j) = \partial(\alpha)z_j + (-1)^i\alpha \in \partial(\alpha)z_j + C_{i-1}(K_n)$, so

$$\partial(\alpha z_j + C_i(K_n)) = \partial(\alpha)z_j + C_{i-1}(K_n)$$

Therefore $\sum_j \alpha_j z_j \in \partial^{-1}(C_{i-1}(K_n))$ if and only if $\alpha_j \in \tilde{Z}_i(\mathrm{Link}_K(z_j))$ for all j. That is

$$Z_i(K, K_n) = \partial^{-1}(C_{i-1}(K_n))/C_i(K_n) = \bigoplus_j Z_{ij}.$$

Similarly the second statement of (1) holds. Of course (1) implies (2) and (2) and 11.2 imply (4). Also a typical element of $H_i(K, K_n)$ is of the form $[\sum_j \alpha_j z_j] = u + B_i(K, K_n)$, where $u = \sum_j \alpha_j z_j + C_i(K_n) \in Z_i(K, K_n)$. Then $\partial(\alpha_j) = 0$ so $\partial(\sum_j \alpha_j z_j) = (-1)^i \sum_j \alpha_j$ by 11.3, and hence

$$\partial([\sum_j \alpha_j z_j]) = \partial(\sum_j \alpha_j z_j) + B_{i-1}(K_n) = (-1)^i \sum_j \alpha_j + B_{i-1}(K_{n-1})$$

establishing (3).

(12.8) *Assume for all $x \in X$ of height n that $\mathcal{O}(\dot{f}(x))$ is homology spherical. Then $H_i(K^n) \cong H_i(K_n)$ for $i \leq n-2$.*

Proof. By 12.7.4, $H_i(K^n, K_n) = 0$ for $i \neq n$. Then by 12.5.3, $H_i(K_n) \cong H_i(K^n)$ for $i \neq n, n-1$.

(12.9) $H_*(K_{n+1}, K^n) = 0$.

Proof. By 11.4.2, $\iota_* : H_*(K^n) \to H_*(K_{n+1})$ is an isomorphism, so the lemma follows from 12.5.2.

(12.10) *Assume for all $m \geq n+2$ and for all $z \in X$ of height m that $\mathcal{O}(\dot{f}(x))$ is homology sphericial. Then $H_i(K) \cong H_i(K^{n+1})$ for $i \leq n$.*

Proof. It suffices to show $H_i(K^m) \cong H_i(K^{m-1})$ for each $m \geq n+2$ and $i \leq n$. As $m \geq n+2$, $\mathcal{O}(\dot{f}(z))$ is homology sphericial for each $z \in X$ of height m. So by 12.8, $H_i(K^m) \cong H_i(K_m)$ for $i \leq m-2$, and hence for $i \leq n$. Therefore by 11.4.2, $H_i(K^m) \cong H_i(K_m) \cong H_i(K^{m-1})$, as desired.

Recall the simplicial cell complex $X(n)$ of the n-simplex and its triangulating complex $K(n)$ discussed in earlier sections.

(12.11) (1) *$K(n)$ has the homotopy type of the n-simplex so it is contractible with trivial reduced homology and fundamental group.*

(2) *$K(n)^m \cong K(m)$ for $m \leq n$.*

(3) *$\tilde{H}_*(K(n)_m) = 0$ for all m, n.*

(4) *$H_*(K(n)^m, K(n)_m) = 0$ for $m > 0$.*

(5) *$H_*(K(n)_m, K(n)_{m-1}) = 0$ for $m \neq 1$.*

Proof. By Theorem 7.6, the n-simplex is triangulated by $K(n)$, so (1) holds. We observed in Section 11 that (2) holds. By 11.4.2, $H_*(K(n)_m) \cong H_*(K(n)^{m-1}) \cong H_*(K(m-1))$ by (2), so (1) implies (3).

To prove (4), we apply 12.5.1 with $K = K(m)$ and $L = K(m)_m$. By (3), $H_i(L) = 0$ for $i \neq 0$ and $\iota_* : H_0(L) \to H_0(K)$ is an isomorphism. As $H_i(L) = 0$ for $i \neq 0$, 12.5.1 says $H_i(K, L) = 0$ for $i \neq 0, 1$. Also

$$H_0(L) \xrightarrow{\iota_*} H_0(K) \xrightarrow{j_*} H_0(K, L) \xrightarrow{\partial_*} H_{-1}(L) = 0$$

is exact with ι_* an isomorphism, so $H_0(K, L) = 0$. Similarly

$$0 = H_1(K) \xrightarrow{j_*} H_1(K, L) \xrightarrow{\partial_*} H_0(L) \xrightarrow{\iota_*} H_0(K)$$

is exact with ι_* and isomorphism, so $H_1(K, L) = 0$. Thus (4) holds.

Finally we prove (5) by applying 12.6.3 with $J = K(n)_{m-1}$, $L = K(n)^{m-1}$, and $K = K(n)_m$. $H_*(L, J) = 0$ by (4) while $H_*(K, L) = 0$ by 12.9, so $H_*(K, J) = 0$ by 12.6.3.

(12.12) *Let $n \geq 1$, $B(n) = B_{n-1}(K(n)_n) + C_{n-1}(K(n)_{n-1})$, and let η and θ be the $(n-1)$-simplices of $K(n)$ defined by*

$$\eta = \{(n, i) : 0 \leq i < n\},$$

$$\theta = \{(n-1, i) : 0 \leq i < n\}.$$

Then $\eta \equiv \theta \mod B(n)$.

Proof. We prove the result by induction on n. Let $K = K(n)$. When $n = 1$, $\eta = \{(1,0)\}$, $\theta = \{(0,0)\}$, and $\sigma = \{(0,0),(1,0)\}$ is a simplex in K_1 with $\partial(\sigma) = \eta - \theta$. Then $\partial(\sigma) \in B_0(K_1) \leq B(1)$, so the lemma holds when $n = 1$ and our induction is anchored.

Let $0 \leq i < n-1$ and

$$L(i) = \{(a, b) \in K : a \neq i \neq b\}$$

If we define $\pi : X(n) - \{i\} \to X(n-1)$ by $\pi(a) = a - 1$ for $a > i$ and $\pi(a) = a$ for $a < i$, then π induces an isomorphism $\pi : L(i) \to K(n-1)$ via $\pi(a, b) = (\pi(a), \pi(b))$. Moreover under this isomorphism, $\pi(\eta^i) =$

$\eta(n-1)$ and $\pi(\theta^i) = \theta(n-1)$ are the simplices playing the role of η and θ in $K(n-1)$. Thus by induction on n, $\eta(n-1) - \theta(n-1) = \partial(e) + f$ for some $e \in C_{n-1}(K(n-1)_{n-1})$ and $f \in C_{n-2}(K(n-1)_{n-2})$. Then $d_i = \pi^{-1}(e) \in C_n(K_n)$ and $c_i = \pi^{-1}(f) \in C_{n-1}(K_{n-1})$ with $\eta^i - \theta^i = \partial(d_i) - c_i$.

Next claim $\eta^{n-1} - \theta^{n-1} = \partial(d_{n-1}) + c_{n-1}$ for some $d_{n-1} \in C_n(K_n)$ and $c_{n-1} \in C_{n-1}(K_{n-1})$. When $n = 2$ take $d_{n-1} = \{(0,1),(0,2)\}$. Then observe that for $n > 2$ and $i < n-1$, $\pi(\eta^{n-1,i})$ and $\pi(\theta^{n-1,i})$ play the role in $K(n-1)$ of η^{n-1} and θ^{n-1}, so by induction on n there is

$$g \in C_{n-1}(K(n-1)_{n-1}) \text{ and } h \in C_{n-2}(K(n-1)_{n-2})$$

with $\pi(\eta^{n-1,i} - \theta^{n-1,i}) = \partial(g) + h$. Then $\delta_i = \pi^{-1}(g) \in C_n(K_n)$ and $\gamma_i = \pi^{-1}(h) \in C_{n-1}(K_{n-1})$ with $\eta^{n-1,i} - \theta^{n-1,i} = \partial(\delta_i) + \gamma_i$. Let $\delta = \sum_i (-1)^i \delta_i$ and $\gamma = \sum_i (-1)^i \gamma_i$. Then $\partial(\delta) + \gamma = \partial(\eta^{n-1} - \theta^{n-1})$, so $\eta^{n-1} - \theta^{n-1} - \delta \in Z_{n-1}(K_n, K_{n-1})$. However by 12.11.5, $H_{n-1}(K_n, K_{n-1}) = 0$, so there is $d_{n-1} \in C_n(K_n)$ and $c_{n-1} \in C_{n-1}(K_{n-1})$ with $\partial(d_{n-1}) = \eta^{n-1} - \theta^{n-1} - c_{n-1}$, completing the proof of the claim.

We now complete the proof of the lemma using the argument of the previous paragraph. Namely let $d = \sum_{i=0}^{n-1} (-1)^i d_i$ and $c = \sum_{i=0}^{n-1} (-1)^i c_i$. Then $\partial(d) + c = \partial(\eta - \theta)$, so as $H_{n-1}(K_n, K_{n-1}) = 0$, there is $\alpha \in C_n(K_n)$ and $\beta \in C_{n-1}(K_{n-1})$ with $\partial(\alpha) = \eta - \theta - \beta$.

In the next lemma we use the map

$$\phi \in \mathrm{Hom}(C_*(K(n)), \mathrm{Hom}(C_{n-1}(\mathrm{Link}_{K^n}(z)), C_*(K^n)))$$

defined in Section 11 for each $z \in V_n^n$.

(12.13) *Let θ be the $(n-1)$-simplex of $K(n)$ defined by*

$$\theta = \{(n-1, i) : 0 \le i < n\}$$

and let $B = B_{n-1}(K_n) + C_{n-1}(K_{n-1})$. Then for each $z \in V_n^n$ and each $\xi \in C_{n-1}(\mathrm{Link}_K(z))$,

$$\xi \equiv \phi_\theta(\xi) \mod B.$$

Moreover for $s = \{v_0, \ldots, v_{n-1}\} \in \Sigma^n(\mathrm{Link}_{K^n}(z))$, $\phi_\theta(s) = f_{v_{n-1}}(s)$.

Proof. The last statement of the lemma is just the definition of ϕ_θ. By 11.7,

$$\phi_{B_{n-1}(K(n)_n)} = \phi_{\partial(C_n(K(n)_n))} = \partial(\phi_{C_n(K(n)_n)}),$$

while $\phi_{C_n(K(n)_n)}(C_{n-1}(L)) \subseteq C_n(K_n)$. Therefore

$$\phi_{B_{n-1}(K(n)_n)}(C_{n-1}(L)) \leq B_{n-1}(K_n).$$

Similarly $\phi_{C_{n-1}(K(n)_{n-1})}(C_{n-1}(L)) \leq C_{n-1}(K_{n-1})$. Therefore the lemma follows from 12.12 and the fact that $\phi_\eta(\xi) = \xi$ for each $\xi \in C_{n-1}$ $(\mathrm{Link}_{K^n}(z))$, where η is the simplex of $K(n)$ defined in 12.12.

We now define the cellular homology of X. Let $D_n(X) = H_n(K^n, K_n)$. Equivalently for $z \in V_n^n$ let

$$D(z) = \{ \sum_{s \in \Sigma^{n-1}(\dot{f}(z))} d_{z,s} zs : \sum_s d_{z,s} s \in \tilde{Z}_{n-1}(\dot{f}(z))\} \leq C_n(f(z))$$

with $d_{z,s} \in \mathbf{Z}$, so that $D(z) \cong \tilde{H}_{n-1}(\dot{f}(z))$. Then by 12.7,

$$D_n(X) = \bigoplus_{z \in V_n^n} D(z) \cong \bigoplus_{z \in V_n^n} \tilde{H}_{n-1}(\dot{f}(z))$$

and this definition of $D_n(X)$ is usually easier to work with. Define the boundary map on the chain complex $D_*(X) = (D_n(X) : 0 \leq n \in \mathbf{Z})$ by

$$\partial_n : D_n(X) \to D_{n-1}(X)$$
$$\sum_{z \in V_n^n} z d_z \mapsto (-1)^n \sum_z \phi_\theta(d_z)$$

where $d_z \in \tilde{Z}_{n-1}(\mathrm{Link}_{K^n}(z))$ and $\theta = \{(n-1, i) : 0 \leq i < n\} \in \Sigma_{n-1}(n)$. Recall from 12.13 that if $d_z = \sum_{s \in \Sigma^{n-1}(\mathrm{Link}_{K^n}(z))} d_{z,s} s$ with $d_{z,s} \in \mathbf{Z}$ then $\phi_\theta(d_z) = \sum_s d_{z,s} f_{w(s)}(s)$, where $w(s)$ is the greatest element of s. Therefore

$$\partial_n : \sum_{z,s} d_{z,s} zs \mapsto (-1)^n \sum_{z,s} d_{z,s} f_{w(s)}(s)$$

which is usually an easier definition to work with. We see in a moment that $\partial_{n-1} \circ \partial_n = 0$, so ∂ is indeed a boundary map. We call $D_*(X)$ the *cellular chain complex* of X and ∂ the *cellular boundary map*. The *cellular homology* of X is $H_*^c(X) = H_*(D_*(X))$.

Now we verify that $\partial_{n-1} \circ \partial_n = 0$ by showing $\partial_{n-1}(\partial_n(z d_z)) = 0$ for each $z \in V_n^n$ and $d_z \in \tilde{Z}_{n-1}(\mathrm{Link}_{K^n}(z))$. First $d_z = \sum_u u d_{z,u}$ and $d_{z,u} = \sum_t d_{z,u,t} t$, where

$$d_{z,u,t} \in \mathbf{Z}, \quad d_{z,u} \in \tilde{Z}_{n-2}(\mathrm{Link}_{K^n}(\{z, u\})),$$

and the sums are over all $u \in \mathrm{Link}_{K^n}(z)$ of height $n-2$ and all $(n-2)$-simplices t in $\mathrm{Link}_{K^n}(\{z,u\})$. As $\partial(d_z) = 0$, 1.1.4 says $0 = \sum_u d_{z,u}$. That is

$$0 = \sum_{u,t} d_{z,u,t}t = \sum_t \Big(\sum_{t<u<z} d_{z,u,t} \Big) t, \text{ so } \sum_{t<u<z} d_{z,u,t} = 0$$

for each t. Now

$$(\partial_{n-1} \circ \partial_n)(zd_z) = (-1)^n \partial_{n-1} \Big(\sum_{u,t} d_{z,u,t} f_u(u) f_u(t) \Big)$$

$$= -\sum_{u,t} d_{z,u,t} f_{w(f_u(t))}(f_u(t)) = -\sum_t \Big(\sum_{t<u<z} d_{z,u,t} \Big) f_{w(t)}(t) = 0.$$

So the cellular boundary map ∂ is indeed a boundary map on the cellular chain complex.

(12.14) *Let*

$$f_m : H_m(K^m, K^{m-1}) \to H_m(K^m, K_m) \text{ and}$$

$$_{n+1}\partial_* : H_{n+1}(K^{n+1}, K^n) \to H_n(K^n, K^{n-1})$$

be the maps induced via 12.6 *by the inclusions*

$$K^{m-1} \subseteq K_m \subseteq K^m \text{ and}$$

$$K^{n-1} \subseteq K^n \subseteq K^{n+1}$$

respectively. Let

$$\partial_{n+1} : H_{n+1}(K^{n+1}, K_{n+1}) \to H_n(K^n, K_n)$$

be the cellular boundary map. Then the diagram

$$\begin{array}{ccc} H_{n+1}(K^{n+1}, K^n) & \xrightarrow{\ _{n+1}\partial_*\ } & H_n(K^n, K^{n-1}) \\ f_{n+1}\downarrow & & \downarrow f_n \\ H_{n+1}(K^{n+1}, K_{n+1}) & \xrightarrow[\partial_{n+1}]{} & H_n(K^n, K_n) \end{array}$$

commutes and f_n is an isomorphism for all n.

Proof. By 12.9, $H_*(K_m, K^{m-1}) = 0$ for all m, so applying 12.6.2 to the chain $K^{m-1} \subseteq K_m \subseteq K^m$, we conclude $f_m : H_m(K^m, K^{m-1}) \to H_m(K^m, K_m)$ is an isomorphism for each m.

To complete the proof it remains to show the following diagram commutes:

$$\begin{array}{ccccc} H_{n+1}(K^{n+1},K^n) & \xrightarrow{\hat{\partial}_*} & H_n(K^n,K^{n-2}) & \xrightarrow{j_*} & H_n(K^n,K^{n-1}) \\ f_{n+1}\downarrow & & \iota_*\downarrow & & \downarrow f_n \\ H_{n+1}(K^{n+1},K_{n+1}) & \xrightarrow[\partial'_*]{} & H_n(K_{n+1},K^{n-2}) & \xrightarrow[\rho]{} & H_n(K^n,K_n) \end{array}$$

that ${}_{n+1}\partial_* = j_* \circ \hat{\partial}_*$, and that $\rho \circ \partial'_* = \partial_{n+1}$. Here ∂'_* and ι_* are the maps induced via 12.6 by the inclusions

$$K^{n-2} \subseteq K_{n+1} \subseteq K^{n+1}$$

$$K^{n-2} \subseteq K^n \subseteq K_{n+1}$$

respectively, $\rho = f_n \circ j_* \circ \iota_*^{-1}$, and

$$\hat{\partial}_* : H_{n+1}(K^{n+1},K^n) \to H_n(K^n,K^{n-2}) \text{ and}$$

$$j_* : H_n(K^n,K^{n-2}) \to H_n(K^n,K^{n-1})$$

are the maps induced via 12.6 by the inclusions

$$K^{n-2} \subseteq K^n \subseteq K^{n+1} \text{ and}$$

$$K^{n-2} \subseteq K^{n-1} \subseteq K^n$$

respectively.

By 12.9, $H_*(K^n,K_{n+1}) = 0$, so applying 12.6 to $K^{n-2} \subseteq K^n \subseteq K_{n+1}$ we conclude $\iota_* : H_n(K^n,K^{n-2}) \to H_n(K_{n+1},K^{n-2})$ is an isomorphism, so ρ is well defined. By definition of ρ, the right hand square in the diagram commutes, so to show the full diagram commutes, it remains to show the left hand square commutes.

From the discussion at the beginning of this section,

$$H_{m+1}(K^{m+1},K^m) = Z_{m+1}(K^{m+1},K^m)/B_{m+1}(K^{m+1},K^m)$$

with

$$Z_{m+1}(K^{m+1},K^m) = \partial^{-1}(C_m(K^m))/C_{m+1}(K^m)$$

and

$$B_{m+1}(K^{m+1},K^m) = (B_{m+1}(K^{m+1}) + C_{m+1}(K^m))/C_{m+1}(K^m).$$

So as $C_{m+1}(K^m) = B_{m+1}(K^{m+1}) = 0$, we have

$$H_{m+1}(K^{m+1},K^m) = \partial^{-1}(C_m(K^m)).$$

That is a typical element of $H_{m+1}(K^{m+1}, K^m)$ is some $d \in C_{m+1}(K^{m+1})$ with $\partial(d) \in C_m(K^m)$. Similarly

$$H_{m+1}(K^{m+1}, K^m) = \partial^{-1}(C_m(K_{m+1})/C_{m+1}(K_{m+1})$$

and $f_{m+1} : d \mapsto d + C_{m+1}(K_{m+1})$.

Also $H_n(K^n, K^{n-2}) = \tilde{Z}_n(K^n)$,

$$H_n(K_{n+1}, K^{n-2}) = H_n(K_{n+1}) = \tilde{Z}_n(K_{n+1})/B_n(K_{n+1}),$$

$\hat{\partial}_* : d \mapsto \partial(d)$, $\partial'_* : d + C_{n+1}(K_{n+1}) \mapsto \partial(d) + B_n(K_{n+1})$, and $\iota_* : z \mapsto z + B_n(K_{n+1})$. Therefore

$$(\iota_* \circ \hat{\partial}_*)(d) = \partial(d) + B_n(K_{n+1}) = (\partial'_* \circ f_{n+1})(d)$$

completing the proof that the diagram commutes.

Finally the proof that $\partial_{n+1} = \rho \circ \partial'_*$ and ${}_{n+1}\partial_* = j_* \circ \hat{\partial}_*$. As

$$\iota_* : H_n(K^n, K^{n-2}) \to H_n(K_{n+1}, K^{n-2})$$

is an isomorphism, $\tilde{Z}_n(K_{n+1}) = \tilde{Z}_n(K^n) \oplus B_n(K_{n+1})$. Thus if $\pi : \tilde{Z}_n(K_{n+1}) \to \tilde{Z}_n(K^n)$ is the projection with respect to this direct sum decomposition, we have $\rho = f_n \circ j_* \circ \iota_*^{-1} = f_n \circ j_* \circ \pi$. Further we have seen $H_n(K^n, K^{n-1}) = \partial^{-1}(C_{n-1}(K^{n-1}))$, $H_n(K^n, K_n) = \partial^{-1}(C_{n-1}(K_n)/C_n(K_n))$, and $f_n : d' \mapsto d' + C_n(K_n)$. Finally j_* is the identity map on $H_n(K^n, K^{n-1})$, so $\rho : e \mapsto \pi(e) + C_n(K_n)$. Therefore

$$\rho \circ \partial'_* : d + C_{n+1}(K_{n+1}) \mapsto \pi(\partial(d)) + C_n(K_n).$$

Also $(j_* \circ \hat{\partial}_*)(d) = \partial(d) = {}_{n+1}\partial_*(d)$, with the last equality following from the discussion at the beginning of this section.

Recall from 12.7 that we can choose are coset representative d so that $d = \sum_{z \in V^{n+1}_{n+1}} z d_z$, with $d_z \in \tilde{Z}_n(\mathrm{Link}_{K^{n+1}}(z))$, and that by 11.3, $\partial(z d_z) = (-1)^{n+1} d_z$. Thus it remains to show that if $d = z d_z$ then $\pi(d_z) \equiv \phi_\theta(z d_z) \mod C_n(K_n)$. Let $e = \phi_\theta(z d_z)$. By 12.13, $d_z = e + b + c$ for some $b \in B_n(K_{n+1})$ and $c \in C_n(K_n)$. Then

$$d_z - b = e + c \in \tilde{Z}_n(K_{n+1}) \cap C_n(K^n) \le \tilde{Z}_n(K^n),$$

so $e + c = \pi(d_z)$ and therefore $\pi(d_z) = e + c \equiv e \mod C_n(K_n)$, completing the proof.

(12.15) *Assume for each $x \in X$ of height n that $\mathcal{O}(\dot{f}(x))$ is homology sphericial. Then $H_i(K^n, K^{n-1}) = 0$ for all $i \neq n$.*

Proof. By 12.9, $H_*(K_n, K^{n-1}) = 0$, so applying 12.6.2 to the sequence $K^{n-1} \subseteq K_n \subseteq K^n$, we conclude $H_*(K^n, K^{n-1}) \cong H_*(K^n, K_n)$. Then as $\mathcal{O}(\dot{f}(x))$ is homology spherical for each $x \in X$ of height n, 12.7.4 completes the proof.

Theorem 12.16. *Let X be a restricted combinatorial cell complex such that $\mathcal{O}(\dot{f}(x))$ is spherical for each $x \in X$. Then the ordinary homology $H_*(X)$ of X is isomorphic to the cellular homology $H_*^c(X)$.*

Proof. In Lemma 12.14 we defined maps

$$_{n+1}j_* \circ {}_{n+1}\hat{\partial}_* = {}_{n+1}\partial_* : H_{n+1}(K^{n+1}, K^n) \to H_n(K^n, K^{n-1})$$

and proved these maps are isomorphic to the cellular maps

$$\partial_{n+1} : D_{n+1}(X) \to D_n(X).$$

Thus the chain complex $(H_n(K^n, K^{n-1}), {}_n\partial_* : 0 \leq n \in \mathbf{Z})$ is isomorphic to the cellular chain complex $D_*(X)$. Therefore it suffices to show

$$H_n(X) \cong ker({}_n\partial_*)/Im({}_{n+1}\partial_*)$$

To do so we use the standard proof for CW- complexes.

First the inclusions

$$K^{n-2} \subseteq K^{n-1} \subseteq K^n$$

$$K^{n-2} \subseteq K^n \subseteq K^{n+1}$$

induce via 12.6 the maps

$$\begin{tikzcd}
H_n(K^{n-1}, K^{n-2}) \arrow[dr] & & H_n(K^{n+1}, K^{n-1}) \arrow[r] & H_n(K^{n+1}, K^n) \\
& H_n(K^n, K^{n-2}) \arrow[ur, "\iota_*"] \arrow[dr, "j_*"] & & \\
H_{n+1}(K^{n+1}, K^n) \arrow[ur, "\hat{\partial}_*"] & & H_n(K^n, K^{n-1}) \arrow[r, "{}_n\partial_*"] & H_{n-1}(K^{n-1}, K^{n-2})
\end{tikzcd}$$

By 12.15, $H_n(K^{n-1}, K^{n-2}) = 0$, so as the sequences of 12.6 are exact, $j_* = {}_{n+1}j_*$ is injective with $Im(j_*) = ker({}_n\partial_*)$. As $\mathcal{O}(\dot{f}(x))$ is homology spherical for each $x \in X$, $H_n(K^{n+1}, K^n) = 0$ by 12.15. Therefore ι_* is a surjection, so

$$\iota_* \circ j_*^{-1} : ker({}_n\partial_*) \to H_n(K^{n+1}, K^{n-2})$$

is a surjection with kernel

$$j_*(ker(\iota_*)) = j_*(Im({}_{n+1}\partial_*)) = Im({}_{n+1}\partial_*)$$

That is

$$H_n(K^{n+1}, K^{n-2}) \cong ker({}_n\partial_*)/Im({}_{n+1}\partial_*)$$

Finally by 12.4, $H_n(K^{n+1}) \cong H_n(K^{n+1}, K^{n-2})$, while as $\mathcal{O}(\dot{f}(x))$ is homology spherical for each $x \in X$, $H_n(K^{n+1}) \cong H_n(K)$ by 12.10. Thus the Theorem is established.

(12.17) *Assume $\dot{f}(x)$ is simply connected for each $x \in X$ with $h(x) \geq 3$. Then*
(1) $\pi_1(K(X^{n+1})) \cong \pi_1(K(X^n))$ *if* $n \geq 2$,
(2) $\pi_1(K(X)) \cong \pi_1(K(X^n))$ *for* $n \geq 2$.

Proof. Evidently (1) implies (2), so we prove (1). Recall the definitions of X^n, K^n, and K_n from Section 11. In particular $K^n = K(X^n)$, so it suffices to show $\pi_1(K^{n+1}) \cong \pi_1(K^n)$.

Let $\iota : K^n \to K_{n+1}$ be the inclusion map. By 11.4.2, ι is a homotopy equivalence, so $\pi_1(K^n) \cong \pi_1(K_{n+1})$. Thus it remains to show $\pi_1(K_{n+1}) \cong \pi_1(K^{n+1})$.

Let $\pi : K_{n+1} \to K^{n+1}$ be the inclusion map and $s = \{v_0, \ldots, v_k\}$ be a simplex in K^{n+1} ordered as in Lemma 5.4. If $v_0 \in K_{n+1}$ then $v_0 \in P = \pi^{-1}(st_{K^{n+1}}(s)) \subseteq st_{K_{n+1}}(v_0)$ and hence P is contractible. Thus we may take $s = \{x\}$ with $h(x) = n+1$. In this case $P = \dot{f}(x)$ is simply connected by hypothesis. Thus by Theorem 1 in [3], $\pi_1(K_{n+1}) \cong \pi_1(K^{n+1})$, as desired.

§13. The torus and the Klein bottle

In this section we consider the torus and the Klein bottle as examples.

Let X be the poset of dimension 2 with a unique maximal element c, and unique minimal element d, and two elements a_1 and a_2 of height 1.

We associate a combinatorial cell $f(x)$ to each $x \in X$. Let $f(c)$ have 4 elements b_i of height 1 and 4 elements v_i of height 0, with $v_i, v_{i-1} < b_i$, where the indices are read modulo 4. Further we let $\zeta(b_i) = a_{[i]}$, where $[i] = i \mod 2$. This forces $f(a_i)$ to have 2 elements u_{ij}, $j = 1, 2$, of height 0.

It remains to describe $f_i = f_{b_i} : f(c)(\leq b_i) \to f(a_{[i]})$. We may choose notation so that $f_2(v_i) = u_{2,i}$ and $f_3(v_{i+1}) = u_{1,i}$. We say that b_2 and b_4 have the *same orientation* if $f_4(v_3) = f_2(v_2)$; here the pair v_2, v_3 is distinguished by $b_3 > v_2, v_3$, whereas no member of $f(c)$ of height 1 is greater than v_2 and v_4. Up to change of notation, we are left with 3 cases:

(i) b_2 and b_4 and b_1 and b_3 both have the same orientation.
(ii) Exactly one pair (say b_2 and b_4) has the same orientation.
(iii) Neither pair has the same orientation.

We will see that the geometric realization of (X, f) in case (i) corresponds to the torus, in case (ii) to the Klein bottle, and in case (iii) the realization is not a manifold.

We have defined our combinatorial cell (X, f). We next consider the polyhedral cell complex $\mathcal{P}(X) = (X, f, F, B)$ of (X, f) and the geometric realization $T(X)$ of (X, f). We can regard $F(c)$ as the unit square with vertices $F(v_i)$ arranged in order as in the diagram below. Then $F(b_i)$ is the edge $[v_{i-1}, v_i]$. Thus we have the picture

$$\begin{tikzcd} v_4 \arrow[d, "F(b_1)"'] & v_3 \arrow[l, "F(b_4)"'] \\ v_1 \arrow[r, "F(b_2)"] & v_2 \arrow[u, "F(b_3)"'] \end{tikzcd}$$

Similarly $F(a_i)$ is the unit interval $[F(u_{i1}), F(u_{i2})]$ with initial point $F(u_{i1})$ and endpoint $F(u_{i2})$. Finally we have

$$F_{b_i} : F(b_i) \to F(a_{[i]})$$
$$tF(v_{i-1}) + (1-t)F(v_i) \mapsto tF(f_i(v_{i-1})) + (1-t)F(f_i(v_i))$$

In cases (i) and (ii), b_2 and b_4 have the same orientation, so in the geometric realization $T(X)$ the edges $F(b_2)$ and $F(b_4)$ are identified with the same orientation, resulting in a tube with ends $F(b_1)$ and $F(b_3)$. In case (i), these ends are identified with the same orientation, resulting in a torus, while in case (ii) they are identified with a twist, resulting in a Klein bottle. Finally in case (iii), $F(b_2)$ and $F(b_4)$ are identified with a twist, resulting in a Möbius strip, and then $F(b_1)$ and $F(b_3)$ are identified with a twist, yielding a space $T(X)$ which is not a manifold, since at a neighborhood of $\tilde{F}(d)$ we get two copies of the 2-ball glued at $\tilde{F}(d)$.

We next discuss the the homology of our cell complexes. First $\dot{f}(c)$ has the homotopy type of the 1-sphere and $\dot{f}(a_i)$ the type of the 0-sphere, so by Theorem 12.15, $H^c(X)_* = H_*(X)$. Further $H_1(\dot{f}(c)) = Z_1(\dot{f}(c))$ has a unique generator

$$\gamma = \sum_{i=1}^{4} v_i b_i - v_{i-1} b_i$$

where the indices are read modulo 4, so by the definition of $D_*(X)$, $D_2(X) = D(c)$ has a unique generator $c\gamma$. Similarly $D_1(X) = D(a_1) \oplus D(a_2)$ has two generators $a_i\alpha_i$, where

$$\alpha_i = u_{i,2} - u_{i,1}$$

and $D_0(X) = D(d)$ is 1-dimensional with generator d.

Next by definition of the cellular boundary map ∂,

$$\partial(a_i\alpha_i) = f_{u_{i,2}}(u_{i,2}) - f_{u_{i,1}}(u_{i,1}) = d - d = 0$$

so that $D_1(X) = Z_1(X)$ and $H_0(X) \cong D_0(X) \cong \mathbf{Z}$. Similarly

$$\partial(c\gamma) = \sum_{i=1}^{4} f_{b_i}(b_i(v_i - v_{i-1})) = \sum_{i=1}^{4} a_{[i]} f_i(v_i - v_{i-1})$$

$$= a_1(f_1(v_1-v_4)+f_3(v_3-v_1))+a_2(f_2(v_2-v_1)+f_4(v_4-v_3)) = a_1A_1+a_2A_2$$

Now if b_2 and b_4 have the same orientation then $f_2(v_2) = f_4(v_3)$, and hence also $f_4(v_4) = f_2(v_1)$, so that $A_2 = 0$. On the other hand if the orientation is opposite then $u_{2,2} = f_2(v_2) = f_4(v_4)$, so $u_{2,1} = f_2(v_1) = f_4(v_3)$ and hence $A_2 = 2\alpha_2$. A similar remark holds for A_1.

Now if X is the torus then both pairs have the same orientation, so $\partial(c\gamma) = 0$ and $B_1(X) = 0$. Thus $H_2(X) = D_2(X) \cong \mathbf{Z}$ and $H_1(X) \cong D_1(X) \cong \mathbf{Z} \oplus \mathbf{Z}$. On the other hand if X is the Klein bottle then $A_2 = 0$ and $A_1 = 2\alpha_1$, so $\partial(c\gamma) = 2\alpha_1$. Therefore for the Klein bottle, $H_2(X) = 0$ and $B_1(X) = 2\mathbf{Z}\alpha_1$, so $H_1(X) \cong \mathbf{Z} \oplus \mathbf{Z}_2$. Finally in case (iii), $\partial(c\gamma) = 2(\alpha_1 + \alpha_2)$, so again $H_2(X) = 0$ and $H_1(X) \cong \mathbf{Z} \oplus \mathbf{Z}_2$.

§14. The dual cell complex of a restricted cell complex

In this section (X, f) is a restricted combinatorial cell complex of finite height whose map ζ is *surjective*; that is $\zeta : f(x) \to X(\leq x)$ is surjective for each $x \in X$. The *dual complex* $(\hat{X}, \hat{f})$ of (X, f) is the combinatorial cell complex defined below. But first an example.

Example. Let (X, f) be a regular cell complex. Then (X, f) is isomorphic to the simplicial cell complex of the poset X, so without loss it is that complex. That is $f(x) = X(\leq x)$ and the maps ζ and f_v, $v \in f(x)$, are the appropriate identity maps. It will turn out that the dual complex for *this* complex is the simplicial cell complex of the dual poset $\hat{X}$ whose ordering $\lesssim$ is obtained by reversing the ordering on X. So the duality operation on cell complexes is a generalization of the duality on posets.

In general the poset of the dual complex is the dual poset $\hat{X}$. In addition to defining the poset of the dual complex we must define the cells $\hat{f}(x)$ for each $x \in \hat{X}$, the zeta map for $\hat{X}$, and the f-v maps for each $v \in \hat{V}$. In our example, $\hat{f}(x) = X(\geq x)$. It will turn out in general that the zeta map for the dual complex $\hat{X}$ is the hat-zeta map $\hat{\zeta}$ for the complex $\hat{X}$. So in our case $\hat{\zeta}(v) = v$ for $v \in V = \hat{V}$. Similarly it will turn out that the f-v map $\hat{f}_v : \hat{f}(x)(\lesssim v) \to \hat{f}(\hat{\zeta}(v))$ for $\hat{X}$ is the hat-f-v map for X, so in our case that map is $\hat{f}_v(u) = u$ for $u \leq v \in f(x)$.

Now the definition of the dual complex. First the poset $\hat{X}$ of the complex is the dual poset of X. That is X and $\hat{X}$ are the same as sets with the ordering $\lesssim$ on $\hat{X}$ defined by $x \lesssim y$ if and only if $y \leq x$. As X has finite height, $\hat{X}$ has the same finite height, so all elements of $\hat{X}$ are of finite height.

Second for $x \in \hat{X}$, define

$$\hat{f}(x) = \{v \in V : \zeta(v) = x\}$$

partially ordered by $u \lesssim v$ if and only if $v \in L(u)$. We check that this is a partial order: By definition the relation is reflexive and antisymmetric. Suppose $u \lesssim v \lesssim w$ in $\hat{f}(x)$. Then $v \in L(u)$ so $f_a(u) \geq v$ where $a = \hat{f}_v(u)$. Indeed as $u, v \in \hat{f}(x)$, $\zeta(u) = x = \zeta(v)$, so $\zeta(f_a(u)) = \zeta(u) = \zeta(v)$ and hence $f_a(u) = v$. Similarly as $v \lesssim w$, $f_b(v) = w$, where $b = \hat{f}_w(v)$. But now $c = f_a^{-1}(b) = \hat{f}_w(u)$ so $u \lesssim w$. For by axiom (ii) for combinatorial cell complexes,

$$f_c(u) = f_{f_a(c)}(f_a(u)) = f_b(v) = w.$$

Observe that by definition $\hat{V} = V$ and $\hat{f}(x)$ is the set of elements $v \in V$ such that $x_\infty \in L(v)$.

Recall the hat-zeta-function $\hat{\zeta}$ for X is defined by $\hat{\zeta}(v)$ is the unique $x \in X$ with $v \in f(x)$. Notice that the hat-zeta-function for $\hat{X}$ is the zeta-function for X. Conversely we define the zeta-function for $\hat{X}$ to be the hat-zeta-function for X. We verify $\hat{X}$ satisfies axiom (i) for combinatorial cell complexes; that is we check that $\hat{\zeta} : \hat{f}(x) \to \hat{X}$ is a map of posets preserving height. Namely if $u, v \in \hat{f}(x)$ with $u \lesssim v$ then $f_a(u) = v$ where $a = \hat{f}_v(u)$, and $\hat{\zeta}(u) \geq \zeta(a) = \hat{\zeta}(v)$, so $\hat{\zeta}$ preserves the order. Next as X is of finite height, there is a chain $y_0 > \cdots > y_m = \hat{\zeta}(u)$ of maximal length m and as $\zeta : f(y_i) \to X(\leq y_i)$ is surjective for each i, we can lift this chain to a chain $w_0 > \cdots > w_m$ in $f(y_0)$ with $\zeta(w_i) = y_i$. Let $a = f_{w_m}^{-1}(u)$ and $a_i = f_{w_i}(a)$. Then $a_0 \lesssim \cdots \lesssim a_m = u$ is a chain of length m in $\hat{f}(x)$, so $\hat{\zeta}$ preserves height. This completes the verification of axiom (i).

Next for $x \in X$ and $v \in \hat{f}(x)$,

$$\begin{aligned}\hat{f}(x)(\lesssim v) =& \{u \in V : \zeta(u) = \zeta(v) \text{ and } v \in L(u)\}\\ =& \{u \in V : v \in L(u) \text{ and } f_{\hat{f}_v(u)}(u) = v\}.\end{aligned}$$

As X is restricted, for each $v \in V$ we have the hat-f-v-function $\hat{f}_v$ defined on the set of $u \in V$ with $v \in L(u)$ by $\hat{f}_v(u)$ the unique $w \in f(\hat{\zeta}(u))$ with $w \geq u$ and $\zeta(w) = \hat{\zeta}(v)$. We define the f-v-function for $\hat{X}$ to be the restriction to $\hat{f}(x)(\lesssim v)$ of the hat-f-v-function $\hat{f}_v$ for X. We next verify that $\hat{X}$ satisfies axiom (ii) for combinatorial cell complexes. By definition $\zeta(\hat{f}_v(u)) = \hat{\zeta}(v) = x$, so

$$\hat{f}_v : \hat{f}(x)(\lesssim v) \to \hat{f}(\hat{\zeta}(v)) = \hat{f}(x).$$

We check that $\hat{f}_v$ is an isomorphism of posets. First if $\hat{f}_v(u_1) = \hat{f}_v(u_2) = w$ then $f_w(u_1) = v = f_w(u_2)$, so as f_w is injective, $u_1 = u_2$. That is $\hat{f}_v$ is injective. Next if $w \in \hat{f}(x)$ then $f_w^{-1}(v) = u \in \hat{f}(x)(\lesssim v)$ with $\hat{f}_v(u) = w$, so $\hat{f}_v$ is a bijection.

Suppose $a, b \in \hat{f}(x)(\lesssim v)$ with $a \lesssim b$. Then $b \in L(a)$ so $f_w(a) = b$ for $w = \hat{f}_b(a)$. Also $f_w(\hat{f}_v(a)) = \hat{f}_v(b)$ so $\hat{f}_v(b) \in L(\hat{f}_v(a))$ and therefore $\hat{f}_v(a) \lesssim \hat{f}_v(b)$. Similarly if $\hat{f}_v(a) \lesssim \hat{f}_v(b)$ then $a \lesssim b$. This completes the verification of axiom (ii).

We next verify axiom (iii); that is we prove that if $a, b \in \hat{f}(x)$ with $a \lesssim b$ then $\hat{f}_a = \hat{f}_{\hat{f}_b(a)} \circ \hat{f}_b$. For let $c \in \hat{f}(x)(\lesssim a)$. Then $\hat{f}_b(c) \lesssim \hat{f}_b(a)$ as $\hat{f}_b$ preserves order, so $w = \hat{f}_{\hat{f}_b(a)}(\hat{f}_b(c)) \geq \hat{f}_b(c) \geq c$ with $\zeta(w) = \hat{\zeta}(\hat{f}_b(a)) = \hat{\zeta}(a)$. Hence as X is restricted, $w = \hat{f}_b(c)$. So axiom (iii) is established.

Next axiom (iv). Let $v \in \hat{f}(x)$ and $a \in \hat{f}(x)(\lesssim v)$. Then $v \in L(a)$ with $\hat{\zeta}(\hat{f}_v(a)) = \hat{\zeta}(a)$ by definition of the hat-zeta map and the hat-f-a-map, so indeed $\hat{\zeta} = \hat{\zeta} \circ \hat{f}_v$, as required in axiom (iv).

Finally we check axiom (v). First $\hat{\infty}_x = \infty_x$ and for $v \in \hat{f}(x)$, $\hat{f}_{\infty_x}(a) = a$ as $\zeta(a) = x = \zeta(\infty_x)$. Thus $\hat{f}_{\infty_x}$ is the identity map. Also $\hat{\zeta}(\infty_x) = x$ again by definition of the hat-zeta-map.

We have shown

(14.1) *Let (X, f) be a combinatorial cell complex of finite height with ζ surjective. Then the dual complex $(\hat{X}, \hat{f})$ is a combinatorial cell* *complex.*

(14.2) $(\hat{X}, \hat{f})$ *is restricted.*

Proof. Let $u_1, u_2, v \in \hat{f}(x)$ with $v \lesssim u_1, u_2$ and $\hat{\zeta}(u_1) = \hat{\zeta}(u_2) = y$ say. Then $\zeta(u_1) = \zeta(u_2) = \zeta(v) = x$ and $u_1, u_2 \in L(v)$, so there exists $w_i \geq v$ with $f_{w_i}(v) = u_i$. But $\zeta(w_i) = \hat{\zeta}(u_i) = y$, so as X is restricted, $w_1 = w_2$ and hence $u_1 = f_{w_1}(v) = f_{w_2}(v) = u_2$.

(14.3) $u \in \hat{L}(v)$ *if and only if* $v \in L(u)$*, in which case the image of* v *under the hat-f-u-map for* $\hat{X}$ *is*

$$f_{f_{\hat{f}_v(u)}(u)}(v) = f_u(f^{-1}_{\hat{f}_v(u)}(v)).$$

Proof. Let $v \in L(u)$. Then $f_w(u) \geq v$ for $w = \hat{f}_v(u)$ and $u \lesssim f_w(u)$ in $\hat{f}(\zeta(u))$. Let $z = f_{f_w(u)}(v)$. Then $z \in L(v)$ with $\hat{f}_z(v) = f_w(u)$ and $\zeta(z) = \zeta(v)$ so $v \lesssim z$ in $\hat{f}(\zeta(v))$. Then as $\hat{f}_z(v) = f_w(u) \gtrsim u$ in $\hat{f}(\zeta(u))$, we conclude $u \in \hat{L}(v)$ and z is the image of v under the hat-f-u-map for $\hat{X}$.

Conversely suppose $u \in \hat{L}(v)$ and let z be the image of v under the hat-f-u-map for $\hat{X}$. That is $u \lesssim \hat{f}_z(v) = y$ say. Thus $y \geq v$ and as $u \lesssim y$ in $\hat{f}(\zeta(u))$, $f_w(u) = y$ for $w = \hat{f}_y(u)$. Therefore $f_w(u) = y \geq v$, so $v \in L(u)$.

(14.4) *Let* (X, f) *be a restricted combinatorial cell complex with* ζ *surjective. Then the dual complex* $(\hat{X}, \hat{f})$ *also satisfies these properties and* (X, f) *is the dual of* $(\hat{X}, \hat{f})$*.*

Proof. We have already observed that $\hat{X}$ is of finite height. Next if $y \in X(\lesssim x)$ then $x \in X(\geq y)$ so as $\zeta : f(y) \to X(\leq y)$ is surjective there is $v \in f(y)$ with $\zeta(v) = x$. Thus $v \in \hat{f}(x)$ and $\hat{\zeta}(v) = y$, so $\hat{\zeta} : \hat{f}(x) \to \hat{X}(\lesssim x)$ is a surjection.

Let (Y, g) be the dual of $(\hat{X}, \hat{f})$. Then $\hat{X} = X$ so $Y = \hat{X} = X$ and as $\lesssim$ is the dual of the ordering on X, the order on Y is the ordering $\leq$ dual to $\lesssim$. That is $Y = X$ as a partially ordered set. Similarly for $x \in X$,

$$g(x) = \{v \in V : \hat{\zeta}(v) = x\} = f(x)$$

and as we observed during the proof of 14.1, the hat-zeta-function for $\hat{X}$ is ζ. The ordering on $g(x)$ is given by $u \leqq v$ if and only if $v \in \hat{L}(u)$ if and only if $u \in L(v)$ (by 14.3) if and only if $u \leq v$ as $u, v \in f(x)$. That is $g(x) = f(x)$ as a partially ordered set.

Next the zeta-function for Y is the hat-zeta function for $\hat{X}$ which is the zeta-function for X. That is X and Y have the same zeta-function. Finally for $u \in V$, the f-u-function g_u for Y is the hat-f-u-function for $\hat{X}$, so by 14.3, $g_u(v) = f_{f_w(u)}(v)$, where $w = \hat{f}_v(u)$. But as $v \leq u$ in $f(x)$, $w = x$ and $f_w(u) = u$, so $g_u = f_u$, completing the proof.

(14.5) *$K(X) = K(\hat{X})$, so X and $\hat{X}$ have the same homology and fundamental group.*

Proof. As observed during the construction of $\hat{X}$, $V = \hat{V}$, so $K(X)$ and $K(\hat{X})$ have the same vertex set. Further $K(X)$ is the clique complex of the symmetric relation $*$ on V define by $u * v$ if and only if $u \in L(v)$ or $v \in L(u)$. Similarly $K(\hat{X})$ is the clique complex of the relation $\hat{*}$. But by 14.3, these two relations are the same.

§15. CW-complexes

Denote by B^n the unit n-ball

$$B^n = \{x \in \mathbf{R}^n : |x| \leq 1\}$$

in $\mathbf{R}^n$ and let

$$S^{n-1} = \{x \in \mathbf{R}^n : |x| = 1\}$$

be the $(n-1)$-sphere. We also write $\dot{B}^n$ for the boundary S^{n-1} of B^n and $I(B^n)$ for the interior $B^n - \dot{B}^n$ of B^n. In particular when $n = 0$, $I(B^0) = B^0$ and $\dot{B}^0 = \varnothing$.

(15.1) *Let X, Y be copies of B^n and $f : I(X) \to I(Y)$ be a homeomorphism. Then f extends to at most one homeomorphism $g : X \to Y$.*

Proof. Suppose $g, h : X \to Y$ are homeomorphisms extending f. Then $h^{-1} \circ g$ is a homeomorphism of X extending the identity map on $I(X)$, and it suffices to show $h^{-1} \circ g$ is the identity map on X. Thus it suffices to show that if $k : X \to X$ is a homeomorphism which is the identity on $I(X)$ then k is the identity. Let $x \in \dot{X}$, $y \in I(X)$ and consider the line segment $[x, y]$. Then $[x, y]$ is the closure of $(x, y]$ so $k([x, y])$ is the closure of $k((x, y]) = (x, y]$. That is $k(x) = x$, as desired.

Recall that a *CW-complex* is a triple (T, Λ, φ), where T is a topological space, Λ is a collection of subspaces of T, and

(CW1) T is Hausdorff.

(CW2) For each $\lambda \in \Lambda$, $\varphi_\lambda : F(\lambda) \to \bar{\lambda}$ is a continuous surjection such that $F(\lambda) = B^{n(\lambda)}$ and $\varphi_\lambda : I(F(\lambda)) \to \lambda$ is a homeomorphism.

(CW3) $\dot{\lambda} = \varphi_\lambda(\dot{F}(\lambda)) \subset \bigcup_{\alpha \in \Lambda(\lambda)} \alpha$ for some finite subset $\Lambda(\lambda)$ of Λ.

(CW4) T is the disjoint union of the subspaces $\lambda \in \Lambda$.

(CW5) A subset C of T is closed in T if and only if $C \cap \bar{\lambda}$ is closed in $\bar{\lambda}$ for all $\lambda \in \Lambda$, and if $C \subset \bar{\lambda}$ then C is closed in $\bar{\lambda}$ if and only if $\varphi_\alpha^{-1}(C)$ is closed in $F(\alpha)$ for all $\alpha \in \Lambda(\lambda) \cup \{\lambda\}$.

Define the CW-complex (T, Λ, φ) to be *normal* if

(R1) $\dot{\lambda} = \bigcup_{\alpha \in \Lambda(\lambda)} \alpha$ for each $\lambda \in \Lambda$.

(R2) For each $\lambda \in \Lambda$ and $\alpha \in \Lambda(\lambda)$, the set $\mathcal{C}_\lambda(\alpha)$ of connected components of $\varphi_\lambda^{-1}(\alpha)$ is finite and for each $C \in \mathcal{C}_\lambda(\alpha)$, the restriction φ_C of φ_λ to C is a homeomorphism of C with α such that $\varphi_\alpha^{-1} \circ \varphi_C$ extends to a homeomorphism of $\bar{C}$ with $F(\alpha)$.

Define the CW-complex (T, Λ, φ) to be *restricted* if $\bar{C}_1 \cap \bar{C}_2 = \varnothing$ for distinct $C_1, C_2 \in \mathcal{C}_\lambda(\alpha)$ and all $\lambda \in \Lambda$ and $\alpha \in \Lambda(\lambda)$.

Example. Recall a CW-complex (T, Λ, φ) is *regular* if Axiom (R1) holds and $\varphi_\lambda : F(\lambda) \to \bar{\lambda}$ is a homeomorphism for each $\lambda \in \Lambda$. Notice that if $\lambda \in \Lambda$ and $\alpha \in \Lambda(\lambda)$ then as φ_λ and φ_α are homeomorphisms, also $\varphi_\alpha^{-1} \circ \varphi_\lambda : \varphi_\lambda^{-1}(\bar{\alpha}) \to F(\alpha)$ is a homeomorphism, so Axiom (R2) is satisfied and (T, Λ, φ) is restricted. Therefore regular CW-complexes are restricted.

In the next few lemmas, assume (T, Λ, φ) is a normal CW-complex.

(15.2) *For each $\lambda \in \Lambda$ and $\alpha \in \Lambda(\lambda)$, $\Lambda(\alpha) \subseteq \Lambda(\lambda)$.*

Proof. As $\alpha \subseteq \dot{\lambda}$, $\bar{\alpha} \subseteq \bar{\lambda}$, so $\beta \subseteq \bar{\lambda}$ for each $\beta \in \Lambda(\alpha)$. Now by axiom (CW4), $\beta \cap \lambda = \varnothing$, so $\beta \subseteq \dot{\lambda}$. Then by Axioms (R1) and (CW4), $\beta \in \Lambda(\lambda)$.

(15.3) *For $\alpha, \beta \in \Lambda$, the following are equivalent*:

(1) $\beta \in \Lambda(\alpha)$.

(2) $\bar{\beta} \subset \bar{\alpha}$.

(3) $\beta \subseteq \dot{\alpha}$.

(4) $\bar{\alpha} \cap \beta \neq \varnothing$ *and* $\beta \neq \alpha$.

Proof. Clearly (1) $\Rightarrow$ (3) $\Rightarrow$ (2) $\Rightarrow$ (4). Assume (4). By Axiom (CW4), $\alpha \cap \beta = \varnothing$, so $\dot{\alpha} \cap \beta \neq \varnothing$. Hence by axiom (R1), $\gamma \cap \beta \neq \varnothing$ for some $\gamma \in \Lambda(\alpha)$. Then by axiom (CW4), $\beta = \gamma$.

Remark 15.4. Let (T, Λ, φ) be a normal CW- complex. Partially order Λ by $\alpha \leq \beta$ if $\bar{\alpha} \subseteq \bar{\beta}$. Then by 15.3, $\Lambda(< \lambda) = \Lambda(\lambda)$ for each $\lambda \in \Lambda$. In particular $\Lambda(\leq \lambda)$ is finite, so $\Lambda \in \mathcal{P}$.

Next for $\lambda \in \Lambda$, define

$$f(\lambda) = \bigcup_{\alpha \leq \lambda} \mathcal{C}_\lambda(\alpha)$$

and partially order $f(\lambda)$ by $C \leq D$ if $\bar{C} \subseteq \bar{D}$. As $\Lambda(\leq \lambda)$ is finite and $\mathcal{C}_\lambda(\alpha)$ is finite for each $\alpha \leq \lambda$, $f(\lambda)$ is finite, so $f(\lambda) \in \mathcal{P}^*$.

Next let V be the disjoint union of the sets $f(\lambda)$, $\lambda \in \Lambda$, and define $\zeta : V \to \Lambda$ by $\zeta(C) = \alpha$ for $C \in \mathcal{C}_\lambda(\alpha)$. Then $\zeta : f(\lambda) \to \Lambda(\leq \lambda)$ is a morphism in $\mathcal{P}$.

Let $\alpha \in \Lambda(\lambda)$ and $C \in \mathcal{C}_\lambda(\alpha)$. By axiom (R2), $\varphi_C : C \to \alpha$ is a homeomorphism and $\varphi_\alpha^{-1} \circ \varphi_C$ extends to a homeomorphism of $\bar{C}$ with $F(\alpha)$. By 15.1, this homeomorphism is unique; denote it by F_C.

Finally $I(F(\lambda))$ is the unique member of $f(\lambda)$ mapping to λ under ζ; we define $F_{F(\lambda)} : f(\lambda) \to f(\lambda)$ to be the identity map.

(15.5) *For $\lambda \in \Lambda$, φ_λ is the unique extension of $\varphi_\lambda : I(F(\lambda)) \to \lambda$ to a continuous map of $F(\lambda)$ to $\bar{\lambda}$.*

Proof. Let $\alpha \in \Lambda(\lambda)$, $C \in \mathcal{C}_\lambda(\alpha)$, $x \in C$, and $y \in I(F(\lambda))$. Let $A = \varphi_\lambda([x, y])$ and $U = \varphi_\lambda((x, y])$. Then $\varphi_\lambda^{-1}(A) = (x, y] \cup \{x_1, \ldots, x_r\}$, where $\{x_i\} = C_i \cap \varphi_\lambda^{-1}(A)$ and $\mathcal{C}_\lambda(\alpha) = \{C_1, \ldots, C_r\}$. Further $[x, y]$ and each of the points x_i is closed in $F(\lambda)$, so $\varphi_\lambda^{-1}(A)$ is closed in $F(\lambda)$. Further if $\beta \in \Lambda(\lambda)$ then either $\varphi_\beta^{-1}(A) = \varnothing$ or $\alpha \leq \beta$ and $\varphi_\beta^{-1}(A) = \varphi_\beta^{-1}(\varphi_\lambda(x))$ is finite and hence closed in $F(\beta)$. So by axiom (CW5), A is closed in T.

Thus A is the closure in T of U. Therefore if $\psi : C^* = C \cup I(F(\lambda) \to \bar{\lambda}$ is a continuous extension of φ_λ then as $[x, y]$ is the closure of $(x, y]$ in C^*, $\psi(x)$ is contained in the closure A of U. Hence $\psi(x) \in \bigcap_{y \in I(F(\lambda))} \varphi_\lambda([x, y]) = \{\varphi_\lambda(x)\}$, so $\psi(x) = \varphi_\lambda(x)$. As this holds for each $C \in f(x)$ and $x \in C$, the lemma follows.

(15.6) *Let $\lambda \in \Lambda$ and $C, D \in f(\Lambda)$ with $C \leq D$. Then*
(1) $\varphi_\lambda = \varphi_{\zeta(C)} \circ F_C$ on $\bar{C}$.
(2) $F_D = F_{F_C(D)} \circ F_C$.
(3) $\zeta = \zeta \circ F_C$ on $f(\lambda)(\leq C)$.
(4) If T is restricted then ζ is injective on $f(\lambda)(\geq C)$.

Proof. Let $C \in \mathcal{C}_\lambda(\alpha)$ and $D \in \mathcal{C}_\lambda(\beta)$. By 15.5, φ_α is the unique extension to $F(\alpha)$ of φ_α restricted to $I(F(\alpha))$, so $\varphi_\alpha \circ F_C$ is the unique extension to $\bar{C}$ of $\varphi_\alpha \circ F_C$ restricted to C. But by construction of F_C, this restriction is equal to the restriction of φ_λ, so (1) holds.

Now by (1), $\varphi_\alpha^{-1}(\beta) = F_C(\varphi_\lambda^{-1}(\beta) \cap \bar{C}$, so if $D_1, \dots, D_r$ are the connected components of $\varphi_\lambda^{-1}(\beta) \cap \bar{C}$ then $F_C(D_1), \dots, F_C(D_r)$ are the connected components of $\varphi_\alpha^{-1}(\beta)$ and hence the members of $\zeta^{-1}(\beta)$. This establishes (3).

Next $F_{F_C(D)}$ is the unique extension of $\varphi_\beta^{-1} \circ \varphi_\alpha$ from $F_C(D)$ to $F_C(\bar{D})$ so $F_{F_C(D)} \circ F_C$ is the unique extension of $\psi = \varphi_\beta^{-1} \circ \varphi_\alpha \circ F_C$ from D to $\bar{D}$. But by (1), $\psi = \varphi_\beta^{-1} \circ \varphi_\lambda$ and by definition F_D is the extension of the latter map, so (2) is established.

Finally if $G, H \in \mathcal{C}_\lambda(\gamma)$ are distinct for some $\gamma \in \Lambda(\lambda)$ then $\bar{G} \cap \bar{H} = \varnothing$ if T is restricted. This proves (4).

(15.7) *Let $\lambda \in \Lambda$. Then*

(1) *If $S \subseteq \Lambda$ then $\bigcap_{\sigma \in S} \bar{\sigma} = \bigcup_{\alpha \leq S} \alpha$.*

(2) *$F(\lambda)$ is partitioned by $\Lambda_\lambda = \{C : C \in f(\lambda)\}$.*

(3) *For $S \subseteq f(\lambda)$, $\bigcap_{C \in S} \bar{C} = \bigcup_{D \leq S} D$.*

(4) *$(F(\lambda), \Lambda_\lambda, \psi)$ is a regular CW-complex, where $\psi = (\psi_C : C \in f(\lambda))$ and $\psi_C = F_C^{-1} : F(\zeta(C)) \to \bar{C}$.*

Proof. Part (1) follows from axiom CW4 and 15.3. Next $F(\lambda)$ is the disjoint union of $I(F(\lambda))$ and $\dot{F}(\lambda)$, so to prove (2) we must show $\dot{F}(\lambda)$ is partition by $\{C : C \in f(\lambda) - \{I(F(\lambda))\}\}$. Let $x \in \dot{F}(\lambda)$. Then $\varphi_\lambda(x) \in \alpha$ for a unique $\alpha \in \Lambda(\lambda)$, so $x \in C$ for some $C \in \mathcal{C}_\lambda(\alpha)$. Further as the members of $\mathcal{C}_\lambda(\alpha)$ are disjoint, $x \notin C'$ for $C \neq C' \in \mathcal{C}_\lambda(\alpha)$, while if $x \in D$ for $D \in f(x)$ then $\varphi_\lambda(x) \in \varphi_\lambda(D) \cap \alpha = \varphi_D(D) \cap \alpha = \zeta(D) \cap \alpha$, so $\zeta(D) = \alpha$. This establishes (2). Then (1) and (2) imply (3).

It remains to prove (4). As $F(\lambda) \cong B^{n(\lambda)}$, $F(\lambda)$ is a Hausdorff space. For $C \in f(\lambda)$, by definition $\bar{C} \cong F(C) = F(\zeta(C)) \cong B^{n(\zeta(C))}$ and $\psi_C : F(C) \to C$ is a homeomorphism. By (2) and (3), $\dot{C} = \bar{C} - C = \bigcup_{D \subset C} D$, and Λ_λ is a partition of $F(\lambda)$. Visibly axiom CW5 is satisfied. Thus (4) is established.

Remark 15.8. We can now associate a topological cell complex $\mathcal{T}(T)$ with the normal CW-complex $T = (T, \Lambda, \varphi)$. Namely let Λ be the poset of $\mathcal{T}(T)$; by Remark 15.4, $\Lambda \in \mathcal{P}$. For $\lambda \in \Lambda$ the cell $f(\lambda)$ associated to λ has as its poset the poset $f(\lambda)$; by Remark 15.4, $f(\lambda) \in \mathcal{P}^*$. The topological space associated to λ is of course $F(\lambda)$. For $C \in f(\lambda)$, $F(C) = \bar{C}$ is a closed subspace of $F(\lambda)$ and as usual with topological cells we take the maps $F(D, C)$, $D, C \in f(x)$, $D \leq C$, to be inclusions. Lemma 15.7 then says $f(\lambda)$ is a topological cell.

Our maps ζ and $f_C : f(\lambda)(\leq C) \to f(\zeta(C))$ were defined in Remark 15.4. That is $f_C(D) = F_C(D)$ for $D \in F(\lambda)(\leq C)$ and $F_C : \bar{C} \to F(\zeta(C))$ is our topological isomorphism. By 15.6, axioms 5 and 6 for cell complexes are satisfied. Axiom 7 for cell complexes holds by definition of $F_{\infty_\lambda} = F_{F(\lambda)}$ as the identity map in Remark 15.4. Thus we have checked that $\mathcal{T}(T)$ is a topological cell complex. By 15.6.4, $\mathcal{T}(T)$ is restricted if T is restricted.

We can extent the map $\mathcal{T}$ to a functor from the category of normal CW- complexes to the category of topological cell complexes. A *morphism of CW-complexes* from $(T_1, \Lambda_1, \varphi_1)$ to $(T_2, \Lambda_2, \varphi_2)$ is a continuous map $\phi : T_1 \to T_2$ such that $\phi(\Lambda_1) \subseteq \phi(\Lambda_2)$, together with homeomorphisms $\phi_\lambda : F(\lambda) \to F(\phi(\lambda))$, $\lambda \in \Lambda_1$, such that $\varphi_{\phi(\lambda)} \circ \phi_\lambda = \phi_{|\bar{\lambda}} \circ \varphi_\lambda$.

Given such a morphism ϕ, we define $\mathcal{T}(\phi) : \mathcal{T}(T_1) \to \mathcal{T}(T_2)$ by $\mathcal{T}(\phi)(\lambda) = \phi(\lambda)$ and $\mathcal{T}(\phi)_\lambda = \phi_\lambda$. Check that $\mathcal{T}(\phi)$ is a morphism of topological cell complexes.

(15.9) *For $\lambda \in \Lambda$ and $X \subseteq \bar{\lambda}$, X is closed in T if and only if $\varphi_\lambda^{-1}(X)$ is closed in $F(\lambda)$.*

Proof. By axiom CW5, X is closed in T if and only if $\varphi_\alpha^{-1}(X)$ is closed in $F(\alpha)$ for all $\alpha \leq \lambda$. Next

$$\varphi_\alpha^{-1}(X) = \coprod_{C \in \mathcal{C}_\lambda(\alpha)} \varphi_\alpha^{-1}(X) \cap \bar{C}$$

and as $\varphi_\lambda = \varphi_\alpha \circ F_C$ on $\bar{C}$, $\varphi_\alpha^{-1}(X)$ is closed in $F(\alpha)$ if and only if $\varphi_\lambda^{-1}(X) \cap \bar{C}$ is closed in $\bar{C}$ for all $C \in \mathcal{C}_\lambda(\alpha)$. Therefore X is closed in T if and only if $\varphi_\lambda^{-1}(X) \cap \bar{C}$ is closed in $\bar{C}$ for all $C \in f(\lambda)$ if and only if $\varphi_\lambda^{-1}(X)$ is closed in $F(\lambda)$.

(15.10) *Let $A = A(\mathcal{T}(T))$, where $\mathcal{T}(T)$ is the topological cell complex supplied by Remark* 15.8. *Then the map*

$$\begin{aligned} \phi : A &\to T \\ \tilde{x} &\mapsto \varphi_\alpha(x) \end{aligned}$$

for $x \in \alpha \in \Lambda$, is a homeomorphism.

Proof. By 4.5, the sets $\tilde{I}(\alpha)$, $\alpha \in \Lambda$, partition A. Further by 4.4 and 4.6, $\lambda_\alpha : F(\alpha) \to A$ is continuous with $\lambda_\alpha : I(\alpha) \to \tilde{I}(\alpha)$ a homeomorphism. Then as $\varphi_\alpha : I(\alpha) \to \alpha$ is a homeomorphism, so is

$\varphi_\alpha \circ \lambda_\alpha^{-1} : \tilde{I}(\alpha) \to \alpha$. But

$$\phi = \bigcup_{\alpha \in \Lambda} \varphi_\alpha \circ \lambda_\alpha^{-1}$$

so $\phi : A \to T$ is a bijection and for each $\alpha \in \Lambda$, $\phi \circ \lambda_\alpha = \varphi_\alpha$.

Next by 15.9, $X \subseteq \bar{\alpha}$ is closed if and only if $\varphi_\alpha^{-1}(X)$ is closed in $F(\alpha)$. But $\varphi_\alpha^{-1}(X) = \lambda_\alpha^{-1}(\phi^{-1}(X))$, so by definition of the topology on $\tilde{F}(\alpha)$, $\varphi_\alpha^{-1}(X)$ is closed in $F(\alpha)$ if and only if $\phi^{-1}(X)$ is closed in $\tilde{F}(\alpha)$. That is X is closed in T if and only if $\phi^{-1}(X)$ is closed in $\tilde{F}(\alpha)$ if and only if $\phi^{-1}(X)$ is closed in A by 4.6. That is ϕ is a homeomorphism.

(15.11) *If (T, Λ, φ) is a regular CW-complex then*

(1) *$\mathcal{T}(T)$ is isomorphic to $\mathcal{P}(\mathcal{X}(T))$, where $\mathcal{X}(T) = (\Lambda, f)$ is the combinatorial cell complex of $\mathcal{T}(T)$.*

(2) *T is homeomorphic to the geometric realization of $\mathcal{O}(\Lambda)$.*

Proof. Let $X = \mathrm{sd}(\Lambda)$ and for $\lambda \in \Lambda$ pick $P(\lambda) \in \lambda$. Let $x = \{\lambda_0, \ldots, \lambda_k\} \in X$ with $\lambda_i < \lambda_{i+1}$ for each i. Given $0 \leq a_i \in \mathbf{R}$ with $\sum_i a_i = 1$, we define $\sum_i a_i P(\lambda_i)$ recursively. Namely let $u = \sum_{i<k} a_i P(\lambda_i)/(1 - a_k)$, $v = P(\lambda_k)$, and define

$$\sum_I a_i P(\lambda_i) = \varphi_{\lambda_k}^{-1}(a_k \varphi_{\lambda_k}(v) + (1 - a_k)\varphi_{\lambda_k}(u))$$

This makes sense as $F(\lambda_k) \subseteq \mathbf{R}^{n(\lambda_k)}$ and $u \in \bar{\lambda}_{k-1} \subseteq \bar{\lambda}_k$ has been defined already via our recursive procedure.

Now define

$$G(x) = \{\sum_i a_i P(\lambda_i) : 0 \leq a_i \in \mathbf{R} \text{ and } \sum_i a_i = 1\}$$

Notice $G(x) \cap G(y) = G(x \cap y)$. Thus if we take (X, g) to be the simplicial cell complex of X, we can regard $G(x)$ as defining a topological cell on $g(x) = X(\leq x)$. Then we extend (X, g) to a topological cell complex by letting G_y be the identity map for $y \subseteq x$.

We next consider the topological cell complex $\chi = \chi(\mathcal{O}(\Lambda))$ and define an isomorphism $\phi : \chi \to (X, g, G)$. Recall from Example 6.4 that the combinatorial cell complex of χ is just (X, g); thus as a map of combinatorial cell complexes we take ϕ to be the identity. Also the topological cell associated to x by χ is the standard simplex $T_{h(x)} = [u_{h(v)} : v \in x]$, so it remains to define $\phi_x : T_{h(x)} \to G(x)$ by

$$\phi_x : \sum_i a_i u_{h(\lambda_i)} \mapsto \sum_i a_i P(\lambda_i)$$

which is of course a homeomorphism. Check that ϕ is a morphism, and then observe that as each ϕ_x is a homeomorphism, ϕ is an isomorphism. Therefore by 4.7, $A(\chi) \cong A(X, g, G)$. Recall also from Example 6.4 that $A(\chi)$ is the geometric realization of $\mathcal{O}(\Lambda)$. Further as (X, g) is the simplicial cell complex of X and $G(x) \subseteq T$ with the maps G_y identities, $G(x) \to \tilde{G}(x)$ for each x and $A(X, g, G) = \bigcup_x G(x) = T$. Thus (2) is established.

Next from Example 7.1, $\mathcal{P}(\mathcal{X}(T))$ is a topological cell complex with combinatorial cell complex $\mathcal{X}(T) = (\Lambda, f)$ and for $\lambda \in \Lambda$, the topological cell $H(\lambda)$ is just $T(f(\lambda))$ together with the subspaces $T(f(\alpha))$ for $\alpha \leq \lambda$. Applying our conclusions of the previous paragraph to $(\bar{\lambda}, \Lambda(\lambda) \cup \{\lambda\}, \varphi)$ in place of (T, Λ, φ), we see that $H(\lambda) \cong \bar{\lambda} \cong F(\lambda)$ via a homeomorphism ϕ_λ preserving the cell structure. Then the identity morphism on the combinatorial cell complex (Λ, f) together with the homeomorphisms ϕ_λ, $\lambda \in \Lambda$, define an isomorphism of $\mathcal{T}(T)$ with $\mathcal{P}(\mathcal{X}(T))$, establishing (1).

(15.12) (1) *$\mathcal{T}(T)$ is isomorphic to $\mathcal{P}(\mathcal{X}(T))$, where $\mathcal{X}(T) = (\Lambda, f)$ is the combinatorial cell complex of $\mathcal{T}(T)$.*

(2) *T is homeomorphic to the geometric realization $T(\mathcal{X}(T))$ of $\mathcal{X}(T)$.*

Proof. As $T(\mathcal{X}(T)) = A(\mathcal{P}(\mathcal{X}(T)))$, (1) and 15.10 imply (2), so it remains to prove (1). The last paragraph of the proof of 15.11 can be repeated virtually verbatim to prove (1).

Define a poset P to be a *n-sphere* if P is of height n and the geometric realization of $\mathcal{O}(P)$ is homeomorphic to the n-sphere S^n.

(15.13) *Let $\lambda \in \Lambda$ and $n = n(\lambda)$. Then $n = h(\lambda)$ and $\mathcal{O}(\dot{f}(\lambda))$ is an $(n-1)$-sphere.*

Proof. Let m be the height of $f(\lambda)$. By 15.7.4, $F(\lambda) = (F(\lambda), \Lambda_\lambda, \psi)$ is a regular CW-complex, so its $(n-1)$-skeleton $\dot{F}(\lambda) = (\dot{F}(\lambda), \Lambda(\lambda), \psi)$ is also a regular CW-complex. Then by 15.11, $S^{n-1} \cong \dot{F}(\lambda)$ is homeomorphic to the geometric realization of $\mathcal{O}(\Lambda(\lambda))$. But $\Lambda(\lambda) \cong \dot{f}(\lambda)$, so the geometric realization of $\mathcal{O}(\dot{f}(\lambda))$ is an $(n-1)$-sphere. In particular $H_{n-1}(\mathcal{O}(\dot{f}(n))) \neq 0$, so $h(\dot{f}(\lambda)) = \dim(\mathcal{O}(\dot{f}(\lambda))) \geq n-1$. Therefore $m = h(f(\lambda)) = h(\dot{f}(\lambda)) + 1 \geq n$.

So to complete the proof it remains to show $m \leq n$. We proceed by induction on m. If $m = 0$ the inequality is trivial. Now $m = h(\dot{f}(\lambda)) + 1$ and $h(\dot{f}(\lambda)) = \max\{h(\alpha) : \alpha \in \Lambda(\lambda)\}$. By induction on m, $h(\alpha) \leq n(\alpha)$, so if $n(\alpha) < n$, we are done. On the other hand $B^{n(\alpha)} \cong F(C)$ is a closed subspace of $\dot{F}(\lambda) \cong S^{n-1}$, and therefore $n(\alpha) < n$, (cf. Exercise H in Chapter 3 of [5]) completing the proof.

Example 15.14. Let $\mathcal{Q}$ be the category of combinatorial cell complexes (X, f) such that for each $x \in X$ of height n, $\dot{f}(x)$ is an $(n-1)$-sphere. Notice that as $T(\dot{f}(x))$ is compact, $\dot{f}(x)$ is finite, a fact we use below without comment several times. We associate to $X = (X, f) \in \mathcal{Q}$ a normal CW-complex $Q(X)$.

The topological space of $Q(X)$ is the geometric realization $A = T(X)$ of (X, f). The set Λ of open cells of A is the set of subspaces $\tilde{I}(x)$, $x \in X$, where we use the notation of Section 4. By definition of $T(X)$, $F(x)$ is the geometric realization of $f(x)$, and as $(X, f) \in \mathcal{Q}$, $\dot{f}(x)$ is an $(n-1)$-sphere, so $F(x)$ is homeomorphic to B^n and $\dot{F}(x)$ is homeomorphic to S^{n-1}. Thus we define

$$\varphi_x : F(x) \to \tilde{F}(x)$$
$$a \mapsto \tilde{a}$$

Our normal CW-complex is $Q(X) = (T(X), \lambda, \varphi)$.

By definition of $A = T(X)$, $\tilde{F}(x)$ is closed in A and φ_x is a continuous surjection. By 4.6, $C \subseteq \tilde{F}(x)$ is closed in A if and only if $\varphi_x^{-1}(C)$ is closed in $F(x)$, so as $F(x)$ is the closure of $I(x)$, $\tilde{F}(x)$ is the closure of $\tilde{I}(x)$. By 4.4, $\varphi_x : I(x) \to \tilde{I}(x)$ is a homeomorphism.

By definition of the topology on A and by remarks in the previous paragraph, axiom CW5 is satisfied. We have also seen that CW2 is satisfied. By 4.5, CW4 is satisfied. Also $\tilde{F}(x) - \tilde{I}(x) = \bigcup_{y<x} \tilde{I}(y)$, so axiom (R1) holds. For $y < x$,

$$\varphi_x^{-1}(\tilde{I}(y)) = \bigcup_{v \in f(x) \cap \zeta^{-1}(y)} I(v)$$

Also for distinct $u, v \in f(x) \cap \zeta^{-1}(y)$, $I(u) \cap I(v) = \varnothing$, so $\{I(v) : v \in f(x) \cap \zeta^{-1}(y)\}$ is the set of connected components of $\varphi_x^{-1}(\tilde{I}(y))$. As $F_v : F(v) \to F(y)$ is a homeomorphism with $\varphi_x = \varphi_y \circ F_v$ on $F(v)$, axiom (R2) holds. Also $F(u) \cap F(v) = \bigcup_{w \leq u,v} F(w)$ and if X is restricted, no such w exists, so $Q(X)$ is restricted. Therefore to complete our proof that $Q(X)$ is a normal CW-complex, it remains to show that $A = T(X)$ is a Hausdorff space, which we leave as an exercise.

So $Q(X)$ is a normal CW-complex. Now we extend Q to a covariant functor from $\mathcal{Q}$ into the category of normal CW-complexes. Namely if $\phi : X_1 \to X_2$ is a morphism of cell complexes in $\mathcal{Q}$, define $Q(\phi) : Q(X_1) \to Q(X_2)$ to be the morphism of CW-complexes whose map of topological spaces is $T(\phi) : T(X_1) \to T(X_2)$ and with $Q(\phi)_x : F(x) \to F(\phi(x))$ defined to be ϕ_x. Check that $Q(\phi)$ is indeed a morphism. Thus we have our functor.

Theorem 15.15. *Let $\mathcal{N}$ be the category of normal CW-complexes. Then we have a covariant functor $\mathcal{X}$ from $\mathcal{N}$ into $\mathcal{Q}$, where $\mathcal{X}(T)$ is the combinatorial cell complex of $\mathcal{T}(T)$. Further the functors $\mathcal{X} : \mathcal{N} \to \mathcal{Q}$ and $Q : \mathcal{Q} \to \mathcal{N}$ are equivalences of categories. Under these equivalences, restricted cell complexes correspond to restricted CW-complexes.*

Proof. First by Remark 15.8, $\mathcal{X}(T)$ is a combinatorial cell complex, and then by 15.13, $\mathcal{X}(T) \in \mathcal{Q}$. So $\mathcal{X}(T)$ is indeed a functor from $\mathcal{N}$ to $\mathcal{Q}$.

By definition of $Q(\mathcal{X}(T))$, the topological space of the CW-complex $Q(\mathcal{X}(T))$ is $T(\mathcal{X}(T))$, the set of open cells is $\Lambda' = \{\tilde{I}(\lambda) : \lambda \in \mathcal{X}(T) = \Lambda\}$, and the characteristic maps are $\varphi'_\lambda : F(\lambda) \to \tilde{F}(\lambda)$ taking a to $\tilde{a}$. By 15.12 and its proof, we have an isomorphism $\phi' : \mathcal{P}(\mathcal{X}(T)) \to \mathcal{T}(T)$ which is the identity map on the combinatorial cell complex $\mathcal{X}(T)$ and with homeomorphisms $\phi'_\lambda : F(x)' \to F(x)$, $\lambda \in \Lambda$. Then ϕ' induces the homeomorphism

$$A(\phi') : T(\mathcal{X}(T)) = A(\mathcal{P}(\mathcal{X}(T))) \to A(\mathcal{T}(T)).$$

We compose $A(\phi')$ with the homeomorphism $\phi : A(\mathcal{T}(T)) \to T$ of 15.10 to obtain a homeomorphism $\iota_T = \phi \circ A(\phi') : T(\mathcal{X}(T)) \to T$. Then we define $\iota_{T,\lambda} = \phi'_\lambda$. Then $\iota_T : Q(\mathcal{X}(T)) \to T = (T, \Lambda, \varphi)$ is an isomorphism of CW-complexes.

In the other direction, $\mathcal{X}(Q(X))$ is the combinatorial cell complex of $\mathcal{T}(Q(X))$, which is by definition (Λ', f'), where the poset Λ' is the set of open cells of $Q(X)$ and $f'(\lambda')$ is the set of connected components of $\varphi_{\lambda'}^{-1}(\alpha')$, where $\alpha' \leq \lambda'$ are open cells in $Q(X)$. Further by definition of $Q(X)$ in Example 15.14, $\Lambda' = \{\tilde{I}(x) : x \in X\}$ and $f'(\tilde{I}(x)) = \{I(v) : V \in f(x)\}$. Thus we define our isomorphism $\iota_X : X \to \mathcal{X}(Q(X))$ by defining $\iota_X : x \mapsto \tilde{I}(x)$ as a map of posets and defining

$$\iota_{X,x} : f(x) \to f'(\tilde{I}(x)) = \{I(v) : v \in f(x)\}$$

by $\iota_{X,x} : v \mapsto I(v)$. Check that ι_X is indeed an isomorphism of combinatorial cell complexes.

Finally check that if $\phi : X_1 \to X_2$ is a morphism in $\mathcal{Q}$ then $\mathcal{X}(Q(\phi)) \circ \iota_{X_1} = \iota_{X_2} \circ \phi$ and similarly if $\psi : T_1 \to T_2$ is a morphism in $\mathcal{R}$ then

$\psi \circ \iota_{T_1} = \iota_{T_2} \circ Q(\mathcal{X}(\psi))$. Therefore the functors $\mathcal{X}$ and Q are inverses of each other on equivalence classes of objects and maps in the two categories and hence equivalences of categories. To illustrate these last checks, observe that as a map of posets, $(\mathcal{X}(Q(\phi)) \circ \iota_{X_1})(x) = \tilde{I}(\phi(x)) = (\phi \circ \iota_{X_2})(x)$ while $(\mathcal{X}(Q(\phi)) \circ \iota_{X_1})_x(v) = I(\phi_x(v)) = (\phi \circ \iota_{X_2})_x(v)$.

§16. Group actions on posets and cell complexes

Let G be a group and $\mathcal{F} = (G_i : i \in I)$ a family of subgroups of G. Let

$$X = \coprod_{i \in I} G/G_i$$

be the disjoint union of the coset spaces G/G_i, $i \in I$. We consider relations defined on X which are preserved by the action of G via right multiplication. Let

$$\mathcal{E} = \bigcup_{(i,j) \in I \times I} \mathcal{E}_{i,j}$$

with $\mathcal{E}_{i,j} \subseteq G_i \backslash G / G_j$, where $G_i \backslash G / G_j$ is the set of double cosets $G_i x G_j$, $x \in G$. Define $\Gamma(G, \mathcal{F}, \mathcal{E})$ to be the relational structure on the set X with relation $G_i x$ related to $G_j y$ if and only if $G_i x y^{-1} G_j \in \mathcal{E}_{i,j}$. Check that this relation is well defined and preserved by the representation of G on X via right multiplication. Observe that there is a type function from X to I defined by $G_i x \mapsto i$ for each $i \in I$ and $x \in G$. Further this type function is preserved by the action of G; that is if h is our type function then $h(xg) = h(x)$ for each $x \in X$ and $g \in G$.

(16.1) *Let $\Gamma(G, \mathcal{F}, \mathcal{E})$ be a relational structure with relation R. Then*

(1) *R is reflexive if and only if $G_i \in \mathcal{E}_{i,i}$ for each $i \in I$.*

(2) *R is symmetric if and only if $G_i u G_j \in \mathcal{E}_{i,j}$ implies $G_j u^{-1} G_i \in \mathcal{E}_{j,i}$ for all i, j.*

(3) *R is antisymmetric if and only if whenever $G_i u G_j \in \mathcal{E}_{i,j}$ and $G_j u^{-1} G_i \in \mathcal{E}_{j,i}$, then $i = j$ and $u \in G_i$, for all i, j.*

(4) *R is transitive if and only if whenever $i, j, k \in I$, $G_i u G_j \in \mathcal{E}_{i,j}$, and $G_j v G_k \in \mathcal{E}_{j,k}$, then $G_i u g v G_k \in \mathcal{E}_{i,k}$ for all $g \in G_j$.*

Proof. First if $G_i x R G_j y$ and $G_j y R G_k z$, then $G_i x y^{-1} G_j \in \mathcal{E}_{i,j}$ and $G_j y z^{-1} G_k \in \mathcal{E}_{j,k}$. Then if $\mathcal{E}$ satisfies the hypotheses of (4), then

$$G_i x z^{-1} G_j = G_i (x y^{-1})(y z^{-1}) G_k \in \mathcal{E}_{i,k},$$

so $G_i x R G_k z$ and the relation R is transitive.

Conversely suppose R is transitive and let $G_iuG_j \in \mathcal{E}_{i,j}$ and $G_jvG_k \in \mathcal{E}_{j,k}$. Then G_iugRG_j for each $g \in G_j$ and $G_jRG_kv^{-1}$, so $G_iugRG_k v^{-1}$, and therefore $G_iugvG_k \in \mathcal{E}_{i,k}$.

Remark. Lemma 16.1 says that the relation defined by a relational structure $\Gamma(G, \mathcal{F}, \mathcal{E})$ is a partial order if and only if $\mathcal{E}$ satisfies the conditions of 16.1.1, 16.1.3, and 16.1.4. Define a *coset poset* over the index set I to be a relational structure $\Gamma(G, \mathcal{F}, \mathcal{E})$ over I such that

(CSP0) I comes equipped with a partial order $\leq$.

(CSP1) $\mathcal{E}_{i,j} = \varnothing$ unless $i \leq j$.

(CSP2) $\mathcal{E}_{i,i} = \{G_i\}$ for each i.

(CSP3) If $i \leq j \leq k$ in I, $G_iuG_j \in \mathcal{E}_{i,j}$, and $G_jvG_k \in \mathcal{E}_{j,k}$, then $G_iugvG_k \in \mathcal{E}_{i,k}$ for each $g \in G_j$.

Thus by Lemma 16.1, the relation defined by a coset poset is a partial order on X. Write $\leq$ for this partial order. Observe that our type function from X to I is a map of posets by axiom CSP1. Moreover the next lemma gives us a characterization of those posets which are coset posets. Notice that condition (*) of the lemma is always satisfied if the poset P is in $\mathcal{P}$, since morphisms in $\mathcal{P}$ preserve height.

(16.2) *Let G be represented as a group of automorphisms of the poset P and let $I = P/G$ be the orbit poset of this representation. That is $aG \leq bG$ if and only if $ag \leq b$ for some $g \in G$. Assume*

(*) *For each $a \in P$, $aG \cap P(< a) = \varnothing$.*

Let $(x_i : i \in I)$ be a set of representatives for the orbits of G on P with $i = x_iG$ and let $G_i = G_{x_i}$, $\mathcal{F} = (G_i : i \in I)$, and $\mathcal{E}_{i,j} = \{G_iuG_j : x_iu \leq x_j\}$. Then $\Gamma = \Gamma(G, \mathcal{F}, \mathcal{E})$ is a coset poset and the map $x_ig \mapsto G_ig$ is a G-equivariant isomorphism of posets.

Proof. By construction, Γ satisfies axioms CSP0 and CSP1. Axiom CSP2 follows from hypothesis (*). Finally if $G_iuG_j \in \mathcal{E}_{i,j}$ and $G_jvG_k \in \mathcal{E}_{j,k}$ then $x_iug \leq x_j$ for each $g \in G_j$ and $x_j \leq x_kv^{-1}$, so as P is a poset, $x_iug \leq x_kv^{-1}$, and hence $x_iugv \leq x_k$. Therefore $G_iugvG_k \in \mathcal{E}_{i,k}$, so axiom CSP3 is satisfied and Γ is a coset poset. As $G_i = G_{x_i}$, our map is a G-equivariant bijection. Finally $x_ig \leq x_jh$ if and only if $G_igh^{-1}G_j \in \mathcal{E}_{i,j}$ if and only if $G_ig \leq G_jh$, so the map and its inverse preserve order.

(16.3) *Let $\Gamma = \Gamma(G, \mathcal{F}, \mathcal{E})$ and $\bar{\Gamma} = \Gamma(\bar{G}, \bar{\mathcal{F}}, \bar{\mathcal{E}})$ be coset posets over I and $\bar{I}$, respectively, let $\beta : I \to \bar{I}$ be a map of posets, and let $\alpha : G \to \bar{G}$ a group homomorphism such that $\alpha(G_i) \subseteq \bar{G}_{\beta(i)}$ and $\alpha(\mathcal{E}_{ij}) \subseteq \bar{\mathcal{E}}_{\beta(i),\beta(j)}$ for each $i, j \in I$. Then the map $G_ix \mapsto \bar{G}_{\beta(i)}\alpha(x)$ is a map of posets from Γ into $\bar{\Gamma}$.*

Proof. As $\alpha(G_i) \subset \bar{G}_\beta(i)$, the map is well defined. As $\alpha(\mathcal{E}_{i,j}) \subseteq \bar{\mathcal{E}}_{\beta(i),\beta(j)}$, the map preserves order.

Let $\alpha : \tilde{P} \to P$ be a map of posets. We say α is a *lower covering* if for all $\tilde{a} \in \tilde{P}$, the restriction $\tilde{\alpha}_{\tilde{a}} : \tilde{P}(\leq \tilde{a}) \to P(\leq \alpha(\tilde{a}))$ is an isomorphism. The lower covering is *restricted* if α is injective on $\tilde{P}(\geq \tilde{a})$.

The *cone* of P is the poset $CP = P \cup \{x_*\}$ obtained by adjoining an element $x_* > a$ for all $a \in P$.

(16.4) *Assume $\alpha : \tilde{P} \to P$ is a lower covering of posets with $P \in \mathcal{P}$ of height n, and let (P, f) be the simplicial cell complex of P. Let $X = P \cup \{x_*\}$ be the cone of P. Then (X, f) is a combinatorial cell complex, where (X, f) has n-skeleton (P, f), $\dot{f}(x_*) = \tilde{P}$, $\zeta_{|\dot{f}(x_*)} = \alpha$, and $f_{\tilde{a}} = \alpha_{\tilde{a}}$ for each $\tilde{a} \in \dot{f}(x_*)$. If α is restricted, so is (X, f).*

Proof. Straightforward.

(16.5) *Assume (X, f) is a restricted combinatorial cell complex of height $n+1$, $\dot{f}(x)$ is homology spherical for each $x \in X$, and $H_n(X^n) = 0$. Then $\dim(H_{n+1}(X)) = \sum_{x\in X} \dim(\tilde{H}_n(\dot{f}(x)))$, the $(n+1)$st cellular boundary map ∂_{n+1} of X is 0, and $H_n(X) = 0$.*

Proof. As $\dot{f}(x)$ is homology spherical for each $x \in X$, $H^c_*(X) = H_*(X)$ by Theorem 12.16.

Suppose $\partial_{n+1} = 0$. Then $H^c_{n+1}(X) = Z^c_{n+1}(X) = D_{n+1}(X)$ has rank

$$\sum_{x\in X} \dim(\tilde{H}_n(\dot{f}(x)))$$

by 12.7. Also $H^c_n(X) = Z^c_n(X)/B^c_n(X) = Z^c_n(X)$ as $B^c_n(X) = \partial_{n+1}(D_{n+1}(X)) = 0$. Therefore $H_n(X) = H^c_n(X) = Z^c_n(X) = Z^c_n(X^n) = H^c_n(X^n) = H_n(X^n) = 0$.

So it remains to show $\partial_{n+1} = 0$. Assume not. Then $0 \neq B^c_n(X) \leq Z^c_n(X) = Z^c_n(X^n) = H^c_n(X^n) = H_n(X^n) = 0$, a contradiction.

Corollary 16.6. *Assume (X, f) is a restricted combinatorial cell complex of height $n+1$ such that $\dot{f}(x)$ is homology spherical for each $x \in X$ and X^n is acyclic. Then X is homology spherical with* $\dim(H_{n+1}(X)) = \sum_{x\in X} \dim(\tilde{H}_n(\dot{f}(x)))$.

Example. We consider the classical example of the Poincaré dodecahedron and the Poincaré dodecahedron disk. We regard the dodecahedron as the poset $\tilde{X}$ of faces partially ordered by inclusion. Then $\tilde{X}$

has one element x_* of height 3, 12 2-dimensional faces of height 2, 30 1-dimensional faces of height 1, and 20 vertices of height 0.

The Coxeter group W of type H_3 is isomorphic to $\mathbf{Z}_2 \times A_5$, where A_5 is the alternating group of degree 5, and we will see that the Coxeter complex of W is the order complex of the boundary $\tilde{X}(< x_*)$ of the dodecahedron. Namely W is regular on 3-chains in $\tilde{X}$ and if $c = (x_0 < \cdots < x_3)$ is a 3-chain and $c_i = c - \{x_i\}$ for $0 \leq i \leq 2$, then $W_{c_i} = \langle r_i \rangle$ is of order 2 and $R = \{r_0, r_1, r_2\}$ is the set of fundamental reflections making (W, R) a Coxeter system of type H_3. Further the stabilizer of the chain $c_J = c - \{x_j : j \in J\}$ is the parabolic $W_J = \langle R_J \rangle$, where $R_J = \{r_J : j \in J\}$. Let $I = \{0, 1, 2\}$ ordered as usual and $M_i = W_{I-\{i\}}$ be the stabilizer of x_i. The Coxeter complex of W is the order complex of the coset poset $\Gamma(W, \mathcal{F}^*, \mathcal{E}^*)$, where $\mathcal{F}^* = (M_i : i \in I)$ and $\mathcal{E}^*_{i,j} = \{M_i M_j\}$. By 16.2, $\tilde{X}(< x_*)$ is isomorphic to this coset poset and hence its order complex is isomorphic to the Coxeter complex.

Next the commutator group of W is the alternating group $G = A_5$ on $\{1, 2, 3, 4, 5\}$ and as $W_J \not\leq G$ for $J \subset I$, $W = W_J G$ for all such J so G is transitive on pairs (x, y) in $\tilde{X}$ with $y \leq x$, $h(x) = i$, and $h(y) = j$, for all $0 \leq j \leq i \leq 3$. On the other hand G has two orbits on 3-chains of $\tilde{X}$. Let $\tilde{G}_i = G \cap M_i = G_{x_i}$, and $\tilde{\mathcal{F}} = (\tilde{G}_i : 0 \leq i \leq 3)$. Then $\tilde{G}_3 = G$, $\tilde{G}_2 \cong \mathbf{Z}_5$, $\tilde{G}_1 \cong \mathbf{Z}_2$, and $\tilde{G}_0 \cong \mathbf{Z}_3$. By 16.2, $\tilde{X} \cong \Gamma(G, \tilde{\mathcal{F}}, \tilde{\mathcal{E}})$, where $\tilde{\mathcal{E}_{j,i}} = \{\tilde{G}_j \tilde{G}_i\}$ for $j \leq i$.

Let $\tilde{G}_i = \langle g_i \rangle$ for $0 \leq i \leq 2$. As the pair $\langle g_0 \rangle$, $\langle g_2 \rangle$ is determined up to conjugation in $Aut(G) = S_5$, we may take $g_2 = (1, 2, 3, 4, 5)$ and $g_0 = (2, 3, 5)$. As r_1 inverts g_0 and g_2, so does its projection p_1 on G, so $p_1 = (1, 4)(2, 3)$. Then we may take the projection p_2 of r_2 to be $p_2 = p_1 g_2 = (1, 5)(2, 4)$. Next the projection p_0 centralizes p_2 and inverts g_0, so $p_0 = (1, 4)(2, 5)$. Finally $g_1 = p_0 p_2 = (1, 2)(4, 5)$.

Next let G_0 be the stabilizer on G of the point 1, G_1 the global stabilizer of $\{1, 2\}$, and $G_2 = N_G(\tilde{G}_2)$. Let $\mathcal{F} = (G_i : 0 \leq i \leq 2)$ and consider the coset poset $P = \Gamma(G, \mathcal{F}, \mathcal{E})$, where $\mathcal{E}_{i,j} = \{G_i G_j\}$. Now $G_1 = G_{0,1} G_{1,2}$, so G is transitive on 2-chains of the poset P. The poset P is the *Poincaré dodecahedron disk.* It is well known that

(16.7) *Let P be the Poincaré dodecahedron disk. Then P is aspherical with $\pi_1(P) \cong SL_2(5)$.*

Proof. See for example [4].

Let $\bar{\mathcal{F}} = (\tilde{G}_i : 0 \leq i \leq 2)$ and $\tilde{P} = \Gamma(G, \bar{F}, \bar{\mathcal{E}})$ the coset poset with $\bar{\mathcal{E}} = \{\mathcal{E}_{j,i} : 0 \leq j \leq i \leq 2\}$. That is $\tilde{P}$ is the 2-skeleton or boundary of the dodecahedron. Thus the geometric realization of $\tilde{P}$ is the 2-sphere.

Define $\alpha : \tilde{P} \to P$ to be the map induced via the identity homomorphism on G and the identity map on $\{0,1,2\}$ as in Lemma 16.3. Observe that $\tilde{G}_i \leq G_i$ and $E \subseteq G_j G_i$ for all $j \leq i$ and all $E \in \bar{\mathcal{E}}_{i,j}$, so α is a map of posets by 16.2. Indeed α is a restricted lower covering. Therefore if we form the cone $X = \{x_*\} \cup P$ of P and make the construction of Lemma 16.4, we obtain a restricted combinatorial cell complex (X, f) whose 2-skeleton X^2 is the simplicial cell complex (P, f) of the Poincaré dodecahedron disc. We call X the *Poincaré dodecahedron*, since the geometric realization of X *is* the Poincaré dodecahedron.

By Lemma 16.7, $X^2 = P$ is acyclic. Also $\dot{f}(x_*) = \tilde{P}$ has the homotopy type of the 2-sphere, so $\dot{f}(x_*)$ is spherical with $H_2(\dot{f}(x_*)) \cong \mathbf{Z}$. Finally $\dot{f}(x)$ is isomorphic to the 5-gon, and the 0-sphere for x of height 2,1, respectively, so $\dot{f}(x)$ is homology spherical for all $x \in X$. Therefore by Corollary 16.6, X is homology spherical with $H_3(X) \cong \mathbf{Z}$; that is X is a homology 3-sphere. Also as $\dot{f}(x_*)$ has the homotopy type of the 2-sphere, it is simply connected, so by 12.17, $\pi_1(X) \cong \pi_1(X^2) \cong SL_2(5)$ by 16.7. We summarize this as:

(16.8) *Let X be the Poincaré dodecahedron. Then X is a homology 3-sphere with $\pi_1(X) \cong SL_2(5)$.*

Notice that G is a group of automorphisms of P transitive on 2-chains and hence also on 3-chains of X.

References

[1] M. Aschbacher, Simple connectivity of p-group complexes, Israel J. Math., **82** (1993), 1–43.

[2] M. Aschbacher and Y. Segev, Extending morphisms of groups and graphs, Ann. of Math., **135** (1992), 297–323.

[3] M. Aschbacher and Y. Segev, Locally connected simplical maps, Israel J. Math., **77** (1992), 285–303.

[4] M. Aschbacher and Y. Segev, A fixed point theorem for groups acting on finite 2-dimensional acyclic simplicial complexes, Proc. London Math. Soc., **67** (1993), 329–354.

[5] E. Spanier, "Algebraic Topology", McGraw-Hill, New York, 1966.

Department of Mathematics
California Institute of Technology
Pasadena, CA 91125-0001
U.S.A.

Advanced Studies in Pure Mathematics 24, 1996
Progress in Algebraic Combinatorics
pp. 81–100

Generalized Generalized Spin Models Associated with Exactly Solvable Models

Tetsuo Deguchi

Abstract.

We show a close connection between generalized generalized spin models (four-weight spin models) and exactly solvable models. The defining relations of four-weight spin models are discussed from the viewpoint of the star-triangle relations. For an illustration we derive symmetric spin models from the self-dual Z_N model, and then construct various four-weight spin models through gauge transformations. We also show that the gauge transformations do not change the values of the link invariants derived from the four-weight spin models.

§1. Introduction

In mathematical physics, there is a close connection between the theories of integrable models and the topological invariants of links [1, 22]. The key is the Yang-Baxter relation, which plays a central role in the 1-dim. many body system [24] and in solvable lattice models in statistical mechanics [5].

The formalism of spin model was introduced by Jones as a device for construction of link invariants and representations of the braid group [15]. We call topological invariant of knots and links link invariant. Various spin models and link invariants have been discussed [4, 10, 12, 13, 18]. Recently, the formalism of spin model due to Jones has been extended into that of generalized spin model [18] and further into that of generalized generalized spin model (four-weight spin model) [4]. Many four-weight spin models have been constructed explicitly [4, 16].

In this paper we discuss four-weight spin models from the viewpoint of the star-triangle relations of exactly solvable models in statistical mechanics. We show that certain limits of the Boltzmann weights of the

Received January 17, 1995.
Revised September 2, 1995.

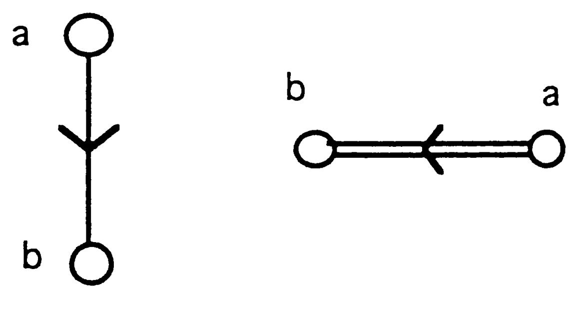
self-dual Z_N model give explicitly the symmetric spin models and the generalized generalized spin models. We introduce two gauge transformations on the Boltzmann weights of solvable models satisfying the star-triangle relations. Various four-weight spin models are constructed by applying the gauge transformations to the self-dual Z_N model. It is thus shown that these four-weight spin models can be generalized into the exactly solvable models, i.e., they can be "Yang-Baxterized". Furthermore, we show that the derived link invariants do not depend on the parameters related to the gauge transformations. We hope that the results of the paper might shed some light on connection of generalized generalized spin models to exactly solvable models in statistical mechanics.

After submission of the manuscript the author was informed that the gauge transformations for four-weight spin models have been independently introduced by F. Jaeger, who also proved that the derived link invariants are independent from the transformations [14].

§2. The star-triangle relations

2.1. The Z_N model

In statistical physics, a lattice model is called solvable if it has an infinite number of commuting transfer matrices [5]. If the Boltzmann weights satisfy the Yang-Baxter relations, then we can show that the transfer matrices commute, and that the system has an infinite number of "symmetries" or "conserved quantities". For Potts models and Z_N models the Yang-Baxter relations are called the star-triangle relations. Here Z_N denotes $\mathbf{Z}/N\mathbf{Z}$.



(1) (2)

Fig. 1. Two types of Boltzmann weights (1) $X(a, b; u)$ and (2) $Y(a, b; u)$. Arrows are often abbreviated.

Let us introduce a general Z_N model with nearest-neighbor interaction on an M-by-$2L$ rectangular lattice [9]. A lattice point is described by a vector $\vec{n} = (n_1, n_2)$, where $1 \le n_1 \le M$ and $1 \le n_2 \le 2L$. We assume the periodic boundary condition: $n_1 = n_1 + M(\mathrm{mod} M)$ and $n_2 = n_2 + 2L(\mathrm{mod} 2L)$.[1] To every lattice point $\vec{n}$ we associate a spin variable $\alpha(\vec{n})$, which takes Z_N values.

We introduce two types of Boltzmann weights $X(\alpha, \beta; u)$ and $Y(\alpha, \beta; u)$, where α and β are spin variables on two nearest-neighboring lattice points and u is a complex parameter called spectral parameter. Mathematically, the Boltzmann weights are functions on the spin variables and spectral parameter: $\mathbf{Z}_N \otimes \mathbf{Z}_N \otimes \mathbf{C} \to \mathbf{C}$. The physical meaning of the Boltzmann weight is that it denotes the probability of the two nearest-neighboring spin variables taking a particular value in $Z_N \otimes Z_N$, where the spectral parameter may play the role of the "temperature".

We now consider the solvability condition of the Z_N model. The star-triangle relations are given by the following.

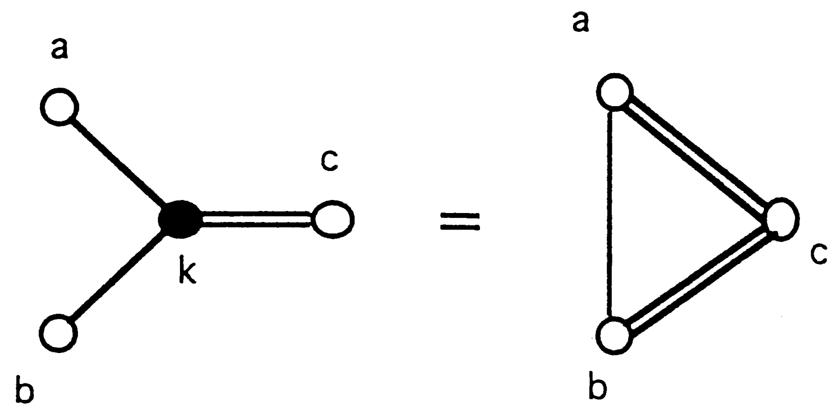


Fig. 2. Star-triangle relations:
$\sum X(a,k;u)X(k,b;v)Y(c,k;w)$
$= C(u,v)X(a,b;w)Y(c,b;u)Y(c,a;v)$

$$\sum_{k \in Z_N} X(\alpha, k; u) X(k, \beta; v) Y(\gamma, k; w) = C(u, v, w) X(\alpha, \beta; w) Y(\gamma, \beta; u) Y(\gamma, \alpha; v), \tag{2.1}$$

$$\sum_{k \in Z_N} X(k, \alpha; v) X(\beta, k; u) Y(k, \gamma; w) = C(u, v, w) X(\beta, \alpha; w) Y(\beta, \gamma; v) Y(\alpha, \gamma; u). \tag{2.2}$$

[1]Under the periodic boundary condition we may assume $\vec{n} \in \mathbf{Z}_M \otimes \mathbf{Z}_{2L}$.

Here the spectral parameter w is to be determined from the parameters u and v under a certain rule, which may be different depending on models.

In order to show the merit of the star-triangle relations we consider the transfer matrices of the Z_N model. Let $\vec{\alpha}$ and $\vec{\beta}$ be the sequences of the spin variables on the k-th and $(k+1)$-th rows of the M-by-$2L$ lattice, respectively, i.e., $\alpha_j = \alpha(k,j)$ and $\beta_j = \alpha(k+1,j)$. We define the transfer matrices $V_M(u)$ and $W_M(u)$ as follows

$$
\begin{aligned}
V_M(u)_{\vec{\alpha},\vec{\beta}} &= \prod_{j=1}^{M} \left(X(\alpha_j, \beta_j; u) Y(\alpha_{j+1}, \beta_j; u) \right), \\
W_M(u)_{\vec{\alpha},\vec{\beta}} &= \prod_{j=1}^{M} \left(X(\alpha_j, \beta_j; u) Y(\beta_{j+1}, \alpha_j; u) \right).
\end{aligned} \tag{2.3}
$$

Here we recall the periodic boundary condition that $\alpha_{M+1} = \alpha_1$ and $\beta_{M+1} = \beta_1$.

The transfer matrices are important in calculation of the partition function of the model. In terms of the transfer matrices the partition function Z of the lattice model is given by $Z = \mathrm{Tr}\left((V_M(u) W_M(u))^L \right)$.

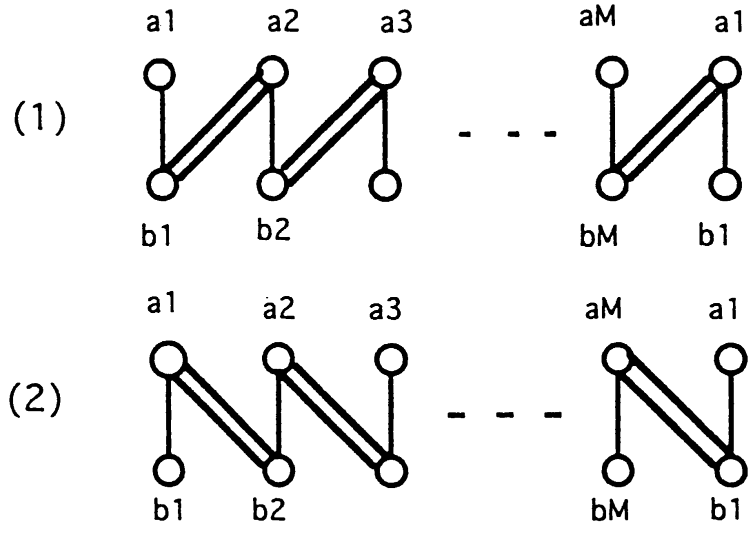


Fig. 3. (1) Transfer matrix
$V(u) = \prod X(\alpha_j, \beta_j; u) Y(\alpha_{j+1}, \beta_j; u)$
(2) Transfer matrix
$W(u) = \prod X(\alpha_j, \beta_j; u) Y(\beta_{j+1}, \alpha_j; u)$
Periodic boundary condition is assumed:
$\alpha_{M+1} = \alpha_1, \beta_{M+1} = \beta_1$.

Proposition 2.1. *If the star-triangle relations* (2.1) *and* (2.2) *hold, then we can show for any M that the transfer matrices commute*:

$$\begin{aligned} V_M(u)W_M(v) &= W_M(v)V_M(u), \\ V_M(u)V_M(v) &= V_M(v)V_M(u), \\ W_M(u)W_M(v) &= W_M(v)W_M(u). \end{aligned} \tag{2.4}$$

(For proof, see for instance, §7.2 of [5].)

2.2. Gauge transformations on solvable models

We consider two transformations on the Boltzmann weights of the models satisfying the star-triangle relations. [2]

For an arbitrary function $f(\alpha)$ on $\alpha \in Z_N$ we introduce the following transformation of weights from X and Y into $\tilde{X}$ and $\tilde{Y}$

$$\tilde{X}(\alpha,\beta;u) = X(\alpha,\beta;u) \times f(\alpha)/f(\beta), \quad \tilde{Y}(\alpha,\beta;u) = Y(\alpha,\beta;u), \tag{2.5}$$

for any $\alpha, \beta \in Z_N$.

Proposition 2.2. *Let us assume that the Boltzmann weights $X(\alpha,\beta;u)$ and $Y(\alpha,\beta;u)$ satisfy the star-triangle relations* (2.1) *and* (2.2), *and that $\tilde{X}$ and $\tilde{Y}$ are obtained from X and Y by the gauge transformation* 1. *Then the weights $\tilde{X}$ and $\tilde{Y}$ also satisfy the star-triangle relations* (2.1) *and* (2.2).

We now introduce another transformation by the following.

$$\tilde{Y}(\alpha,\beta;u) = Y(\alpha,\beta+\gamma;u), \quad \tilde{X}(\alpha,\beta;u) = X(\alpha,\beta;u), \tag{2.6}$$

for any α and $\beta \in Z_N$. We call the transformation gauge transformation 2.

Proposition 2.3. *Let us assume that the Boltzmann weights $X(\alpha,\beta;u)$ and $Y(\alpha,\beta;u)$ satisfy the star-triangle relations* (2.1) *and* (2.2), *and that $\tilde{X}$ and $\tilde{Y}$ are obtained from X and Y by the gauge transformation* 2. *If there is a function $g(m;u)$ on $Z_N \otimes \mathbf{C}$ such that $Y(\alpha,\beta;u) = g(\alpha-\beta;u)$, then the transformed Boltzmann weight $\tilde{Y}$ and the weight X satisfy the star- triangle relations* (2.1) *and* (2.2).

---

[2]For four-weight spin models essentially the same transformations were discussed independently by F. Jaeger [14].

§3. The self-dual Z_N model

3.1. The Boltzmann weights

We now introduce the Boltzmann weights of the self-dual Z_N model. Let r denote an integer coprime to N, $(N,r)=1$. For a positive integer m we introduce weight $x_m(u)$ by the following. [3]

(3.1)
$$x_m(u) = \prod_{k=0}^{m-1} \sin\left(\frac{r-N}{2N}u + \frac{\pi r k}{N}\right) \Big/ \sin\left(\frac{r-N}{2N}u - \frac{\pi r(k+1)}{N}\right).$$

For $m=0$ we assume $x_0(u)=1$. We may define x_{-m} for $m>0$ by $x_{-m}=x_m(u)$, since from (3.1) we have $x_{N-m}=x_m(u)$ for $0<m<N$.

We define $y_{\alpha-\beta}(u)$ by $y_{\alpha-\beta}(u) = x_{\alpha-\beta}(\pi-u)$. The Boltzmann weights $\{x_m(u)\}$ satisfy the star-triangle relation [9]:

(3.2)
$$\sum_{k=0}^{N-1} x_{\alpha-k}(u)x_{k-\beta}(v)y_{\gamma-k}(u+v) = C(u,v)x_{\alpha-\beta}(u+v)y_{\gamma-\beta}(u)y_{\gamma-\alpha}(v),$$

where $C(u,v)=F_0(u)F_0(v)/F_0(u+v)$. $F_j(u)$ denotes the Fourier transform of $x_m(u)$

$$F_j(u) = \sum_{k=0}^{N-1} x_k(u)\omega^{jk}, \tag{3.3}$$

where ω is given by $\omega = \exp(2\pi\sqrt{-1}/N)$. The star-triangle relations (3.2) correspond to (2.1) and (2.2) by

$$X(\alpha,\beta;u) = x_{\alpha-\beta}(u), \qquad Y(\alpha,\beta;u) = y_{\alpha-\beta}(u) = x_{\alpha-\beta}(\pi-u). \tag{3.4}$$

We make a remark on the proof of the star-triangle relations (3.2) of the model. The self-dual Z_N model corresponds to the cyclotomic solution in[17], whose star-triangle relations are explicitly proved in Ref. [17].

One of the most characteristic properties of the self-dual Z_N model is the following relation called self-duality [9]:

$$x_m(\pi-u) = \tilde{x}_m(u) = F_m(u)/F_0(u). \tag{3.5}$$

The Boltzmann weights $x_m(u)$ also satisfy the following basic properties.

[3] The parametrization in Ref. [9] corresponds to the case $r=N+1$.

1. The initial conditions:

$$x_m(0) = \delta_{m0}, \qquad x_m(\pi) = 1. \tag{3.6}$$

2. The inversion relations:

$$x_m(\pi + u)x_m(\pi - u) = 1, \tag{3.7}$$

$$\sum_k x_{\alpha-k}(u)x_{k-\beta}(-u) = C(u,\pi)\delta_{\alpha,\beta}. \tag{3.8}$$

It is straightforward to see the initial conditions (3.6) and the inversion relation (3.7) from the parametrization (3.1). We can show (3.8) by putting $v = \pi$ in the star-triangle relation (3.2) and using (3.6) and (3.7).

We give a comment on solvable models associated with the self-dual Z_N model. Two families of generalizations are known for the model. The chiral Potts model [2] and the broken Z_N symmetric model [17]. The latter model has been reconstructed from the representation of the Sklyanin algebra [11], and has also been generalized into $Z_N^{\otimes n-1}$-symmetric model [21]. We may say that the chiral Potts model and the broken Z_N symmetric model are associated with the generalized generalized spin models which will be discussed in §4. For the chiral Potts model a method for constructing multicomponent models has been given [3].

3.2. Modification of the self-dual Z_N model

We now apply the gauge transformations (2.5) and (2.6) to the Boltzmann weight $x_m(u)$ (3.1) of the self-dual Z_N model so that we obtain a modified solvable model. Let us introduce two types of Boltzmann weights $y(\alpha,\beta;u)$ and $x(\alpha,\beta;u)$ by $y(\alpha,\beta;u) = y_{\alpha-\beta+s}(u) = x_{\alpha-\beta+s}(\pi-u)$ and $x(\alpha,\beta;u) = x_{\alpha-\beta}(u) \times f(\alpha)/f(\beta)$, respectively. Here $\alpha,\beta \in Z_N$. Then the Boltzmann weights satisfy the star-triangle relations (2.1) and (2.2) in §2 and the following basic relations.

$$\begin{gathered} x(\alpha,\beta;0) = \delta_{\alpha,\beta}, \qquad x(\alpha,\beta;\pi) = 1, \\ x(\alpha,\beta;\pi+u)x(\beta,\alpha;\pi-u) = 1, \\ \sum_k x(\alpha,k;u)x(k,\beta;-u) = C(u,\pi)\delta_{\alpha,\beta}, \end{gathered} \tag{3.9}$$

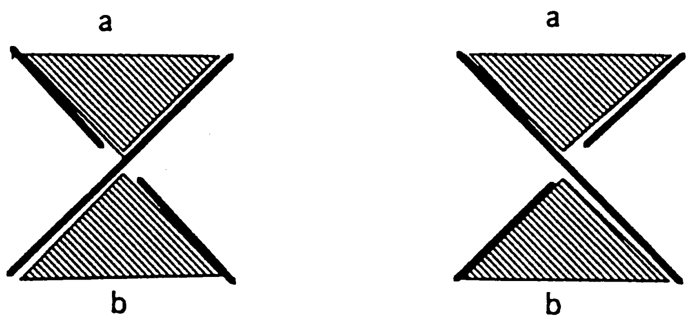
$$\begin{gathered} y(\alpha,\beta;0) = \delta_{\alpha,\beta}, \qquad y(\alpha,\beta;\pi) = 1, \\ y(\alpha,\beta;\pi+u)y(\alpha,\beta;\pi-u) = 1, \\ \sum_k y(\alpha,k;u)y(\beta,k;-u) = C(u,\pi)\delta_{\alpha,\beta}, \end{gathered} \tag{3.10}$$

for any $\alpha,\beta \in Z_N$.

§4. Spin models

4.1. Symmetric spin models

Let us introduce the symmetric spin models [15].



(1) (2)

Fig. 4. Double points (1) of type $+$ and (2) of type $-$. Spin variables a and b are defined on the shaded regions.

Let S denote a finite set $S = \{1, \ldots, n\}$. We consider two types of weights w_j $(j = \pm)$, which are complex functions on $S \otimes S$. For α and $\beta \in S$, the weight w_j gives a complex number $w_j(\alpha, \beta)$ for $j = \pm$.

Definition 4.1 ([15]). A set of weights $\{w_j; j = \pm\}$ is called a symmetric spin model, if the weights satisfy the following relations.

(4.1) $w_+(\alpha, \beta) = w_+(\beta, \alpha), \qquad w_-(\alpha, \beta) = w_-(\beta, \alpha),$

(4.2) $w_+(\alpha, \beta) w_-(\alpha, \beta) = 1,$

(4.3) $\sum_{k \in S} w_+(\alpha, k) w_-(k, \beta) = n \delta_{\alpha\beta},$

(4.4) $\sum_{k \in S} w_+(\alpha, k) w_+(k, \beta) w_-(\gamma, k) = \sqrt{n} w_+(\alpha, \beta) w_-(\gamma, \alpha) w_-(\gamma, \beta).$

Let us derive symmetric spin models from the self-dual Z_N model by taking advantage of the star-triangle relation (3.2). The set S is given by $S = Z_N$, i.e., n is given by N. Let $G(N, r)$ denote the Gaussian sum for integers N and r with $(N, r) = 1$: $G(N, r) = \sum_{k=0}^{N-1} \exp(\pi \sqrt{-1} r k^2 / N)$. We now define weights $w_\pm(\alpha, \beta)$ by the limit of sending the spectral parameter to infinity: [4]

$$w_{\pm}(\alpha,\beta)=\lim_{u\to\infty}x_{\alpha-\beta}(\pm u/\sqrt{-1})\left(\frac{\sqrt{N}}{G(N,r)}\right)^{\pm 1/2} \tag{4.5}$$

Then we have

$$w_{\pm}(\alpha,\beta)=\exp(\pm\pi\sqrt{-1}r(\alpha-\beta)^2/N)\left(\frac{\sqrt{N}}{G(N,r)}\right)^{\pm 1/2} \tag{4.6}$$

Here we have used the fact that

$$\lim_{u\to\infty}x_m(\pm u/\sqrt{-1})=\exp(\pm\pi\sqrt{-1}rn^2/N)$$

for $r>N$.

Let us show that the weights (4.6) defined by the limit (4.5) satisfy the defining relations of symmetric spin models. We first note the following

$$\lim_{u\to\infty}F_0(u/\sqrt{-1})=\lim_{u\to\infty}\sum_k x_k(u/\sqrt{-1})=G(N,r), \tag{4.7}$$

$$\lim_{u\to\infty}C(u/\sqrt{-1},\pi)=\lim_{u\to\infty}\frac{F_0(u/\sqrt{-1})F_0(\pi)}{F_0(u/\sqrt{-1}+\pi)}=N. \tag{4.8}$$

Here $F_0(\pi)=N$ since $x_k(\pi)=1$ (see (3.6)). We next note that the weights $w_{\pm}(\alpha,\beta)$ defined by (4.5) are symmetric, since $x_m(u)=x_{-m}(u)$. The inversion relation (3.7) leads to the relation (4.2). The relation (4.3) is derived from the inversion relation (3.8). Finally, from the star-triangle relation (3.2) we have the relation (4.4), since we have $\lim_{u,v\to\infty}C(u,v)=G(N,r)$ from (4.7) .

Thus we have shown that the weights defined by (4.5) give a symmetric spin model. From the derivation we see that the relations (4.3) and (4.4) are consequences of the star-triangle relation (3.2).

A version of symmetric spin model has been constructed in terms of the IRF model (the cyclotomic solution) and the link invariants derived from it have been extensively investigated [19]. It is interesting to note that the symmetric spin model is related to the invariants of 3 dimensional manifolds [20]. A certain class of the spin models associated with $sl(n)$ generalization of the Chiral Potts model [7] has been also shown to be derived from the invariants of 3 dimensional manifolds [20].

---

[4] Hereafter we assume that $r>N$ without loss of generality.

Recently, from the Boltzmann weights of the chiral Potts model an extended version of spin model has been constructed [23]. We can regard it as a four-weight spin model.

4.2. Four-weight spin models

Let us introduce four-weight spin model (generalized generalized spin model) [4]. Recall that S denotes a finite set $S = \{1, \ldots, n\}$. We consider four types of weights w_j ($j = 1, \ldots, 4$), which are complex functions on $S \otimes S$. For α and $\beta \in S$, the weight w_j gives a complex number $w_j(\alpha, \beta)$.

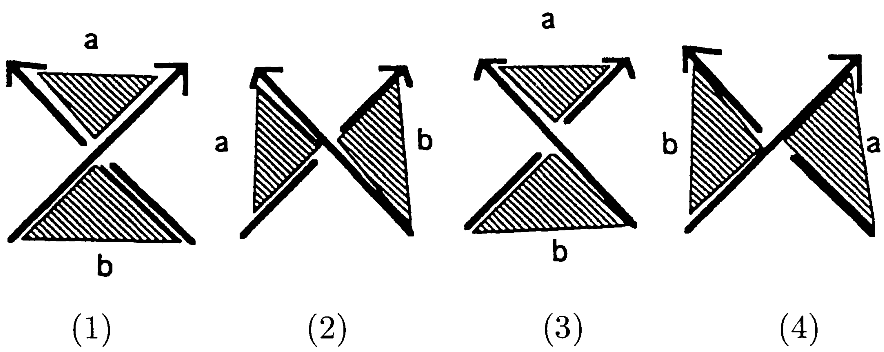


Fig. 5. Double point with index j is defined by figure (j) for $j = 1, 2, 3, 4$. Weight $w_j(a, b)$ is associated with figure (j).

Definition 4.2 ([4]). A set of weights $\{w_j; j = 1, \ldots, 4\}$ is called a four-weight spin model, if the weights satisfy the following relations.

$$w_1(\alpha, \beta) w_3(\beta, \alpha) = 1, \qquad w_2(\alpha, \beta) w_4(\beta, \alpha) = 1, \tag{4.9}$$

$$\begin{aligned} \sum_k w_1(\alpha, k) w_3(k, \beta) &= n \delta_{\alpha,\beta}, \\ \sum_k w_2(\alpha, k) w_4(k, \beta) &= n \delta_{\alpha,\beta}, \end{aligned} \tag{4.10}$$

$$\begin{aligned} \sum_k w_1(\alpha, k) w_1(k, \beta) w_4(\gamma, k) &= \sqrt{n} w_1(\alpha, \beta) w_4(\gamma, \alpha) w_4(\gamma, \beta), \\ \sum_k w_1(k, \alpha) w_1(\beta, k) w_4(k, \gamma) &= \sqrt{n} w_1(\beta, \alpha) w_4(\alpha, \gamma) w_4(\beta, \gamma), \end{aligned} \tag{4.11}$$

for any α, β and γ in S.

Let us introduce graph of an oriented link [6].

Definition 4.3 (graph of an oriented link). Let a regular link diagram be chess-board colored with colors black (or shade) and white. Assign to every double point A^k of the diagram an index 1, 2, 3, or 4 with respect to the colorings as defined by Fig. 5. Denote by S_j the black-colored (or shaded) regions of a link diagram. Define a graph Γ whose vertices P_i correspond to the S_i, and whose edges a^k_{ij} correspond to the double points A^k that is the intersection between regions S_i and S_j, where a^k_{ij} connects P_i and P_j and has the index of A^k.

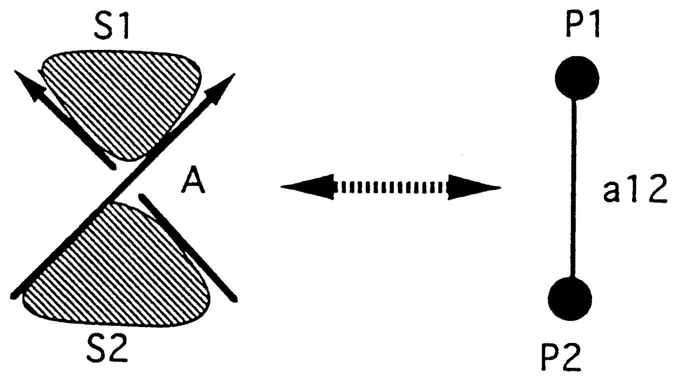


Fig. 6. Double point A with regions $S1$ and $S2$ corresponds to edge $a12$ with vertices $P1$ and $P2$. The index of A is 1.

We now define configuration of spin variables. To each region S_j we associate a spin variable which takes its values in Z_N. A configuration of the spin variables on the graph is given by assigning an element in Z_N to each spin variable on the graph. There are N^m configurations if the graph has m vertices.

We shall define the partition function of a four-weight spin model. Let us consider a link diagram, and then by chess-board coloring of the diagram we make the graph of the link with spin variables assigned on the shaded regions. We assign the weight w_j to each edge in the graph with index j. We make the product of the weights w_j for all the edges in the graph under a given configuration of the spin variables. Then we take the sum of the product over all the configurations of the spin variables. By the sum we define the partition function of the four-weight spin model for the graph of the link diagram.

Proposition 4.4 ([4]). *The partition function of a four-weight model leads to a link invariant.*

We now construct a generalized generalized spin model from the modified solvable model given in §3.2. We define the weights $w_j(\alpha,\beta)$ for $j=1,2,3,4$ by the following limit:

$$(4.12)\qquad \begin{aligned}
w_1(\alpha,\beta) &= \lim_{u\to\infty} x(\alpha,\beta;u/\sqrt{-1})\left(\frac{\sqrt{N}}{G(N,r)}\right)^{1/2},\\
w_2(\alpha,\beta) &= \lim_{u\to\infty} y(\beta,\alpha;-u/\sqrt{-1})\left(\frac{\sqrt{N}}{G(N,r)}\right)^{1/2},\\
w_3(\alpha,\beta) &= \lim_{u\to\infty} x(\alpha,\beta;-u/\sqrt{-1})\left(\frac{\sqrt{N}}{G(N,r)}\right)^{-1/2},\\
w_4(\alpha,\beta) &= \lim_{u\to\infty} y(\alpha,\beta;u/\sqrt{-1})\left(\frac{\sqrt{N}}{G(N,r)}\right)^{-1/2},
\end{aligned}$$

where $\alpha,\beta\in Z_N$. Then we have

$$(4.13)\qquad \begin{aligned}
w_1(\alpha,\beta) &= \frac{f(\alpha)}{f(\beta)}\exp(\frac{\pi\sqrt{-1}}{N}r(\alpha-\beta)^2)\left(\frac{\sqrt{N}}{G(N,r)}\right)^{1/2},\\
w_2(\alpha,\beta) &= \exp(\frac{\pi\sqrt{-1}}{N}r(\alpha-\beta-s)^2)\left(\frac{\sqrt{N}}{G(N,r)}\right)^{1/2},\\
w_3(\alpha,\beta) &= \frac{f(\alpha)}{f(\beta)}\exp(-\frac{\pi\sqrt{-1}}{N}r(\alpha-\beta)^2)\left(\frac{\sqrt{N}}{G(N,r)}\right)^{-1/2},\\
w_4(\alpha,\beta) &= \exp(-\frac{\pi\sqrt{-1}}{N}r(\alpha-\beta+s)^2)\left(\frac{\sqrt{N}}{G(N,r)}\right)^{-1/2},
\end{aligned}$$

for $\alpha,\beta\in Z_N$.

It is a consequence of the star-triangle relations that the weights defined by (4.12) satisfy the relations (4.11). In the same way as the symmetric spin model, we can derive the other defining relations from the basic relations (3.9) and (3.10). Thus all the defining relations are shown by using the star-triangle relations and the basic relations of the solvable model.

§5. Discussion

5.1. Gauge transformations on spin models

In terms of spin models we discuss the gauge transformations in §2 defined on the exactly solvable models.

We first consider the gauge transformation 1. For an arbitrary function $f(\alpha)$ we consider the following transformation of the weight w_j into $\tilde{w}_j$:

$$\begin{aligned} \tilde{w}_1(\alpha,\beta) &= w_1(\alpha,\beta) \times f(\alpha)/f(\beta), \quad \tilde{w}_2(\alpha,\beta) = w_2(\alpha,\beta), \\ \tilde{w}_3(\alpha,\beta) &= w_3(\alpha,\beta) \times f(\alpha)/f(\beta), \quad \tilde{w}_4(\alpha,\beta) = w_4(\alpha,\beta), \end{aligned} \tag{5.1}$$

for any α and $\beta \in S$. Then we have the following.

Proposition 5.1. *Assume that $\{w_j; j = 1,\dots,4\}$ is a four-weight spin model, and also that $\tilde{w}_j$ is obtained from w_j by the gauge transformation* 1. *Then $\{\tilde{w}_j; j = 1,\dots,4\}$ gives a four-weight spin model.*

We now consider the gauge transformation 2 given by the following.

$$\begin{aligned} \tilde{w}_1(\alpha,\beta) &= w_1(\alpha,\beta), \quad \tilde{w}_2(\alpha,\beta) = w_2(\alpha+s,\beta), \\ \tilde{w}_3(\alpha,\beta) &= w_3(\alpha,\beta), \quad \tilde{w}_4(\alpha,\beta) = w_4(\alpha,\beta+s), \end{aligned} \tag{5.2}$$

for any α and $\beta \in S$. It is easy to see the following.

Proposition 5.2. *Let us assume that $\{w_j; j = 1,\dots,4\}$ is a four-weight spin model, and that $\tilde{w}_j$ is obtained from w_j by the gauge transformation* 2. *If there are functions z_1 and z_2 such that $w_2(\alpha,\beta) = z_1(\alpha-\beta)$ and $w_4(\alpha,\beta) = z_2(\alpha-\beta)$, then $\{\tilde{w}_j; j = 1,\dots,4\}$ gives a four-weight spin model.*

We recall that the transformations 1 and 2 were independently introduced by F. Jaeger [14].

5.2. Link invariants being independent of the gauge transformations

Let us discuss the link invariants derived from those four-weight spin models that are obtained from the symmetric spin model by applying the gauge transformations 1 or 2. We show that these link invariants are independent of any parameter related to the gauge transformations.

In order to show the independence explicitly, we introduce Seifert cycle and smoothing operation [6]. Let $p(L)$ be a regular projection of an oriented link L. By altering $p(L)$ in the neighborhood of double points as shown in Fig. 7, $p(L)$ dissolves into a number of disjoint oriented

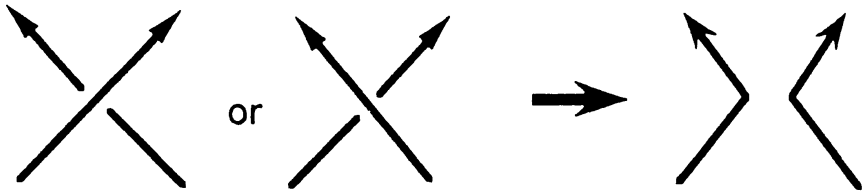


Fig. 7. Smoothing operation.

simple closed curves which are called Seifert cycles. We call this altering procedure smoothing operation.

We first consider the gauge transformation 1. The weight $\tilde{w}_j$ may depend on complex-valued parameters. For instance, we can set $f(\alpha) = a^{\alpha}$, where a is a complex variable. However, we can show the following.

Proposition 5.3. *Let us assume that weights* $\{\tilde{w}_j; j = 1, 2, 3, 4\}$ *are obtained from a four-weight spin model* $\{w_j; j = 1, 2, 3, 4\}$ *through the gauge transformation 1 with the function* $f(\alpha)$. *Then the link polynomial derived from the weights* $\{\tilde{w}_j; j = 1, 2, 3, 4\}$ *coincides with that of the spin model* $\{w_j; j = 1, 2, 3, 4\}$.

Proof. To a given link diagram we apply the smoothing operation. Suppose that the diagram (or the graph of the link diagram) has at least one double point with index 1 or 3. Then the two black-colored regions separated by the double point are connected by the smoothing operation. Thus we see that all the edges with indices 1 or 3 make cycles in the graph. If we take the product of the weights $\{\tilde{w}_j; j = 1, 2, 3, 4\}$ for all the edges, then the factors $f(\alpha)$ cancel each other out for each such cycles

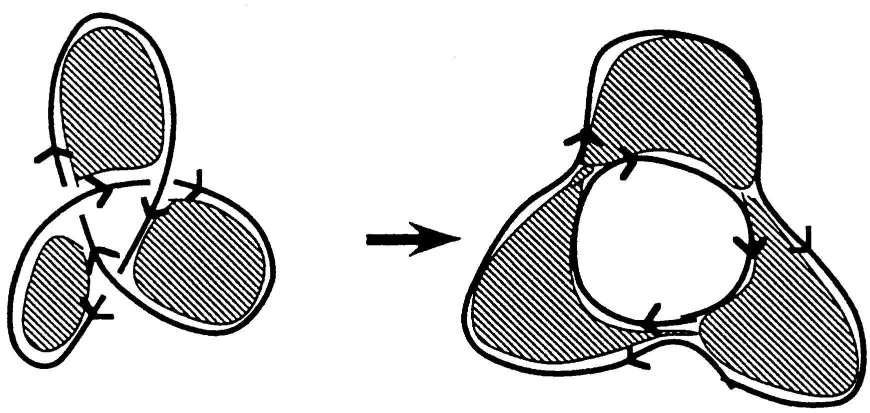


Fig. 8. Smoothing operation connects regions separated by the double points with index 1 or 3.

under any given configuration of spin variables. Therefore the partition function is independent of the function $f(\alpha)$. If the link diagram does not have any double point with index 1 or 3, then by definition the partition function does not depend on the gauge transformation 1.

We next consider the gauge transformation 2. Let us introduce special projection [6].

Definition 5.4. Let $P(L)$ be a regular projection of a link L on a plane. Choose a chess-board coloring on $P(L)$ such that the infinite region is white-colored. $P(L)$ is called a special projection if the union of black-colored regions is the image of a Seifert surface of L under the projection.

Using the same techniques in the proof of Prop. 13.15 of [6], we have the following.

Proposition 5.5. *Every link has a special projection.*

If we apply the smoothing operation to a special projection, then the image of the Seifert surface is divided into sub-regions encircled by the Seifert cycles. These sub-regions are disjoint and do not overlap each other, and they correspond to the black-colored regions by the definition of special projection. Thus we see that for a special projection every black-colored region is encircled by a Seifert cycle.

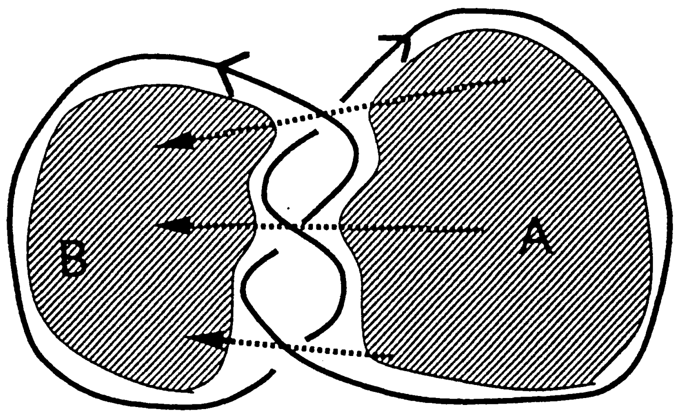


Fig. 9. A special projection of trefoil knot. Arrows denote transverse directions at double points. Positive region A and negative region B.

Note that a Seifert cycle has either a clockwise or counterclockwise orientation. Let us call a black-colored region positive (negative), if it is encircled by a clockwise (counterclockwise) Seifert cycle. We note the following.

Lemma 5.6. *If two Seifert cycles in a link diagram share a double point, then they have different orientations; if one has counterclockwise (clockwise) orientation then the other has clockwise (counterclockwise) orientation.*

We now consider the graph Γ of an oriented link L for its special projection $P(L)$. Any vertex of the graph Γ corresponds to either a positive or negative region of $P(L)$. We call a vertex positive (negative) if it corresponds to a positive (negative) region of $P(L)$. From Lemma 5.6 and the discussion in the previous paragraphs we have the following.

Lemma 5.7. *The graph of a special projection is a bipartite graph such that any of the vertices is either positive or negative and that every vertex connected to a positive (negative) vertex is negative (positive).*

For a special projection all the double points are of index 2 or 4. If there exists a double point of index 1 or 3, then the boundaries of the black-regions around the double point are not Seifert cycles, which is not consistent with the definition of special projection. Thus, every edge of the graph of a special projection is of index 2 or 4.

Let us introduce two directions with respect to a double point as shown in Fig. 10. For a double point we define smoothing direction by the direction of the tangential vector of the Seifert cycle at the double point. Transverse direction is defined by the direction obtained by rotating the smoothing direction 90 degree counterclockwise around the double point.

For a special projection the transverse direction at a double point is consistent with the direction from the positive to negative regions around the double point. Denote $S(\alpha_i)$ $(S(\beta_j))$ the positive (negative) regions of a special projection. We also denote by α_i (β_j) the spin variables defined on the positive (negative) regions. If a double point A^k is in the intersection of the boundaries of positive region $S(\alpha_i)$ and negative region $S(\beta_j)$ in a special projection, then the transverse direction at A^k is consistent with the direction from $S(\alpha_i)$ to $S(\beta_j)$. (See Fig. 9.)

Let us introduce weights y_2 and y_4 given by the following

(5.3)
$$y_2(\gamma_1,\gamma_2)=w_2(\gamma_2,\gamma_1), \quad y_4(\gamma_1,\gamma_2)=w_4(\gamma_1,\gamma_2), \qquad \gamma_1,\gamma_2\in Z_N.$$

Then at the double point A^k the weight $y_k(\alpha_i,\beta_j)$ with k=2 or 4 is assigned.

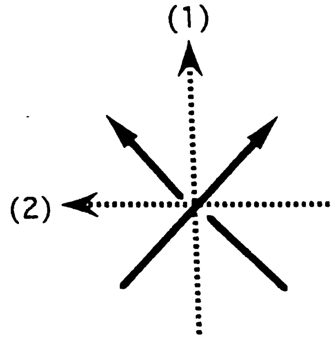


Fig. 10. (1) Smoothing direction. (2) Transverse direction.

Under the gauge transformation 2 the weights y_2 and y_4 are transformed into $\tilde{y}_2$ and $\tilde{y}_4$ defined by the following

$$\tilde{y}_k(\alpha,\beta) = y_k(\alpha,\beta+s), \quad \alpha,\beta \in Z_N, \tag{5.4}$$

for k=2 and 4.

For the gauge transformation 2 we can show the next proposition.

Proposition 5.8. *Let $\{w_j; j = 1,\dots,4\}$ be a such four-weight spin model that has functions z_1 and z_2 such that $w_2(\alpha,\beta) = z_1(\alpha-\beta)$ and $w_4(\alpha,\beta) = z_2(\alpha-\beta)$. Assume also that $\tilde{w}_j$ is obtained from w_j by the gauge transformation 2. Then the link polynomial derived from the transformed weights $\{\tilde{w}_j; j = 1,2,3,4\}$ coincides with that of the spin model $\{w_j; j = 1,2,3,4\}$.*

Proof. Let us consider calculation of the partition function of the four-weight spin model $\{\tilde{w}_j; j = 1,2,3,4\}$ for a special projection $P(L)$ of a link L. We first recall Lemma 5.7. We denote by $S(\alpha_i)$ $(S(\beta_j))$ the positive (negative) regions of $P(L)$ and by α_i (β_j) the spin variables assigned on $S(\alpha_i)$ $(S(\beta_j))$. Note that all the weights in the partition function are given in the form $\tilde{y}_k(\alpha_i,\beta_j)$ for k=2 or 4. Then we see that the partition function of the four-weight spin model $\{\tilde{w}_j; j = 1,2,3,4\}$ is reduced to that of the spin model $\{w_j; j = 1,2,3,4\}$ by replacing all the spin variables β_j on $S(\beta_j)$ with $\beta_j - s$. Here we have used the eq. (5.4) and the fact that the summation of the partition function is invariant under such a replacement that $\gamma \to \gamma - s$ (mod N) for any spin variable γ. Thus we have the proposition.

Let us consider the four-weight spin models constructed by Kac and Wakimoto in Ref. [16] from the viewpoint of the gauge transformations. Let $\mathbf{L} = \mathbf{Z}^m$ and $<,>$ denote a symmetric bilinear form on $\mathbf{L}$. We set $< e_i, e_j >= a_{ij}$, where e_i is the standard basis of $\mathbf{L}$, $a_{ij} \in \mathbf{Q}$ and

$a_{ii} = p/q$, $\gcd(p,q) = 1$ with pq is even. Let $\mathbf{M} = \{\alpha \in \mathbf{L}; <\alpha, \beta> \in \mathbf{Z}$, for all $\alpha \in \mathbf{M}\}$. We define the set S by $S = \mathbf{L}/\mathbf{M}$. We have $n = |S|$. Let us define t_α and D by

$$D = \sum_{\beta \in S} t_\beta, \quad t_\alpha = \exp(\pi i < \alpha, \alpha >), \quad \alpha \in S. \tag{5.5}$$

Then the weights are given by [16]

$$\begin{aligned} &w_1(a,b) = At_{\xi+a-b}, w_2(a,b) = Bt_{\eta+a-b}, \\ &w_3(a,b) = A^{-1}t^{-1}_{\xi+a-b}, w_4(a,b) = B^{-1}t^{-1}_{\eta+a-b}, \end{aligned} \tag{5.6}$$

where $AB = \sqrt{n}/(Dt_\xi)$.

By applying the Propositions 5.3 and 5.8 we can show that the variables ξ and η correspond to the parameters of the gauge transformations 1 and 2, respectively, and that the link invariant derived from the four-weight spin model (5.6) does not depend on ξ or η, essentially. The invariant may depend on the normalization factors of the weights. However, such dependence can be determined by the writhe or the numbers of the positive and negative crossing points of the link diagram. Thus as far as evaluation of link invariants is concerned, we may only consider the case $\xi = \eta = 0$.

Acknowledgments. The author would like to thank Prof. Eiichi Bannai for his invitation to the International Conference on Algebraic Combinatorics (Fukuoka 1993) November 22–26, 1993, in Kyushu University, and also to the volume. The author would also like to thank Prof. F. Jaeger for helpful discussion and valuable comments on spin models. From the discussion the author came to know that the gauge transformations for four-weight spin models in this paper were independently introduced by F. Jaeger, and that the "proof" of Prop. 5.8 in the previous version of the manuscript was incomplete. Prof. F. Jaeger had proved the same result independently and kindly pointed out to the author the importance of the notion of special projection. Thanks to this comment the author was able to complete his proof.

References

[1] Y. Akutsu and M. Wadati, Knot invariants and critical statistical systems, J. Phys. Soc. Japan, **56** (1987), 839–842.

[2] H. Au-Yang, B. M. M. McCoy, J. H. H. Perk, S. Tang and M. Yan, Commuting transfer matrices in the chiral Potts model: solution of

star-triangle equation with genus greater than 1, Phys. Lett., **123A** (1987), 219–223.
[3] H. Au-Yang and J. H. H. Perk, Star-triangle equations and multicomponent chiral Potts models, Physica A, **177** (1991), 139–145. (See also, RIMS-839, preprint 1991.)
[4] E. Bannai and E. Bannai, Generalized generalized spin models, to appear in Pacific J.
[5] R. J. Baxter, " Exactly Solved Models in Statistical Mechanics", Academic Press, 1982.
[6] G. Burde and H. Zieschang, " Knots", Walter de Gruyter, 1985.
[7] E. Date, M. Jimbo, K. Miki, and T. Miwa, Braid group representations arising from the generalized chiral Potts models, RIMS-729, preprint, 1990.
[8] T. Deguchi, A note on generalized spin models, preprint, 1995.
[9] V. A. Fateev and A. B. Zamolodchikov, Self-dual solutions of the star-triangle relations in Z_N-models, Phys. Lett., **92A** (1982), 37–38.
[10] P. De La Harpe, Spin models for link polynomials, strongly regular graphs and Jaeger's Higman-Sims model, preprint, 1992.
[11] K. Hasegawa and Y. Yamada, Broken Z_N-symmetric model, Phys. Lett. **146A** (1990) 387– 396.
[12] F. Jaeger, Strongly regular graphs and spin models for the Kauffman polynomial, Geom. Dedicata, **44** (1992), 23–52.
[13] F. Jaeger, On spin models, triply regular association schemes, and duality, J. Algebraic Combin., **4** (1995), 103–144.
[14] F. Jaeger, private communication, August 1995.
[15] V. F. R. Jones, On knot invariants related to some statistical mechanical models, Pacific J. of Math., **137** (1989), 311–334.
[16] V. G. Kac and M. Wakimoto, A construction of generalized spin models, in "Perspectives in Mathematical Physics", Conf. Proc. Lecture Notes Math. Phys., III, Internat. Press, 1994, pp. 125–150.
[17] M. Kashiwara and T. Miwa, A class of elliptic solutions to the star-triangle relation, Nucl. Phys., **B275**[FS] (1986), 121–134.
[18] K. Kawagoe, A. Munemasa and Y. Watatani, Generalized Spin Models, J. Knot Theory Ramifications, **3** (1994), 465–476.
[19] T. Kobayashi, H. Murakami and J. Murakami, Cyclotomic invariants for links, Proc. Japan Acad. Ser. A Math. Sci., **64** (1988), 235–238.
[20] H. Murakami, T. Ohtsuki and M. Okada, Invariants of three-manifolds derived from linking matrices of framed links, Osaka J. Math., **29** (1992), 545–572.
[21] Y. Quano and A. Fujii, Yang-Baxter equation for broken $Z_N^{\otimes n-1}$ models, Modern Phys. Lett. A, **8** (1993), 1585–1597.
[22] V. G. Turaev, The Yang-Baxter equation and invariants of links, Invent. Math., **92** (1988), 527–553.
[23] F. Y. Wu, P. Pant and C. King, A new invariant from the chiral Potts model, Phys. Rev. Lett., **72** (1994), 3937–3940.

[24] C. N. Yang, Some exact results for the many-body problem in one dimension with repulsive delta-function interaction, Phys. Rev. Lett., **19** (1967), 1312–1315.

Department of Physics
Ochanomizu University
Ohtsuka, Bunkyo-ku
Tokyo 112
Japan

Advanced Studies in Pure Mathematics 24, 1996
Progress in Algebraic Combinatorics
pp. 101–136

Vertex Operator Algebras and Moonshine: A Survey

Chongying Dong and Geoffrey Mason

§1. Introduction

This survey is, in many regards, a reprise of the paper [Ma1]. That paper reviewed what was known at the time about the connections between finite groups and elliptic modular forms. Shortly after, Borcherds presented the notion of a *vertex (operator) algebra* (VOA), and the work of Frenkel–Lepowsky–Meurman showed that the Monster was the automorphism group of a particular VOA—the *Moonshine module.* Although the Conway–Norton conjectures were not proved till later, it became clear that *in principle*, one can associate modular functions to the elements of a finite group, following the dictates of *Moonshine*, whenever the group operates as automorphisms of a VOA. Indeed, the bulk of the Conway-Norton conjectures, together with Norton's generalization, can be understood in the framework of the representation theory of so-called *holomorphic orbifold models.* One sees that the theory applies to any finite group, and only the *genus zero* condition marks the Moonshine module as special.

In this survey we present the main ideas which led to the framework in which one can understand Moonshine as a general and natural phenomenon as opposed to an apparent miracle. Thus it is about VOAs, their representations, and their automorphisms. There are no proofs—indeed there are no lemmas or theorems either! Rather, we try and explain the main ideas, tell what is known, what remains to be done. Thus the reader will find 26 "problems" in the course of the paper. These may be considered guides to what still needs to be proved in order to make the theory more complete. A few of them are rather specialized, but most are of a general nature and (it seems to us) quite difficult!

There is no pretense at presenting a general survey of algebraic conformal field theory—that would be a vast undertaking. Rather, we have

Received March 31, 1995.
Revised August 17, 1995.
Supported by grants from the National Science Foundation and the Committee on Research at Santa Cruz.

included only those topics in VOA theory which impinge directly on the explication of "Moonshine." On the other hand, our notion of Moonshine is broader than that of others. Thus we have included some topics of a more cohomological and topological nature with the expectation (and hope) that equivariant elliptic cohomology—whatever this turns out to really mean—will eventually become closely identified with Moonshine.

We thank the referee for helpful comments.

Table of Contents

§2. Axioms

We discuss the axioms for a *vertex operator algebra* (VOA). A fuller account may be found in [FLM], [FHL], [DL].

Let us begin with the definition of a VOA. This entails a $\mathbb{Z}$-graded vector space V over the complex numbers $\mathbb{C}$:

$$V = \coprod_{n \in \mathbb{Z}} V_n \tag{2.1}$$

such that $V_n = 0$ for $n \ll 0$ and $\dim V_n$ is finite for all n. If $v \in V_n$ we say that v is *homogeneous of (conformal) weight* n and write $wt(v) = n$. Roughly speaking, V is endowed with a countable infinity of bilinear products $u *_n v$ (u, $v \in V$, $n \in \mathbb{Z}$) which are required to satisfy infinitely many identities.

For fixed $u \in V$ we assemble the endomorphisms $u_n \in \operatorname{End} V$, defined by $u_n v = u *_n v$, into a generating function (or *vertex operator*)

$$Y(u, z) = \sum_{n \in \mathbb{Z}} u_n z^{-n-1} \in \operatorname{End} V[[z, z^{-1}]] \tag{2.2}$$

which provides a linear map $Y \colon V \to \operatorname{End} V[[z, z^{-1}]]$, $u \mapsto Y(u, z)$.

The other VOA axioms are as follows:

(i) For $u,\ v \in V$, there is $n = n(u,v) \in \mathbb{Z}$ such that $u_m v = 0$ for $m > n$,

This tells us that the sum $Y(u,z)v := \sum_m u_m v z^{-m-1}$ is a Laurent series, i.e., only finitely many non-zero negative powers occur.

(ii) *Existence of vacuum*: There is a distinguished element $1 \in V$ which satisfies

(a) $Y(1,z) = \mathrm{id}$,

(b) $Y(u,z)1 = u + \sum_{n=2}^{\infty} u_{-n} 1 z^{n-1}$.

Part (b) tells us that $u_n 1 = 0$ for $n \geq 0$ and $u_{-1}1 = u$. In particular, the map $u \mapsto Y(u,z)$ is an injection. One shows that in fact $1 \in V_0$.

(iii) *Existence of conformal vector*: There is a distinguished element $\omega \in V$ with generating function $Y(\omega, z) = \sum_n L_n z^{-n-2}$ such that the component operators L_n generate a copy of the *Virasoro algebra* Vir represented on V with *central charge c*. This means that there are relations

$$[L_m, L_n] = (m-n)L_{m+n} + \frac{m^3 - m}{12}\delta_{m,-n} c \tag{2.3}$$

for some constant c. It turns out that c is an important invariant associated to V. And inasmuch as V is a module for Vir, the representation theory of the Lie algebra Vir (cf. [KR]) will be valuable for VOA theory. One has $\omega \in V_2$.

(iv) V_n is the L_0-eigenspace of V with eigenvalue n. That is, if $v \in V_n$ then $L_0 v = nv$.

(v) $\dfrac{d}{dz} Y(v,z) = Y(L_{-1}v, z)$.

Note from (2.3) that the span of L_0, L_{-1}, L_1 is a Lie subalgebra of Vir isomorphic to $sl(2)$. The last two axioms suggest that L_0 and L_{-1} are particularly important operators on V.

(vi) *Jacobi Identity*: we have

$$\begin{aligned} &z_0^{-1}\delta\left(\frac{z_1 - z_2}{z_0}\right) Y(u,z_1)Y(v,z_2) \\ &\qquad - z_0^{-1}\delta\left(\frac{z_2 - z_1}{-z_0}\right) Y(v,z_2)Y(u,z_1) \\ &= z_2^{-1}\delta\left(\frac{z_1 - z_0}{z_2}\right) Y(Y(u,z_0)v, z_2). \end{aligned} \tag{2.4}$$

This is the VOA embodiment of commutativity and associativity. With a suitable interpretation of the delta function $\delta(\ )$ (cf. [FLM], [FHL]), when applied to $w \in V$, both the l.h.s. and r.h.s. of (2.4) are expressions of the form $\sum P_{ijk} z_0^i z_1^j z_2^k$ where P_{ijk} is a *finite* sum of vectors in V (recall axiom (i)). In this way, (2.4) represents a threefold infinity of identities which must be satisfied.

We observe that if $v \in V_m$ then

$$v_n \colon V_p \to V_{p+m-n-1}. \tag{2.5}$$

In particular, there is a linear map $V \to \mathrm{End}(V_p)$ such that $v \mapsto v_{wt(v)-1}$ if v is homogeneous.

There are alternate ways to view the Jacobi identity—the so-called *duality* axioms. Not only are they conceptually easier to swallow, but they are important in a number of situations. To set the stage, let $V = \coprod V_n$ be a VOA, let V_n^* be the dual space $\mathrm{Hom}(V_n, \mathbb{C})$, and set $V' = \coprod V_n^*$. Thus there is a canonical pairing $\langle\ ,\ \rangle \colon V' \times V \to \mathbb{C}$ so that $\langle V_m^*, V_n \rangle = 0$ for $m \neq n$.

Next, let R be the subring of $\mathbb{C}(z_1,z_2)$ obtained by localizing $\mathbb{C}[z_1,z_2]$ at the monoid generated by homogeneous polynomials of degree 1. For each of the two orderings (i_1, i_2) of $\{1, 2\}$ there is an injective ring map

$$\iota_{i_1,i_2} \colon R \to \mathbb{C}[[z_1, z_1^{-1}, z_2, z_2^{-1}]] \tag{2.6}$$

by which $(az_1 + bz_2)^{-1}$ is expanded as a power series in non-negative powers of z_{i_2}. Then we have

(vii) *Duality*: let u, v, $w \in V$, $w' \in V'$. There is $f \in R$ of the form $f = h(z_1, z_2)/z_1^r z_2^s (z_1 - z_2)^t$ with $h(z_1, z_2) \in \mathbb{C}[z_1, z_2]$, r, s, $t \in \mathbb{Z}$ such that

(a) *Rationality*: $\langle w', Y(u, z_1)Y(v, z_2)w\rangle = \iota_{12} f(z_1, z_2)$
(b) *Commutativity*: $\langle w', Y(v, z_1)Y(u, z_2)w\rangle = \iota_{21} f(z_1, z_2)$
(c) *Associativity*: $\langle w', Y(Y(u, z_0)v, z_2)w\rangle = \iota_{20} f(z_0 + z_2, z_2)$.

It transpires ([FLM], [FHL]) that the duality axioms are *equivalent* to the Jacobi identity. In fact, there is an even simpler variant:

(viii) If u, $v \in V$ there is $n = n(v, v) \geq 0$ such that

$$(z_1 - z_2)^n [Y(u, z_1), Y(v, z_2)] = 0 \tag{2.7}$$

(cf. [DL, chptr.7]). In conjunction with the equality $[L_{-1}, Y(v, z)] = Y(L_{-1}v, z)$, this is also equivalent to (vi). This formulation is closely related to the approach via *local systems*, which is frequently used in the physics literature (cf. [G], for example). We discuss this further below. To illustrate (2.7) we observe that the Virasoro axioms (2.3) imply that

$$(z_1 - z_2)^4 [Y(\omega, z_1), Y(\omega, z_2)] = 0. \tag{2.8}$$

This concludes our discussion of the VOA axioms, but there is much more to be said. In particular there are a number of more or less obvious variations that are important. For example, it is clear that we can define VOAs over *any* field, or even more general rings of coefficients. Moreover, one often wants to *weaken* the above axioms in various ways. For example one may allow the $\mathbb{Z}$-grading on V to be unbounded below, and/or allow the homogeneous spaces to be of infinite dimension. One may also dispense with the Virasoro algebra and keep just the degree operator L_0 and derivation L_{-1}. This is essentially the original approach of Borcherds in [Bo1], where he studies so-called *vertex algebras.*

One can also add *additional* axioms. In the physics literature, for example, it is often an implicit axiom that V carries the structure of a Hilbert space. In the approach of [BPZ](see also [MS]), even assumptions concerning modular-invariance of characters (see Section 10) are contemplated. We will always stick with our axiom system (i)-(vi).

A *morphism* of VOAs V_1, V_2 is a linear map $f\colon V_1 \to V_2$ which preserves the vacuum, the conformal vector and each of the binary products $*_n$. The latter condition is equivalent to

$$fY_1(u,z) = Y_2(fu,z)f. \tag{2.9}$$

(Here and below, we frequently refer to a VOA via its underlying space, say V_1, rather than through the pair (V_1, Y_1) consisting of V_1 and its Y-map, or even to $(V_1, Y_1, 1_1, \omega_1)$ etc.)

An *automorphism* of the VOA V is an invertible endomorphism of V. The set of these form a group $\mathrm{Aut}(V)$. Since $\mathrm{Aut}(V)$ leaves ω invariant, (2.9) shows that $\mathrm{Aut}(V)$ commutes with each component L_n of $Y(\omega, z)$. Thus $\mathrm{Aut}(V)$ commutes with the Virasoro algebra Vir in their joint action on V. This can have useful consequences (see [DLM], for example).

Note in particular that since $\mathrm{Aut}(V)$ commutes with L_0 then each V_n affords a representation of $\mathrm{Aut}(V)$, thanks to axiom (iv). It is through this mechanism that group representation theory enters into the study of VOAs.

§3. Examples of VOA

One of the difficulties of the subject is that examples are not easy to describe. In our exposition we will generally only discuss how to construct the *Fock space*, that is the underlying graded space V (2.1), and eschew a description of the vertex operators themselves.

(i) *Local systems*: Let M be a $\mathbb{C}$-vector space and let $F(M)$ be the subspace of $\mathrm{End}\, M[[z, z^{-1}]]$ consisting of power series $u(z) = \sum_m u_m z^{-m-1}$ which satisfy (i) of Section 2. Define a relation on $F(M)$ as follows: $u(z)$ and $v(z)$ are related if (2.7) holds, that is $(z_1-z_2)^n[u(z_1),v(z_2)] = 0$ for some $n \geq 0$.

This relation is symmetric, but not necessarily reflexive. One usually says that $u(z)$ and $v(z)$ are *mutually local* if they are related as above. We may call $u(z)$ a local vertex operator if it is mutually local with itself. From Section 2 we know that if V is a VOA then V consists of a set of mutually local power series $u(z) = Y(u, z)$ in which M is V itself. The idea is now that VOAs and local systems are essentially equivalent. This is the point-of-view taken in Goddard's work [G] and in his joint work with Dolan and Montague ([DGM1], [DGM2], for example).

There is an axiomatic treatment due to Li ([Li1]). If we start with a Virasoro module M such that $L(z) = \sum_n L_n z^{-n-2}$ satisfies (i) of Section 2 (we call M *restricted* if this holds), then $L(z)$ is a local vertex operator (2.8). We define a *local system* on M to be a maximal set of mutually local operators $u(z)$ on M which contain $L(z)$. These exist by dint of Zorn's Lemma. Li shows that a local system in this sense is a VA (vertex algebra); that is, the axioms for a VOA are satisfied except that the homogeneous spaces may have infinite dimension.

If S is any set of mutually local operators which includes $L(z)$ we may define $\langle S \rangle$ to be the smallest VOA containing S.

This is a rather abstract way to define VOAs, but it has the advantage that *existence* is not in doubt. We refer to the papers of Li and Goddard et al. for application of these ideas.

(ii) *Virasoro Modules*: For reference here we suggest [KR]. It is natural, given the requirement that a VOA admit an action of Vir, that various Vir-modules carry the structure of a VOA. This is indeed the case.

To describe the Fock space, recall (*loc. cit.*) that for a pair of complex numbers (c, h) there is a *Verma module* $M(c, h)$ which is the universal highest weight module for the Lie algebra Vir of central charge c, and such that $L_0 v = hv$ for maximal vector $0 \neq v \in M(c, h)$. Moreover $M(c, h)$ has a unique maximal submodule with quotient denoted by $L(c, h)$. Thus $L(c, h)$ is the unique simple Vir-module of highest weight (c, h).

To construct a VOA from $M(c, h)$ in which v is the vacuum we must take $h = 0$ (by axiom (iv) of Section 2) and we need $L_{-1}v \equiv 0$ (by axioms (ii)(a) and (v)). Therefore, let M_c be the quotient of $M(c, 0)$ by the Vir-submodule generated by $L_{-1}v$. Then indeed M_c has the structure of a VOA, as shown by Frenkel–Zhu [FZ]. There is also a proof due to Li [Li1] using the technique of local systems.

$L(c, 0)$ also gets the structure of VOA, being a VOA quotient of M_c. These are the physicist's *minimal models* [BPZ].

(iii) *Kac–Moody modules*: Let $\mathfrak{J}$ be a finite-dimensional, complex, simple Lie algebra with $\hat{\mathfrak{J}} = \mathfrak{J} \otimes \mathbb{C}[t, t^{-1}] \oplus \mathbb{C}c$ the corresponding affine Lie algebra [K]. There is a theory parallel to (ii) in which Vir is replaced by $\hat{\mathfrak{J}}$, and again there are mathematical descriptions due to Frenkel-Zhu [FZ] and Li [Li1].

Briefly, let $L(\lambda)$ be the simple $\mathfrak{J}$-module with highest weight λ and let* $M(c, \lambda) = \mathcal{U}(\hat{\mathfrak{J}}) \otimes_B L(\lambda)$ where $B = \mathfrak{J} \otimes \mathbb{C}[t] \oplus \mathbb{C}c$, c acts on $L(\lambda)$ as multiplication by c (the *level*) and $g \otimes t^n$ acts as 0 if $n > 0$ and as g if $n = 0$ ($g \in \mathfrak{J}$). Then $M(c, 0)$ has the structure of VOA if $c + g_0 \neq 0$, where g_0 is the dual Coxeter number—this condition is necessary for the existence of a conformal vector. Again as in Example (ii), $M(c, \lambda)$ has a unique maximal $\hat{\mathfrak{J}}$-submodule with quotient $L(c, \lambda)$, and $L(c, 0)$ is a quotient VOA of $M(c, 0)$ *as long as* $c + g_0 \neq 0$ and $c \neq 0$.

When c is a positive integer, $L(c, 0)$ is a (highest weight) integrable $\hat{\mathfrak{J}}$-module and the theory in this case is essentially the physicist's *WZW-model.*

(iv) *Lattice Theories*: Let L be a *positive-definite even* lattice; that is, $L \cong \mathbb{Z}^d$ is a free abelian group equipped with an integral, positive-definite, symmetric, bilinear form $(\ ,\ ) \colon L \times L \to \mathbb{Z}$ such that $(x, x) \in 2\mathbb{Z}$ for $x \in L$. Let $H = \mathbb{C} \otimes_{\mathbb{Z}} L$ and set

$$V_L = S(H_1 \oplus H_2 \oplus \cdots) \otimes \mathbb{C}[L]. \tag{3.1}$$

Here, each H_i is linearly isomorphic to H, $S(\ )$ denotes symmetric algebra, and $\mathbb{C}[L]$ is the group algebra of L with basis e^α, $\alpha \in L$. Then we can give V_L the structure of VOA in which the central charge c is equal to $d/24$. This result was announced in [Bo1], with a complete discussion provided in [FLM].

If L is the root lattice of a simple Lie algebra (of type ADE) then we recover the level 1 WZW theory of (iii) above. The advantage of the present approach is that it is quite explicit and computable, and because of this the lattice theories are of the utmost importance.

(v) *Spinor Constructions*: The theory of (iv) is often called a *bosonic* theory, being based on symmetric algebras. Spinor constructions, on the other hand, are called *fermionic theories* and are based on exterior and Clifford algebras. When available, spinor constructions are very useful and often very amenable to calculation. We give just a sample of the idea, based on [FFR] and [DM1].

Let $A \cong \mathbb{C}^{2\ell}$ be a complex linear space equipped with a non-degenerate, symmetric, bilinear form $(\ ,\ )$ and with a polarization

---

*$\mathcal{U}$ denotes universal enveloping algebra.

$A = A^+ \oplus A^-$ into maximal isotropic subspaces. Let $A(\mathbb{Z}) = A^+ \oplus A_1 \oplus A_2 \oplus \cdots$ and $A(\mathbb{Z} + 1/2) = A_{1/2} \oplus A_{3/2} \oplus \cdots$ where each A_i, $A_{i/2}$ is linearly isomorphic to A. (As in (3.1), the indexing will be useful for later discussions.) Then one can endow both $\Lambda^{\mathrm{even}}(A(\mathbb{Z}+1/2))$ and $\Lambda^{\mathrm{even}}(A(\mathbb{Z}+1/2)) \oplus \Lambda^{\mathrm{even}}(A(\mathbb{Z}))$ with the structure of VOA in a quite canonical way which makes use of the fact that both $\Lambda(A(\mathbb{Z}))$ and $\Lambda(A(\mathbb{Z}+1/2))$ may be viewed as modules for certain infinite-dimensional Clifford algebras. The *boson-fermion correspondence* ([F], [DM1]) says that in fact $\Lambda^{\mathrm{even}}(A(\mathbb{Z}+1/2))$ is VOA-isomorphic to the level 1 WZW model of type D_ℓ (alternatively, to V_L in which L is the D_ℓ-root lattice). Similarly, $\Lambda^{\mathrm{even}}(A(\mathbb{Z}+1/2)) \oplus \Lambda^{\mathrm{even}}(A(\mathbb{Z}))$ is VOA-isomorphic to $V_{L'}$ where L' is the spin lattice of rank ℓ.

These constructions are also important with regard to the so-called Witten-genus. See [Ta].

(vi) *Moonshine Module*: Perhaps the most famous VOA, its existence was announced by Borcherds [Bo1] and established in [FLM]. It is somewhat harder to construct than previous examples, and we will return to this point later. It is usually denoted $V^\natural$, and of course we have $\mathrm{Aut}(V^\natural)$ equal to the Fischer–Griess Monster group $\mathbb{M}$. $V^\natural$ has central charge 24.

(vii) *Tensor Products*: Given VOAs $V_1, \ldots, V_n$, the tensor product $V_1 \otimes \cdots \otimes V_n$ can be given the structure of a VOA in a canonical way. See [FHL] for details. We just observe here that the central charge of the tensor product is the *sum* of the central charges of the V_i.

(viii) For a different sort of example we go back to the Moonshine module $V^\natural$ and note that its grading (2.1) is such that $V_0^\natural = 0$ for $n < 0$ or $n = 1$, while $V_0^\natural = \mathbb{C}1$ is 1-dimensional and $V_2^\natural$ is of dimension 196,884 (and contains the conformal vector ω). $V_2^\natural$ is often denoted by B, and called the *Griess algebra* after its originator ([Gr]). B is a commutative, non-associative algebra with identity $1/2\omega$, the product being given by $u \circ v = u_1 v$ for u_1, $v \in B$ (cf. (2.5)).

There are idempotents e in B where the corresponding vertex operator $Y(e,z) = \sum_n K_n z^{-n-2}$ is such that the components K_n span a Virasoro algebra of central charge $1/2$, namely $L(1/2, 0)$ (see Example (ii)). In fact, one can find 48 linearly independent e's which span a 48-dimensional associative subalgebra of B and such that the corresponding vertex operators have components which commute. They generate a vertex operator subalgebra L of $V^\natural$ isomorphic to $L(1/2, 0)^{\otimes 48}$ (and of central charge 24). See [DMZ], [MN], [Mi] for more information. Frequently one can answer questions about $V^\natural$ by studying L. For examples of this technique see [DLiM3], [Hu].

Note that the idempotent e is associated with a certain involution of type $2A$ in the Monster $\mathbb{M}$ ([C], [MN], [Mi]). Similarly, there are idempotents f in B associated with elements of type $3A$ and such that the components of $Y(f,z)$ span a copy of $L(4/5,0)$.

Problem 1. *Is there a theory of f's analogous to that for e's? In particular, can one find* 30 *f's spanning a* 30*-dimensional associative subalgebra of B, with a corresponding sub VOA of $V^\natural$ isomorphic to $L(4/5,0)^{\otimes 30}$?*

§4. Representations of VOAs

To a great extent, the study of VOAs is the study of their representations. So in this regard, VOA theory is similar to more classical algebraic systems such as groups, associative algebras and Lie algebras.

We let V be a VOA, and adopt the notation of Section 2. Roughly speaking, a *module* for V is a $\mathbb{C}$-linear space M such that the operators v_n ($v \in V$, $n \in \mathbb{Z}$) operate on M in such a way that the relations among the v_n implicit in the Jacobi identity are preserved. More precisely, in contrast to (2.1) we require M to possess a $\mathbb{C}$-grading

$$M = \coprod_{x \in \mathbb{C}} M_x \tag{4.1}$$

such that for any $z \in \mathbb{C}$ we have $M_{z+n} = 0$ for integral $n \ll 0$, and $\dim M_x < \infty$ for each x.

Note that this extends the definition of [FLM], say, who require that the grading (4.1) be rational. Under suitable circumstance one expects that in fact V-modules are rationally graded, but it seem better not to assume this at the outset, but rather try to *prove* it (cf. Problem 5). Furthermore, there is a linear map $Y_M\colon V \to \operatorname{End} M[[z,z^{-1}]]$, $Y_M\colon v \mapsto Y_M(v,z) = \sum_{n\in\mathbb{Z}} v_n z^{-n-1}$. Once again, various axioms are imposed analogous to axioms (i)-(vi) of Section 2. The analogues of (i), (ii) are still required, though there is no analogue of the vacuum vector. The conformal vector ω is such that the components of $Y_M(\omega,z)$ generate a representation of the Virasoro algebra (2.3) on M with the *same* central charge c. Again, M_x is required to be the L_0-eigenspace with eigenvalue

x, and (v) still holds. The analogue of Jacobi now reads

$$\begin{aligned}
&z_0^{-1}\delta\left(\frac{z_1-z_2}{z_0}\right)Y_M(u,z_1)Y_M(v,z_2)\\
&\qquad - z_0^{-1}\delta\left(\frac{z_2-z_1}{-z_0}\right)Y_M(v,z_2)Y_M(u,z_1)\\
&= z_2^{-1}\delta\left(\frac{z_1-z_0}{z_2}\right)Y_M(Y(u,z_0)v,z_2).
\end{aligned} \tag{4.2}$$

Now suppose that $g \in \mathrm{Aut}(V)$ has finite order N. There is the notion of a *g-twisted V-module*, or *g-twisted sector*, which, as the name suggests, is some sort of "twisted" module for V. But it would be incorrect to compare it to the notion of twisted module in group representation theory; the correct analogy is with representations of twisted Kac–Moody algebras [K]. The details are as follows: M is a g-twisted V-module if $M = \coprod_{n\in\mathbb{C}} M_n$ is a $\mathbb{C}$-grade $\mathbb{C}$-linear space with the usual conditions on the grading. We also have a linear map $Y_g\colon V \to \mathrm{End}\, M[[z^{1/N}, z^{-1/N}]]$, $v \mapsto Y_g(v,z)$ satisfying the following: if $v \in V$ is such that $gv = e^{2\pi i j/N}v$ then

$$Y_g(v,z) = \sum_{n\in j/N+\mathbb{Z}} v_n z^{-n-1}, \tag{4.3}$$

and

$$\begin{aligned}
&z_0^{-1}\delta\left(\frac{z_1-z_2}{z_0}\right)Y_g(v,z_1)Y_g(w,z_2)\\
&\qquad - z_0^{-1}\delta\left(\frac{z_2-z_1}{-z_0}\right)Y_g(w,z_2)Y_g(v,z_1)\\
&= z_2^{-1}\left(\frac{z_1-z_0}{z_2}\right)^{-j/N}\delta\left(\frac{z_1-z_0}{z_2}\right)Y_g(Y(v,z_0)w,z_2).
\end{aligned} \tag{4.4}$$

Note that if $g = 1$, a g-twisted V-module is precisely a V-module in the earlier sense. One can define twisted sectors even for automorphisms of infinite order, but we will not be concerned with them in this paper.

If (M, Y_g), (M', Y_g') are two g-twisted V-modules, a map $f\colon (M, Y_g) \to (M', Y_g')$ is a linear map $f\colon M \to M'$ which intertwines the Y-maps, i.e., $Y_g'(v,z)f = fY_g(v,z)$, $v \in V$. It is then apparent that there is a *category g-V-Mod* of g-twisted V-modules with objects being g-twisted V-modules and with morphisms as above.

A (g-twisted) submodule of the g-twisted V-module M is, rather obviously, a linear subspace $N \subseteq M$ such that for each $v \in V$ and

$Y_g(v,x) = \sum v_n z^{-n-1}$, each v_n leaves N invariant. There are then the usual notions of *simple* module, *semi-simple* (= completely reducible) module, *indecomposable* module, etc. As in Lie theory, it seems that the issue of complete reducibility is quite crucial. But it is fair to say that it has not been studied much in the present context. So we suggest

Problem 2. *Investigate the notion of complete reducibility in the category* g-V-Mod.

It turns out that the notion of *weak* (g-twisted) V-module is relevant (mainly because of the theory of Zhu algebras—see Section 6). We obtain a definition of a weak (g-twisted) V-module if we allow the homogeneous subspaces M_x ((4.1)) to be *infinite-dimensional*, but otherwise retain the previous definition. There is a corresponding category W-g-V-Mod of weak g-twisted V-modules.

We will be mainly concerned with certain classes of VOAs for which the category V-Mod (plain V-modules: we habitually drop the "g" in case $g = 1$) satisfies certain finiteness conditions. We call V *rational* in case V satisfies the following conditions:

(4.5)
(a) V has only a finite number of inequivalent simple modules.
(b) Every V-module is completely reducible.
(c) Every weak simple V module is an (ordinary) V-module.

Similarly, V is called *g-rational* in case V satisfies (a)–(c) of (4.5) with "g-twisted modules" in place of "modules."

It seems likely that (a)–(c) are not independent conditions, and an affirmative solution of the next problem would be very powerful.

Problem 3. *Does* (4.5)(a) *imply the other two conditions? Same problem for g-twisted modules.*

We call V *holomorphic* in case it satisfies the following conditions:

(4.6)
(a) V has a unique simple V-module, namely itself.
(b) Every V-module is complete reducible.

To clarify the meaning of (4.6)(a), it is clear that any VOA V is itself a V-module, which we will call the *adjoint module* for V. To say that the adjoint module is simple is to say that V is a *simple VOA*, that is, it has no non-trivial ideals.

The group $\mathrm{Aut}(V)$ induces equivalences among the various categories (W-)g-V-Mod, $g \in \mathrm{Aut}(V)$. Namely, suppose that (M, Y_M) is an object in (W-)g-V-Mod and that $h \in \mathrm{Aut}(V)$. Then h induces a categorical equivalence $\varepsilon(h) = \varepsilon_g(h)$:

$$g\text{-}V\text{-Mod} \xrightarrow{\varepsilon(h)} h^{-1}gh\text{-}V\text{-Mod} \tag{4.7}$$

as follows: the image of (M, Y_M) is the pair $(M, Y_M \circ h)$ where by definition $Y_M \circ h(v, z) = Y_M(hv, z)$, $v \in V$; and $\varepsilon(h)f = f$ for a morphism f. This assertion follows from the definitions: see [DM2] for more information.

For a group G and $g \in G$ we define the *centralizer* of g in G via $C_G(g) = \{x \in G | gx = xg\}$. If we set $C(g) = C_{\mathrm{Aut}(V)}(g)$ for $g \in \mathrm{Aut}(V)$ then it is clear from (4.7) that $C(g)$ induces automorphisms of the category g-V-Mod, and in particular $\mathrm{Aut}(V)$ acts on V-mod.

For $g \in \mathrm{Aut}\, V$ of finite order, let $V\text{-Mod}^g$ denote the full subcategory of g-invariants of V-Mod, so that objects of $V\text{-Mod}^g$ are V-modules (M, Y_M) satisfying $\varepsilon(g)(M, Y_M) \cong (M, Y_M)$. We would like to believe that the number of (isomorphism classes of) simple g-invariant V-modules is equal to the number of (isomorphism classes of) simple g-twisted V-modules—at least under some finiteness assumptions. We state a sharp form of this conjecture:

Problem 4. *Let $g \in \mathrm{Aut}(V)$ have finite order, and assume that V has only a finite number of isomorphism classes of simple modules. Show that there is a $C(g)$-equivariant categorical equivalence*

$$V\text{-Mod}^g \xrightarrow{\cong} g\text{-}V\text{-Mod}. \tag{4.8}$$

This conjecture contains a lot of information (assuming it is true!), and we regard it as one of the basic problems concerning twisted sectors. In particular, if (4.8) holds then being g-equivariant, we may expect that $\varepsilon_g(g)$ leaves invariant (up to isomorphism) each object of g-V-Mod. It is useful to observe that this is indeed the case for simple g-twisted V-modules (M, Y_M). It follows readily from the definition that such M have a grading of the form

$$M = \coprod_{n=0}^{\infty} M_{c+\frac{n}{N}}. \tag{4.9}$$

Here, N is the order $o(g)$ of g, $M_c \neq 0$, and c is some constant.* Then

*Not to be confused with central charge.

by defining $\phi(g) \colon M \to M$ via

$$\phi(g)|M_{c+\frac{n}{N}} = e^{-2\pi i n/N}$$

we find that $\phi(g)$ induces an isomorphism $\phi(g) \colon (M, Y_M) \overset{\cong}{\longmapsto} (M, Y_M \circ g)$ in g-V-Mod.

We note here that the constant c of (4.9), which we call the *top weight* (of M) is an important invariant. So too is the space M_c itself—the *top level* of M—and its dimension.

Problem 5. *Suppose that V has only finitely many simple modules. Prove that the top weights of all simple (twisted) V-modules are rational.*

We refer to [AM] for further information on this subject.

To complete this section we introduce the notion of *extended automorphisms* of g-twisted V-modules. To keep the exposition simple we assume that (V, Y) is a simple VOA and that (M, Y_M) is a simple g-twisted V-module.

Then an extended automorphism of M is a pair $(x, \alpha(x))$ where x, $\alpha(x)$ are invertible linear maps of M, V respectively, and which satisfy

$$Y_M(\alpha(x)v, k) = x Y_M(v, z) x^{-1} \tag{4.11}$$

for $v \in V$. We also require that $\alpha(x)\omega = \omega$ and that $\alpha(x)$ commutes with g.

One can show that $x \mapsto \alpha(x)$ is a group homomorphism into the group of $C(g)$. The kernel is the group $\mathbb{C}^*$ of invertible scalar operators on M. Let $\mathrm{Aut}^e(M)$ denote the group of extended automorphisms, where $(x, \alpha(x))(y, \alpha(y)) = (xy, \alpha(xy))$. Generally there is no canonical complement to $\mathbb{C}^*$ in $\mathrm{Aut}^e(M)$—indeed there may be no complement at all. This is why we refer to "extended automorphism."

As an example, with $\phi(g)$ as in (4.10) the assertion that $\phi(g)$ induces an isomorphism from (M, Y_M) to $(M, Y_M \circ g)$ can alternately be viewed as saying that $(\phi(g), g)$ is an extended automorphism of M.

Similarly, if we set $C(M) = \{h \in C(g) | \varepsilon(h)(M, Y_M) \cong (M, Y_M)\}$ (cf. (4.7)), then $C(M)$ is a group (which contains g). And if $\phi(h) \colon M \to M$ is the morphism which realizes the equivalence of (M, Y_M) and its image under $\varepsilon(h)$, then $(\phi(h), h) \in \mathrm{Aut}^e(M)$.

Thus $\mathrm{Aut}^e(M)$ fits into a short exact sequence $1 \to \mathbb{C}^* \to \mathrm{Aut}^e(M) \to C(M) \to 1$ and hence defines an element of $C^2(C(M), \mathbb{C}^*)$.

§5. Examples of V-modules

We give examples of (twisted) V-modules which by-and-large track the examples of VOAs given in Section 3. Before proceeding, however, we want to emphasize that there are some serious gaps in our understanding of (twisted) modules insofar as *existence* is concerned. Here is a sample problem:

Problem 6. *Let V be a VOA with $g \in \mathrm{Aut}(V)$ of finite order. Prove that V has a (non-zero) weak g-twisted V-module. Better, prove that V has a (non-zero) g-twisted V-module.*

Now we turn to some examples.

(i) Let the notation be as in Example (i) of Section 3. Then the vector space M is a module for the VOA $\langle S \rangle$ generated by a set of mutually local operators S.

(ii) Let the notation be as in Example (ii) of Section 3. Then both $M(c,h)$ and $L(c,h)$ are modules for the VOA $L(c,0)$. It is in fact known when $L(c,0)$ is rational in the sense of (4.5), [W], [DMZ]. Namely, $L(c,0)$ is rational precisely when c has the form $c = 1 - 6(p-q)^2/pq$ for a pair of coprime integers p, q each at least 2. In this case, the simple modules for $L(c,0)$ are the Fock spaces $L(c,h)$ for $h = (np-mq)^2 - (p-q)^2/4pq$, $0 < m < p$, $0 < n < q$.

Again this result is obtained by studying the Zhu algebra associated to $L(c,0)$ [W]. For more information on these remarkable families see, for example, [KR], [GKO], [FQS].

(iii) Let the notation be as in Example (iii) of Section 3. We limit our discussion to the *WZW*-models $L(c,0)$, so that c here is a positive integer. Then $L(c,0)$ is indeed rational, and the simple modules are those $L(c,\lambda)$ for which λ is an integrable weight satisfying $\langle \lambda, \mu \rangle \leq c$ for μ the root of maximal height. For a proof, see [FZ] or [Li1].

(iv) Let V_L be the Fock space of the VOA associated to an even lattice L as in Example (iv) of Section 3. The rationality of V_L is established by Dong [Do1], where it is shown that the simple modules are precisely the Fock spaces $V_{L+\lambda}$ for λ in the dual lattice $L^0 = \mathrm{Hom}_{\mathbb{Z}}(L, \mathbb{Z})$. Here $V_{L+\lambda}$ is defined as in (3.1), except that the group algebra $\mathbb{C}[L]$ is now replaced by the "group algebra" of the coset: $\mathbb{C}[L+\lambda] = \coprod_{\alpha \in L} \mathbb{C}e^{\alpha+\lambda}$. In particular, V_L is holomorphic (4.6) if, and only if, L is a self-dual lattice.

There is a canonical action of the torus $\mathbb{R}^n/L^0$ as automorphisms of V_L. (Here, n is the rank of L.) Namely if, in the notation of (3.1), we take $u \in S(H_1 \oplus H_2 \oplus \cdots)$ and $\alpha \in L$ then set

$$g(\nu)\colon u \otimes e^{\alpha} \mapsto e^{2\pi i(\nu,\alpha)} u \otimes e^{\alpha} \tag{5.1}$$

for $\nu \in \mathbb{R}^n = \mathbb{R} \otimes_{\mathbb{Z}} L$. Now $g(\nu)$ has finite order when ν lies in $Q^n = Q \otimes_{\mathbb{Z}} L$. Then it is shown in [DM1] that the simple $g(\nu)$-twisted V_L-modules arise precisely from the Fock species $V_{L+\lambda-\nu}$, where λ runs over coset representatives L^0/L. So for such automorphisms of V_L, (4.8) is indeed an equivalence of categories.

We give one more example along these lines: let $\theta\colon L \to L$ be the automorphism which acts as -1. Then θ naturally induces an automorphism of V_L, which we also denote by θ. There is then a θ-twisted V_L-module $V_{L,\theta}$ with Fock space defined via

$$V_{L,\theta} = S(H_{1/2} \oplus H_{3/2} \oplus \cdots) \otimes T. \tag{5.2}$$

Here, each $H_{i/2}$ is linearly isomorphic to $H = \mathbb{C} \otimes_{\mathbb{Z}} L$. The space T is finite of dimension $2^{n/2}$ (n is necessarily even since L is an even lattice). It is best regarded as an irreducible module for a so-called extra-special 2-group Q (see [Gor, Chapter 5]) of order 2^{n+1} which is constructed from L as follows. The map $x \mapsto (-1)^{(x,x)/2}$ defines a quadratic form $q\colon L \to \pm 1$ with associated bilinear form $\beta\colon L \times L \to \pm 1$. We may regard β as a 2-cocycle in the group $C^2(L, \pm 1)$ and as such it defines a 2-fold central extension $\hat{L}$ of L. Then T is a simple module for $\hat{L}$ where the kernel is isomorphic to $2L$ (which splits over the center of $\hat{L}$), and $Q = \hat{L}/2L$.

This example is discussed at length in [FLM]. Generalizations may be found in various places, e.g., [Do3], [Le], [DGM2]. Historically, this has been an important example, both for understanding the Moonshine module and for coming to grips with the "philosophy" of twisted sectors.

(v) Parallel to the spinor construction of V_L for L the D_ℓ root lattice, which was discussed in Example (v) of Section 3, existence and uniqueness results for twisted sectors corresponding to certain kinds of automorphisms are given in [DM1]. The Fock spaces can again be described as Clifford modules for "twisted" Clifford algebras of a certain kind. The only point we want to mention here is that, in some senses, the only difference between the twisted and *untwisted sectors* (this latter being the VOA itself) is in the grading of the Fock spaces.

(vi) We can now say something more about the Moonshine module $V^\natural$ (Example (vi), Section 3). Let Λ be the Leech lattice ([CS]) and let θ be as in the previous example. Then θ is an automorphism of V_Λ, whence the subspace V_Λ^θ of θ-invariants is itself a VOA. (More on this sort of thing later!) Similarly θ acts on $V_{\Lambda,\theta}$ (5.2) if we take θ to act trivially on T. If we let $V_{\Lambda,\theta}^{\pm}$ be the ± 1 eigenspace of θ on this space, then the Fock space of the Moonshine module is

$$V^\natural = V_\Lambda^\theta \oplus V_{\Lambda,\theta}^-. \tag{5.3}$$

Dong has established [Do2] the important fact that $V^\natural$ is a holomorphic VOA. Moreover every simple weak $V^\natural$-module is itself isomorphic to $V^\natural$: This follows from the existence of the VOA subalgebra $L(1/2,0)^{\otimes 48}$ discussed in Example (viii) of Section 3 together with Example (ii) of the present section and the theory of tensor products.

Less is known about twisted modules for $V^\natural$. Of course $\mathrm{Aut}(V^\natural) \cong \mathbb{M}$ (the Monster), and as a special case of Problem 4 one has

Problem 7. *For each $g \in \mathbb{M}$, prove that there is a unique simple g-twisted $V^\natural$-module. Prove also that any simple weak g-twisted $V^\natural$-module necessarily has finite-dimensional homogeneous subspaces.*

An affirmative answer to Problem 7 is provided in [DLiM3] in case g is one of the elements $2A$, $2B$, $4A$. (This is standard notation for certain elements in $\mathbb{M}$ of orders, 2, 2, 4, respectively. See [ATLAS].) Again the theory of $L(1/2,0)^{\otimes 48}$-modules is the crucial ingredient. Huang also solves Problem 7 for g of type $2B$ in [Hu], where he does much more, including a fresh approach to the VOA structure on the Fock space (5.3). The work of Tuite [Tu1], [Tu2] on generalized Moonshine implicitly assumes the uniqueness of twisted sectors for $V^\natural$.

The extended automorphism groups of the twisted sectors of $V^\natural(2A)$ and $V^\natural(2B)$ are known [DLiM3]. We have $\mathrm{Aut}^e(V^\natural(2A)) \cong \mathbb{C}^* \times 2Baby$. Here $2Baby$ is the non-split extension of the Baby Monster ([ATLAS]) by $\mathbb{Z}_2$, and is the centralizer of $2A$ in the Monster $\mathbb{M}$. The group $\mathrm{Aut}^e(V^\natural(2B))$ is a central product $\mathbb{C}^* \cdot \hat{C}$, where $\hat{C}$ is the non-split extension of the (non-simple) Conway group $\cdot O$ (= Automorphism group of the Leech lattice) by the extra-special group 2^{1+24}. The subgroup of $\mathbb{C}^*$ of order 2 is identified in $\mathbb{C}^* \cdot \hat{C}$ with a certain central subgroup of $\hat{C}$. Note in this case that the centralizer of $2B$ in $\mathbb{M}$ is *not* isomorphic to a subgroup of $\mathbb{C}^* \cdot \hat{C}$, i.e., the map α which intervenes in the definition of extended automorphism (4.11) is not an embedding in this case.

(vii) Given VOAs V_i and V_i-modules M_i, $1 \le i \le n$, the tensor product $M_1 \otimes \cdots \otimes M_n$ is a module for the VOA $V_1 \otimes \cdots \otimes V_n$ in a canonical way. See [FHL] for details.

(viii) *Contragredient Module.* Let $M = \coprod M_n$ be a g-twisted V-module, with corresponding Y-map Y_M. Let $M_n^* = \mathrm{Hom}_{\mathbb{C}}(M_n, \mathbb{C})$

be the dual of M_n and set $M' = \coprod M_n^*$, the *restricted dual* of M. Then there is a g^{-1}-twisted V-module $(M', Y_{M'})$ where $Y_{M'} : V \to (\operatorname{End} M')[[z^{1/N}, z^{-1/N}]]$ is defined by the condition

$$\langle Y'_M(v,z)m', m\rangle = \left\langle m', Y_M(e^{zL_1}(-z^{-2})^{L_0}v, z^{-1})m\right\rangle . \tag{5.4}$$

Here $\langle\ ,\ \rangle : M' \times M \to \mathbb{C}$ is the canonical pairing and m', m are arbitrary elements in M', M respectively. The case $g = 1$ is discussed in [FHL], following the original remark of Borcherds [Bo1].

Note that the assignment $(M, Y_M) \to (M', Y_{M'})$ defines a contravariant equivalence

$$g\text{-}V\text{-Mod} \xrightarrow{\cong} g^{-1}\text{-}V\text{-Mod}. \tag{5.5}$$

§6. Zhu algebras

One approach to establishing equivalences such as (4.8) is this: define associative algebras whose module categories are equivalent to categories of V-modules (of various kinds). Then prove that the associative algebras are themselves Morita equivalent. Zhu algebras are associative algebras which arise from one attempt in this direction and we have already referred to them on several occasions. The ideas leading to their construction were first explained in [Z], and recently generalized in [DLiM2], which we follow here.

Let (V, Y) be a VOA and $g \in \operatorname{Aut} V$ be of finite order. If $u \in V$ is homogeneous and satisfies $gu = e^{2\pi i r}u$, $0 < r \leq 1$, define a bilinear product $*_g$ on V to be the linear extension of the following:

$$u *_g v = \begin{cases} \operatorname{Res}_z Y(u,z)\dfrac{(z+1)^{wt(u)}}{z}v, & r = 1, \\ 0, & r < 1. \end{cases} \tag{6.1}$$

Explicitly, we have if $r = 1$,

$$u *_g v = \sum_{i \geq 0} \binom{wt(u)}{i} u_{i-1}(v). \tag{6.2}$$

Similarly, let

$$u \circ_g v = \begin{cases} \operatorname{Res}_z Y(u,z)\dfrac{(z+1)^{wt(u)}}{z^2}v, & r = 1, \\[2ex] \operatorname{Res}_z Y(u,z)\dfrac{(z+1)^{wt(u)+r-1}}{z}v, & r < 1. \end{cases} \tag{6.3}$$

Let $O_g(V)$ be the linear span of all $u \circ_g v$ over all such u, v, r. Note that $u \circ_g 1 = u$ if $gu = e^{2\pi i r}u$ and $0 < r < 1$. Thus we have $V = V^{\langle g \rangle} + O_g(v)$ where $V^{\langle g \rangle}$ is the space of g-invariants of V. Define

$$A_g(V) = V/O_g(V). \tag{6.4}$$

At first blush it appears as though $A_g(V)$ is merely a linear space, but in fact the product $*_g$ defined by (6.1) descends to an associative product on $A_g(V)$. More precisely, formula (2.5) continues to hold for the action of $Y_M(v,z) = \sum_n v_n z^{-n-1}$ on a weak g-twisted V-module (M, Y_M). In particular, if we set $o_g(v) = v_{wt(v)-1}$ then, as mentioned following (2.5), there is a linear map $V \to \mathrm{End}(M_c)$, $v \mapsto o_g(v)$ where M_c is the top level of M (which may be of infinite dimension). It turns out that $O_g(V)$ lies in the kernel of this map, and that the resulting map $A_g(V) \to \mathrm{End}(M_c)$ is a morphism of associative algebras.

What is crucial is the following result, due to Zhu [Z] if $g = 1$. If M is a *simple weak* g-twisted V-module then M_c affords a simple module for $A_g(V)$. Moreover, the map $M \mapsto M_c$ induces a bijection between equivalence classes of simple weak g-twisted V-modules and simple $A_g(V)$-modules. Finally if weak g-twisted V-modules are completely reducible then $A_g(V)$ is semi-simple. So in this latter case, there is an equivalence of categories

$$W\text{-}g\text{-}V\text{-Mod} \xrightarrow{\cong} A_g(V)\text{-Mod}. \tag{6.5}$$

It is these results that make weak twisted V-modules important, even though we may be primarily interested in (ordinary) twisted V-modules. It is hardly necessary to say that one tries to use the Zhu algebras to reduce questions about twisted V-modules to questions about representations of associative algebras, which are presumably more manageable. Actually calculating $A_g(V)$ is difficult; we refer to [FZ], [DMZ], [W] for examples.

To give a different example, we know from [Do2] that $V^\natural$ is both rational and holomorphic (cf. Example (iv), Section 5). It follows that $A(V^\natural)$ is a matrix algebra, and since the top level of $V^\natural$ is spanned by the vacuum vector then in fact $A(V^\natural) \cong \mathbb{C}$. Similarly from [DLiM3] we find that if $g \in \mathbb{M} = \mathrm{Aut}(V^\natural)$ is of type $2A$ then $A_{2A}(V^\natural) \cong \mathbb{C}$; on the other hand $A_{2B}(V^\natural) \cong \mathrm{Mat}_{24}(\mathbb{C})$, the algebra of 24×24 matrices over $\mathbb{C}$. None of these facts seems to be readily deducible from the definitions (6.1)-(6.3) themselves.

One of the drawbacks (at present) with the algebras $A_g(V)$ is that there are problems in trying to make the connection between g-twisted

V-modules and $A_g(V)$ functorial. Note also that even though $A(V)$ has a unit (the image of the vacuum vector), it is not clear if $A_g(V)$ is non-zero in general. We suggest some problems along these lines.

Problem 8. *Prove that $A_g(V)$ is non-zero. That is, V possesses non-zero weak g-twisted V-modules.*

Problem 9. *Suppose that $A_g(V)$ is semi-simple. Is it true that W-g-V-Mod is semi-simple?*

§7. Group cohomology

We discuss the connections between group cohomology and VOAs.

In the following we fix a VOA V and finite group of automorphisms $G \leq \mathrm{Aut}(V)$. We assume given a family $\{(V(g), Y_g)\}$ of simple twisted modules, one for each $g \in G$, such that $V(1) = V$ and the following condition holds:

$$\varepsilon(h)(V(g)) \cong V(h^{-1}gh) \tag{7.1}$$

for $h \in G$ (cf. (4.7)). Thus the group G permutes the twisted modules $V(g)$ via the action (4.7). (7.1) is equivalent to choosing, for a fixed g in each conjugacy class, a $V(g)$ with the property that $\varepsilon(h)V(g) \cong V(g)$ for $h \in C_G(g)$. By way of example, if G is a group of *inner* automorphisms of V in the sense that G leaves invariant each (untwisted) V-module, then the truth of conjecture (4.8) would tell us that every $V(g)$ has the desired property. The case we are mainly interested in is when V is holomorphic (cf. (4.10)). Then according to (4.8) there is a unique simple g-twisted V-module, in which case we take the family $\{V(g)\}$ to consist of the simple g-twisted V-modules, $g \in G$.

By definition, there are linear maps $\phi_g(h)\colon V(g) \to V(hgh^{-1})$ satisfying $\phi_g(h)Y_g(v,z)\phi_g(h)^{-1} = Y_{hgh^{-1}}(hv, z)$. (Compare with (4.11) and the accompanying discussion.) Then there are $\alpha_g(h,k) \in \mathbb{C}^*$ (or S^1) satisfying $\phi_g(h,k) = \alpha_g(h,k)^{-1}\phi_{kgk^{-1}}(h)\phi_g(k)$.

Set

$$V^* = \bigoplus_{g \in G} V(g) \tag{7.2}$$

and introduce the group algebra $\mathbb{C}[G]$ of G and its dual $\mathbb{C}[G]^*$. The latter algebra is spanned by $e(g)$, $g \in G$ with relations $e(g)e(h) = e(g)\delta_{g,h}$. As in [Ba1], [Ba2], one may think of $e(g)$ in this context as projection of V^* onto $V(g)$. The group of units U of $\mathbb{C}[G]^*$ arises in this context because of the following: if $\alpha\colon G \times G \to U$ is given by $\alpha(h,k) = \sum_g \alpha_g(h,k)e(g)$ then we have $\alpha \in C^2(\mathbb{Z}G, U)$. That is, α is a 2-cocycle on G with coefficients in U, where U is a multiplicative $\mathbb{Z}G$ module with right G-action $e(g)\cdot h = e(h^{-1}gh)$.

Let $HH^*(\mathbb{Z}G)$ denote the Hochschild cohomology of $\mathbb{Z}G$ (coefficients in $\mathbb{Z}G^*$ by assumption). We have $H^*(\mathbb{Z}G, U) \cong HH^{*+1}(\mathbb{Z}G)$, so that the compatible family $\{V(g)\}$ provides us with an element $[\alpha] \in HH^3(\mathbb{Z}G)$. Now we have ([Lo], [Be]) $HH^3(\mathbb{Z}G) \cong \bigoplus_* H^3(C_G(g), \mathbb{Z}) \cong \bigoplus_* H^2(C_G(g), \mathbb{C}^*)$ where here and below, $*$ denotes a sum over one element from each conjugacy class of G. From (4.10) and the accompanying discussion we see that the restriction of $[\alpha]$ to $H^3(C_G(g), \mathbb{Z}) = H^2(C_G(g), \mathbb{C}^*)$ defines a central extension $\widehat{C_G(g)}$ of $C_G(g)$ on $V(g)$ which has the following property: a pre-image of g in $\widehat{C_G(g)}$ lies in the center of $\widehat{C_G(g)}$.

Let $HH_0^3(\mathbb{Z}G)$ be the group of Hochschild 3-cohomology classes satisfying this condition. Thus $\{V(g)\}$ defines an element $[\alpha] \in HH_0^3(\mathbb{Z}G)$.

Let $HC(\mathbb{Z}G)$ denote cyclic cohomology ([Lo], [Be], [Bu]). We have $HC^3(\mathbb{Z}G) \cong \bigoplus_* H^2(XC_G(g), \mathbb{C}^*)$. Here, $XC_G(g)$ is the so-called *extended centralizer* defined as the push-out

$$\begin{array}{ccc} \mathbb{Z} & \xrightarrow{1 \mapsto g} & C_G(g) \\ \downarrow & & \downarrow \\ \mathbb{R} & \longrightarrow & XC_G(g) \end{array}$$

If $H_0^2(XC_G(g), \mathbb{C}^*)$ is the subgroup of cohomology classes in $H^2(XC_G(g), \mathbb{C}^*)$ whose restriction to $\mathbb{R}$ (or rather its image in $XC_G(g)$) is trivial, and if $HC_0^3(\mathbb{Z}G) = \bigoplus_* H_0^2(XC_G(g), \mathbb{C}^*)$, then there is a natural isomorphism $HC_0^3(\mathbb{Z}G) \xrightarrow{\cong} HH_0^3(\mathbb{Z}G)$ given by restriction to $C_G(g)$. Conclusion: our family $\{V(g)\}$ determines an element $[\alpha] \in HC_0^3(\mathbb{Z}G)$. Note that the discussion in Example (vi), Section 5, shows that in the case of the Moonshine module $V^\natural$, the class $[\alpha]$ will be non-zero.

Given an element of $HC^3(\mathbb{Z}G)$, one can follow the recipe of Kontsevich [Ko] (see also [PS]) to derive certain topological invariants of moduli spaces. This is an example of how one can expect to derive topological invariants from VOAs.

Elliptic cohomology is a newer subject (see [La], [Hi]). In the work of Baker and Thomas ([Bak], [Th1], [Th2]) and others one can see how it may well be reasonable to expect that families $\{V(g)\}$ of the type we are considering will lead to elements in $\mathrm{Ell}^*(BG)$, the elliptic cohomology of the classifying space BG of G. The true geometric meaning of this theory remains a closed book, so we will simply recast our preceding discussion in bundle-theoretic language which might be part of the prescription for $\mathrm{Ell}^*(BG)$.

First, if $V = V(1)$ is graded as in (2.1), we may form the Borel space $EG \times_G V_n$ for each n (here, $EG \to BG$ is the universal G-bundle), so that we may consider $EG \times_G V \to BG$ as a $\mathbb{Z}$-graded vector bundle over BG with fiber V. Next, if LBG is the free loop space of BG then LBG is (homotopic to) a space with components $BC_G(g)$, one for each conjugacy class. Thus $H^*(LBG, \mathbb{Z}) \cong \bigoplus_* H^*(BC_G(g), \mathbb{Z}) = \bigoplus_* H^*(C_G(g), \mathbb{Z}) = HH^*(\mathbb{Z}G)$. Moreover we may form the Borel construction $EC_G(g) \times_{C_G(g)} \mathbb{P}V(g)$ where $\mathbb{P}V(g)$ is the projective space of $V(g)$ (recall that $C(g)$ only acts projectively on $V(g)$). We could also form $EC_G(g) \times_{\widehat{C_G(g)}} V(g)$ in the obvious way. In this way we get a vector bundle $\xi_V \colon E \to LBG$ which is naturally associated with $\{V(g)\}$.

In this picture, $HC^*(\mathbb{Z}G)$ arises as the integral cohomology of the Borel construction $ES^1 \times_{S^1} LBG$, where S^1 acts on LBG by rotation of loops. If $V(g) = V = \coprod_{n=0}^{\infty} V(g)_{c+\frac{n}{N}}$ is graded as in (4.9) then the action (4.10), essentially that of g on $V(g)$, is part of an action of $\mathbb{R}$ given by $t|V(g)_{c+\frac{n}{N}} = e^{-2\pi i n t/N}$, $t \in \mathbb{R}$. So $1 \in \mathbb{R}$ acts as g does, in other words we have an action of $XC_G(g)$ on $V(g)$, at least projectively. Since $ES^1 \times_{S^1} LBG$ is homotopic to $\bigcup_* BXC_G(g)$ we can repeat the above construction to get a bundle $\pi_V \colon M \to ES^1 \times_{S^1} LBG$ whose fiber over a point in $BXC_G(g)$ is $\mathbb{P}V(g)$. Note that in this situation, the grading on $V(g)$ is now indexed by a character of S^1. This is related to modular-invariance properties, which we discuss further in Section 10. This aspect of the theory is crucial to the very existence of elliptic cohomology, though our own manipulations have been purely formal. The reader might compare our discussion with Segal's notion of "elliptic object" [Seg2].

Problem 10. *Clarify the relationship between elliptic cohomology and VOA theory.*

To complete this section we consider another aspect of group cohomology in VOA theory, which we call the *Dijkgraaf–Witten cocycle.* To be on safe(r) ground, let us assume that V is holomorphic and that $\{V(g)\}$ constitutes all the simple twisted V-modules.

In [DW] it is argued on topological grounds that there is a cohomology class $[\delta] \in H^4(G, \mathbb{Z})$ which determines $[\alpha]$, and thereby puts conditions on the nature of $[\alpha]$ which go beyond the containment $[\alpha] \in HC_0^3(\mathbb{Z}G)$ which we have already established. We call δ the Dijkgraaf–

Witten cocycle. To be precise, if we think of δ as an element of $C^3(G, \mathbb{C}^*)$ then one expects that the following holds (up to coboundaries):

$$\alpha_g(h,k) = \frac{\delta(hkgk^{-1}h^{-1}, h, k)\delta(h,k,g)}{\delta(h, kgk^{-1}, k)}. \tag{7.3}$$

See* [DPR] for more information on this point. We note here only that δ is related to the issue of tensor products of V-modules, which we address in Section 9. Note that one see easily that the r.h.s. of (7.3) indeed defines an element of $HC_0^3(\mathbb{Z}G)$, so that it defines a map

$$DW\colon H^4(G, \mathbb{Z}) \to HC_0^3(\mathbb{Z}G). \tag{7.4}$$

In this notation we have

Problem 11. *Prove that $[\alpha]$ lies in the image of the map DW. Can one describe $[\delta]$ in terms of G and its action on V?*

An affirmative answer to the second part of Problem 11 would thus provide a canonical element of $H^4(G, \mathbb{Z})$ associated with the pair (V, G). A further topological significance of this is that one knows from the work of a number of authors ([DW], [FQ], [F], [Y] for example) that there is a canonical topological quantum field theory (TQFT) associated with a finite group G and an element of $H^4(G, \mathbb{Z})$. So one would get a map

$$\text{VOA} \to \text{TQFT}. \tag{7.5}$$

This is in accord with G. Segal's definition of conformal field theory (CFT) in [Segl], which is categorical in nature. From this point-of-view (7.5) is a forgetful functor from CFT to TQFT in which the conformal structure is forgotten, leaving only a topological theory.

Finally, we mention

Problem 12. *Let $[\alpha]$ lie in the image of DW. Is there a holomorphic VOA V with $G \leq \operatorname{Aut} V$ such that $[\alpha]$ arises as described above?*

§8. Fixed-point theory

We return to the theme of Section 6. It is a basic philosophy of the subject that the category of V-modules should be equivalent to the module category of some quasi-quantum group. In the case that V is

*Some of our formulae differ in appearance from those of [DPR]. See the footnote on page 123 in this regard.

a holomorphic VOA and G is a finite subgroup of $\mathrm{Aut}(V)$ one may consider the space V^G of G-invariants, which is evidently itself a VOA. In this case there is a precise conjecture, essentially described in [DPR], for what the quasi-quantum group is. We describe what is known about this situation, following [DM2]. But first we mention

Problem 13. *Let V be a holomorphic VOA with a finite group $G \leq \mathrm{Aut}(V)$. Prove that V^G is rational.*

We will always assume that V is holomorphic and that there is a unique simple g-twisted module $V(g)$ for each $g \in G$. Thus the results of Section 7 show that there is an element $[\alpha] \in HC_0^3(\mathbb{Z}G)$ canonically associated to this situation.

Let $D_\alpha(G)$ denote the smash product of $\mathbb{C}[G]$ with $\mathbb{C}[G]^*$ modified by α, where G acts on $\mathbb{C}[G]^*$ by right conjugation. Thus $D_\alpha(G) = \mathbb{C}[G] \otimes \mathbb{C}[G]^*$ as vector space with product*

$$a \otimes e(x) \cdot b \otimes e(y) = \alpha_y(a,b)ab \otimes e(b^{-1}xb)e(y). \tag{8.1}$$

So $D_\alpha(G)$ is essentially a twisted version of the quantum double of G (cf. [Dr]).

The space V^* ((7.2)) becomes a left $D_\alpha(G)$-module via the action $a \otimes e(x) \cdot m = \delta_{x,g}\phi_x(a)m$ for $m \in V(g)$ and $\phi_x(a)\colon V(x) \to V(axa^{-1})$ as in Section 7. Now from definition (4.3) we see that the restriction of the action of V on each $V(g)$ to V^G yields an ordinary V^G-module, so that V^* is in fact a V^G-module. Moreover the actions of V^G and $D_\alpha(G)$ commute, so that what obtains is an action of $D_\alpha(G) \otimes V^G$ (in the obvious sense) on V^*.

The simple modules for $D_\alpha(G)$ are easily classified ([DPR], [Ma2]), and it is known ([DM2],[DM3]) that *every* such module occurs in V^*. Thus we may write

$$V^* = \bigoplus_\chi M_\chi \otimes V_\chi \tag{8.2}$$

where χ ranges over the simple characters of $D_\alpha(G)$ and M_χ is a $D_\alpha(G)$-module affording χ. We may take V_χ to be a V^G-module, and if it has a grading of the form $V_\chi = \coprod_{n\in\mathbb{C}} V_{\chi,n}$ then we may think of $\sum_n(\dim V_{\chi,n})$ q^n as the *graded multiplicity* of M_χ in V^*. (For now, q is merely an indeterminate.)

The two main problems are the following:

*The map $a \otimes e(x) \mapsto a \otimes e(axa^{-1})$ will transform (8.1) into the definition used in [DPR].

Problem 14. *Prove that each V_χ in (8.2) is a simple V^G-module.*

Problem 15. *Prove that every simple V^G-module is contained in V^*.*

Problem 14 is proved in [DM2] for the case that G is a nilpotent group, and a similar proof will almost certainly deal with G solvable. But the general case is more difficult.

Now $D_\alpha(G)$ is semi-simple. So if we know also that V^G-modules are completely reducible, then affirmative solutions of Problems 14 and 15 provide us with a categorical equivalence

$$D_\alpha(G)\text{-Mod} \xrightarrow{\cong} V^G\text{-Mod} \tag{8.3}$$

$$M_\chi \mapsto V_\chi.$$

Suppose that α is trivial. Then as explained in [DPR] or [Ma2], the simple $D_\alpha(G) = D(G)$-modules are naturally indexed by pairs (g, χ_g) where g ranges over one element in each conjugacy class of G and χ_g ranges over the simple characters of $C_G(g)$. Hence, granted the equivalence (8.3), there are bijections between the following sets:

$$\begin{aligned} &\{\text{simple } D(G)\text{-modules}\},\\ &\{\text{simple } V^G\text{-modules}\},\\ &\mathrm{Hom}(\mathbb{Z}^2, G)/\text{conjugation by } G,\\ &\{\text{principle } G\text{-bundles over } S^1 \times S^1\}/G,\\ &\{\text{homotopy classes of maps } S^1 \times S^1 \to BG\}/G. \end{aligned} \tag{8.4}$$

Thus again we observe the convergence of ideas from topology and VOA theory.

§9. Tensor products

We discuss the quasi-quantum structure of $D_\alpha(G)$ and its conjectured meaning for V. We retain the notation of the preceding section.

Suppose to begin with that the cocycle α of the last section is trivial. Then $D_\alpha(G) = D(G)$ becomes a Hopf algebra with comultiplication $\Delta\colon D(G) \to D(G) \otimes D(G)$ defined by

$$\Delta\colon a \otimes e(x) \mapsto \sum_{\substack{h,k \in G\\ hk = x}} a \otimes e(h) \otimes a \otimes e(k). \tag{9.1}$$

(There is also a counit and antipode which, however, will not concern us here. See [DPR].) With this structure we may define a (strictly associative) tensor product of $D(G)$-modules, so that $D(G)$-Mod becomes a *strict monoidal category.* (For this and other categorical ideas in this section, we refer the reader to [SJ] or [SS].)

Problem 16. *Assume that α is trivial. Show that there is a notion of tensor product of V^G-modules such that the equivalence* (8.3) *is one of strict monoidal categories.*

In fact, the problem of defining tensor products of (twisted) modules for general VOAs is fundamental, and impinges on a number of the problems that we discuss in this paper. Given two modules X, Y for a VOA V, the tensor product $X \boxtimes Y$ should again be a V-module, but $X \boxtimes Y$ will not resemble the ordinary tensor product. For example, if V is holomorphic, then one expects that $V \boxtimes V \cong V$.

Huang and Lepowsky ([HL1], [HL2]) have considered this problem at great length. The paper [Li2] should also be consulted, along with the work [KL] which announces a significant advance concerning tensor products of Kac–Moody modules.

There are complications if the cocycle α is non-trivial, and we consider this case next. The main point is that $D_\alpha(G)$ should now be a *quasi-Hopf* algebra ([Dr], [SS]). This entails a modification of the coproduct (9.1) of the form

$$\Delta\colon a \otimes e(x) \mapsto \sum_{\substack{h,k \in G \\ hk=x}} \gamma_a(h,k) a \otimes e(h) \otimes a \otimes e(k) \tag{9.2}$$

where $\gamma_a(h,k) \in \mathbb{C}^*$ can be expressed in terms of the Dijkgraaf–Witten cocycle δ in a way that is entirely analogous to (7.3). One still requires Δ to be an algebra morphism from $D_\alpha(G)$ to $D_\alpha(G) \otimes D_\alpha(G)$ but it will no longer be associative. Instead one requires only that there is $\varphi \in D_\alpha(G)^{\otimes 3}$, φ invertible, such that for $d \in D_\alpha(G)$ we have

$$(\Delta \otimes \mathrm{id}) \circ \Delta(d) = \varphi((\mathrm{id} \otimes \Delta) \circ \Delta(d))\varphi^{-1}. \tag{9.3}$$

When this holds, the tensor product of $D_\alpha(G)$-modules is *quasi-associative* in the sense that $(X \otimes Y) \otimes Z$ and $X \otimes (Y \otimes Z)$ are isomorphic by an isomorphism which depends on X, Y, Z. One takes φ to be a linear combination of element of the form $1 \otimes e(g) \otimes 1 \otimes e(h) \otimes 1 \otimes e(k)$ as (g,h,k) range over G^3, with coefficients given by the Dijkgraaf–Witten cocycle (thought of as an element of $C^3(G,\mathbb{C}^*)$).

There is even more structure available: $D_\alpha(G)$ has a so-called *R-matrix*, which in this case is the element

$$R = \sum_{g,h\in G} 1 \otimes e(g) \otimes 1 \otimes e(h) \tag{9.4}$$

in $D_\alpha(G) \otimes D_\alpha(G)$. The first property of R is that it satisfies for $d \in D_\alpha(G)$:

$$\tau \circ \Delta(d) = R\Delta(d)R^{-1} \tag{9.5}$$

where τ is the twist operator on $D_\alpha(G)\otimes D_\alpha(G)$. Equation (9.5) ensures that $X \otimes Y \cong Y \otimes X$ for $D_\alpha(G)$-modules via an isomorphism that depends once again on X and Y. Two further properties of R allow us to deduce the so-called *quasi-Yang–Baxter equations*, and together these axioms give $D_\alpha(G)$ the structure of a *quasi-triangular quasi-Hopf algebra.*

Needless to say, all of this structure on $D_\alpha(G)$ and its module category is expected to reflect analogous properties of the category V^G-Mod. We will merely state the main problem, the extension of Problem 17 to the general case.

Problem 17. *Show that there is a notion of tensor product of V^G-modules such that the equivalence* (8.3) *is one of braided monoidal categories.*

§10. Modular-invariance

We discuss the conjectured modular-invariance properties of the characters of modules of rational VOAs, concentrating on the case of holomorphic orbifolds, i.e., VOAs of the form V^G where V is holomorphic and $G \leq \operatorname{Aut} V$ is finite (cf. Section 8). We assume some familiarity with elliptic modular functions.

Assume to begin with that V is a VOA and M a simple V-module. Then according to (4.5), M is graded with shape $M = \coprod_{n=0}^{\infty} M_{c+n}$ with $M_c \neq 0$. The (graded) character of M is defined to be

$$ch\, M = q^c \sum_{n=0}^{\infty} (\dim M_{c+n}) q^n. \tag{10.1}$$

According to convenience, we take q to be either an indeterminate or of the form $q = e^{2\pi i \tau}$ with $\tau \in \mathbb{C}$ satisfying $\mathrm{im}\tau > 0$. Thus τ lies in the upper-half plane $\mathfrak{H}$.

Problem 18. *Is it true that* $ch\, M$ *is a holomorphic function on* $\mathfrak{H}$*?*

Let us now take V to be a rational VOA (cf. (4.5)) with simple V-modules $M_1, \ldots, M_k$. Let c' denote the central charge of V, and set

$$ch'\, M_i = q^{-c'/24}\, ch\, M_i. \tag{10.2}$$

Inclusion of the factor $q^{-c'/24}$ may seem rather strange, but it considerably enhances the invariance properties of the character. If $\Gamma = SL(2, \mathbb{Z})$ is the modular group then Γ acts on $\mathfrak{H}$ via Möbius transformations in the usual way, and thence on holomorphic functions on $\mathfrak{H}$.

Problem 19. *Let E be the $\mathbb{C}$-space spanned by the functions $ch'\, M_i$, $1 \leq i \leq k$. Prove that E is invariant under the action of Γ. Prove, moreover, that the kernel of the action is a congruence subgroup of Γ.*

The work of Kac–Peterson [KP], [K] on the analogue of Problem 19 for Kac–Moody Lie algebras is very important. And the paper of Zhu [Z] establishes the truth of Problem 19 for VOAs satisfying certain additional finiteness conditions which are, so far as we know, satisfied in all of the known rational VOAs.

As illustration, suppose that V is holomorphic so that $k = 1$ and $M_1 = V$. Problem 19 suggests that $ch'\, V$ is then a (semi-)invariant of Γ, i.e., if we set $Z(\tau) = ch'(V)$ then $Z(\gamma\tau) = \varepsilon(\gamma)Z(\tau)$ for some character ε of Γ. This expectation is borne out in practice: for example, if L is a self-dual lattice and V_L the corresponding VOA (Section 3, Example (iv)) then V_L is holomorphic (Section 5, Example (iv)) and $ch'\, V_L = \theta_L(\tau)/\eta(\tau)^{\dim L}$. Here $\theta_L(\tau)$ is the theta-function of L and $\eta(\tau)$ the Dedekind eta-function. It is well-known [Se] that $\theta_L(\tau)$ is a form level one and that $\eta(\tau)^{\dim L}$ transforms under Γ with a character of order 1 or 3 (since 8 divides $\dim L$). If $V^\natural$ is the Moonshine module (Sections 3 and 5, Example (vi)) then $ch'\, V^\natural = J(q)$ is the modular function $q^{-1} + 0 + 196884q + \cdots$ of level one.

Now let V again be holomorphic of central charge c' with finite $G \leq \operatorname{Aut} V$. In this case there is a beautiful and sharp reformulation of Problem 19 for the VOA V^G which goes to the heart of "Moonshine." To describe this we need to assume that the simple V^G-modules are as described in Problems 13–15. Thus after choosing a g in each conjugacy class of G, we assume existence and uniqueness of a simple g-twisted V-module, say $V(g)$. It is convenient to assume that the cocycle α is trivial—the general case involves only cosmetic changes.

Now $V(g)$ will be graded as in (4.9), that is $V(g)=\coprod_{n=0}^{\infty} V(g)_{c(g)+\frac{n}{N}}$ with $N = o(g)$ and $c(g)$ a scalar (presumably rational). As $V(g)$ admits $C(g)$ we may define for $h \in C(g)$,

$$Z(g,h,\tau) = q^{c(g)} \sum_{n=0}^{\infty} (\operatorname{tr} h | V(g)_{c(g)+\frac{n}{N}}) q^{\frac{n}{N}}. \tag{10.3}$$

According to (8.2) the simple V^G-modules M contained in $V(g)$ are parameterized bijectively by the simple modules (or characters) of $C(g)$, and the character *ch M* of M is the graded multiplicity of the corresponding character of $C(g)$. By the orthogonality relations for the characters of $C(g)$, the space spanned by *ch M* as M ranges over the simple V^G-modules in $V(g)$ is that spanned by the functions $Z(g,h,\tau)$ as h ranges over one element in each conjugacy class in $C(g)$. Thus one sees that the space E of characters of V^G (Problem 19) is also spanned by the functions

$$Z'(g,h,\tau) = q^{-c'/24} Z(g,h,\tau) \tag{10.4}$$

as (g,h) ranges over the set $P = P(G)$. Here P is a set of representatives of $\{(x,y) \in G \times G | xy = yx\}$ modulo simultaneous conjugation by G. (Hence we have $|P(G)| = \#$simple V^G-modules—cf. (8.4).)

Now Γ acts on the right of P via $(x,y)\cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = (x^a y^c, x^b y^d)$. An element of P may be thought of as an element of $\operatorname{Hom}(\pi_1(S^1 \times S^1), G)$, where the prescribed action of Γ is essentially that on a homology basis of the torus. Now we have

Problem 20. *Let E be the space spanned by the functions $ch'\, M$ as M ranges over the simple V^G-modules. Then E is also spanned by the trace functions $Z'(g,h,\tau)$, $(g,h) \in P$. Moreover E is a Γ-module and the action of $\gamma \in \Gamma$ satisfies*

$$Z'(g,h,\gamma\tau) = \varepsilon(g,h,\gamma) Z'((g,h)\gamma,\tau) \tag{10.5}$$

for some root of unity $\varepsilon(g,h,\gamma)$.

The condition (10.5) was introduced by Norton in an appendix to the paper [Ma1] on an empirical basis, and was subsequently extended to more general contexts in [Ma3], [Ma4]. Its origins in CFT can be traced through [DHW], [DGH]. In [DM1], (10.5) is actually established for certain fermionic orbifolds (Section 3, Example (v)). Little else is *known* to be true save for some calculations in [DLiM3] concerning the Moonshine module $V^\natural$ and its involutorial automorphisms.

Included in Problem 20 is the conjecture that (taking $g = 1$), the graded characters of $h \in G$ on V, more precisely the functions $Z'(1, h, \tau)$, are modular functions. Moreover one expects them to be semi-invariants of $\Gamma_0(M)$ where $M = o(h)$. Of course this is precisely how Monstrous Moonshine got started: the paper of Conway and Norton [CN] made very precise conjectures about the functions $Z'(1, h, \tau)$ in the case $V = V^\natural$ and $G = \mathbb{M}$ (Monster). These conjectures were subsequently established by Borcherds [Bo2], although his methods went outside the theory of VOAs as described here and utilized his theory of generalized Kac–Moody Algebras. Perhaps this is to be expected; after all, the Conway–Norton conjectures specify that in the case of $V^\natural$ the trace functions $Z'(1, h, \tau)$, are so-called *hauptmoduls.* These are very special kinds of modular functions, and certainly for other VOAs the corresponding trace functions are *not* hauptmoduls.

Problem 21. *Can Borcherds' theorem (Conway–Norton conjectures) be established via the representation theory of VOAs?*

Michael Tuite has studied this problem from the viewpoint of conformal field theory [Tu1], [Tu2]. His very interesting ideas suggest a specific path towards a solution to Problem 21 based on so-called orbifolding.

Problem 22. *Let $V = V^\natural$ be the Moonshine Module. Prove that each $Z'(g, h, \tau)$ is either a hauptmodul or is constant.*

This is Norton's Generalized Moonshine Conjecture. Actually, we have already pointed out in a previous section that for $V^\natural$ the cocycle α is not trivial. Thus one has to suitably interpret $Z'(g, h, \tau)$ in order to make the problem meaningful. Borcherds' methods do not appear to extend to this situation, whereas in principle those of Tuite do.

Problem 23. *What is the significance of the hauptmodul property of the trace functions on the Moonshine Module?*

One can say that the hauptmodul property is a deep-seated consequence of properties of the Leech lattice, but this doesn't seem to help much. We also note that in [DM4] it is shown that a certain fermionic orbifold, with group G being the large Mathieu group M_{24}, enjoys the hauptmodul property. In fact it satisfies the conclusions of Problem 22. The Leech lattice plays no rôle in this latter work though M_{24} is of course intimately related to it.

§11. Quantum Galois theory

It is well-known that VOAs behave in some ways like commutative, associative algebras. So a *simple* VOA V (i.e., one with no proper V-submodule) may be likened to a commutative field. Quantum Galois Theory (QGT) is a manifestation of this analogy.

We fix a simple VOA (V, Y) with finite group $G \leq \operatorname{Aut} V$. A sub VOA of V is a subspace $W \subseteq V$ preserved by all component operators w_n of $Y(w, z)$ for $w \in W$, and with the same vacuum vector and conformal vector as V.

Problem 24. *Prove that there is a Galois correspondence (i.e., a containment-reversing bijection) between subgroups of G and sub VOAs of V containing V^G given by*

$$H \mapsto V^H. \tag{11.1}$$

This was conjectured in [DM3], and the case in which G is *nilpotent* or *dihedral* is settled affirmatively in [DM2], [DM3]. It is also known (*loc. cit.*) that (11.1) is an injection, and that if G is *solvable* then (11.1) gives a bijection between normal subgroups of G and G-invariant sub VOAs of V which contain V^G. An interesting point is that we know (*loc. cit.*) that if V is a simple VOA and $G \leq \operatorname{Aut} V$ is finite then V^G is also simple. This enables one to proceed inductively and, if G is solvable at least, to reduce certain questions to the case in which G is abelian. But if G is not solvable we have no clue at present; what is lacking is any perspective which would at least explain why there *should* be a Galois correspondence.

Another point is this: if in (8.2) we restrict our attention to $V(1) = V$, then we obtain

$$V = \bigoplus_{\chi} M_\chi \otimes V_\chi \tag{11.2}$$

where χ ranges over the simple characters of G, M_χ a simple G-module affording χ, and V_χ a certain V^G-module. Now (8.2) was a decomposition for holomorphic V, but (11.2) holds for a simple VOA V.

Problem 25. *Prove that in (11.2) with V simple, each V_χ is a non-zero simple V^G-module.*

This is known for solvable groups (*loc. cit.*). We may regard (11.2) as the QGT analogue of the normal basis theorem, which says that if

E/K is a normal extension of fields with (finite) Galois group G, then E is a regular KG-module.

Armed with a character ψ of G and normal extension E/K with Galois group G, one may construct the Artin L-series $L_{E/K}(\psi, s)$ and ask after its analytic properties. The QGT analogue might then be the analytic properties of the characters of the V_χ in (11.2): of course this is an *additive* theory, whereas $L_{E/K}(\psi, s)$ is, by definition, an Euler product.

Let us take V holomorphic to be on safe ground (a churlish phase at this stage of the proceedings!). Then the character $ch' V$ (cf. Section 10) is the analogue of the zeta-function $\zeta_{E/K}(s)$, and one expects $ch' V$ to be a modular function of level 1 (possibly with character). The characters $ch' V_\chi$ are the analogues of the L-series $L_{E/K}(\chi, s)$. The decomposition of $\zeta_{E/K}(s)$ into factors $\zeta_{E/K}(s) = \prod_\chi L_{E/K}(\chi, s)^{\deg \chi}$ is just a formal consequence of the basis theorem, and its QGT analogue comes from (11.2), namely $ch' V = \sum_\chi (\deg \chi)\, ch' V_\chi$.

Problem 26. *Is there a QGT-analogue of the Euler products which define the L-series?*

Put another way, one might ask if there is a local theory of VOAs. Recent remarkable work of Borcherds and Ryba [R], [BR], [Bo3] suggests that the answer is "yes"! Not only that, but there seem to be connections between local VOAs (whatever they are) and twisted sectors.

Added in Proof (November 3rd. 1995): Since this paper was written there has been some progress towards solving several of the 26 problems. We briefly discuss these newer results here.

Following Problem 2 we introduced the notion of a weak g-twisted V-module. There is a useful variant of this which is intermediate between weak module and ordinary module, namely 'admissible' module. An admissible (g-twisted) V-module is a weak module such that the component operators of $Y(v, z)$ for homogeneous v preserve the homogeneous spaces of the module in the usual way. This idea is important in Zhu's work [Z]. It is now established in [DLiM2] that if all admissible g-twisted V-modules are completely reducible (a strengthening of (4.5)(b)), then (4.5)(a) holds and also (4.5)(c) holds for admissible modules. Compare these results with Problem 3. Further results in this direction are contained in another preprint (Regularity of rational vertex operator algebras) of the authors and H. Li. Let us now call a VOA g-rational if indeed all of its admissible g-twisted modules are completely reducible.

We referred to certain finiteness conditions in [Z] following Problem 19. Essentially, these are (g-)rationality in the sense of the last paragraph together with a certain "C_2 Condition"(loc.cit.). In forthcoming work of the authors and Li (Elliptic functions and orbifold theory) the following are established: affirmative solution of Problem 4 in case V is both rational, g-rational, and satisfies C_2; affirmative solution of Problem 5 under the same conditions; affirmative solution of Problems 6 and 8 if V is rational and satisfies C_2; affirmative solution of Problem 7. Problem 20 is also solved under similar conditions on V, though with the proviso that there is no information available with regard to the intervening constant — which should be a root of unity.

References

[AM] G. Anderson and G. Moore, Rationality in conformal field theory, Comm. Math. Phys., **117** (1988), 441–450.

[ATLAS] J. Conway et al., "Atlas of Finite Groups", Clarendon Press, Oxford, 1985.

[Ba1] P. Bantay, Orbifolds, Hopf algebras, and the moonshine, Lett. Math. Phys., **22** (1991), 187–194.

[Ba2] ———, Orbifolds and Hopf algebras, Phys. Lett. B, **245** (1990), 477–479.

[Bak] A. Baker, An introduction to elliptic cohomology of classifying spaces, preprint.

[Bo1] R. Borcherds, Vertex algebras, Kac–Moody algebras, and the Monster, Proc. Nat. Acad. Sci. U.S.A., **83** (1986), 3068–3071.

[Bo2] ———, Monstrous moonshine and monstrous Lie superalgebras, Invent. Math., **109** (1992), 405–444.

[Bo3] ———, Modular moonshine III, preprint.

[BR] R. Borcherds and A. Ryba, Modular moonshine II, preprint.

[Be] D. Benson, "Representations and Cohomology I, II", Cambridge studies in advanced mathematics, C. U. P., Cambridge, 1991.

[BPZ] A. Belavin, A. Polykov and A. Zamolodchikov, Infinite conformal symmetry in two-dimensional quantum field theory, Nuclear Phys. B, **241** (1984), 333–380.

[Br] J.-J. Brylinski, Representations of loop groups, Dirac operators on loop spaces, and modular forms, Topology, **29** (1990), 461–480.

[Bu] D. Burghelea, The cyclic homology of the group rings, Comment. Math. Helv., **60** (1985), 354–365.

[C] J. Conway, A simple construction for the Fischer–Griess Monster group, Invent. Math., **79** (1985), 513–540.

[CN] J. Conway and S. Norton, Monstrous moonshine, Bull. London Math. Soc., **11** (1979), 308–339.

[CS] J. Conway and N. Sloane, "Sphere Packing, Lattices and Groups", Springer-Verlag, New York, 1988.

[DGH] L. Dixon, P. Ginsparg and J. Harvey, Beauty and the beast: Superconformal symmetry in a Monster module, Comm. Math. Phys., **119** (1988), 221–241.

[DHVW] L. Dixon, J. Harvey, C. Vafa and E. Witten, Strings on Orbifolds I, II, Nuclear Phys. B, **261** (1985), 678–686; Nuclear Phys. B, **274** (1986), 285–314.

[DPR] R. Dijkgraaf, V. Pasquier and P. Roche, Quasi-quantum groups related to orbifold models, in "Modern quantum field theory (Bombay, 1990)", World Scientific, 1991, pp. 375–383.

[DLM] C. Dong, Z. Lin and G. Mason, On vertex operator algebra as sl_2-modules, Proc. Monster Conf. at O.S.U., to appear.

[DLiM1] C. Dong, H. Li and G. Mason, Simple currents, abelian intertwining algebras and extended vertex operator algebras, preprint.

[DLiM2] ———, Twisted representations of vertex operator algebras, preprint.

[DLiM3] ———, Some twisted sectors for the moonshine module, Proc. Monster Conf. at Mount Holyoke, to appear.

[DL] C. Dong and J. Lepowsky, "Generalized Vertex Algebras and Relative Vertex Operators", Birkhäuser, Boston, 1993.

[DM1] C. Dong and G. Mason, Non-abelian orbifolds and the boson-fermion correspondence, Comm. Math. Phys., **163** (1994), 523–559.

[DM2] ———, On the operator content of nilpotent orbifold models, preprint.

[DM3] ———, Quantum Galois theory, preprint.

[DM4] ———, An orbifold theory of genus zero associated to the sporadic group M_{24}, Comm. Math. Phys., **164** (1994), 87–104.

[DMZ] C. Dong, G. Mason and Y. Zhu, Discrete series of the Virasoro algebra and the moonshine Module, Proc. Sympos. Pure Math., **56** (1994), 295–316, Part 2.

[Do1] C. Dong, Vertex algebras associated with even lattices, J. Algebra, **161** (1993), 245–265.

[Do2] ———, Representations of the moonshine module vertex operator algebra, Contemp. Math., **175** (1994), 27–36.

[Do3] ———, Twisted modules for vertex algebras associated with even lattices, J. Algebra, **165** (1994), 91–112.

[DGM1] L. Dolan, P. Goddard and P. Montague, Conformal field theory of twisted vertex operators, Nuclear Phys. B, to appear.

[DGM2] ———, Conformal field theories, representations and lattice construction, preprint.

[Dr] V. Drinfeld, Quantum Groups, in "Proc. Int'l. Cong. Math. Berkeley", American Math. Soc., 1986, pp. 798–820.

[DW] R. Dijkgraaf and E. Witten, Topological gauge theories and group cohomology, Comm. Math. Phys., **129** (1990), 393–429.

[F] I. Frenkel, Two constructions of affine Lie algebra representations and boson-fermion correspondence in quantum field theory, J. Funct. Anal., **44** (1984), 259–327.

[FLM] I. Frenkel, J. Lepowsky and A. Meurman, "Vertex Operator Algebras and the Monster", Academic Press, New York, 1988.

[FHL] I. Frenkel, Y.-Z. Huang and J. Lepowsky, On axiomatic approaches to vertex operator algebras and modules, Mem. Amer. Math. Soc., **104** (1993), no. 494.

[Fe] K. Ferguson, Link invariants associated to TQFT's with finite gauge groups, J. Knot Theory Ramifications, **2** (1993), 11–36.

[FFR] A. Feingold, I. Frenkel and J. Ries, Spinor construction of vertex operator algebras, triality and $E_8^{(1)}$, Contemp. Math., **121** (1991).

[FQ] D. Freed and F. Quinn, Chern–Simons theory with finite gauge group, Comm. Math. Phys., **156** (1993), 435–472.

[FQS] D. Friedan, Z. Qiu and S. Shenker, Conformal invariance, unitarity and two-dimensional critical exponents, in "M.S.R.I. Publ. No.3", Springer-Verlag, New York, 1985, pp. 419-449.

[FZ] I. Frenkel and Y. Zhu, Vertex operator algebras associated to representations of affine and Virasoro algebras, Duke Math. J., **66** (1992), 123–168.

[G] P. Goddard, Meromorphic conformal field theory, in "Infinite-Dimensional Lie Algebras and Lie groups", World Scientific, Singapore, 1989.

[Gor] D. Gorenstein, "Finite Groups", Harper & Row, New York, 1968.

[Gr] R. Griess, The friendly giant, Invent. Math., **69** (1982), 1–102.

[GKO] P. Goddard, A. Kent and D. Olive, Unitary representations of the Virasoro algebra and super-Virasoro algebras, Comm. Math. Phys., **103** (1986), 105–119.

[Hu] Y.-Z. Huang, A nonmeromorphic extension of the moonshine module vertex operator algebra, Proc. Monster Conf. at Mount Holyoke, to appear.

[Hi] F. Hirzebruch, "Manifolds and Modular forms", Vieweg, 1992.

[HL1] Y.-Z. Huang and J. Lepowsky, Towards a theory of tensor product for representations for a vertex operator algebra, in "Proc. 20th International Conference on Differential Geometric Methods in Theoretical Physics", World Scientific, Singapore, 1992.

[HL2] ———, A theory of tensor product for module category of a vertex operator algebra I, II, preprint.

[JS] A. Joyal and R. Street, "Braided Monoidal Categories", Macquarie Math. Reports, Macquarie Univ., 1986.

[Ko] M. Kontsevich, Feynman diagrams and low dimensional topology, preprint.

[K] V. Kac, "Infinite-dimensional Lie Algebras", 3rd ed., C.U.P., Cambridge, 1990.

[KL] D. Kazhdan and G. Lusztig, Affine Lie algebras and quantum groups, Intl. Math. Research Not. (Duke Math. J.), **2** (1991), 21–29.

[KP] V. Kac and D. Peterson, Infinite-dimensional Lie algebras, theta-functions and modular forms, Adv. in Math., **53** (1984), 125–264.

[KR] V. Kac and A. Raina, "Bombay Lectures on Highest Weight Representations", World Scientific, Singapore, 1987.

[Le] J. Lepowsky, Calculus of twisted vertex operators, Proc. Nat. Acad. Sci. U.S.A., **82** (1985), 8295–8299.

[La] P. Landweber (Ed.), "Elliptic Curves and Modular Forms in Algebraic Topology", Lecture Notes in Math., no.1326, Springer-Verlag, Berlin, 1988.

[Lo] J.-L. Loday, "Cyclic Homology", Grundlehren, Springer-Verlag, Berlin, 1992.

[Li1] H. Li, Local systems of vertex operators, vertex superalgebras and modules, J. Pure Appl. Algebra, to appear.

[Li2] ———, Representation theory and tensor product theory for vertex operator algebras, Ph.D. thesis, Rutgers.

[Ma1] G. Mason, Finite groups and modular functions, Proc. Sympos. Pure Math., **47** (1987), 181–210.

[Ma2] ———, The quantum double of a finite group and its rôle in conformal field theory, Proc. L.M.S. Conf. in Galway, to appear.

[Ma3] ———, Elliptic systems and the eta-function, Notas Soc. Mat. Chile, **8** (1989), 37–53.

[Ma4] ———, Hecke operators and conformally-invariant orbifold models, preprint.

[MN] W. Meyer and W. Neutsch, Associative subalgebras of the Griess algebra, J. Algebra, **158** (1993), 1–17.

[Mi] M. Miyamoto, Griess Algebras and conformal vectors in vertex operator algebras, preprint.

[MS] G. Moore and N. Seiberg, Classical and quantum conformal field theory, Comm. Math. Phys., **123** (1989), 177–254.

[PS] M. Penkava and A. Schwarz, A_∞ algebras and the cohomology of moduli spaces, preprint.

[R] A. Ryba, Modular moonshine?, Proc. Monster Conf. at Mount Holyoke, to appear.

[Se] J.-P. Serre, "A Course in Arithmetic", Graduate Texts in Mathematics Vol.7, Springer-Verlag, New York, 1973.

[Seg1] G. Segal, The definition of conformal field theory, preprint.

[Seg2] ———, "Elliptic cohomology", Sem. Bourbaki no. 695, 1987-88.

[SS] S. Shnider and S. Sternberg, "Quantum Groups", Graduate Texts in Mathematical Physics vol. V, International Press Publications, Hong Kong, 1993.

[Ta] H. Tamanoi, Elliptic genera and vertex operator superalgebras, preprint.

[Tu1] M. Tuite, Monstrous moonshine from orbifolds, Comm. Math. Phys., **146** (1992), 277–309.

[Tu2] ——, On the relationship between monstrous moonshine and the uniqueness of the moonshine module, Comm. Math. Phys., to appear.

[Th1] C. Thomas, Elliptic cohomology of the classifying space of the Mathieu group M_{24}, Amer. Math. Soc., Contemp. Math., **158** (1991), 307–318.

[Th2] ——, "O.S.U. Lecture Notes on Elliptic Cohomology and Moonshine", to appear.

[W] W. Wang, Rationality of Virasoro vertex operator algebras, Intl. Math. Res. Notices (Duke Math. J.), **7** (1993), 197–211.

[Y] D. Yetter, Topological quantum field theories associated to finite groups and crossed G-sets, J. Knot Theory Ramifications, **1** (1992), 1–20.

[Z] Y. Zhu, Modular invariance of characters of vertex operator algebras, preprint.

Department of Mathematics
University of California
Santa Cruz, CA 95064-1099
U.S.A.

Advanced Studies in Pure Mathematics 24, 1996
Progress in Algebraic Combinatorics
pp. 137–163

Covers of Complete Graphs

Chris D. Godsil

Abstract.

We study antipodal distance-regular covers of complete graphs. The first part of the paper gives an introduction to the basic theory and the main constructions. In the second part, we turn to *linear* covers, which can be described geometrically.

§1. Covers

If we view a graph as a 1-dimensional simplicial complex then a cover of a graph is a cover in the topologist's sense. Fortunately there is a simple combinatorial approach.

Let X be a graph with vertex set $V(X)$. An *arc* of X is an ordered pair of adjacent vertices, and an *arc function of index* r is any function f from the arc set of X into $\mathrm{Sym}(r)$, the symmetric group of degree r, such that

$$f(a,b)f(b,a) = 1 \tag{1.1}$$

for any arc (a, b) of X. Given an arc function f of index r on X we construct a covering graph X^f as follows. The vertex set of X^f is the Cartesian product $V(X) \times \{1, \dots, r\}$ and the vertices (a, i) and (b, j) are adjacent if

$$if(a,b) = j.$$

Our condition on arc functions guarantees that X^f is an undirected graph.

If f is the identity on each arc of X then X^f is just r disjoint copies of X. If $r = 2$ and f is equal to the non-identity element of $\mathrm{Sym}(2)$ on each arc then X^f can be seen to be the direct product $X \times K_2$. The mapping p from $V(X^f)$ to $V(X)$ such that $p(a,i) = a$ for all i in $\{1, \dots, r\}$ is called

Received March 28, 1995.

Support from grant OGP0009439 of the National Sciences and Engineering Council of Canada is gratefully acknowledged.

a *covering map*. It can be shown to be a local isomorphism from X^f to X, although this will play no role in our work here. The set of vertices $\{(a,i) : i = 1, \ldots, r\}$ is called a *fibre* of the cover. The subgraph of X^f induced by a fibre is an independent set, and the subgraph induced by two fibres is either an r-matching or an independent set, according as p maps the two fibres to adjacent or non-adjacent vertices in X.

If f and g are arc functions with index r and s respectively, we can define their *product* as follows. The ordered pair (f,g) can be viewed as an arc function taking values in Sym(r) $\times$ Sym(s), viewed as a permutation group on rs points. Hence Sym(r) $\times$ Sym(s) is a subgroup of Sym(rs), in its natural action, and therefore (f,g) is an arc function of index rs. (The corresponding cover might not be connected, even if X^f and X^g are.)

The subgroup of $\mathrm{Aut}(X^f)$ consisting of the automorphisms that fix each fibre is called the *covering group*. It is an easy exercise to show that, if X^f is connected, this group acts semi-regularly. We call a cover *regular* if the covering group acts regularly on the vertices in a fibre. It is another easy exercise to show that a 2-fold cover is always regular; its covering group is cyclic with index two. The size of a fibre is equal to the index of f, and is called the *index* of the cover. A cover of index r is also called an r-fold cover.

If X is connected and T is a spanning tree in X then it can be shown that any cover of X is isomorphic to a cover X^f, where f is the identity map on each arc of T. Arc functions of this form will be said to be *normalised* with respect to T. If f is normalised with respect to some spanning tree of X we define $C(f)$ to be the subgroup of Sym(r) generated by the elements in the image of f. We call it the *connection group* of f. The important properties of the connection group are summarised in the following combination of results from Section 7 of [18].

Lemma 1.1. *Let f be a normalised arc function with index r on a connected graph X and let C be the connection group of f. Then X^f is connected if and only if C is transitive. The covering group of X^f is isomorphic to the centraliser of C in* Sym(r).

A graph with diameter d is *antipodal* if the relation "is equal or at distance d from" is an equivalence relation on its vertices. An antipodal cover is a cover such that the fibres are the antipodal classes; a covering graph may be antipodal without being an antipodal cover. Two vertices in distinct fibres of an antipodal cover of diameter d must be at distance at most $d-1$. By way of example, if f is an arc function of index two on K_4 and f is not equal to the identity on any arc, then X^f is the cube, which is an antipodal cover of K_4. The line graph of the Petersen graph can also be seen to be an antipodal cover of K_5 with index three.

Antipodal covers arise naturally in the study of distance regular graphs. If X is a graph with vertex set V, let X_i denote the graph with vertex set V, where two vertices are adjacent in X_i if and only if they are at distance i in X. Suppose that X is distance regular with diameter d. Then X is *primitive* if the graphs $X_1, \dots, X_d$ are all connected, and *imprimitive* otherwise. If X is imprimitive then either X_2 is not connected, and X is bipartite, or X_d is not connected and X is antipodal. Further, it can be shown that there is a distance regular graph Y such that X is a covering graph of Y. If Y has diameter d then the diameter of X is $2d$ or $2d + 1$. The fibres of the cover are the antipodal classes of X, that is, the vertex sets of the components of X_d. The cube and the line graph of the Petersen graph provide two examples.

§2. Parameters of covers

We now come to the main topic of this paper—antipodal distance regular covers of complete graphs. We will usually refer to these simply as covers. In this section we describe some of their properties. Any un-attributed results are taken from [18].

An antipodal graph with diameter two is complete multipartite, and consequently cannot be a covering graph unless each antipodal class is a single vertex. Thus it cannot be a non-trivial covering graph. Therefore an antipodal distance regular cover of K_n must have diameter three. In addition to its index, such a cover has two further important parameters. If X is distance regular then there are constants a_1 and c_2 such that any two adjacent vertices in X have exactly a_1 common neighbours, and two vertices at distance two have c_2 common neighbours. Any covering graph has the same valency as the graph it covers, hence a cover of K_n has valency $n - 1$.

For antipodal distance regular covers of K_n and index r we have

$$n - 2 - a_1 = (r - 1)c_2. \tag{2.1}$$

To prove this, let u be a fixed vertex in a cover of K_n and let v be adjacent to u. Then there are a_1 edges from v to a neighbour of u and hence there are $n-1-(a_1+1)$ edges joining v to a vertex at distance two from u. On the other hand there are exactly $r - 1$ vertices at distance three from u, each of these is at distance two from v and so has c_2 neighbours in common with v. This yields the identity. We will refer to the triple (n, r, c_2) as the *parameter set* of the cover.

If A is the adjacency matrix of a (n, r, c_2) cover then the eigenvalues of A are $n-1$, -1 and the two zeros θ and τ of the quadratic

$$x^2 - (a_1 - c_2)x - (n-1) = 0. \tag{2.2}$$

Because $\theta\tau = 1 - n$ it follows that, when $n \geq 2$, one of these zeros is positive and the other negative. When discussing covers of K_n, we will assume always that $n \geq 3$, and let θ be the positive zero of (2.2). The eigenvalues $n-1$ and -1 have multiplicities 1 and $n-1$ respectively. If m_θ and m_τ denote the multiplicities of θ and τ respectively then we have

$$m_\theta = \frac{n(r-1)\tau}{\tau - \theta}, \qquad m_\tau = \frac{n(r-1)\theta}{\theta - \tau}.$$

The fact that these multiplicities are integers is a strong constraint on the possible parameter sets (n, r, c_2). The eigenvalues θ and τ are integers if $m_\theta \neq m_\tau$; if $m_\theta = m_\tau$ then we must have $\theta = -\tau = \sqrt{n-1}$.

There are three basic conditions on the parameters of a cover with parameter set (n, r, c_2):

(F1) $1 \leq (r-1)c_2 \leq n-2$,
(F2) if n is even, c_2 is even,
(F3) m_θ and m_τ are integers.

Here (F1) is an easy consequence of (2.1). For (F2), see Section 3 in [18]. The theory of association schemes provides the Krein conditions, which in this setting reduce to the inequality

$$((r-1)^2 - 1)((n-1)^2 + \tau^3) \geq 0. \tag{2.3}$$

If $r > 2$, this reduces to

$$\theta^3 \geq n - 1.$$

There are a number of other conditions that a parameter set must satisfy. Most of these are given in [18], see also [16], [17].

It is natural to partition the parameter sets of covers by the value of $a_1 - c_2$, which we denote by δ. Evidence for this is provided by the following, which combines Theorem 3.6 and Lemma 3.7 from [18].

Theorem 2.1. *Let (n, r, c_2) be a parameter set satisfying conditions* (F1) *and* (F2). *If $\delta = 0$ then* (F3) *is satisfied, if $\delta = \pm 2$ then* (F3) *is satisfied if and only if n is square. For fixed r and δ, there are only finitely many parameter sets satisfying* (F1), (F2) *and* (F3), *unless $\delta = -2$, 0 or 2.*

§3. Examples

In this section we provide a number of examples of antipodal distance-regular covers of complete graphs.

Let H be the Hoffman-Singleton graph. This a 7-regular graph on 50 vertices with girth five and diameter two, and this implies each path of length three in H is contained in a unique 5-cycle. Let u be a fixed vertex in H and let $v_1, \dots, v_7$ be the seven neighbours of u. Let Y be the graph formed by the 42 vertices at distance two from u. Because X has diameter two and contains no 4-cycles, each vertex in Y is adjacent to exactly one of the neighbours of u. Therefore Y is regular with valency six. If two vertices in Y are adjacent to the same neighbour of u then they must lie at distance at least three in Y, or we would have a triangle or a 4-cycle in X.

Let F_i denote the set of six vertices in Y adjacent to v_i. If x in H is adjacent to v_i and $j \neq i$ then (x, v_i, u, v_j) is a path of length three; the fifth vertex y in the unique 5-cycle that contains it must lie in F_j. This shows that a vertex in F_i has exactly one neighbour in F_j. Accordingly Y is a covering graph of a 6-regular graph on seven vertices, necessarily K_7, with the sets F_i as fibres. With more effort, it can be shown that this is an antipodal distance-regular cover. Conversely, any antipodal distance-regular 6-fold cover of K_7 is the set of vertices at distance from a fixed vertex in a copy of the Hoffman-Singleton graph.

A Moore graph is a graph with diameter d and girth $2d+1$. From (F1) we know that if an (n, r, c_2)-cover exists then $r - 1 \leq n - 2$. If equality holds, $c_2 = 1$ and so, using (2.1), we also have $a_1 = 0$. Gardiner [12] noted that an antipodal distance-regular $(n-1)$-fold cover of K_n is equivalent to the existence of a Moore graph with diameter two and valency n. (So we would very much like to have a $(57, 56, 1)$-cover!)

Next we consider *regular two-graphs.* These have a fairly extensive literature. A good access point is provided by Seidel's selected works [25], which contains a number of papers on them, including two surveys.

A regular two-graph is essentially just an antipodal distance-regular cover of K_n with index two. To introduce them, we first consider arbitrary 2-fold covers of K_n. Let f be an arc-function of index two on K_n. Then, for any arc (i, j) we have $f(i, j) = f(j, i)$ and so f is determined by specifying the edges of K_n on which it takes the value -1. This set of edges is just a graph, which we will denote by $G(f)$. As noted in the introduction, we may always assume that an arc-function is equal to the identity on some spanning forest. In K_n choose T to be a copy of $K_{1,n-1}$ and only use arc-functions that are equal to the identity on T; this means that $G(f)$ will have an isolated vertex. More thought reveals that the subgraph obtained from $G(f)$ by deleting this isolated vertex is isomorphic to the neighbourhood of a vertex in K_n^f.

Thus covers of K_n with index two can be specified by giving the neighbourhood of some vertex in the cover. It is not hard to see that a 2-fold cover of K_n is distance-regular if each vertex neighbourhood is a regular graph and, if distance-regular, it is antipodal. Although any graph on $n-1$ vertices can occur as the neighbourhood of a vertex in 2-fold cover of K_n, not all regular graphs can occur as vertex neighbourhoods in distance-regular 2-fold covers.

Lemma 3.1. *A graph X can be the neighbourhood of a vertex in a distance regular double cover of a complete graph if and only it is empty, or is strongly regular with $k = 2c_2$.*

If we take X to be complete on r vertices, the corresponding 2-fold cover is two disjoint copies of K_n. This is not distance-regular, because it is not connected. If X is empty then the cover is the graph obtained by deleting a perfect matching from a complete bipartite graph $K_{n,n}$. We call this a *trivial* distance regular cover of K_n.

The cycle on five vertices has $k = 2$ and $c_2 = 1$; the corresponding double cover is the icosahedron. If X is the line graph of $K_{3,3}$ then $k = 4$ and $c_2 = 2$, this provides a non-trivial double cover of K_{10}. It can be shown that if a non-trivial distance-regular double cover of K_n exists then n must be even.

If X is strongly regular and $k = 2c_2$ then the same relation holds for the complement $\overline{X}$ of X. This gives a second distance-regular double cover of K_n. The sum of the values of δ for these two covers is zero, and if the first has non-trivial eigenvalues θ and τ,the second has $-\theta$ and $-\tau$. We refer to these two graphs as a *complementary pair.* There does not appear to be anything analogous to this when $r > 2$.

A set $\mathcal{L}$ of m lines in $\mathbf{R}^n$ is *equiangular* if the angle between any two distinct lines is the same. Lemmens and Seidel [20] (or [25]) show that, if the cosine of this angle is α and $n\alpha^2 < 1$, then

$$m \leq \frac{n(1-\alpha^2)}{1-n\alpha^2}.$$

Equality holds if and only if there is a distance-regular 2-fold cover of K_n.

If u is a vertex in the graph X, let $X(u)$ denote the neighbourhood of u in X. Let X be an (n, r, c_2) cover with $\delta = 0$. Then any two vertices in different fibres have the same number of common neighbours, $a_1 = c_2$. Define an incidence structure $\mathcal{D}$ with the vertices of X as its points and the sets $X(u)$, for u in $V(X)$, as its blocks. Then $\mathcal{D}$ is a square divisible design and the map σ such that $u^\sigma = X(u)$ is a polarity. The classes are the fibres of X, so two points in the same class do not lie in any block. Two vertices in different classes lie in exactly a_1 blocks.

If X is a $(n, n-2, c_2)$-cover then (F1) implies that $c_2 = 1$ and then (2.1) yields that $a_1 = 1$. Hence $\delta = 0$, and we can construct a divisible design $\mathcal{D}$. Extend $\mathcal{D}$ to a new incidence structure $\mathcal{D}'$, as follows. Adjoin $n+1$ points $p_1, \ldots, p_n$ together with a point ∞. Let $F_1, \ldots, F_n$ be the fibres of X. Define $\ell_\infty = \{p_1, \ldots, p_n\}$ to be one new block, along with the blocks

$$\ell_i = \infty \cup p_i \cup F_i, \quad i = 1, \ldots, n.$$

Then $\mathcal{D}'$ has $n^2 - n + 1$ points, $n^2 - n + 1$ lines and every pair of distinct points lies on exactly one line, and we have constructed a projective plane of order $n-1$. The polarity σ extends to a polarity of $\mathcal{D}'$ if we define $\sigma(\infty) = \ell_\infty$ and $\sigma(\ell_i) = p_i$. Thus our plane has a polarity with n collinear absolute points; this implies that $n-1$ is even. Conversely, every projective plane of even order n that has a polarity with exactly n absolute points, all collinear, determines an $(n, n-2, 1)$-cover. This construction is ascribed by Biggs [1] to Bondy, who remembered none of it when I spoke to him about it. (Which is not to say that I doubt the attribution!)

The line graph of the Petersen graph is a $(5, 3, 1)$-cover; if we carry out the construction just described we get the projective plane of order four.

Covers with $\delta = -2$ also lead to divisible designs. We still take the vertices of X to be the points of the design, this time the blocks are the sets $u \cup X(u)$. If $u^\sigma := u \cup X(u)$ then σ is, once more, a polarity. The divisible designs here are *transversal designs*, each block contains exactly one point from each class.

If A is the adjacency matrix of an r-fold cover of K_n and $\delta = -2$ then

$$\begin{pmatrix} 0 & A+I \\ A+I & 0 \end{pmatrix}$$

can be shown to be the adjacency matrix of an antipodal distance-regular r-cover of $K_{n,n}$. Distance-regular covers of complete bipartite graphs are another interesting class of objects; for more information about them, see [13].

If X is an (n, r, c_2)-cover with $\delta = 2$ we can, once again, construct a divisible design from it. Let X_i denote the graph with the same vertex set as X, where two vertices are adjacent if and only if they are at distance i in X. Since X is antipodal, X_3 consists of n vertex disjoint copies of K_r. Let A_i be the adjacency matrix of X_i. Then $A_1 = A$ (in our usual notation) and, as

$$A_1^2 = (n-1)I + a_1 A_1 + c_2 A_2.$$

If we substitute $J - I - A_1 - A_3$ for A_2 and observe that $A_2 = A_1 A_3$, this implies that

$$(A_1 + A_3)^2 = (n - c_2 + r - 4)I + (c_2 + 2)(J - A_3) + (r-2)A_3.$$

Therefore $A_1 + A_3$ is the incidence matrix of a square divisible design (with a polarity). When $r = c_2 + 4$ we actually get a symmetric design, but in this case we have stumbled onto a known construction of symmetric designs from a generalised quadrangle with parameters $(q+1, q-1)$. (We will see in the next section that covers such that $\delta = 2$ and $r = c_2+4$ can be built from a $GQ(q+1, q-1)$ with a spread.) For more information on these designs, see [23, Ch. 3.6].

The possible parameter sets of a square group divisible design are constrained by the Bose-Connor conditions [2]. These imply that covers with parameter sets $(11, 3, 3)$ and $(7, 5, 1)$ do not exist. The latter would, of course, give rise to a projective plane of order six. The Bose-Connor conditions only provide extra information when $\delta = 0$; if $\delta = \pm 2$ then we already have that n must be a perfect square.

§4. Constructions

We will describe some of the main techniques for constructing families of distance-regular covers of K_n. The relations between the various known classes of covers are discussed in [5], which might even appear one day.

Standard Covers: We call a cover *standard* if $\delta = 0$. We present a construction of a class of standard covers due to Neumaier, as described in [4, 12.5.3]. The graphs produced were first found by Mathon, using a different technique. Let q be a prime power, let K be the subgroup of r-th powers in the multiplicative group of F and suppose that $-1 \in K$. Let V be a vector space of dimension two over F, provided with a non-degenerate symplectic form B. Let X be the graph with vertices the sets Kv, where v ranges over the non-zero vectors in V, and where Ku and Kv are adjacent if and only if $B(u, v) \in K$. Then X is a cover with $\delta = 0$ and parameters $(q+1, r, (q-1)/r)$; the largest possible value of r is $q-1$ when q is even and $(q-1)/2$ if q is odd.

It is shown in [18] that an antipodal cover of K_n with diameter three is distance-regular if and only if there is a constant, μ say, such that any two non-adjacent vertices in different fibres have exactly μ common neighbours. (If μ is well-defined then it is equal to c_2.) This is a trick that makes it easier to verify that the graphs we have just defined are distance-regular, and it can be also applied to the constructions that follow.

In [9], Peter Cameron provides a number of constructions of covers with $\delta = 0$. It appears that most of these are isomorphic to quotients of the covers just described. (We will discuss the operation of taking quotients of covers towards the end of this section.) However the two classes of graphs he constructs from unitary forms cannot be obtained in this way, although in one case $\delta = 0$.

Symplectic Covers: Next, let p be a prime, and let V and U be vector spaces over $GF(p)$, with dimensions m and s respectively. Let B be a $GF(p)$-linear alternating form from $V \times V$ to U. Thus, if v and w belong to V then $B(v, w) \in U$ and

$$B(v, v) = 0, \quad B(v, w) = -B(v, w).$$

Assume further that, for each element a in V, the linear mapping B_a for V to U given by

$$B_a(v) = B(a, v)$$

is onto. Clearly this implies that $\dim U \leq \dim V$ and, since $B(a, a) = 0$, it follows that $\dim U < \dim V$. Let $X(B)$ be the graph with vertex set $V \times U$, where (v, a) and (w, b) are adjacent in X if and only if they are distinct and

$$B(v, w) = a - b.$$

A straightforward computation shows that this yields a (p^m, p^s, p^{m-s}) cover. As $\delta = -2$ for this cover, m must be even. If U is 1-dimensional, this construction reduces to one due to Somma, that in turn generalises a construction of Thas. For more on these, see Section 12.5 of [4].

We describe a class of examples. Let Σ be a set of invertible $m \times m$ skew-symmetric matrices over $GF(p)$ that forms a subspace over $GF(p)$. Let $A_1, \dots, A_s$ be a basis for this subspace and define $B(v, w)$ to be the vector with entries $v^T A_1 w, \dots, v^T A_s w$. We must verify that B_a is onto. Let $M(a)$ be the matrix with $a^T A_i$ $(i = 1, \dots, s)$ as its rows. Then $M(a)w = B(a, w)$ and, if $x = (x_1, \dots, x_s)$, then

$$x^T M(a) = a^T \sum_{i=1}^{s} x_i A_i.$$

If $x \neq 0$ then $x_i A_i$ is invertible; this implies that the rows of $M(a)$ are linearly independent. Consequently $M(a)$ has rank s and B_a is onto.

It is not to difficult to see that changing the basis of Σ is equivalent to replacing B by a form B', where $B' = gB$ for some invertible $s \times s$ matrix g over $GF(p)$. Further $X(B)$ and $X(B')$ are isomorphic.

A *spread set* Σ is a set of q^n matrices of order $n \times n$ over $GF(q)$, such that the difference of any two matrices in it is invertible. Subtracting a fixed matrix from each element gives rise to a spread set that contains the zero matrix, in this case all non-zero matrices in the set are invertible. Spread sets are equivalent to translation planes; if Σ is closed under addition then the translation plane is a *semifield plane.* (If q is a power of the prime p then Σ is closed under addition if and only if it is a subspace over $GF(p)$.) For some more information on this subject, see [21]. Suppose then that Σ is spread set such that is a vector space over $GF(p)$. If $S_i \in \Sigma$ let $\hat{S}_i$ be the skew symmetric matrix

$$\begin{pmatrix} 0 & S_i \\ S_i^T & 0 \end{pmatrix}.$$

Then the matrices $\hat{S}_i$, form a subspace of invertible skew symmetric matrices.

De Caen (private communication) and Cameron [19, Thm. 3.8] have observed that a subspace of invertible $2m \times 2m$ matrices over a finite field has dimension at most m. This bound is realised by the semifield examples we have just described.

The graph $X(B)$ is the Cayley graph for a p-group. Let P be the group with vertex set $V \times U$ and multiplication given by

$$(v, a) \oplus (w, b) = (v + w, a + b + B(v, w)).$$

Then P is elementary abelian if $p = 2$, otherwise it is special and has exponent p. In all cases it acts as a regular group of automorphisms of X and, if p is odd, the fibres of $X(B)$ are the cosets of the centre of P.

Spreads in Generalised Quadrangles: Our third construction uses generalised quadrangles and is due to Brouwer [4, Prop. 12.5.2]. A *spread* in a generalised quadrangle is a set of lines that partitions its points. Let $\mathcal{Q}$ be a $GQ(s, t)$ with a spread Σ. Let X be the graph with the points of $\mathcal{Q}$ as its vertices, with two points adjacent in X if and only if there are collinear in $\mathcal{Q}$, but the line joining them is not in Σ. This yields an $(st + 1, s + 1, t - 1)$-cover.

More generally, it is possible to construct covers from strongly regular graphs that have the same parameters as the point graph of a generalised quadrangle, and have a spread. For details, see [4, Prop. 12.5.2] or [18]. Note also that a generalised quadrangle may have more than one spread, and hence may give rise to more than one cover.

For information about generalised quadrangles and their spreads, see [23]. The parameter sets of the generalised quadrangles that are known to have spreads, and the parameters of the related covers, are listed in the following table.

(s,t)	(n,r,c_2)	δ
(q,q)	$(q^2+1,q+1,q-1)$	0
(q,q^2)	$(q^3+1,q+1,q^2-1)$	$q-q^2$
$(q-1,q+1)$	(q^2,q,q)	-2
$(q+1,q-1)$	$(q^2,q+2,q-2)$	2

Here q is a prime power in the first three cases, and a power of two in the fourth.

Preparata Covers: Recently de Caen, Mathon and Moorhouse [7] found a new infinite family of covers which are related to the Preparata codes and have parameter sets $(2^{2t}, 2^{2t-1}, 2)$. These can be described quite succinctly. Assume $q = 2^{2t-1}$ and $s = 2^e$, where e and $2t-1$ are coprime. Let $X(q,s)$ be the graph with vertex set

$$GF(q) \times GF(2) \times GF(q),$$

where (a,i,α) and (b,j,β) are adjacent if and only if

$$\alpha + \beta = a^s b + ab^s + (i+j)(a^{s+1} + b^{s+1}).$$

Then $X(q,s)$ is a $(2^{2t}, 2^{2t-1}, 2)$-cover. The fibres are the sets of the form

$$\{(a,i,u) : u \in GF(q)\}.$$

The maps

$$(a,i,\alpha) \mapsto (a,i,\alpha+u), \quad u \in GF(q)$$

generate the covering group of $X(q,s)$, which is therefore regular and isomorphic to the additive group of $GF(q)$. The graphs $X(q,2^e)$ and $X(q,2^f)$ are isomorphic if and only if $e+f$ is zero, modulo $2t+1$. These covers are all triangle-free; prior to their construction the only examples of triangle-free cover known were the $(7,6,1)$-cover coming from the Hoffman-Singleton graph and the cycle on six vertices.

Another feature of these covers is that corresponding association schemes are formally dual to the schemes formed by *linked systems of symmetric designs*. For information about the connection, see [7]; for information about these systems see, e.g., [10, p. 148].

Schur Groups: Let f be an arc function of index r on K_n, normalised with respect to some spanning tree and suppose that the covering group of X^f is abelian. Then the connection group C is abelian and has r distinct irreducible characters $\chi_1, \dots, \chi_r$, where χ_1 is the trivial character. Let H_s be the $n \times n$ matrix with diagonal entries equal to 0 and

$$(H_s)_{i,j} = \chi_s(f(i,j)).$$

The set of matrices $\mathcal{H} = \{H_1, \dots, H_r\}$ has the following properties:

(a) $H_1 \circ H_i = H_i$ for all i.
(b) $H_i^* = H_i$.
(c) $H_i \circ H_j \in \mathcal{H}$, for all i and j.
(d) The adjacency matrix A of X^f is similar to a block-diagonal matrix, with $H_1, \dots, H_r$ as its diagonal blocks.

Here $A \circ B$ denotes the *Schur* or pointwise product of matrices A and B. A set of r matrices satisfying (a) and (c) is called a *Schur group*, because it is closed under Schur multiplication. Note the entries of the matrices in a Schur group of order r must be r-th roots of unity. As the irreducible characters of an abelian group C form a multiplicative group isomorphic to C, the Schur group of a cover with regular abelian covering group is isomorphic to the covering group.

Now suppose that X is an (n, r, c_2)-cover with abelian connection group, and Schur group $\mathcal{H} = \{H_1, \dots, H_r\}$. Then $H_1 = J$ and, when $2 \le i \le r$,

$$H_i^2 - \delta H_i - (n-1)I = 0. \tag{4.1}$$

Assume further that the covering group of X has exponent two and $i \ge 2$. If $\delta = 0$, then (4.1) implies that H_i is a conference matrix for $i = 2, \dots, r$. If $\delta = -2$ then $H_i + I$ is a Hadamard matrix, and the matrices $H_i + I$ form a Schur group with identity element J. Finally, if $\delta = 2$ then $H_i - I$ is a Hadamard matrix. (If we drop the assumption that the covering group has exponent two, "generalised" Hadamard matrices result.)

The converse is also true, and is proved in Section 9 of [18].

Theorem 4.1. *Let $\mathcal{H}$ be a Schur group of $n \times n$ Hermitian matrices with order r, such that*

$$H_i^2 - \delta H_i - (n-1)I = 0$$

when $2 \le i \le r$. Then $\mathcal{H}$ determines an (n, r, c_2)-cover of K_n with cover group isomorphic to $\mathcal{H}$.

By way of example, let q be a power of the prime p, let V be a vector space of dimension $2m$ over $GF(q)$, let A be a non-degenerate alternating bilinear form on V and let ξ be a primitive q-th root of unity. Let $H(A)$ be the matrix with rows and columns indexed by the entries of V, such that

$$H(A)_{u,v} := \xi^{u^T A v}.$$

Then $H_A^2 = q^{2m} I$ and, if B is a second alternating form, then

$$H(A) \circ H(B) = H(A + B).$$

It follows that, if S is subspace of invertible skew-symmetric matrices over $GF(p)$, then the matrices $H(S)$ form a Schur group.

The Preparata covers have a regular covering group that is an elementary abelian 2-group. Hence they give rise to a Schur group of 2^{2n-1} symmetric $2^{2n} \times 2^{2n}$ Hadamard matrices with constant diagonal.

If $\mathcal{H} = \{H_1, \dots, H_r\}$ is the Schur group of an (n, r, c_2)-cover X and $i \neq 1$ then H_i determines a distance-regular cover, X_i say, of K_n, with index equal to the order of H_i. It follows that X can be expressed as the product of distance-regular covers of K_n, each with regular cyclic covering group. The condition that the matrices H_i form a Schur group is equivalent to the condition that the product of the covers X_i be distance-regular. Note that any cover with regular abelian covering group can always be expressed as a product of covers with cyclic covering groups.

Quotients: There is a useful technique which often enables us to construct covers of lower index from a given cover of index r. We describe a special case. Let X be an r-fold cover of K_n and let H be a group of automorphisms of X that fixes each fibre as a set. As noted before, H must act semi-regularly and so $|H|$ must divide r. Construct a graph with the orbits of H as its vertices, with two orbits adjacent if and only if some vertex in one is adjacent to a vertex in the second. It is easy to show that this new graph is a cover of K_n with index $r/|H|$; it can also be shown that is antipodal and distance regular. For details, see Section 6 of [18]. The quotienting procedure just described can be generalised. The only role H plays in the construction is to provide a nice partition of the vertex set of X. Any equitable partition of X with each cell contained in a fibre of X is "nice" enough. For details, see Section 6 of [18]. In all our applications, we will use the group version.

If X is an (n, r, c_2)-cover and $h = |H|$ then the quotient of X over the orbits of H is a $(n, r/h, hc_2)$-cover. This means that the value of δ is not changed when we form quotients.

But which of our examples admit quotients? Most! The covering groups of the standard covers obtained from Neumaier's construction are regular and cyclic, being quotient groups of the multiplicative group of the underlying field. This set of covers is closed under taking quotients. The covering groups of the symplectic covers are regular and elementary abelian and, once again, this class is closed under taking quotients. The covering groups of the Preparata covers of are regular elementary abelian 2-groups. In this we obtain new covers by taking quotients. (To be more precise, any quotient with $r^2 > n$ has not been constructed in any other way.) The unitary generalised quadrangles, with parameters (q, q^2), give rise to covers with regular cyclic covering group. The corresponding quotients were first found by Cameron [9]. (For more on these, see also [16].) We cannot realise any new parameter sets by taking quotients of the covers coming from generalised quadrangles with parameters (q, q) and $(q-1, q+1)$. I do not if the covers coming from generalised quadrangles with parameters $(q+1, q-1)$ have any quotients; it will follow from our next theorem that these covers cannot be abelian.

Suppose S is the Schur group determined by an (n, r, c_2)-cover with abelian regular covering group C. Then any non-identity subgroup of S determines a cover with index equal to its order. The Schur group is isomorphic to the dual group of C, hence its subgroups correspond to homomorphic images of C. This implies that taking the cover formed from a Schur subgroup is equivalent to forming a quotient of the cover.

In combination with taking quotients, the following result is often useful.

Theorem 4.2 ([18, Thm. 9.2]). *Let X be an antipodal distance-regular cover of K_n with index r, and suppose that the covering group of X is regular and cyclic. If X is not a trivial 2-fold cover of K_n, then r divides n.*

For example, this can be used to show that the covering group of a 6-fold cover of K_7 is trivial. Suppose that C is the covering group of such a cover X. Then X has parameters $(7, 6, 1)$; hence $\delta = -1$. If C has a subgroup of order two then, quotienting over this yields a $(7, 3, 2)$-cover while, if C has a subgroup of order three, quotienting yields a $(7, 2, 3)$-cover. In both cases we find that the eigenvalue multiplicities are not integers, hence these quotients do not exist. Because C has to act semi-regularly on a fibre, $|C|$ must divide six. Hence we have shown that C must be trivial. (Or, to put it another way, the stabiliser of a vertex in the Hoffman-Singleton graph acts faithfully on the neighbours of that vertex.)

There is a 4-fold cover of K_{10} with regular abelian covering group; this shows that Theorem 4.2 does not hold if we replace "cyclic" by "abelian".

§5. Linear association schemes

A graph X is called a *translation graph* if there is an abelian group of automorphisms acting regularly on its vertices. For our purposes the most useful abelian groups are the additive groups of a vector space $V = V(n,q)$ over a field F of order q. We can construct graphs relative to V as follows. Let $\mathcal{C}$ be a subset of non-zero vectors in V that is closed under multiplication by -1. Then $X(\mathcal{C})$ is the graph with vertex set V, and vectors u and v are adjacent in x if and only if $u - v \in \mathcal{C}$. Thus X is a Cayley graph for the additive group of V. We say that $X(\mathcal{C})$ is *linear* over F is $\mathcal{C}$ is closed under multiplication by the non-zero elements of F.

There is an attractive geometric construction for linear graphs. Let $\mathcal{P}$ be a hyperplane in n-dimensional projective space $PG(n,q)$, and view $\mathcal{P}$ as the hyperplane at infinity for the affine space $AG(n,q)$ formed by the points not on $\mathcal{P}$. Choose a subset Ω in $\mathcal{P}$ and define $X(\Omega)$ to be the graph with $AG(n,q)$ as its vertex set, where two points u and v are adjacent if the line $u \vee v$ meets $\mathcal{P}$ at a point in Ω. The correspondence with the previous description is trivial: the subset Ω determines a set of lines through the origin in $AG(n,q)$, and the non-zero points in these lines form the set $\mathcal{C}$. The valency of $X(\Omega)$ is $(q-1)|\Omega|$, and this one of its eigenvalues. The remaining eigenvalues of $X(\Omega)$ are also determined geometrically by the following result from [15, Lemma 12.9.3].

Lemma 5.1. *Let $X(\Omega)$ be a linear graph, where Ω is a subset of the projective space $\mathcal{P}$ over $GF(q)$, and let n_x denote the number of hyperplanes of $\mathcal{P}$ that meet Ω in exactly x points. Then $qx - |\Omega|$ is an eigenvalue of X with multiplicity $(q-1)n_x$.*

We need yet a third version of our construction. Let C be a linear code in $V(n,q)$. The *coset graph* of C has the cosets of C as its vertices, where two cosets are adjacent if there is a vector in one at Hamming distance one from a vector in the second. We may prevent multiple edges occurring by assuming that the minimum distance of C is at least three. It is an easy exercise to show that the coset graph of a linear code with minimum distance three is linear. The following result is essentially Theorem 11.1.10 from Brouwer, Cohen and Neumaier [4].

Theorem 5.2. *Let X be a distance-regular graph that is not a self-complementary conference graph. If X is a translation graph with respect to an elementary abelian p-group then it is the coset graph of a completely regular linear code with minimum distance at least three.*

Because we will only apply this result to graphs with diameter three, all we need to say about conference graphs is that they have diameter two. And it will suffice to define a linear code to be completely regular if its coset graph is distance-regular. (For more information, see Chapter 11 in either [4] or [15].) Our application of Theorem 5.2 is as follows. Suppose X is an antipodal distance-regular graph that is linear over $GF(p^s)$, and let G be a proper, non-identity, subgroup of the covering group of X. As X is linear, there is an elementary abelian p-group, T say, acting regularly on X. The fibres of X are a system of imprimitivity for T and, as T is abelian, the subgroup of T fixing each fibre is non-trivial and acts regularly on the vertices in a fibre. Thus this subgroup is the covering group of X. Accordingly, if G is a subgroup of the covering group of X then the quotient of X over the partition formed by the orbits of G is antipodal and distance-regular. It is also clearly a translation graph for the elementary abelian quotient group T/G, whence we deduce, using Theorem 5.2, that it is linear over some finite field of characteristic p. (Although possibly not over the field we started with.)

We turn to linear association schemes. We may view an association scheme with d classes on the vertex set V as a set of graphs $X_1, \dots, X_d$ on V whose edges partition the edge set of the complete graph on V. We call an association scheme *linear* if each graph in it is linear, relative to the same vector space of course, Therefore a linear association scheme is determined by giving a partition of the points of a projective space. But then we are faced with the question: which partitions give rise to association schemes? There is a nice answer to this.

Now suppose that $\omega = (\Omega_1, \dots, \Omega_d)$ is a partition of the points of $\mathcal{P}$ with d cells. Define two hyperplanes H and K of $\mathcal{P}$ to be ω-equivalent if

$$|H \cap \Omega_i| = |K \cap \Omega_i|, \qquad i = 1, \dots, d.$$

This gives rise to a partition of the hyperplanes of $\mathcal{P}$ into ω-equivalence classes, that we will call the partition *induced* by ω.

For a proof of the next result, due to Bridges and Mena, see [3] or [15, Thm. 12.10.1].

Theorem 5.3. *Let $\mathcal{P}$ be the projective space formed by the points at infinity of an affine space of dimension n over $GF(q)$. Let $\omega = (\Omega_1, \dots, \Omega_d)$ be a partition of the points of $\mathcal{P}$ and let ω^* be the induced partition on the hyperplanes of $\mathcal{P}$. Then ω^* has at least as many cells as ω and, equality holds if and only if the graphs $X(\Omega_1), \dots, X(\Omega_d)$ form an association scheme with d classes.*

Call a partition ω of the points of $\mathcal{P}$ *equitable* if the dual partition ω^* has the same number of cells as ω. If ω is equitable then it follows (see, e.g., [15, p. 250]) that ω is the partition induced by ω^* on the hyperplanes of the space dual to $\mathcal{P}$. Hence ω^* determines a second association scheme that is dual, in the sense of Delsarte [11, §2.6], to the scheme determined by ω. (We discuss duality in a little more detail in Section 7.) The hyperplanes in a given cell Ω_i^* of ω^* determine an eigenspace with dimension $(q-1)|\Omega_i^*|$ for the association scheme given by ω.

As one example, we may take ω to be the orbits of a group Γ of collineations of $\mathcal{P}$; this is always equitable. The cells of ω^* are the orbits of Γ on hyperplanes.

Let $\Omega_1, \dots, \Omega_d$ be a partition of the points of $\mathcal{P}$ and let X_i denote $X(\Omega_i)$. If a and b are affine points, let $p_{i,j}(a,b)$ denote the number of affine points x such that a is adjacent to x in X_i and b is adjacent x in X_j. Then the line $a \vee x$ must meet $\mathcal{P}$ in a point p of Ω_i, while $b \vee x$ meets it in a point q of Ω_j. As a, b and x are coplanar, p, q and the point $z = (a \vee b) \cap \mathcal{P}$ must be collinear. It follows that, if $i \neq j$ and z does not lie in Ω_i or Ω_j, then

$$p_{i,j}(a,b) = \sum_{\ell} |\ell \cap \Omega_i||\ell \cap \Omega_j|, \tag{5.1}$$

where the sum is over all lines in $\mathcal{P}$ on z. If ω is equitable then $X_1, \dots, X_d$ form an association scheme, and the value $p_{i,j}(a,b)$ is determined by the index k such that $\{a,b\}$ is an edge of X_k. In this case the sum in (5.1) is independent of the choice of z in Ω_k, and therefore it is equal to

$$|\Omega_k|^{-1} \sum_{\ell \in \mathcal{P}} |\ell \cap \Omega_i||\ell \cap \Omega_j||\ell \cap \Omega_k|. \tag{5.2}$$

If ω is equitable and $\{a,b\}$ is an edge of X_k then we write $p_{i,j}(k)$ in place of $p_{i,j}(a,b)$. These are the *intersection numbers* of the association scheme.

Theorem 5.4. *Let $\mathcal{P}$ be a hyperplane in a projective space over $GF(q)$, let ω be an equitable partition of $\mathcal{P}$ with cells $\Omega_1, \dots, \Omega_d$. Assume i, j and k are distinct and non-zero. Then the intersection numbers $p_{i,j}(k)$ are given as follows.*
(a) $p_{i,j}(0) = \delta_{ij}(q-1)|\Omega_i|$.
(b) $p_{i,j}(k) = |\Omega_k|^{-1} \sum_{\ell \in \mathcal{P}} |\ell \cap \Omega_i||\ell \cap \Omega_j||\ell \cap \Omega_k|$.
(c) $p_{i,i}(k) = |\Omega_k|^{-1} \sum_{\ell \in \mathcal{P}} (|\ell \cap \Omega_i| - 1)|\ell \cap \Omega_i||\ell \cap \Omega_k|$.
(d) $p_{k,k}(k) = q - 2 + |\Omega_k|^{-1} \sum_{\ell \in \mathcal{P}} |\ell \cap \Omega_k|(|\ell \cap \Omega_k| - 1)(|\ell \cap \Omega_k| - 2)$.

An association scheme $X_1, \dots, X_d$ is *primitive* if each graph X_i is connected. If this scheme is given by an equitable partition ω of $\mathcal{P}$, it is easy to show that it is primitive if and only if each cell of ω spans $\mathcal{P}$. In particular, $X(\Omega)$ is a disjoint union of cliques if and only if Ω is a subspace of $\mathcal{P}$.

Linear association schemes are discussed at some length in Chapter 12 of [15].

§6. Linear covers

We are going to study linear antipodal distance-regular covers of K_n. To begin, we present a class of examples due to Thas. (For further information, see [4, §12.5].) Let $\mathcal{P}$ be a hyperplane in $PG(3,q)$, where q is even. Let Ω_1 be an oval in $\mathcal{P}$, that is, a set of $q+1$ points with no three collinear. Let Ω_3 be the nucleus of the oval, and let Ω_2 be the remaining points. We claim that $X(\Omega_1)$ is a (q^2, q, q)-cover. This will follow readily from Theorem 5.4 if we show that our partition is equitable. But this is easy. All lines on the nucleus meet the oval in exactly one point. Any line not on the nucleus is not a tangent, hence it is either a secant, meeting the oval in exactly two points, or a passant with no point in common with the oval. Thus the induced partition on lines has exactly three classes, and therefore $(\Omega_1, \Omega_2, \Omega_3)$ is equitable.

The fibres of the cover we get are the sets of points on affine lines that meet the nucleus of the oval. The transvections in $PG(3,q)$ with axis $\mathcal{P}$ and Ω_3 as centre form the covering group, which is thus regular and isomorphic to the additive group of $GF(q)$. Hence we may take quotients and obtain covers with parameters $(q^2, q/t, qt)$, for any divisor t of q.

Our next task is to characterise the equitable partitions of projective space that give rise to such graphs.

Theorem 6.1. *Let $\mathcal{P}$ be a hyperplane in $PG(n,q)$ and let $\omega = (\Omega_1, \Omega_2, \Omega_3)$ be an equitable partition of $\mathcal{P}$. Then $X(\Omega_1)$ is an antipodal distance-regular cover of a complete graph if and only if:*
(a) *Ω_1 is not a subspace,*
(b) *any line of $\mathcal{P}$ that meets Ω_1 in two or more points is disjoint from Ω_3 and*
(c) *Ω_3 is a subspace.*

Proof. Because ω is equitable, the graphs $X(\Omega_i)$ do form an association scheme $\mathcal{A}$. As Ω_1 is not a subspace, there is line ℓ such that $|\ell \cap \Omega_1| \geq 2$ but ℓ is not contained in Ω_1. From condition (b) it follows that $\ell \cap \Omega_2 \neq \emptyset$, and now Theorem 5.4 yields that $p_{1,1}(2) > 0$. Now choose ℓ to be a line meeting both Ω_1 and Ω_3. Then (b) implies that ℓ intersects Ω_1 in exactly one point, and (c) implies that it intersects Ω_3 in exactly one point. Therefore ℓ must contain at least one point from Ω_2, whence Theorem 5.4 yields that $p_{1,2}(3) > 0$. This implies that $X(\Omega_1)$ has diameter three, and consequently $\mathcal{A}$ is metric relative to $X(\Omega_3)$, that is, $X(\Omega_3)$ is distance regular.

Because Ω_3 is a subspace, being "equal or at distance three" in $X(\Omega_1)$ is an equivalence relation on its vertices. Therefore X is antipodal with diameter three, and must be an antipodal cover of a complete graph. Q.E.D.

Suppose Ω_3 has dimension t. Then condition (b) in the theorem is equivalent to the condition that each $(t+1)$-dimensional subspace of $\mathcal{P}$ on Ω_3 must intersect Ω_1 in exactly one point. A subspace U of $\mathcal{P}$ is a *complement* to a subspace V if it is disjoint from it and

$$\dim U + \dim V = \dim \mathcal{P}.$$

If U is a complement to V then the partition of $\mathcal{P}$ with cells U, V and $\mathcal{P} \setminus (U \cup V)$ is equitable. Further any line that meets U in at least two points is contained in U,and consequently is disjoint from V. If V is a subspace of $\mathcal{P}$, define a subset Ω of $\mathcal{P}$ to be a *pseudo-complement* for V if:

(1) The partition $(\Omega, V, \mathcal{P} - (\Omega \cup V))$ is equitable.
(2) Any line that meets Ω in at least two points is disjoint from V.

Then Theorem 6.1 implies that if Ω_3 is a subspace of $\mathcal{P}$ and Ω_1 is a pseudo-complement to it then it is a subspace of $X(\Omega_1)$ is an antipodal distance-regular graph with diameter three.

Suppose that $\mathcal{P}$ is a hyperplane in $PG(d+1,q)$, $(\Omega_1, \Omega_2, \Omega_3)$ is an equitable partition of $\mathcal{P}$ such that Ω_3 is a subspace of dimension t and Ω_1 is a pseudo-complement to Ω_3 that is not a subspace. Then

$$\begin{aligned}
|\Omega_1| &= \frac{q^{d-t}-1}{q-1}, \\
|\Omega_2| &= (q-1)\,\frac{q^{d-t}-1}{q-1}\,\frac{q^{t+1}-1}{q-1}, \\
|\Omega_3| &= \frac{q^{t+1}-1}{q-1}.
\end{aligned}$$

Further $X(\Omega_1)$ is a cover of a complete graph, having parameters

$$n = q^{d-t}, \quad r = q^{t+1}.$$

As $r \leq n-1$, we see that

$$d \geq 2t+2.$$

In the next section we will find more restrictions on the parameters of a linear cover.

§7. Parameters of linear covers

Assume $X_1, \dots, X_d$ form an association scheme $\mathcal{A}$ on n vertices. The valency of X_i is denoted by v_i. Set v_0 equal to 1. The Bose-Mesner algebra of the scheme has $d+1$ distinct eigenspaces, with dimensions $m_0, \dots, m_d$, where $m_0 = 1$. These are the *multiplicities* of the scheme. The *Frame quotient* of the scheme is defined to be

$$n^{d-1} \prod_{i=1}^{d} \frac{v_i}{m_i}.$$

It is always an integer. (For details, see [4, p. 46].) Our scheme has a $(d+1) \times (d+1)$ matrix of eigenvalues P, and a matrix of dual eigenvalues Q, of the same order, and $PQ = nI$. A second association scheme is *formally dual* to $\mathcal{A}$ if its matrix of eigenvalues is Q; this scheme has valencies $m_0, \dots, m_d$ and multiplicities $v_0, \dots, v_d$.

If $\mathcal{A}$ is a translation scheme then we may identify its vertices with the elements of an abelian group G. In this case there is a well-defined construction of a *dual scheme*, with the irreducible characters of G as its vertices. A dual scheme is formally dual in the sense just defined. If ω is an equitable partition of a projective space with induced partition ω^* then the schemes determined by ω and ω^* are dual, so all linear schemes have duals. (For more background on duality, see [4, §2.10(B)], [15, Ch. 12] and [11, §2.6].)

Now we are in a position to prove the main result of this section.

Theorem 7.1. *Let X be an (n, r, c_2)-cover with p^a vertices, where p is a prime, and suppose there is an association scheme that is formally dual to the scheme determined by X. Then $p = 2$, n is a perfect square and either $\delta = -2$ or $r = 2$ and $\delta = 2$.*

Proof. Let θ and τ be the non-trivial eigenvalues of X. We first prove that $\theta - \tau$ divides n, which implies that it is power of p.

We compute the Frame quotient F^* of the formally dual scheme. We have

$$m_\theta m_\tau = -\frac{n^2(r-1)^2\theta\tau}{(\theta-\tau)^2} = \frac{n^2(r-1)^2(n-1)}{(\theta-\tau)^2}. \tag{7.1}$$

and so

$$\begin{aligned} F^* &= (nr)^2 \frac{(n-1)m_\theta m_\tau}{(n-1)\cdot(r-1)(n-1)\cdot(r-1)} \\ &= (nr)^2 \frac{n^2(r-1)^2(n-1)}{(r-1)^2(n-1)(\theta-\tau)^2} \\ &= \frac{n^4 r^2}{(\theta-\tau)^2}. \end{aligned}$$

By hypothesis, both n and r are powers of q and so we deduce that $\theta - \tau$ is a power of the prime p. As $m_\theta m_\tau$ is an integer it now follows from (7.1) that $(\theta-\tau)^2$ divides n^2.

As n and r are both powers of p, it follows from the identity

$$n - 2 - \delta = rc_2$$

that p divides $\delta + 2$. Because θ and τ are zeros of the quadratic $x^2 - \delta x - (n-1)$, we find that $\delta = \theta + \tau$ and

$$(\theta - \tau)^2 = \delta^2 - 4 + 4n, \tag{7.2}$$

consequently $\theta - \tau$ divides $\delta^2 - 4$. If $\delta = -2$, then (7.2) yields that $4n$ is a perfect square, $\theta - \tau$ is even and $p = 2$.

Hence we assume that $\delta \neq -2$. Suppose that p is odd. Then, as $\theta - \tau$ is a power of p and as $\theta - \tau$ divides $\delta^2 - 4$, it follows that $\theta - \tau$ divides $\delta - 2$ or $\delta + 2$. Accordingly

$$\theta - \tau \le \delta + 2 = \theta + \tau + 2.$$

This implies that τ, the least eigenvalue of X, is at least -1 and therefore X is disjoint union of cliques.

So $p = 2$ if $\delta \neq -2$. Assume $n = 2^a$ and $\theta - \tau = 2^s$, where $s \le a$. As $\theta\tau = 1 - n$, we find that

$$(\theta + \tau)^2 = (\theta - \tau)^2 + 4\theta\tau = 4(2^{2s-2} - 2^a + 1), \tag{7.3}$$

whence $2^{2s-2} - 2^a + 1$ is a perfect square. Therefore there is an integer λ such that

$$2^{2s-2} - 2^a + 1 = (2^{s-1} - \lambda)^2 = 2^{2s-2} - 2^s\lambda + \lambda^2. \tag{7.4}$$

Consequently

$$\lambda 2^s - 2^a = \lambda^2 - 1,$$

implying that 2^s divides $(\lambda-1)(\lambda+1)$. Hence $2^{s-1} \leq \lambda+1$, so $2^{s-1}-\lambda \leq 1$ and (7.4) now yields that

$$2^{2s-2} - 2^a + 1 = 1.$$

Now (7.3) implies that $(\theta + \tau)^2 = 4$. As we have assumed $\delta \neq -2$, we must have that $\delta = 2$.

The identity $n - 2 - \delta = rc_2$ now yields that r divides 4. If $r = 4$ then

$$c_2 = \frac{n}{4} - 1$$

and so c_2 is odd, or $n = 4$. As $r \leq n - 1$, we cannot have $n = 4$ and, by our feasibility Condition (F2) in Section 2, we know that c_2 is even if n is even.

We have shown that, if $\delta \neq -2$ then $\delta = 2$ and $r = 2$. A two-fold cover of K_n such that $\delta = 2$ is one of a complementary pair with a cover with $\delta = -2$. Therefore n is a perfect square in this case too. Q.E.D.

One consequence of this result is that, in odd characteristic, any pseudo-complement to a subspace is a subspace.

We derive another consequence of our work. Let us say that the association scheme determined by an equitable partition of $PG(d, q)$ has *dimension* d over $GF(q)$. Any partition of a projective line is equitable, so there is little to say in that case.

Lemma 7.2. *Let X be antipodal distance-regular cover of a complete graph that is linear over $GF(q)$ with dimension two and has $\delta = -2$. Then $X = X(\Omega)$, where Ω is an oval.*

Proof. Suppose X is constructed from the equitable partition

$$(\Omega_1, \Omega_2, \Omega_3).$$

Because $d = 2$, we have $t = 0$, so Ω_3 is a point and $|\Omega_1| = q+1$. Further $r = q$ and $|X| = q^3$. As $\delta = -2$, we have $c_2 = q$ and $a_1 = q - 2$. Now Theorem 5.4(d) yields that no line meets Ω_1 in more than two points, and therefore Ω is an oval. Q.E.D.

The Hermitian forms graph Her(2, 3) has an antipodal distance-regular cover with diameter four and index three. (See [4, §11.3H].) In analogy to our last result, Tilla Schade [24] has shown that there is a unique linear cover of this graph, and also given a new construction of it. Both these tasks are much more difficult than what we have just done.

§8. Codes

Let $\mathcal{P}$ be a hyperplane in $PG(d+1,q)$, let ω be an equitable partition of it, with cells $\Omega_1, \ldots, \Omega_s$ and let ω^* be the induced partition on the hyperplanes of $\mathcal{P}$, with ${\Omega^*}_1, \ldots, {\Omega^*}_s$ as its cells. Let Ω denote Ω_i and let M be a matrix whose columns are homogeneous coordinate vectors for the points in Ω. We call the row space of Ω its *code*, and denote it by $C(\Omega)$. Each code word can be written in the form $x^T M$. As x^T can be viewed as the homogeneous coordinate vector of a hyperplane, H say, in $\mathcal{P}$, it follows that the weight of $x^T M$ is equal to

$$|\Omega| - |H \cap \Omega|.$$

All hyperplanes in the same cell of Ω^* give rise to words of the same weight, w_i say. Hence the weight enumerator of $C(\Omega)$ is

$$1 + (q-1)\sum_{i=1}^{s} |\Omega_i^*| x^{w_i}.$$

The code words in $C(\Omega)$ form an s-class association scheme, where a pair of words (u,v) is in the i-th class if $u - v$ has weight w_i. By a result of Delsarte [4, Thm. 2.10.13], its dual code is completely regular. The coset graph of the dual code is $X(\Omega)$.

We now consider the codes obtained from subsets Ω such that $X(\Omega)$ is an antipodal distance-regular cover of complete graph. Assume that $X(\Omega)$ has dimension d over $GF(2)$, that its index r is 2^{t+1} and that $d - t = 2e$. Then X has parameters

$$n = 2^{2e},\ r = 2^{t+1}; \quad \theta = 2^e - 1,\ \tau = -2^e - 1$$

and the multiplicities of the eigenvalues of $X(\Omega)$ are 1, $2^{2e} - 1$ and

$$2^{e-1}(2^{t+1}-1)(2^e-1), \quad 2^{e-1}(2^{t+1}-1)(2^e+1).$$

Hence the weights and frequencies of the non-zero words in $C(\Omega)$ are as follows:

weight	frequency
$2^{2e-1} - 2^{e-1}$	$2^{e-1}(2^{t+1}-1)(2^e-1)$
2^{2e-1}	$2^{2e} - 1$
$2^{2e-1} + 2^{e-1}$	$2^{e-1}(2^{t+1}-1)(2^e+1)$

Our code $C(\Omega)$ has odd length, if we add the binary complement of each code word and then add a parity check to each word, the weights and frequencies of the non-zero words in the resulting code are as follows:

weight	frequency
$2^{2e-1} - 2^{e-1}$	$2^{2e}(2^{t+1} - 1)$
2^{2e-1}	$2^{2e+1} - 2$
$2^{2e-1} + 2^{e-1}$	$2^{2e}(2^{t+1} - 1)$
2^{2e}	1

Denote this code by $K(\Omega)$. As $r < n$ we must have $t \leq 2e - 2$. If $t = 2e - 2$ then $K(\Omega)$ has the same weight distribution as the Kerdock code of the same length. (See, e.g., [22, p. 456].) Since it follows from Brouwer and Tolhuizen [6] that there is no linear code with the same weight distribution as a Kerdock code, we deduce that a cover with $t = 2e - 2$ cannot be linear.

The Preparata covers have parameter sets of the form $(2^{2e}, 2^{2e-1}, 2)$; the above argument shows that that these covers cannot be linear. But more is true: if $e > 2$ then from [7] we know that the automorphism group of a Preparata cover is not vertex transitive.

§9. Questions

A. As $\delta = \pm 2$ for a linear (n, r, c_2)-cover, it follows that r divides n and, as n must be a power of two, $r \leq n/2$. Our work in Section 8 shows that we cannot have r equal to $n/2$. Therefore $r \leq n/4$. We ask:

(1) Is there a linear (n, r, c_2)-cover with $r^2 > n$?

The smallest open case is when $n = 64$. Here there are quotients of a Preparata cover with $r = 16$, but we do not know if any are linear. The linear covers we do have are obtained as quotients of graphs $X(\Omega)$, where Ω is an oval in $PG(2, q)$ and q is even; for these $r^2 \leq n$.

B. A number of covers arise as the set of vertices at distance two from a vertex in a strongly regular graph. Examples are the 6-fold cover of K_7, and the covers obtained from generalised quadrangles with parameters $(q - 1, q + 1)$, which arise as second neighbourhoods in the point graphs of the symplectic generalised quadrangles. The strongly regular graphs such that each second neighbourhood is an antipodal distance-regular graph are determined in [14]. However there are examples of strongly regular graphs where just one second neighbourhood is a cover. This prompts the following:

(2) Are there covers which arise as second neighbourhoods of strongly regular graphs, with parameter sets different from those given in [14]?

The smallest open cases arise from covers with parameter sets

$$(33, 3, 9),\ (36, 4, 8),\ (36, 6, 6). \tag{9.1}$$

If X is a (n, r, c_2)-cover and occurs as the second neighbourhood of a strongly regular graph Y, then the parameters of Y are determined as follows. Set γ equal to $c_2(c_2 - 1)/(n - c_2)$. Then Y has $(r+1)n + c_2 + \gamma$ vertices, valency $n - 1 + c_2 + \gamma$, two adjacent vertices have exactly $a_1 + \gamma$ common neighbours and two vertices at distance two have exactly $c_2 + \gamma$ common neighbours. It is clearly necessary that γ be an integer; it is also necessary to check that the output of the formulas for the multiplicities of the eigenvalues of Y are integers. In the three cases in (9.1), I do not even know that strongly regular graphs with the required parameters exist, let alone have the required second neighbourhoods.

Remark. Don de Caen (private communication) has shown that a $(33, 3, 9)$-cover cannot occur as the second neighbourhood of a strongly regular graph.

C. We have seen that covers with δ in $\{-2, 0, 2\}$ are special, and we have large classes of examples with δ equal to 0 or -2. The supply with $\delta = 2$ is much more limited, the only examples arise from generalised quadrangles with parameters $(q+1, q-1)$ and then n is a power of two. Thus we ask:

(3) Is there a cover with $\delta = 2$ where n is divisible by an odd prime?

All the covers we know with $\delta = -2$ have a prime power number of vertices; hence we also ask:

(4) Is there a cover of K_n such that $\delta = -2$ and n is not a prime power?

There are a number of parameter sets such that $\delta = \pm 2$ and $n = 36$ for which no covers have been found, but all known feasibility conditions are satisfied. In particular, the second and third parameter sets in (9.1) have $\delta = 2$ and -2 respectively.

Finally we list the parameter sets with $n \leq 25$ where we do not know if a cover exists.

	n	r	c_2	δ
1.	13	11	1	0
2.	16	7	2	0
3.	19	17	1	0
4.	21	4	5	-1
5.	21	10	2	-1
6.	22	6	4	-4
7.	25	3	7	2
8.	25	7	3	2
9.	25	21	1	2

Of course, the existence of covers 1. or 3. would imply the existence of a projective plane of order 12 or 18 respectively. These entries are included only for completeness. The cover of K_{22} is interesting in that it would be triangle-free.

References

[1] N. L. Biggs, Distance-regular graphs with diameter three, Ann. Discrete Math., **15** (1982), 69–80.

[2] R. C. Bose and W. S. Connor, Combinatorial properties of group divisible incomplete block designs, Ann. Math. Stat., **32** (1952), 367–383.

[3] W. G. Bridges and R. A. Mena, On the rational spectra of graphs with abelian Singer groups, Linear Algebra Appl., **46** (1982), 51–60.

[4] A. E. Brouwer, A. M. Cohen and A. Neumaier, "Distance-Regular Graphs", Springer, Berlin, 1989.

[5] A. E. Brouwer, C. D. Godsil and H. A. Wilbrink, Isomorphism between antipodal distance-regular graphs with diameter three, unpublished manuscript.

[6] A. E. Brouwer and L. M. G. M. Tolhuizen, A sharpening of the Johnson bound for binary linear codes and the non-existence of linear codes with Preparata parameters, Des. Codes Cryptogr., **3** (1993), 95–98.

[7] D. de Caen, R. Mathon and G. E. Moorhouse, A family of antipodal distance-regular graphs related to the classical Preparata codes, J. Algebraic Combin., to appear.

[8] A. R. Calderbank and W. M. Kantor, The geometry of two-weight codes, Bull. London Math. Soc., **18** (1986), 97–122.

[9] P. J. Cameron, Covers of graphs and EGQ's, Discrete Math. **97** (1991) 83–92.

[10] P. J. Cameron and J. H. van Lint, "Designs, Graphs, Codes and their Links", London Math. Soc. Student Texts 22, Cambridge U. P., Cambridge, 1991.

[11] P. Delsarte, An algebraic approach to the association schemes of coding theory, Philips Research Reports Supplements 1973, No. 10.

[12] A. D. Gardiner, Antipodal covering graphs, J. Combin. Theory Ser. B, **16** (1974) 255–273.

[13] A. D. Gardiner, Imprimitive distance-regular graphs and projective planes, J. Combin. Theory Ser. B, **16** (1974) 274–281.

[14] A. D. Gardiner, C. D. Godsil, A. D. Hensel and G. F. Royle, Second neighbourhoods of strongly regular graphs, Discrete Math. **103** (1992) 161–170.

[15] C. D. Godsil, "Algebraic Combinatorics", Chapman and Hall, New York, 1993.

[16] C. D. Godsil, Krein covers of complete graphs, Australas. J. Combin., **6** (1992) 245–255.

[17] C. D. Godsil, Geometric distance-regular covers, New Zealand J. Math., **22** (1993) 31–38.

[18] C. D. Godsil and A. D. Hensel, Distance regular covers of the complete graph, J. Combin. Theory Ser. B, **56** (1992), 205–238.

[19] W. M. Kantor, Projective planes of order q whose collineation groups have order q^2, J. Algebraic Combin., **3** (1994), 405–425.

[20] P. W. H. Lemmens and J. J. Seidel, Equiangular lines, J. Algebra, **24** (1973), 494–512.

[21] H. Lüneburg, "Translation Planes", Springer, Berlin, 1980.

[22] F. J. MacWilliams and N. J. A. Sloane, "The Theory of Error-Correcting Codes", North-Holland, Amsterdam, 1977.

[23] S. Payne and J. A. Thas, "Finite Generalised Quadrangles", Pitman, New York, 1985.

[24] T. Schade, Antipodal Distance-Regular Graphs, Ph. D. Thesis, Universität Gießen, 1994.

[25] J. J. Seidel, "Geometry and Combinatorics, Selected works of J. J. Seidel", Eds. D. G. Corneil and R. Mathon, Academic Press, San Diego, 1991.

Combinatorics and Optimization
University of Waterloo
Waterloo, Ontario
Canada N2L 3G1

Advanced Studies in Pure Mathematics 24, 1996
Progress in Algebraic Combinatorics
pp. 165–196

The Character Table of ${}^2E_6(2)$ Acting on the Cosets of Fi_{22}

A. A. Ivanov[1] **and Jan Saxl**

Abstract.

We consider the permutation action of $E \cong {}^2E_6(2)$ on the cosets of its maximal subgroup $F \cong Fi_{22}$. We calculate the intersection matrices and character table of the centralizer algebra corresponding to this action. There are three reasons for the interest in this particular representation. Firstly, it is a sporadic multiplicity-free action of a simple group of exceptional Lie type. Secondly, E and F are Y-groups Y_{333} and Y_{332}, respectively, factorized over their centers. We believe that the intersection matrices we have calculated might be useful for a computer-free identification of Y_{333} with $2^3 \cdot {}^2E_6(2)$. Thirdly, the permutation group considered is the one induced by the involution centralizer on the set of points fixed by an involution in the action of the Baby Monster F_2 on the cosets of the Fischer group Fi_{23}. The latter action has the largest rank (namely 23) among the primitive multiplicity-free actions of the sporadic simple groups and the calculation of its character table is an open problem.

§1. Introduction

Let us recall some basic facts concerning permutation groups and their centralizer algebras from [BI] and [BCN]. Let X be the set of (right) cosets of a subgroup H in a finite group G. Then G induces a transitive action on X by translations and H coincides with the stabilizer $G(x_0)$ of the coset $x_0 \in X$ containing the identity (that is, of H itself). We assume that the action is faithful, that is H does not contain a non-trivial normal subgroup of G. Let χ be the permutation character of G acting on X, that is, $\chi(g) = \#\{x \mid x \in X, x^g = x\}$ for $g \in G$.

Received January 17, 1995.
Revised October 23, 1995.
[1] The research was carried out while the author was visiting the University of Cambridge under an SERC Visiting Fellowship. Partial support from ISF, grant NMC000 is also gratefully acknowledged.

Let V be the space of complex valued functions defined on X and let $v(x)$ denote the characteristic function of $x \in X$. Then G acts naturally on V, preserving on it the inner product $\langle . \mid . \rangle$ with respect to which $\mathcal{B} = \{v(x) \mid x \in X\}$ is an orthonormal basis. In this basis the linear transformation of V induced by $g \in G$ is given by the matrix $M(g)$ whose (x, y)-entry is 1 if $y^g = x$ and 0 otherwise. Clearly $\chi(g) = tr(M(g))$.

Let $\Gamma_0, \Gamma_1, \ldots, \Gamma_{r-1}$ be the orbits of G on the set of ordered pairs of elements of X. Then the Γ_i are called the orbitals or 2-orbits and r is known as the rank of the permutation group (G, X). It is standard to assume that $\Gamma_0 = \{(x, x) \mid x \in X\}$ is the diagonal orbital. In what follows we assume that all Γ_i are symmetrical, that is, for $x, y \in X$, $(x, y) \in \Gamma_i$ if and only if $(y, x) \in \Gamma_i$. Let $\chi = \psi_0 + \psi_1 + \cdots + \psi_{s-1}$ be the decomposition of χ into a sum of G-irreducibles. The condition assumed is equivalent to the following: $r = s$, the irreducibles ψ_i are pairwise different (that is, χ is multiplicity-free) and the Frobenius – Schur indicator of every ψ_i is 1. It is also standard to take ψ_0 to be the principal character of G.

Let $\Gamma_i(x) = \{y \mid (x, y) \in \Gamma_i\}$, $0 \leq i \leq r-1$ and let $k_i = |\Gamma_i(x)|$. One can consider Γ_i as the set of edges of an undirected graph on X and we will identify Γ_i with this graph. Then $\Gamma_i(x)$ is the set of vertices adjacent to x in Γ_i and k_i is the valency of Γ_i. Let A_i be the adjacency matrix of Γ_i, that is, a matrix, whose rows and columns are indexed by the elements of X and the (x, y)-entry is 1 if $(x, y) \in \Gamma_i$ and 0 otherwise. We can consider A_i as a linear transformation of V written in the basis $\mathcal{B}$. The matrices A_i for $1 \leq i \leq r-1$ form a linear basis of the algebra $\mathcal{C}$ (the centralizer algebra) consisting of the matrices which commute with $M(g)$ for every $g \in G$. In particular

$$A_i \cdot A_j = \sum_{k=0}^{r-1} p_{ij}^k A_k.$$

The structure constant p_{ij}^k is equal to the number of vertices in $\Gamma_j(x)$ adjacent to a fixed vertex from $\Gamma_k(x)$ in the graph (determined by) Γ_i. Let B_i denote the $r \times r$ matrix whose (j, k)-entry is equal to p_{ij}^k, $0 \leq i, j, k \leq r-1$. Then B_i is called the intersection matrix of the graph Γ_i. We will always append such a matrix by the column $(k_0, k_1, \ldots, k_{r-1})^t$. The mapping $A_i \mapsto B_i$ $(0 \leq i \leq r-1)$ induces a faithful linear representation of $\mathcal{C}$; in particular A_i and B_i have the same minimal polynomial.

Let us consider the centralizer algebra from a different point of view. The G-module V possesses a decomposition $V = V_0 \oplus V_1 \oplus \cdots \oplus V_{r-1}$ into the direct sum of minimal G-invariant subspaces. These subspaces are pairwise orthogonal with respect to the inner product $\langle . \mid . \rangle$ and they support pairwise non-isomorphic irreducible representations of G. We can assume that the irreducible constituent ψ_j of χ is the character of G acting on V_j, $0 \leq j \leq r-1$. Let E_j be the linear transformation of V which acts as the identity on V_j and maps every $w \in V_k$ for $k \neq j$ to the zero vector. Then E_j belongs to $\mathcal{C}$, moreover $E_0, E_1, \ldots, E_{r-1}$ is a linear basis of $\mathcal{C}$ consisting of primitive idempotents (that is, $E_i \cdot E_j = \delta_{ij} E_i$).

Since A_i commutes with the action of G on V, by Schur's lemma it preserves every V_j as a whole and multiplies every vector of V_j by a scalar, which we denote by $p_i(j)$. This number is an entry of the transformation matrix between the two bases of $\mathcal{C}$:

$$A_i = \sum_{j=0}^{r-1} p_i(j) E_j.$$

So $p_i(0), \ldots, p_i(r-1)$ are the eigenvalues (non-necessarily distinct) of A_i and hence of the intersection matrix B_i as well. It is standard to write the inverse transformation as follows:

$$E_j = (1/n) \sum_{i=0}^{r-1} q_j(i) A_i,$$

where $n = |X|$. It is known that $q_j(i) = p_i(j) \cdot m_j / k_i$. As above, k_i is the valency of Γ_i and m_j is the dimension of V_j that is the rank of E_j and can be computed by the formulae:

$$m_j = n \cdot (\sum_{i=0}^{r-1} p_i(j)^2 / k_i)^{-1}.$$

If $(x, y) \in \Gamma_i$ then $(1/n) q_j(i)$ is the (x, y)-entry of E_j written in the basis $\mathcal{B}$. This implies an important geometrical interpretation of the $q_j(i)$. As above, let $v(x) \in \mathcal{B}$ be the characteristic function of $x \in X$ and let $v_j(x)$ be the projection of $v(x)$ into V_j, that is $v_j(x) = E_j v(x)$. If $(x, y) \in \Gamma_i$ then the inner product $\langle v_j(x) \mid v_j(y) \rangle$ equals to $(1/n) q_j(i)$. So after rescaling we obtain a realization of the elements of X as unit vectors $w_j(x) = (n/m_j) v_j(x)$ in V_j such that for $(x, y) \in \Gamma_i$ the inner product $\langle w_j(x) \mid w_j(y) \rangle$ equals to $q_j(i)/m_j = p_i(j)/k_i$ for $0 \leq i, j \leq r-1$.

The vector $w_j(x)$ is fixed by $G(x)$. Since ψ_j appears in χ with multiplicity 1, the Frobenius reciprocity rule implies that the subspace of V_j fixed by $G(x)$ is 1-dimensional. This determines $w_j(x)$ up to multiplication by a scalar. Since the action of G on every V_i can be realized by real matrices, the scalar is plus or minus one.

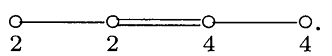
The matrix, whose (i, j)-entry is $p_i(j)$ is known as the character table of the centralizer algebra $\mathcal{C}$. We will append such a matrix by the column $(m_0, \ldots, m_{r-1})^t$.

In the present paper we compute the intersection matrices and character table of the centralizer algebra corresponding to the action of ${}^2E_6(2)$ on the cosets of Fi_{22}. We use a considerable amount of unpublished information on this action and on smaller configurations. The permutation character and 2-point stabilizers were determined by S. P. Norton using the fact that ${}^2E_6(2)$ is a section in the Monster. Later the character was independently computed by T. Breuer and K. Lux. In our work we rely on the information on the permutation character. At the same time we present a self-contained identification of the 2-point stabilizers. To meet the needs of the present project we asked L. H. Soicher to compute the intersection matrices of the primitive action of $O_8^+(2) : S_3$ of degree 11 200. Also, at our request S. A. Linton has computed the sizes of double cosets in Fi_{22} of a particular 2-point stabilizer, isomorphic to $2^{10} : M_{22}$, and all other such stabilizers. This is a very delicate information which has played a crucial role in our arguments. We have also used a computer program by D. V. Pasechnik which calculates the complete set of intersection matrices and the character table of a centralizer algebra from a single intersection matrix (having pairwise distinct eigenvalues). Finally, S. V. Shpectorov has suggested many improvements of the exposition of the paper. We are very grateful to all these people for their helpful cooperation.

Throughout the paper, given a group G we write $\bar{G}$ to denote $G/O_2(G)$.

§2. Preliminaries

The group $E \cong {}^2E_6(2)$ is a flag-transitive automorphism group of a Tits building $\mathcal{E}$ with the diagram



The elements of $\mathcal{E}$ will be called points, lines, planes and symplecta, respectively (nodes from the left to the right in the diagram). There is a natural bijection between the point set Δ of $\mathcal{E}$ and the conjugacy class of central involutions in E. We will not distinguish between these two sets. The following lemma describes the action of E on Δ (see for instance [Ivn]).

Lemma 2.1. *The group E acts transitively on Δ. Let $u \in \Delta$ and let $E(u)$ be the stabilizer of u in E. Then $E(u) \cong 2_+^{1+20} : U_6(2)$ is the* *centralizer of u as a central involution in E. $E(u)$ has five orbits $\Sigma_1(u) = \{u\}$, $\Sigma_2(u)$, $\Sigma_2'(u)$, $\Sigma_4(u)$ and $\Sigma_3(u)$ on Δ with lengths 1, 1 782, 44 352, 1 824 768 and 2 097 152, respectively. If $v \in \Sigma_i^{(')}(u)$ then the product of u and v (as involutions in E) has order i. The permutation character $\mathbf{1}_{E(u)}^E$ of E acting on Δ is $1a + 1\,938a + 48\,620a + 1\,828\,332a + 2\,089\,164a$.*

The subdegree 1 782 of E acting on Δ corresponds to the collinearity graph of $\mathcal{E}$ i.e., to the graph where two points are adjacent if they are incident to a common line. The intersection matrix of this graph is the following:

$$\begin{array}{rrrrrr}
0 & 1 & 0 & 0 & 0 & 1 \\
1782 & 85 & 27 & 1 & 0 & 1782 \\
0 & 672 & 27 & 42 & 0 & 44352 \\
0 & 1024 & 1728 & 715 & 891 & 1824768 \\
0 & 0 & 0 & 1024 & 891 & 2097152
\end{array}$$

In the above notation let $L = O_2(E(u))$, $U = E(u)/L \cong U_6(2)$ and Π be the residue of u in $\mathcal{E}$ on which U acts flag-transitively. Then Π is a rank 3 dual polar space of unitary type. The orbits of L on $\Sigma_2(u)$ are of length 2 and they correspond to 891 lines of $\mathcal{E}$ incident to u. The orbits of L on $\Sigma_2'(u)$ have length 64 and they correspond to the 693 symplecta incident to u. Every symplecton incident to u is also incident to 54 points from $\Sigma_2(u)$ and to 64 points from $\Sigma_2'(u)$. The subgraph in the collinearity graph induced by the points incident to a symplecton is strongly regular with parameters $v = 119$, $k = 54$, $\lambda = 21$, $\mu = 27$.

§3. The permutation character

The fact that Fi_{22} is a subgroup of ${}^2E_6(2)$ was first established by B. Fischer and his arguments were published in [Coo], Section 6. In terms of Y-groups the Fi_{22}-subgroups in ${}^2E_6(2)$ where classified by S. P. Norton in [Nor]. The permutation character of E acting on the cosets of F was computed by S. P. Norton and independently by T. Breuer and K. Lux at Aachen.

Lemma 3.1. *The group $E \cong {}^2E_6(2)$ contains three classes of maximal subgroups $F \cong Fi_{22}$ which are permuted transitively by outer automorphisms of E. The permutation character $\mathbf{1}_F^E$ is the following: $1a + 1\,938a + 48\,620a + 1\,828\,332a + 2\,909\,907x + 29\,099\,070a + 278\,555\,200a + 872\,972\,100x$, where $x = a$, b or c depending on the choice of the E-conjugacy class containing F.*

In the above lemma the permutation character is given in its decomposition into irreducibles and each irreducible is presented by its degree. In case there are several characters with the same degree we use letters a, b, etc. according to the ordering of the characters in [Atlas].

In order to simplify the references we present below the values of $\mathbf{1}_F^E$ on elements of certain classes. We use the upper case letters to name E-conjugacy classes and the lower case letters to name F-classes.

E-classes	E-normalizers	F-classes	F-normalizers	the values of $\mathbf{1}_F^E$
$1A$	${}^2E_6(2)$	$1a$	Fi_{22}	1 185 415 168
$2A$	$2^{1+20} : U_6(2)$	$2a$	$2 \cdot U_6(2)$	1 048 576
$3B$	$(3 \times O_8^+(2) : 3)2$	$3c$	$3.3^4 : 2^3.S_4.S_3$	11 200
$3C$	$3^{1+6}.2^{3+6}.(S_3 \times 3)$	$3b, 3d$		$4 + 576$
$5A$	$(D_{10} \times A_8).2$	$5a$	$F_{20} \times S_5$	168
$7B$	$(F_{21} \times L_2(7)).2$	$7a$	$F_{42} \times S_3$	28
$9A$		$9a$		1
$11A$		$11a$		3

It is straightforward to reconstruct the fusion of F-classes into E-classes (at least for elements of small order) directly from the permutation character $\mathbf{1}_F^E$. The structure of E- and F-normalizers in the above table is taken from [Atlas] for elements of order 1, 2 and 3. For the elements of order 5 and 7 the relevant information is obtained below, in Lemmas 6.4 and 6.3, respectively.

§4. Some properties of Fi_{22}

We will make use of the description of maximal subgroups in the Fischer group Fi_{22} obtained in [KW] (see also [Atlas]).

Lemma 4.1. *The Fischer group $F \cong Fi_{22}$ contains (up to conjugacy in its automorphism group) 12 classes of maximal subgroups with representatives H_i, $1 \le i \le 12$ as given on the next page. The $Aut(F)$-conjugacy classes containing H_2 and H_{11} split into two F-classes each.*

Let Ξ be the transposition graph of F. The vertices of Ξ are the Fischer transpositions of F ($2a$-involutions); two of them are adjacent if they commute. H_1 is the stabilizer of a transposition $\alpha \in \Xi$. The action induced by H_1 on the set $\Xi(\alpha)$ of transpositions adjacent to α is similar to the action of $\bar{H}_1 \cong U_6(2)$ on the set of planes of the dual polar space Π. Two transpositions from $\Xi(\alpha)$ are adjacent if and only if the corresponding planes in Π are incident to a common line. This implies that every maximal clique of Ξ has size 22. H_4 is the stabilizer of such

i	$[F:H_i]$	H_i	Remarks
1	3 510	$2\cdot U_6(2)$	$N(2a)$
2	14 080	$O_7(3)$	
3	61 776	$O_8^+(2):S_3$	
4	142 155	$2^{10}:M_{22}$	$N(2a_{22}2b_{231}2c_{770})$
5	694 980	$2^6:Sp_6(2)$	$N(2b^6)$
6	1 216 215	$(2\times 2_+^{1+8}:U_4(2)):2$	$N(2b)$
7	1 647 360	$S_3\times U_4(3):2_2$	$N(3a)$
8	3 592 512	${}^2F_4(2)'$	
9	3 648 645	$2^{5+8}:(S_3\times A_6)$	$N(2b^4)$
10	12 812 800	$3_+^{1+6}:2^{3+4}:3^2:2$	$N(3b)$
11	17 791 488	S_{10}	
12	679 311 360	M_{12}	

a clique and it is known to contain exactly 22 transpositions (the ones in the clique). In what follows we will denote this subgroup simply by H.

Let Ω denote the set of maximal cliques in the transposition graph Ξ naturally identified with the cosets of $H\cong 2^{10}:M_{22}$ in $F\cong Fi_{22}$. The following result was proved in [RW].

Lemma 4.2. *The subgroup $H\cong 2^{10}:M_{22}$ acting on Ω has 8 orbits $\Omega_0,\dots,\Omega_7$ with lengths 1, 154, 1 024, 3 696, 4 928, 11 264, 42 240 and 78 848 and stabilizers $2^{10}:(M_{22})$, $2^9.(2^4:A_6)$, $1.(M_{22})$, $2^6.(2^4:S_5)$, $2^4.(2^4:A_6)$, $2.(L_3(4))$, $2^3.(2^3:L_3(2))$ and $1.(2^4:A_6)$, respectively. Here when a stabilizer Z is written as $X.(Y)$, we mean that $X=Z\cap O_2(H)$ and Y is the image of Z in $\bar{H}\cong M_{22}$. The permutation character $\mathbf{1}_H^F$ is the following: $1a+78a+429a+1\,430a+3\,080a+30\,030a+32\,032a+75\,075a$.*

Let $z_i\in\Omega_i$, $0\le i\le 7$, so that $y=z_0$ is the clique stabilized by H. The intersection $y\cap z_i$ has size 22, 6, 2 and 1 for i =0, 1, 4 and 5, respectively and is empty in the remaining cases. As above, let α be the transposition centralized by H_1. Then α stabilizes $z\in\Omega$ if and only if $\alpha\in z$. This implies the following.

Lemma 4.3. *The set $\Omega(\alpha)$ of elements in Ω fixed by α has size 891; H_1 induces on $\Omega(\alpha)$ the action of $\bar{H}_1$ as on the points of Π. Assume that $\omega_0\in\Omega(\alpha)$. Then $\Omega(\alpha)$ intersects Ω_i in 1, 42, 336 and 512 elements for $i=0$, 1, 4 and 5, respectively.*

Let $\mathcal{S}$ be the Steiner system $S(3,6,22)$ defined on z_0 and acted on naturally by $\bar{H} \cong M_{22}$. We see from Lemma 4.2 that the orbits of $O_2(H)$ on Ω_i have lengths 1, 2, 2^{10}, 2^4, 2^6, 2^9, 2^7 and 2^{10} for $i = 0$ to 7, respectively. Moreover, the action of $\bar{H} \cong M_{22}$ on the set of $O_2(H)$-orbits on Ω_i is trivial for $i = 0$ and 2; as on the points of $\mathcal{S}$ for $i = 5$; as on the blocks of $\mathcal{S}$ for $i = 1$, 4 and 7; as on the duads for $i = 3$ and as on the special octets for $i = 6$. The intersection matrix of F acting on the cosets of H which correspond to the subdegree 154 and the character table of the centralizer algebra are given below (cf. [ILLSS]).

0	1	0	0	0	0	0	0	1
154	1	5	1	0	0	0	0	154
0	120	5	0	21	7	0	0	3696
0	32	0	1	0	14	77	1	4928
0	0	64	0	21	0	0	16	11264
0	0	80	120	0	21	0	60	42240
0	0	0	16	0	0	0	1	1024
0	0	0	16	112	112	77	76	78848

1	154	3696	4928	11264	42240	1024	78848	1
1	−77	924	1232	−1408	−5280	−320	4928	78
1	49	546	−532	1184	−960	−176	−112	429
1	−35	294	−364	−400	120	160	224	1430
1	37	186	248	32	120	88	−712	3080
1	13	6	−28	−112	120	−32	32	30030
1	−17	24	32	32	120	−20	−172	32032
1	1	−30	−4	32	−96	16	80	75075

We will make use of the following two lemmas. The former comes from [ILLSS] and direct calculations with the permutation character, while the latter follows from calculations performed by S. A. Linton (cf. remark before Lemma 10.1).

Lemma 4.4. *The group $F \cong Fi_{22}$ acting on the cosets of $H_5 \cong 2^6 : Sp_6(2)$ has rank* 10 *with subdegrees* 1, 135, 1 260, 2 304, 8 640, 10 080, 45 360, 143 360 *and* 241 920 (*twice*). *An element of type* 5a *from F fixes exactly five cosets of H_5 in F.*

Lemma 4.5. *The subgroup $H_5 \cong 2^6 : Sp_6(2)$ acting on Ω has* 6 *orbits with lengths* 135, 756, 8 640, 15 120, 48 384 *and* 69 120.

§5. The action of F on the central involutions of E

Lemma 5.1. *Let Δ be the set of central involutions in E. Then F acting on Δ has 4 orbits Δ_i, $1 \le i \le 4$ with lengths 3 510, 142 155, 3 127 410 and 694 980. The stabilizer in F of an involution from Δ_i is isomorphic to $2 \cdot U_6(2)$, $2^{10} : M_{22}$, $2^{10}.L_3(4)$ and $2^6 : Sp_6(2)$, for $i = 1$, 2, 3 and 4, respectively.*

Proof. It follows from the table, given in Section 3, that Δ_1 defined as $\Delta \cap F$ is the class of $2a$-involutions in F and hence it forms an F-orbit with stabilizer $H_1 \cong 2 \cdot U_6(2)$. We will use notation introduced in and after Lemma 2.1 and assume that $u \in \Delta_1$. Then $F(u) \cong 2 \cdot U_6(2)$ intersects L in a subgroup of order 2 and $F(u)L/L = U \cong U_6(2)$. The subgroup $F(u)$ has two orbits, say $\Delta_1^2(u)$ and $\Delta_1^3(u)$ on $\Delta_1 - \{u\}$ consisting of transpositions whose products with u have order 2 and 3, respectively. By Lemma 2.1 we see that $\Delta_1^3(u) \subseteq \Sigma_3(u)$ and either $\Delta_1^2(u) \subseteq \Sigma_2(u)$ or $\Delta_1^2(u) \subseteq \Sigma_2'(u)$. The intersection matrix given after Lemma 2.1 shows that in the collinearity graph of $\mathcal{E}$ (which corresponds to $\Sigma_2(u)$) every vertex $v \in \Sigma_3(u)$ is at distance 3 from u and hence $\Delta_1^2(u) \subseteq \Sigma_2'(u)$. Since $F(u)/\langle u \rangle \cong U_6(2)$ acts on $\Delta_1^2(u)$ as it acts on the symplecta incident to u and for every $u' \in \Sigma_2(u)$ there is a unique symplecton incident to both u and u', we conclude that for every symplecton of $\mathcal{E}$ incident to u there is exactly one point in $\Delta_1^2(u)$ which is also incident to this symplecton. Consider the action of $F(u)$ on $\Sigma_2(u)$. Clearly $\langle u \rangle$ is the kernel of the action and since the stabilizer in $U_6(2)$ of a point from Π (isomorphic to $2^9.L_3(4)$) does not contain subgroups of index 2, $F(u)$ has two orbits, say $\Sigma_2^1(u)$ and $\Sigma_2^2(u)$ on $\Sigma_2(u)$. Every line of $\mathcal{E}$ incident to u is also incident to one point from $\Sigma_2^1(u)$ and to one point from $\Sigma_2^2(u)$.

Let $w \in \Delta_1^2(u)$ and let $\{u, x_1, x_2\}$ be a line incident to the symplecton containing the pair $\{u, w\}$, where $x_i \in \Sigma_2^i(u)$, $i = 1, 2$. Then, in the collinearity graph of $\mathcal{E}$, the point w is incident to exactly one of x_1 and x_2, say to x_1. Now $(F(u) \cap F(x_1))L/L$ is the stabilizer in U of a point in Π and hence it acts (doubly) transitively on the 21 planes of Π incident to this point. Hence in the collinearity graph of $\mathcal{E}$, the vertex x_1 is adjacent to exactly 22 vertices from Δ_1 while x_2 is adjacent to only one such vertex, namely to u. Let Δ_2 and Δ_3 be the orbits of F on Δ which contain $\Sigma_2^1(u)$ and $\Sigma_2^2(u)$, respectively. Then the above arguments show that $F(x_2) \cong 2.2^9.L_3(4)$ is contained in $F(u)$ while $F(x_1)$ contains $F(x_2)$ as a subgroup of index 22 and it is straightforward to see that $F(x_1) \cong 2^{10}.M_{22}$ (a conjugate of H). Finally, comparison of the characters $\mathbf{1}_{E(u)}^E$ and $\mathbf{1}_F^E$ shows that F has four orbits on Δ. Hence Δ_4 defined as $\Delta - \Delta_1 - \Delta_2 - \Delta_3$ is an F-orbit of length 694 980 and if follows from Lemma 4.1 that for $v \in \Delta_4$ we have $F(v) \cong 2^6 : Sp_6(2)$ (a conjugate of H_5). Q.E.D.

Let X be the set of cosets of F in E and let x be the coset fixed by F (that is F itself). For a subset Y of X let $E(Y)$ denote the elementwise stabilizer of Y in E (we write $E(a, b, \dots)$ instead of $E(\{a, b, \dots\})$). By Lemma 3.10 the character $\mathbf{1}_F^E$ has 8 irreducible components, so F has 7 orbits on $X - \{x\}$. We will use Lemma 5.1 to identify some of them.

Lemma 5.2. *Let M be a maximal subgroups in $F = E(x)$ conjugate to H or H_5. Then there is an element $z \in X - \{x\}$ such that $E(x, z) = M$. Moreover, the setwise stabilizer of $\{x, z\}$ in E is $M \times \langle \tau \rangle$ where τ is a central involution in E.*

Proof. By Lemma 5.1 there is a unique involution $\tau \in \Delta - F$ which commutes with M. This implies that $E(x, z) \geq M$ for $z = x^\tau$. Since M is maximal in F and F does not fix cosets in X other than x, we obtain $E(x, z) = H$. Q.E.D.

Notice that the orbit Δ_3 of F on Δ can not be used in similar way to produce a new F-orbit on X. In fact, an involution from Δ_3 (say x_2 as in Lemma 5.1) conjugates the coset x into the orbit with the stabilizer $2^{10} : M_{22}$. Indeed, $x_2 = ux_1$ and since u is contained in F, it stabilizes x.

The above lemma shows that H and H_5 are 2-point stabilizers in the action of E on X. From the properties of E as a Lie type group we can deduce another 2-point stabilizer.

Lemma 5.3. *Let $H_8 \cong {}^2F_4(2)'$ be a maximal subgroup in $F \cong E(x)$, isomorphic to the Tits group. Then H_8 stabilizes a vertex from $X - \{x\}$.*

Proof. Let V be the natural 27-dimensional $GF(4)$-module for E. It follows from [JLPW] or Proposition 5.4.12 in [KL], that V, restricted to H_8, has two composition factors V_1 and V_2 with dimensions 1 and 26, respectively. Substituting V by its dual, if necessary, we assume that H_8 fixes a 1-dimensional subspace V_1 in V. Comparing the ordinary and modular character tables of H_8, we conclude that V_2 is the reduction modulo 2 of a real, irreducible 26-dimensional representations of H_8. Let $N \cong 13.3$ be the normalizer of a Sylow 13-subgroup in H_8. Computing the inner product of the principal character of N and its character on V_2 (or, rather, on the real version of V_2) we obtain 0. Hence N fixes in V no 1-dimensional subspaces besides V_1. On the other hand a maximal subgroup $F_4(2)$ in E fixes a 1-dimensional subspace in V and contains

N. Hence H_8 is contained in the full stabilizer of V_1 in E, isomorphic to $F_4(2)$. The list of maximal subgroups in $F_4(2)$ obtained in [NW] shows that the normalizer in $F_4(2)$ of every subgroup ${}^2F_4(2)'$ is isomorphic to ${}^2F_4(2)$. This means that $N_E(H_8)$ contains H_8 properly and the result follows. Q.E.D.

§6. Fixed points subgraphs

Let τ be the unique involution from Δ which commutes with H (compare Lemma 5.2) and let $y = x^\tau$. Let Γ be the graph on X with the edge set $\{(x,y)^g \mid g \in E\}$. For $z \in X$ let $\Gamma(z)$ denote the set of vertices of Γ adjacent to z. Then $F = E(x)$ acts on $\Gamma(x)$ as on the cosets of $H = E(x,y)$. For an ordered pair (a,b) of adjacent vertices in Γ let $\Gamma(a,b;n)$ denote the orbit of length n of $E(a,b) \cong 2^{10} : M_{22}$ on $\Gamma(a)$. The possible values of n are listed in Lemma 4.2.

Let $\Gamma_2(x)$ be the orbit of F on $X - \{x\}$ which contains the image of x under the unique involution from Δ which commutes with H_5 (compare Lemma 5.2). Then the action of F on $\Gamma_2(x)$ is similar to its action on the cosets of H_5.

We are going to determine the structure of subgraphs in Γ induced by vertices fixed by certain (prime order) elements $d \in E$. Since we are interested in non-trivial subgraphs, we only take d from $E(x,y)$. This leaves us with a number of possibilities, among which we find the classes $11A$, $7B$, $5A$, $3B$ and $2A$ of E. Let cl be the conjugacy class of E containing d; let $D = \langle d \rangle$ be the cyclic subgroup generated by d; $M(cl) = N_E(D)$ and $\Phi(cl)$ be the set of elements from X fixed by d. Finally, let $\Gamma(cl)$ be the subgraph of Γ induced by $\Phi(cl)$. Notice that the size of $\Phi(cl)$ (i.e., the number of cosets from X fixed by d) is equal to the value of $\mathbf{1}_F^E$ on d; so we can use the values from the table in Section 3. This table shows that in each of the five cases we consider, $M(cl)$ acts transitively on $\Phi(cl)$ since only one conjugacy class of F fuses to cl. In addition, if $cl = 11A$, $7B$ or $5A$, then D is a Sylow subgroup in F and hence $N_F(D)$ acts transitively on $\Phi(cl) \cap \Theta$ for every orbit Θ of F on X.

Lemma 6.1. *$\Gamma(11A)$ is the complete graph on 3 vertices and $M(11A)$ induces on it the group S_3.*

Proof. The value of $\mathbf{1}_F^E$ on elements of type $11A$ is 3, hence $|\Phi(11A)| = 3$. Since $|\Gamma(x)| = 2 \bmod 11$, we have $\Phi(11A) \subseteq \{x\} \cup \Gamma(x)$. By the paragraph before the lemma $M(11A)$ is transitive on $\Phi(11A)$ and $N_F(D)$ is transitive on $\Phi(11A) \cap \Gamma(x)$, so the result follows. Q.E.D.

By our choice of d we have $x, y \in \Gamma(11A)$. Since $\Gamma(x, y; 1024)$ is an orbit of $E(x, y)$ on $\Gamma(x) - \{y\}$ whose length is not divisible by 11, the third vertex of $\Gamma(11A)$ must be in this orbit and we have the following.

Corollary 6.2. $\Gamma(x, y; 1024) \subseteq \Gamma(y)$.

Lemma 6.3. *$M(7B)$ induces on $\Phi(7B)$ a primitive action of degree 28 which is similar to the action of $PGL_2(7)$ on the cosets of $2 \times S_3$; the action has rank 5 with subdegrees 1, 3, 6 (twice) and 12. The intersection matrix of $\Gamma(7B)$ is the following.*

$$\begin{array}{cccccc} 0 & 1 & 0 & 0 & 0 & 1 \\ 6 & 2 & 2 & 1 & 0 & 6 \\ 0 & 1 & 0 & 1 & 0 & 3 \\ 0 & 2 & 4 & 2 & 4 & 12 \\ 0 & 0 & 0 & 2 & 2 & 6 \end{array}$$

Proof. The value of $\mathbf{1}_F^E$ on elements of type $7B$ is 28. By the paragraph before Lemma 6.1 $M(7B)$ induces on $\Phi(7B)$ a transitive action (of degree 28) and $N_F(D)$ acts transitively on $\Phi(7B) \cap \Theta$ for every orbit Θ of F on X. From the order we observe that $N_F(D)$ is contained in a maximal subgroup $H_{11} \cong S_{10}$ of F and hence $N_F(D) \cong F_{42} \times S_3$ and $C_F(D) \cong 7 \times S_3$. From the permutation character $\mathbf{1}_H^F$ given in Lemma 4.2, we get $|\Phi(7B) \cap \Gamma(x)| = 6$. Since $\bar{H} \cong M_{22}$ acting on the non-trivial elements of $O_2(H)$ has three orbits with lengths 22, 231 and 770, one can see (compare [Atlas]) that $N_H(D) \cong F_{21} \times 2$. Comparing the structures of $N_F(D)$, $C_F(D)$ and $N_H(D)$, we conclude: (a) if K is the elementwise stabilizer in $M(7B)$ of the connected component of $\Gamma(7B)$ containing x, then $K \cong F_{21}$; (b) $N_F(D)/K \cong 2 \times S_3$; (c) $C_F(D)$ acting on $\Phi(7B) \cap \Gamma(x)$ has two orbits of length 3 each; (d) there is a subgroup T of order 3 in $N_F(D) - K$ which commutes with K. Let S be the setwise stabilizer in $M(7B)$ of the connected component of $\Gamma(7B)$ which contains x and let $\hat{S} = S/K$. The connected component has 7, 14 or 28 vertices and hence $|\hat{S}| = 2^a \cdot 3 \cdot 7$ for $a = 2$, 3 or 4, respectively. By Sylow theorem $\hat{S}$ contains a subgroup of order 21. Suppose this subgroup is cyclic. Then by (d) F contains a subgroup T of order 3 such that $N_E(T)$ has order divisible by 49. Since there are no such subgroups, $\hat{S}$ contains F_{21}. Since $\hat{S}$ also contains $N_F(D)/K \cong 2 \times S_3$, it is easy to show that $\hat{S}$ is non-solvable. The non-abelian composition factor of $\hat{S}$ must be $PSL_2(7)$ and since $N_F(D)/K \cong 2 \times S_3$ contains no normal subgroups of $\hat{S}$, we obtain $\hat{S} \cong PGL_2(7)$. This means that $M(7B)$ acts on $\Phi(7B)$ as $PGL_2(7)$ acts on the antiflags of the projective plane of order 2. There are two orbitals of valency 6 with respect to this action and $\Gamma(7B)$ is characterized by the property that it splits under the restriction to $PSL_2(7)$. Q.E.D.

Lemma 6.4. *$M(5A)$ induces on $\Phi(5A)$ an action of degree* 168 *which is similar to the action of S_8 on the cosets of $S_5 \times 2$ (having two orbits with lengths* 2 *and* 6 *in the natural action of S_8); the action has rank* 6 *with subdegrees* 1, 5, 12, 30 *and* 60 *(twice). The intersection matrix corresponding to $\Gamma(5A)$ is as given below. The* 2*-point stabilizers of the action (ordered in accordance with the rows of the intersection matrix) are $S_5 \times 2$, $D_8 \times 2$, Z_4, $S_4 \times 2$, 2^2 and F_{20}.*

0	1	0	0	0	0	1
30	8	8	6	2	0	30
0	16	9	0	12	5	60
0	1	0	0	2	0	5
0	4	12	24	10	20	60
0	0	1	0	4	5	12

Proof. We know from [Atlas] the orders of the normalizers of D in F and E. This tells us that $N_F(D) < S_{10} < F$ and $N_E(D) < (A_5 \times A_8).2 < O_{10}^-(2) < {}^2E_6(2)$. In particular $N_F(D) \cong F_{20} \times S_5$. On the other hand we can check that $N_F(D)$ induces on $\Phi(5A) \cap \Gamma(x)$ an action of $S_5 \times 2$ of degree 30. Hence the kernel of $N_E(D)$ acting on $\Phi(5A)$ is D_{10} and the action must be isomorphic to S_8. There are two conjugacy classes of subgroups in S_8 isomorphic to $S_5 \times 2$. Let R_1 and R_2 be their representatives such that R_1 has two orbits with lengths 2 and 6 on the 8-element set Y naturally permuted by S_8 and R_2 has three orbits on Y with lengths 1, 2 and 5. Then the action of S_8 on the cosets of R_1 preserves an imprimitivity system with blocks of size 6 and the action on the blocks is similar to the action on 2-element subsets of Y with subdegrees 1, 12 and 15. The action on the cosets of R_2 preserves an imprimitivity system with blocks of size three and the action on the blocks is similar to the action on 3-element subsets of Y with subdegrees 1, 10, 15 and 30. Let $z \in X - \{x\}$ be a point stabilized by the subgroup $H_8 \cong {}^2F_4(2)'$ in $F = E(x)$ (cf. Lemma 5.3). Then F acts on the suborbit $\{z^f \mid f \in F\}$ as it acts on the cosets of H_8. The permutation character of this action is given in [ILLSS] and it has value 12 on $5A$-elements. This means that 12 is a subdegree of $M(5A)$ acting on $\Phi(5A)$ and from above we conclude that the action is of S_8 on the cosets of R_1. Using the information about the action on the imprimitivity blocks it is easy to calculate the intersection matrix and 2-point stabilizers. Q.E.D.

0	1	0	0	0	0	0	0	1
243	26	18	14	3	2	0	0	243
0	24	9	6	9	6	0	0	324
0	112	36	43	42	46	162	0	1944
0	48	108	84	81	84	0	108	3888
0	32	72	92	84	80	81	108	3888
0	0	0	4	0	1	0	0	48
0	0	0	0	24	24	0	27	864

Lemma 6.5. *$M(3B)$ induces on $\Phi(3B)$ a primitive action of $O_8^+(2).S_3$ of degree* 11 200; *the action has rank* 8 *with subdegrees* 1, 48, 243, 324, 864, 1 944 *and* 3 888 (*twice*). *The intersection matrix of* $\Gamma(3B)$ *is as given above.*

Proof. All elements of order 3 in $E(x,y)$ are conjugate and the character $\mathbf{1}_H^F$ (cf. Lemma 4.2) shows that such an element d belongs to the class $3c$ in F and hence to the class $3B$ in E. It is easy to calculate that d fixes 243 points in $\Gamma(x)$ and $N_F(D)$ acts transitively on these points. The intersection matrices of $O_8^+(2):S_3 \cong N_E(D)/D$ in its primitive action of degree 11 200 were computed by L. H. Soicher. There is only one orbital of valency 243 and we identify it with $\Gamma(3B)$. Q.E.D.

Lemma 6.6. *$M(2A)$ induces on $\Phi(2A)$ a primitive action of degree* 1 048 576, *similar to the action of* $2^{20}:U_6(2)$ *on the cosets of* $U_6(2)$; *the action has rank* 6 *with subdegrees* 1, 891, 24 948, 228 096, 295 680 *and* 498 960. *The intersection matrix of* $\Gamma(2A)$ *and the character table of the corresponding centralizer algebra are as given below.*

0	1	0	0	0	0	1
891	42	12	2	0	0	891
0	336	15	42	24	0	24948
0	512	384	147	192	216	228096
0	0	480	420	419	432	498960
0	0	0	280	256	243	295680

1	891	24948	228096	498960	295680	1
1	−133	3444	−9472	−2800	8960	891
1	123	628	768	−240	−1280	24948
1	−37	84	512	−560	0	228096
1	27	−108	0	−432	512	295680
1	−5	−12	−256	528	−256	498960

Proof. A transposition from F is a central involution in E. By Lemma 4.3 $\Gamma(2A)$ has valency 891, so the corresponding orbital is uniquely determined. Q.E.D.

The set $\Phi(2A) \cap \Gamma(x)$ coincides with the set $\Omega(\alpha)$ from Lemma 4.3. The entry 42 in the second row and second column of the intersection matrix of $\Gamma(2A)$ shows that y is adjacent to 42 vertices from $\Phi(2A) \cap \Gamma(x)$. By Lemma 4.3 these 42 vertices are contained in Ω_1 which is the same as $\Gamma(x, y; 154)$ and we have the following.

Corollary 6.7. $\Gamma(x, y; 154) \subseteq \Gamma(y)$.

Let us look more closely at the subgraph $\Gamma(5A)$.

Lemma 6.8. *Let d be an element of type $5A$ from $E(x, y)$. Then the number of vertices fixed by d in $\Gamma(x, y; n)$ is 1 for $n = 1$ and 3696; 4 for $n = 154$, $1\,024$ and $11\,264$; 8 for $n = 4\,928$ and $78\,848$.*

Proof. By Lemma 6.4 and its proof, $N_{E(x,y)}(D)$ induces on $\Phi(5A) \cap \Gamma(x)$ an action of degree 30 and order 8. By the Frattini argument $N_{E(x,y)}(D)$ is transitive on $\Phi(5A) \cap \Gamma(x, y; n)$ for every n and the result follows from elementary congruences. Q.E.D.

As above let d be an element of type $5A$ in $E(x, y)$ and $D = \langle d \rangle$. Because of the obvious symmetry between x and y, Lemma 6.8 describes the orbits of $N_{E(x,y)}(D)$ on $\Phi(5A) \cap \Gamma(y)$ and we can locate them in $\Gamma(y, x; n)$ for suitable values of n. The intersection matrix of $\Gamma(5A)$, given in Lemma 6.4 shows that y is adjacent to 8 vertices from $\Phi(5A) \cap \Gamma(x)$; we have 4 of them in $\Gamma(y, x; 154) = \Gamma(x, y; 154)$ (compare Corollary 6.7) and 4 in $\Gamma(y, x; 1024) = \Gamma(x, y; 1024)$ (compare Corollary 6.2). From the intersection matrix of $\Gamma(5A)$ we see that y is adjacent to a single vertex in the orbit of length 5 of $N_F(D)$ on $\Phi(5A)$. By Lemma 6.8 this vertex is in $\Gamma(y, x; 3696)$ (since it can not be $\{x\}$). We observed in Lemma 4.4 that $\Phi(5A) \cap \Gamma_2(x)$ has size 5 and clearly $N_F(D)$ acts on it transitively. This gives the following.

Corollary 6.9. $\Gamma(y, x; 3696) \subseteq \Gamma_2(x)$.

Let 60_1 and 60_2 be the orbits of length 60 of $N_F(D)$ on $\Phi(5A)$. We assume that y is adjacent to 4 vertices from 60_1. Then these 4 vertices form an orbit of $N_{E(x,y)}(D)$ and are contained in $\Gamma(y, x; 11264)$. There are 16 vertices in 60_2 adjacent to y. By what we already know, Lemma 6.8 imply that these 16 vertices split under the action of $N_{E(x,y)}(D)$ into two orbits of length 8 each and these two orbits are contained in $\Gamma(y, x; 4928)$ and in $\Gamma(y, x; 78848)$. This gives us the following important conclusion.

Lemma 6.10. *$\Gamma(y,x;4928)$ and $\Gamma(y,x;78848)$ are in the same orbit of $F = E(x)$ on X.*

Notice that if $z \in 60_2$ then the stabilizer of z in $N_F(D)$ acting on $\Gamma(x) \cap \Gamma(z) \cap \Gamma(5A)$ has two orbits of length 4 each.

§7. The eigenspace V_1

The information on the local structure of Γ established in the previous section turns out to be sufficient to calculate a non-trivial eigenvalue of the intersection matrix of Γ (without knowing the matrix itself).

Let $\Gamma_0, \Gamma_1, \ldots, \Gamma_7$ be the orbitals of E acting on X, so that Γ_0 is the diagonal orbital and $\Gamma_1 = \Gamma$. The valency of Γ_i will be denoted by k_i and we will write k instead of k_1. Let V be the space of complex valued functions defined on X and let $V = V_0 \oplus V_1 \oplus \cdots \oplus V_7$ be the decomposition of V into minimal E-invariant subspaces. We denote by m_i the dimension of V_i and assume that $m_i > m_j$ for $i > j$. The values of m_j can be read from Lemma 3.1. The subspace V_0 supports the trivial representation of E and V_1 is 1938-dimensional. For $a \in X$ let $v(a)$ be the characteristic function of a, $v_1(a)$ be the projection of $v(a)$ into V_1 and $w_1(a) = (n/m_1)v_1(a)$ where $n = |X|$. Then for $(a,b) \in \Gamma_i$ the inner product $\langle w_1(a) \mid w_1(b)\rangle$ equals to $p_i(1)/k_i$ and $(p_0(1)/k_0, p_1(1)/k_1, \ldots, p_7(1)/k_7)$ is a left eigenvector of the intersection matrix of Γ. We emphasize that the $w_1(a)$ are unit vectors.

We are going to write down certain expressions for the inner products $\langle w_1(y) \mid w_1(z)\rangle$ for various vertices $y, z \in \Gamma(x)$ but before doing so we introduce some notation concerning the action of $F = E(x)$ on $\Gamma(x)$.

Let Ω be the set of cosets of $H \cong 2^{10} : M_{22}$ in F, which can be identified with $\Gamma(x)$. Let $\Omega_0, \ldots, \Omega_7$ be the orbits of F on Ω as in Lemma 4.2, in particular $\Omega_0 = \{y\}$. Let U denote the space of complex valued functions on Ω and $U = U_0 \oplus U_1 \oplus \cdots \oplus U_7$ be its decomposition into minimal F-invariant subspaces. The dimension of U_i will be denoted by d_i and we assume that $d_i > d_j$ for $i > j$. The values of d_i one can get from Lemma 4.2. For $z \in \Omega$ let $u(z)$ be the characteristic function of z, $u_j(z)$ be its projection into U_j and $t_j(z) = (k/d_j)u_j(z)$ (where $k = |\Omega|$). For $z \in \Omega_i$ put $\pi_i(j) = \langle t_j(y) \mid t_j(z)\rangle$. These numbers can be calculated from the character table corresponding to the action of F on the cosets of H and given after Lemma 4.3.

The following result can be checked by straightforward calculations using [Atlas].

Lemma 7.1. *The* 1938*-dimensional E-irreducible module V_1 when restricted to F decomposes as $U_0 \oplus U_1 \oplus U_2 \oplus U_3$ where the irreducibles U_j, $j = 0, 1, 2$ and* 3 *are those involved in the permutation character $\mathbf{1}_H^F$.*

Let us consider V_1 as a module for $F = E(x)$, possessing the decomposition into F-irreducibles from Lemma 7.1. Notice that the U_j are pairwise orthogonal with respect to an inner product on V_1 preserved by E and that the restriction of this inner product to U_j is the unique one (up to a scalar) preserved by F.

It is clear that $w_1(x) \in U_0$. Let us locate the vectors $w_1(z)$ for $z \in \Gamma(x)$ in the above decomposition of V_1. Since U_j appears in the permutation character of F acting on the cosets of $F(z)$ with multiplicity 1, we conclude that the vectors in U_j fixed by $F(z)$ form a 1-dimensional subspace. Since both $t_j(z)$ and the projection of $w_1(z)$ to U_j are fixed by $F(z)$, they must differ by a scalar multiple and hence

$$w_1(z) = \sum_{j=0}^{3} t_j(z) \cdot \alpha_j,$$

for some scalars α_j, which are independent of the choice of $z \in \Gamma(x)$. Since the decomposition of V_1 we are dealing with is orthogonal, for $z \in \Omega_i$ we have

$$\langle w_1(y) \mid w_1(z) \rangle = \sum_{j=0}^{3} \langle t_j(y) \mid t_j(z) \rangle \cdot \alpha_j^2 = \sum_{j=0}^{3} \pi_j(i) \cdot \alpha_j^2.$$

In addition the vector $t_0(z)$ is a unit vector in U_0 independent of the choice of $z \in \Gamma(x)$, so we can assume that they all coincide with $w_1(x)$. In this case $\langle w_1(x) \mid w_1(z) \rangle = \alpha_0$ for every $z \in \Gamma(x)$.

Thus we have four unknowns α_j, $0 \leq j \leq 3$ which determine the inner products $\langle w_1(y) \mid w_1(z) \rangle$ (depending on i such that $z \in \Omega_i$). Now we are going to turn the structural results on Γ proved in the previous section into equations on the α_j.

Let $z_0, \ldots, z_7 \in \Gamma(x)$ be such that $z_i \in \Omega_i$. In these terms Lemmas 6.2 and 6.7 mean that $z_1, z_2 \in \Gamma(y)$. Hence $\langle w_1(y) \mid w_1(z_1) \rangle = \langle w_1(y) \mid w_1(z_2) \rangle = \langle w_1(y) \mid w_1(x) \rangle = \alpha_0$ and we come to the following two equations:

$$\sum_{j=0}^{3} \pi_j(1) \cdot \alpha_j^2 = \alpha_0, \tag{7.1}$$

$$\sum_{j=0}^{3} \pi_j(2) \cdot \alpha_j^2 = \alpha_0. \tag{7.2}$$

Lemma 6.10 means that (y, z_4) and (y, z_7) are in the same E-orbit. Hence $\langle w_1(y) \mid w_1(z_4)\rangle = \langle w_1(y) \mid w_1(z_7)\rangle$ and we have

$$\sum_{j=0}^{3} (\pi_j(4) - \pi_j(7)) \cdot \alpha_j^2 = 0. \tag{7.3}$$

Finally $w_1(y)$ is a unit vector and hence

$$\sum_{j=0}^{3} \alpha_j^2 = 1. \tag{7.4}$$

Thus we have obtained a system of four equations in four unknowns which turns out to have a unique meaningful solution. Let us substitute in the equations the values of $\pi_j(i)$ computed from the character table given after Lemma 4.3.

$$\alpha_0^2 - (1/2)\alpha_1^2 + (7/22)\alpha_2^2 - (5/22)\alpha_3^2 = \alpha_0, \tag{7.5}$$

$$\alpha_0^2 - (5/16)\alpha_1^2 - (11/64)\alpha_2^2 + (5/32)\alpha_3^2 = \alpha_0, \tag{7.6}$$

$$(3/16)\alpha_1^2 - (75/704)\alpha_2^2 - (27/352)\alpha_3^2 = 0, \tag{7.7}$$

$$\alpha_0^2 + \alpha_1^2 + \alpha_2^2 + \alpha_3^2 = 1. \tag{7.8}$$

Comparing the left sides of (7.5) and (7.6), and using (7.7) and (7.8), we obtain a system of three linear equations in four unknowns α_j^2, $0 \le j \le 3$, from which we deduce the following expressions:

$$\alpha_1^2 = (1 - \alpha_0^2)(60/181), \tag{7.9}$$

$$\alpha_2^2 = (1 - \alpha_0^2)(66/181), \tag{7.10}$$

$$\alpha_3^2 = (1 - \alpha_0^2)(55/181). \tag{7.11}$$

Substituting these expressions in (7.5) we obtain a quadratic equation on α_0. There is an obvious and meaningless solution of this equation, namely $\alpha_0 = 1$ (which imply $w_1(z) = w_1(x)$ for all $z \in \Gamma(x)$, definitely not the case). The second solution is $\alpha_0 = -43/405$.

As soon as we know α_0, the expressions (7.9)–(7.11) give us α_1^2, α_2^2 and α_3^2 and using the values $\pi_j(i)$ computed from the character table of F acting on $\Gamma(x)$, we determine the inner products $\langle w_1(y) \mid w_1(z_i)\rangle$ for all $0 \le i \le 7$.

Lemma 7.2. *Let $z_i \in \Omega_i$, $0 \le i \le 7$. Then the inner product $\langle w_1(y) \mid w_1(z_i)\rangle$ equals to* 1, $-43/405$, $-43/405$, $69/405$, $13/405$, $-1/405$, $-15/405$ *and* $13/405$, *respectively.*

§8. The subdegrees

In this section we determine the subdegrees and 2-point stabilizers of E acting on X. We follow the notation introduced above.

By Lemma 4.2 there are 7 orbits of E on the paths of length 2 in Γ with representatives (y, x, z_i) for $1 \le i \le 7$. By Corollaries 6.2 and 6.7 z_1 and z_2 are adjacent to y. By Lemma 6.10 z_4 and z_7 are in the same $E(y)$-orbit. On the other hand by Lemma 7.2 for $i = 3$, 4, 5 and 6 the inner products $\langle w_1(y) \mid w_1(z_i)\rangle$ are pairwise different and none of then is equal to the inner product $-43/405$ of vectors representing adjacent vertices. So we have the following.

Lemma 8.1. *The subgroup $F = E(x)$ has four orbits on vertices at distance* 2 *from x in Γ.*

Let $\Gamma_0(x) = \{x\}$, $\Gamma_1(x) = \Gamma(x)$ and let $\Gamma_2(x)$, $\Gamma_3(x)$, $\Gamma_4(x)$ and $\Gamma_5(x)$ be the orbits of F on vertices at distance 2 from x in Γ. We will assume that $\Gamma(y, x; 3696) \subseteq \Gamma_2(x)$; $\Gamma(y, x; 4928) \cup \Gamma(y, x; 78848) \subseteq \Gamma_3(x)$; $\Gamma(y, x; 42240) \subseteq \Gamma_4(x)$ and $\Gamma(y, x; 11264) \subseteq \Gamma_5(x)$. Then by Corollary 6.9 $\Gamma_2(x)$ is as defined in Section 6, in particular $|\Gamma_2(x)| = 694\,980$. Since the rank of E on X is 8, there are 2 orbits of F on the set of vertices at distance more than 2 from x in Γ. We denote these orbits by $\Gamma_6(x)$ and $\Gamma_7(x)$.

Let us choose a family of representatives $y_i \in \Gamma_i(x)$, $0 \le i \le 7$. We are going to introduce for every i from 0 to 7 a subgroup K_i in $E(x, y_i)$. In some cases it will be clear from the very beginning that $K_i = E(x, y_i)$. For the remaining cases this equality will come at the end because of the equality

$$[E : E(x)] = \sum_{i=0}^{7} [E(x) : K_i]. \tag{8.1}$$

Clearly, we take $K_0 = E(x) \cong Fi_{22}$. Because of Lemma 5.2 we can take $K_1 = E(x, y_1) \cong 2^{10} : M_{22}$ and $K_2 = E(x, y_2) \cong 2^6 : Sp_6(2)$.

Consider $E(x, y_3)$. We assume that $y_3 \in \Gamma(y, x; 4928)$ and hence $E(x, y, y_3) \cong 2^4.2^4.A_6$ (compare Lemma 4.2). Let d be an element of type $5A$ from $E(x, y, y_3)$ and $D = \langle d \rangle$. This means that $x, y, y_3 \in \Gamma(5A)$. Let Θ be the orbit of y under $E(x, y_3)$. Clearly $|E(x, y_3)| = |\Theta| \times |E(x, y, y_3)|$. By the remark after Lemma 6.10, the intersection Θ' of Θ and $\Gamma(5A)$ is of size 4 and $N_{E(x,y_3)}(D)$ acts transitively on Θ'. We define K_3 to be the setwise stabilizer of Θ' in $E(x, y_3)$. We can write that $K_3 \cong 2^4.2^4.A_6.2.2$. The precise structure of K_3 will be established later.

Let us turn to $E(x, y_4)$ and start by calculating the normalizer in E of $Q = O_2(E(x, y_2)) \cong 2^6$.

Lemma 8.2. *Let $P = N_E(Q)$. Then $P/Q \cong [2^9].Sp_6(2)$.*

Proof. We claim that there is a unique central involution τ in E such that $Q \leq O_2(C_E(\tau))$ (the uniqueness will immediately imply that $P \leq C_E(\tau)$). Consider the action of F on the set of central involutions in E. We will follow the notation of Lemma 5.1. Let F_i denote the centralizer in F of $\tau_i \in \Delta_i$ for $1 \leq i \leq 4$. Suppose that $Q \leq O_2(F_i)$. Certainly $i \neq 1$. Since $O_2(F_2)$ and $O_2(F_3)$ are abelian and Q is self-centralized in F, $i \neq 2, 3$. Of course we can make $Q = O_2(F_4)$, but then τ_4 becomes uniquely determined since $F_4 = N_F(Q)$. Thus $Q \leq C_E(\tau_4)$. Since $U_6(2) \cong C_E(\tau_4)/O_2(C_E(\tau_4))$ does not contain a 2-local subgroup with a section $Sp_6(2)$, we conclude that $Q \leq O_2(C_E(\tau_4))$ and the claim follows.

Let S be the image of P in Out $Q \cong L_6(2)$. Then S contains the image of $N_F(Q)$ isomorphic to $Sp_6(2)$. Since $Sp_6(2)$ is maximal in $L_6(2)$ and the latter is not involved in $C_E(\tau)$, $S \cong Sp_6(2)$. So we only have to consider $D = C_E(Q)$. Since every subgroup of $U_6(2)$ having $Sp_6(2)$ as a factor group is $Sp_6(2)$ itself, $D \leq O_2(C_E(\tau))$ and lemma follows from the basic properties of extraspecial groups. Q.E.D.

Corollary 8.3. *Let Σ be the orbit of x under $N_E(Q)$. Then $|\Sigma| =$ 512.*

Clearly Q fixes Σ elementwise. Let us locate some elements of Σ. The setwise stabilizer of $\{x, y_2\}$ clearly normalizes Q and hence $y_2 \in \Sigma$. By Lemma 4.5 $E(x, y_2) = H_5$ has an orbit Ψ of length 135 on $\Gamma(x)$. Since $Q = O_2(E(x, y_2))$, Q fixes Ψ elementwise. Without loss of generality we assume that $y \in \Psi$, which means that $Q \leq E(x, y)$. Let τ be the central involution in E, such that $\langle \tau \rangle \times E(x, y)$ is the setwise stabilizer of $\{x, y\}$.

Then τ normalizes (even commutes with) Q which means that $\Psi \subseteq \Sigma$. The group $E(x, y_2)$ acts on Ψ as $Sp_6(2) \cong E(x, y_2)/Q$ acts on the set of points of the symplectic dual polar space. In particular $E(x, y, y_2) \cong 2^6.2^6.L_3(2)$ acting on Ψ has 4 orbits Ψ_0, Ψ_1, Ψ_2 and Ψ_3 with lengths 1, 14, 56 and 64, respectively. Now one can see (compare Lemma 4.2) that $\Psi_1 \subseteq \Gamma(x, y; 154)$, $\Psi_2 \subseteq \Gamma(x, y; 3696)$ and $\Psi_3 \subseteq \Gamma(x, y; 42240)$. This means that for $i = 2$ and 4 the set $\Sigma \cap \Gamma_i(x)$ contains more than one vertex. By Lemma 4.4 and Corollary 8.3 we have $|\Sigma \cap \Gamma_2(x)| \geq 136$ and hence $|\Sigma \cap \Gamma_4(x)| \leq 240$. Assuming that y_4 is in $\Sigma \cap \Gamma_4(x)$, we define K_4 to be the normalizer of Q in $E(x, y_4)$. By the above K_4 is a subgroup in $K_2 = N_{E(x)}(Q) \cong 2^6 : Sp_6(2)$, which contains Q and the index $[K_2 : K_4]$ is at most 240.

By Lemmas 4.2 and 6.6 one can see that the subgroup $E(x, y, y_5) \cong 2.L_3(4)$ contains a unique transposition τ from $E(x)$, and this transposition is in the center of $E(x, y, y_5)$. We define K_5 to be $C_{E(x,y_5)}(\tau)$. It follows from Lemmas 4.2, 6.6 and the intersection matrix of $\Gamma(2A)$ that $K_5 \cong 2.L_3(4).2$.

Now we are going to deduce some information on stabilizers in $E(x)$ of vertices at distance more than 2 from x in Γ. By Lemma 5.3 there is a point $z \in X - \{x\}$ whose stabilizer in $F = E(x)$ is $H_8 \cong {}^2F_4(2)'$. By Lemma 4.2 every pair of vertices from $\Gamma(x)$ is stabilized by an element of type $3c$ from $F = E(x)$. On the other hand the permutation character of F on the cosets of H_8 given in [ILLSS] is zero on elements of type $3c$. Hence z is at distance more than 2 from x in Γ. Without loss of generality we assume that $z = y_6$ and put $K_6 = E(x, y_6) \cong {}^2F_4(2)'$.

Let d be an element of type $3b$ in F, $D = \langle d \rangle$, so that $N_F(D) = H_{10}$. The character $\mathbf{1}_F^E$ shows that d is of type $3C$ in E and $M(3C)$ has two orbits on $\Phi(3C)$, say $\Phi_1(3C)$ and $\Phi_2(3C)$ with lengths 4 and 576, respectively. We assume that $x \in \Phi_1(3C)$. Let S be a Sylow 3-subgroup of F contained in $N_F(D)$. Since $|\Phi(9A)| = 1$, S acts non-trivially on $\Phi_1(3C)$ and hence the action of $M(3C)$ on this set is doubly transitive. Let $u \in \Phi_1(3C) - \{x\}$. Then $K = N_F(D) \cap E(x, u)$ is an index 3 subgroup in H_{10}. The list of maximal subgroups of F shows that every subgroup L, such that $K < L \leq F$ contains a Sylow 3-subgroup of F and since $|\Phi(9A)| = 1$, L can not be a 2-point stabilizer. Hence $E(x, u) = K$ and it is a $\{2, 3\}$-group having index 3 in H_{10}. Notice that for $i \leq 6$ the order of $E(x, y_i)$ is divisible by a prime greater than 3. So we can assume that $u = y_7$ and put $K_7 = K$.

Now by direct calculations we see that $\sum_{i=0}^{7} [E(x) : K_i]$ is less than $[E : F]$ unless K_4 has index exactly 240 in K_2. In the latter case the equality (8.1) holds and we conclude the following.

Lemma 8.4. *K_4 has index 240 in K_2 and $E(x, y_i) = K_i$ for every $0 \le i \le 7$.*

So we have calculated the subdegrees of the action of E on X and proceed to determination of the precise structure of the 2-point stabilizers $E(x, y_i)$. For $i = 0$, 1, 2, 5 and 6 everything is clear.

The group $Sp_6(2)$ contains a unique conjugacy class of subgroups of index 240 (cf. [Atlas]) and we obtain the following.

Lemma 8.5. $E(x, y_4) \cong 2^6 : U_3(3)$.

Let us turn to $E(x, y_3)$. Let $\Theta = \Theta(x, y_3) = \{z \mid z \in \Gamma(x), y_3 \in \Gamma(z, x; 4928)\}$. Clearly Θ is the same as defined before Lemma 8.2. By now we have the following information on Θ (cf. Lemmas 4.2 and 6.4).

Lemma 8.6. *The set Θ is of size 4, the normalizer in $E(x, y_3)$ of a Sylow 5-subgroup induces on Θ the regular action of Z_4 and for $z \in \Theta$ we have $E(x, y_3, z) \cong 2^4.2^4.A_6$.*

Clearly $E(x, y_3)$ is a subgroup in the setwise stabilizer of Θ in $E(x)$. Let us specify Θ and its stabilizer in $E(x)$.

Let Ω be the graph of valency 154 on $\Gamma(x)$ in which $z \in \Gamma(x)$ is adjacent to the vertices from $\Gamma(x, z; 154)$. The intersection matrix of Ω is given after Lemma 4.3. Let $z \in \Gamma(x, y; 4928)$. Then $z \in \Gamma_3(y)$ and $x \in \Theta(y, z)$. The intersection matrix of Ω shows that there is a unique vertex w, adjacent to both y and z in Ω. Since $E(y, x, z) = E(y, x, z, w) \le E(y, z, w)$, it is clear that $w \in \Theta(y, z)$ and we obtain the following.

Lemma 8.7. *The subgraph of Ω induced by Θ is of valency* 1.

Let $\Theta = \{z_1, z_2, z_3, z_4\}$ and assume that $\{z_1, z_2\}$ and $\{z_3, z_4\}$ are edges of Ω. Let z_0 be the unique vertex adjacent in Ω to both z_1 and z_2. Put $L = E(x, z_1, z_2) = E(x, z_0, z_1, z_2)$ and let $\Lambda(z_i)$ be the orbit of z_i under L, $i = 3$ or 4. Since $L \cong 2^9.2^4.A_6$ and $E(x, y_3, z_1) \cong 2^4.2^4.A_6$ is contained in the elementwise stabilizer of Θ, we conclude that the length of $\Lambda(z_i)$ divides 2^5.

Let S be the Steiner system $S(3, 6, 22)$ naturally associated with $E(x, z_1) \cong 2^{10} : M_{22}$. Then $\{z_0, z_2\}$ is an orbit of $O_2(E(x, z_1))$ which corresponds to a block B of S. Let Λ_1 be the union of all orbits of $O_2(E(x, z_1))$ on $\Gamma(x, z_1; 154)$ which correspond to S-blocks disjoint from B. Then Λ_1 is an orbit of L of length 2^5. For $i = 0$ and 2 let Λ_i denote the similar orbit of L in $\Gamma(x, z_i; 154)$. Now from Lemma 4.2 and basic properties of the Steiner system S it is not difficult to show the following.

Lemma 8.8. *Every orbit of L on $\Gamma(x) - \{z_0, z_1, z_2\}$ whose length divides 2^5 coincides with Λ_i for $i = 0$, 1 or 2.*

The intersection matrix of Ω shows that Λ_i and Λ_j are disjoint for $i \neq j$. Hence $\Lambda(z_3) = \Lambda(z_4) = \Lambda_i$ for some $0 \leq i \leq 2$. Since Θ contains only two Ω-edges (cf. Lemma 8.7), i must be equal to 0. Remembering that the setwise stabilizer in M_{22} of a pair of disjoint blocks of S is isomorphic to M_{10}, we obtain the following.

Lemma 8.9. *There is a unique vertex $z_0 \in \Gamma(x)$ such that $\Theta \subseteq \Gamma(x, z_0; 154)$. The pairs $\{z_1, z_2\}$ and $\{z_3, z_4\}$ are orbits of $O_2(E(x, z_0))$ and correspond to disjoint blocks of the Steiner system associated with $E(x, z_0)$. The setwise stabilizer of Θ in $E(x)$ is the full preimage in $E(x, z_0)$ of an M_{10}-subgroup from $\bar{E}(x, z_0) \cong M_{22}$.*

Thus the setwise stabilizer R of Θ in $E(x)$ contains $E(x, y_3)$ as a subgroup of index 2. Since R induces D_8 on Θ while $E(x, y_3)$ induces Z_4, $E(x, y_3)$ is uniquely specified.

Lemma 8.10. *$E(x, y_3) \cong 2^4.2^4.A_6.Z_4 \cong 2^9.M_{10}$ is a subgroup in a conjugate of $E(x, y_1)$.*

In order to identify $E(x, y_7)$ we need the size of the intersection of $\Gamma(3B)$ and $\Gamma_7(x)$. We prove a more general proposition which will be used later.

Lemma 8.11. *Let d be an element of type cl from E contained in $E(x) = F$, $D = \langle d \rangle$ and let $\Gamma(cl)$ be the subgraph of Γ induced by the vertices fixed by d. Then the orbits of $N_{E(x)}(D)$ on $\Gamma(cl)$ (ordered as the rows of the intersection matrix of $\Gamma(cl)$ given in Section 6) are contained in*

(i) $\Gamma_0(x)$, $\Gamma_1(x)$, $\Gamma_5(x)$, $\Gamma_4(x)$ *and* $\Gamma_2(x)$ *for* $cl = 7B$;
(ii) $\Gamma_0(x)$, $\Gamma_1(x)$, $\Gamma_3(x)$, $\Gamma_5(x)$ *and* $\Gamma_6(x)$ *for* $cl = 5A$;
(iii) $\Gamma_0(x)$, $\Gamma_1(x)$, $\Gamma_3(x)$, $\Gamma_4(x)$, $\Gamma_5(x)$, $\Gamma_7(x)$ *and* $\Gamma_7(x)$ *for* $cl = 3B$;
(iv) $\Gamma_0(x)$, $\Gamma_1(x)$, $\Gamma_2(x)$, $\Gamma_5(x)$, $\Gamma_7(x)$ *and* $\Gamma_3(x)$ *for* $cl = 2A$.

Proof. Using the permutation characters of $E(x) = F$ on the cosets of $H = E(x, y_1)$, $H_5 = H(x, y_2)$, $H_8 = H(x, y_6)$ (cf. [ILLSS]) and the embeddings $E(x, y_3) < E(x, y_1)$, $E(x, y_4) < E(x, y_2)$, we calculate the number of points in $\Gamma_i(x)$ fixed by an element of type cl contained in F for $i \leq 6$. In the case (iii) there are two orbits of length 3 888. Since there are two paths of length 2 joining x and y_5, we have a unique way to locate these orbits and the result follows. Q.E.D.

Lemma 8.12. *$E(x, y_7)$ is an index 3 subgroup in the maximal subgroup H_{10} of $F = E(x)$. In the permutation action of F on the cosets of $E(x, y_7)$ an element of type $3C$ fixes 912 points. These two properties characterize $E(x, y_7)$ up to conjugacy in F.*

Proof. The first sentence follows from the definition of $E(x, y_7)$. The second one comes from Lemma 8.11 (iii), so all we have to prove is the last sentence. The GAP system [GAP] contains the character table of H_{10} along with the fusion pattern of the H_{10}-classes into F-classes. There is one non-principal H_{10}-character of degree 1 and four characters of degree 2 (with pairwise different kernels). So we have five candidates for the permutation character of H_{10} of degree 3 (on the cosets of $E(x, y_7)$). Inducing these five characters to F we observe that the one involving irreducibles of degree 1 only, does not lead to a permutation character at all (the induced character involves negative values). Among the other four, only one (denote it by ϕ) when induced gives the character value 912 on $3C$-elements. It was checked that the kernel of ϕ is a normal subgroup of index 6 in H_{10} and the result follows.
Q.E.D.

We can summarize this section by the following.

Proposition 8.13. *The 2-point stabilizers of $E \cong {}^2E_6(2)$ acting on the cosets of $F \cong Fi_{22}$ are Fi_{22}, $2^{10} : M_{22}$, $2^6 : Sp_6(2)$, $2^9.M_{10}$, $2^6 : U_3(3)$, $2.L_3(4).2$, ${}^2F_4(2)'$ and $3_+^{1+6} : 2^{3+4} : 3 : 2$.*

§9. The eigenvector

Lemma 7.2 gives six entries $p_i(1)/k_i$, $0 \le i \le 5$ in the left eigenvector of the intersection matrix of Γ corresponding to the idempotent of rank 1938. In this section we calculate the missing two entries. We will use for this purpose the character table of the centralizer algebra corresponding to the action of $M(2A)$ on $\Phi(2A)$ given after Lemma 6.6.

For the rest of the section we assume that $\Gamma(2A)$ contains y_i for $i = 0$, 1, 2, 3, 5 and 7 (cf. Lemma 8.11 (iv)). Let $C = C_E(d) = M(2A)$.

Lemma 9.1. *The 1938-dimensional E-module V_1 decomposes into C-irreducibles as $W_0 \oplus W_1 \oplus W_2 \oplus W_3$ where the W_j have dimensions 1, 22, 891 and 1024, respectively.*

Proof. Since $Q = O_2(C)$ is extraspecial of order 2^{21}, V_1 restricted to Q must involve the faithful component of dimension 1024 and only one such component fits. Since the central involution d of C is conjugate in E to an involution from $Q - \langle d \rangle$, there must be an irreducible component whose kernel is exactly $\langle d \rangle$. Such a component has dimension equal to the length of an orbit of C on the non-zero vectors of the module (dual to) $Q/\langle d \rangle$. Only one such component of dimension 891 fits. For the remaining components Q is in the kernel and the conclusion comes from the character table of $U_6(2) \cong C/Q$ in [Atlas]. Q.E.D.

As above for $z \in \Gamma$ let $w_1(z)$ be the unit vector in V_1 realizing z. Suppose that $z \in \Gamma(2A)$. The projection of $w_1(z)$ into W_j can be non-zero only if $C \cap E(z) \cong 2 \cdot U_6(2)$ fixes a vector in W_j. Such a vector is fixed if and only if W_j is involved in the permutation character of $C = M(2A)$ acting on the vertices of $\Gamma(2A)$. The degrees of the irreducible constituents in the permutation character can be read from the character table of the corresponding centralizer algebra given after Lemma 6.6 and we observe that only W_0 and W_2 can be involved. On the other hand W_0 supports the principal character and supports W_2 the unique faithful character of degree 891, so both W_0 and W_2 are involved.

For $j = 0$ or 2 and $z \in \Gamma(2A)$ let $s_j(z)$ be the unit vector in W_j realizing z with respect to the centralizer algebra of C acting on $\Gamma(2A)$, clearly $s_j(z)$ is fixed by $E(z) \cap C$. Then

$$w_1(z) = s_0(z) \cdot \beta_0 + s_2(z) \cdot \beta_2 \tag{9.1}$$

for some β_0 and β_2 independent of z and satisfying

$$\beta_0^2 + \beta_2^2 = 1. \tag{9.2}$$

Then Lemmas 7.2, 8.11 (iv) and the character table given after Lemma 6.6 give us the following (we assume that $x, y_1 \in \Gamma(2A)$):

$$-43/405 = \langle w_1(x) \mid w_1(y_1) \rangle = \beta_0^2 + (-133/891)\beta_2^2. \tag{9.3}$$

Now (9.2) and (9.3) imply $\beta_0^2 = 3/80$, $\beta_2^2 = 77/80$ and using the character table of C acting on $\Gamma(2A)$ we can compute the inner product $\langle w_1(u) \mid w_1(v) \rangle$ for every pair $u, v \in \Gamma(2A)$. In particular

$$p_7(1)/k_7 = \langle w_1(x) \mid w_1(y_7) \rangle = 1/15.$$

In order to compute the last unknown entry in the eigenvector, namely $p_6(1)/k_6$ we use the first orthogonality relation [BI]. This relation applied to the vector we are studying and the one corresponding to the trivial idempotent of rank 1 gives

$$\sum_{i=0}^{7} p_i(1) = 0.$$

This immediately implies $p_6(1)/k_6 = -1/27$. Thus we have proved the following.

Proposition 9.2. *The eigenvector of the intersection matrix of* Γ *corresponding to the idempotent of rank* 1938 *is the following*: $(1, -43/405, 69/405, 13/405, -15/405, -1/405, -15/405, 27/405)$.

§10. The intersection matrix

In this section we calculate the intersection matrix B_1 of the graph Γ. By the definition the (j,i)-entry p^i_{1j} of B_1 is the number of vertices in $\Gamma_j(x)$ adjacent to y_i in Γ. Since all orbitals of E acting on Γ are symmetrical, p^i_{1j} is also the number of vertices in $\Gamma(x)$ contained in $\Gamma_j(y_i)$. So the entries of the intersection matrix are the sizes of parts in the partitions

$$\Gamma(x) = \bigcup_{j=0}^{7} (\Gamma(x) \cap \Gamma_j(y_i)), \qquad 0 \le i \le 7. \tag{10.1}$$

It is clear that for every i and j the set $\Gamma(x) \cup \Gamma_j(y_i)$ is a union of orbits of $E(x, y_i)$ on $\Gamma(x)$. So the partitions of $\Gamma(x)$ into $E(x, y_i)$-orbits are refinements of the partitions (10.1). In this context it is quite helpful (and crucial for our approach) to know the lengths of $E(x, y_i)$-orbits on $\Gamma(x)$. For $i = 1$ the information is contained in Lemma 4.2. For $i = 6$ the orbit lengths follow from the inner product of the permutation characters and elementary congruences. For the remainder i the information is much more non-trivial and was obtained by S. A. Linton using explicit calculations in the Fischer group Fi_{22}. The result is contained in Lemma 4.5 for $i = 2$ and in the following lemma for $3 \le i \le 7$.

Lemma 10.1. *The orbits lengths of* $E(x, y_i)$, $3 \le i \le 7$ *on* $\Gamma(x)$ *are the following.*

(i) *The subgroup* $E(x, y_3) \cong 2^9.M_{10}$ *acting on* $\Gamma(x)$ *has* 26 *orbits with lengths* 1, 4, 60, 64^2, 90, 480, 576, 720, $1\,024^3$, $1\,920^4$, $2\,560^2$, 2 880, 3 840, 6 144, 11 520, $15\,360^2$, 23 040, *and* 46 080.

(ii) *The subgroup* $E(x, y_4) \cong 2^6 : U_3(3)$ *acting on* $\Gamma(x)$ *has* 25 *orbits with lengths* 36^2, 63, 126^2, 288^2, 504, $2\,016^4$, $2\,304^2$, 3 024, $4\,032^3$, $8\,064^2$, $16\,128^4$ *and* 32 256.

(iii) *The subgroup* $E(x, y_5) \cong 2.L_3(4).2$ *acting on* $\Gamma(x)$ *has* 35 *orbits with lengths* 2, 42^2, 105, 112, 224^2, 280, 420, 480, 504, 840^3, 1 120, $1\,344^2$, $2\,520^2$, $3\,360^6$, 4 032, $5\,040^2$, 6 720, $10\,080^2$, $13\,440^2$ *and* $20\,160^2$.

(iv) *The subgroup* $E(x, y_6) \cong {}^2F_4(2)'$ *acting on* $\Gamma(x)$ *has* 3 *orbits with lengths* 1 755, 28 080 *and* 112 320.

(v) *The subgroup* $E(x, y_7) \cong 3_+^{1+6}.2^{3+4} : 3 : 2$ *acting on* $\Gamma(x)$ *has* 6 *orbits with lengths* 2 187, 5 832, 11 664, 17 496, 34 992 *and* 69 984.

The following lemma is obvious.

Lemma 10.2. *Suppose that* $E(x, y_i)$ *acting on* $\Gamma(x) \cap \Gamma_j(y_i)$ *has an orbit of length* l. *Then* $E(x, y_j)$ *acting on* $\Gamma(x) \cap \Gamma_i(y_j)$ *has an orbit of length* $l \cdot k_i/k_j$.

In order to simplify our terminology we present a matrix (denoted by D) and will prove below that it is equal to the intersection matrix B_1 of Γ. First of all, it is straightforward to check that the vector in Proposition 9.2 is a left eigenvector of D.

$$D = \begin{array}{rrrrrrrrr}
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
142155 & 1178 & 756 & 68 & 36 & 2 & 0 & 0 & 142155 \\
0 & 3696 & 135 & 60 & 324 & 42 & 0 & 0 & 694980 \\
0 & 83776 & 15120 & 17899 & 27720 & 20692 & 28080 & 11664 & 175134960 \\
0 & 42240 & 77760 & 26400 & 19035 & 18120 & 0 & 34992 & 166795200 \\
0 & 11264 & 48384 & 94592 & 86976 & 98315 & 112320 & 93312 & 800616960 \\
0 & 0 & 0 & 576 & 0 & 504 & 1755 & 0 & 3592512 \\
0 & 0 & 0 & 2560 & 8064 & 4480 & 0 & 2187 & 38438400
\end{array}$$

Clearly the first row and column in D are as in B_1.

Lemmas 8.1, 8.4 and straightforward calculations imply the following.

Lemma 10.3. *The second row and column in* D *are as in* B_1.

Lemma 10.4. *The third row and column of* D *are as in* B_1.

Proof. The lengths of orbits of $E(x, y_2)$ on $\Gamma(y_2)$ are given in Lemma 4.5. By Lemma 10.3 the orbit of length 756 is contained in $\Gamma(x)$. The intersection matrix of $\Gamma(7B)$ and Lemma 8.11 (i) show that $p_{12}^2 \neq 0$. Clearly the length of an orbit of $E(x, y_2)$ on $\Gamma(y_2) \cap \Gamma_2(x)$ must be equal to a subdegree of $E(x) = F$ acting on the cosets of $E(x, y_2) = H_5$. By Lemmas 4.4 and 4.5 this implies that $p_{12}^2 = 135$. Now the intersection matrix of $\Gamma(7B)$ and Lemma 8.11 (i) imply that each of the remaining orbits of $E(x, y_2)$ on $\Gamma(y_2)$ whose length is not divisible by 7 must be in $\Gamma_4(x)$. There are two such orbits with lengths 8 640 and 69 120. By now there are two orbits left and from the intersection matrix of $\Gamma(2A)$ and Lemma 8.11 (iv) it follows that both p_{13}^2 and p_{15}^2 are non-zero. The divisibility condition forces the orbit of length 15 120 to be in $\Gamma_3(x)$ and so the one of length 48 384 is in $\Gamma_5(x)$. Q.E.D.

Lemma 10.5. *The seventh row and column of D are as in B_1.*

Proof. By Lemma 10.1 (iv) there are at most three indices i for which $p^6_{1i} \neq 0$. On the other hand the intersection matrix of $\Gamma(5A)$ and Lemma 8.11 (ii) show that for $i \in \{3,5,6\}$ these parameters must be non-zero. Now the result follows from straightforward calculations using Lemmas 10.2 and 10.1 (i), (iii), (iv). Q.E.D.

Before proceeding to the next case we prove the following easy lemma.

Lemma 10.6. *Let $E(x, y_5) \cong 2.L_3(4).2$ act transitively on a set Θ of size 280, 480, 840, 3 360 or 6 720. Let σ be a subgroup of order 3 in $E(x, y_5)$. Then the number of elements fixed by σ in Θ is 1, 12, 2^a3 for $a \leq 2$, 2^b3 for $b \leq 2$ or 2^c3 for $c \leq 3$, respectively.*

Proof. The list of maximal subgroups of $L_3(4)$ [Atlas] shows that the action of degree 280 is unfaithful primitive with character value 1 on 3-elements while the one of degree 480 is on the cosets of $L_3(2)$ with the character value 12 on 3-elements. The normalizer N of σ in $E(x, y_5)$ has order 72. Let $a \in \Theta$ be an element fixed by σ. Then σ is a Sylow 3-subgroup of $E(x, y_5, a)$ and by Frattini argument N acts transitively on the set of all elements from Θ fixed by σ. Now the conclusion comes from an easy observation that in a group of order 96 or 24 the normalizer of a 3-subgroup has order at least 6. Q.E.D.

Lemma 10.7. *The eighth row and column in D are as in B_1.*

Proof. The intersection matrix of $\Gamma(3B)$ and Lemma 8.11 (iii) show that for $i \in \{3,4,5,7\}$ the number p^7_{1i} is non-zero. The lengths of orbits of $E(x, y_7)$ on $\Gamma(y_7)$ are given in Lemma 10.1 (v). Lemmas 10.2, 10.1 (iii) show that only three orbits with lengths 5 832, 17 496 and 69 984 might be contained in $\Gamma_5(x)$. These orbits then would correspond to orbits of $E(x, y_5)$ on $\Gamma_7(x)$ with lengths 280, 480 and 3 360, respectively. By Lemma 10.6 an element of order 3 from $E(x, y_5)$ fixes in the union of the latter three orbits at most 1+12+12=25 elements. On the other hand it follows from the intersection matrix of $\Gamma(3B)$ and Lemma 8.11 (iii) that such an element fixes exactly 1+24=25 vertices in $\Gamma(y_5) \cap \Gamma_7(x)$ so all the above three orbits of $E(x, y_7)$ are in $\Gamma_5(x)$. Applying Lemmas 10.1 and 10.2 to the remaining orbits we see that $p^7_{13} = 11\,664$, $p^7_{14} = 34\,992$ and finally $p^7_{17} = 2\,187$. Q.E.D.

It only remains to evaluate the submatrix in B_1 consisting of the elements p^i_{1j} with $3 \leq i, j \leq 5$. We show first that with the information available it is sufficient to determine just one entry in this submatrix, say p^5_{14}.

Lemma 10.8. *Suppose that p^5_{14} is equal to the corresponding entry in D. Then D coincides with B_1.*

Proof. Let $C = B_1 - D$ and let the (i,j)-entry of C be denoted by c_{ij}. By Lemmas 10.3, 10.4, 10.5 and 10.7 $c_{ij} = 0$ unless $3 \leq i \leq 5$ and $3 \leq j \leq 5$. Since both B_1 and D have constant column sum (equal to $k = 142\,155$),

$$c_{3j} + c_{4j} + c_{5j} = 0, \qquad 3 \leq j \leq 5. \tag{10.2}$$

The vector given in Proposition 9.2 is a left eigenvector for both B_1 and D. Hence C applied to this vector gives zero and we have the following:

$$13c_{3j} - 15c_{4j} - c_{5j} = 0, \qquad 3 \leq j \leq 5. \tag{10.3}$$

Now (10.2) and (10.3) imply that $c_{3j} = c_{4j}$, $c_{5j} = -2c_{4j}$. Finally, $k_i p^i_{1j} = k_j p^j_{1i}$ and similar relations hold for the entries of D, hence $k_i c_{ji} = k_j c_{ij}$ for $0 \leq i,j \leq 7$ and the result follows. Q.E.D.

Let us proceed with determination of p^5_{14}. The center of $E(x, y_5)$ is of order 2 and it is generated by a $2A$-involution. Hence we conclude (assuming that $y_5 \in \Gamma(2A)$) that $\Gamma(y_5) \cap \Gamma_j(x) \cap \Gamma(2A)$ is a union of orbits of $E(x, y_5)$ on $\Gamma(y_5)$ for every $0 \leq j \leq 7$. This observation along with the intersection matrix of $\Gamma(2A)$ and Lemmas 10.1 (iii) and 8.11 (iv) enable us to locate some further orbits of $E(x, y_5)$ on $\Gamma(y_5)$.

Lemma 10.9. *The orbits of $E(x, y_5)$ on $\Gamma(y_5)$ with lengths* 2, 42, 42, 105, 280 *and* 420 *are contained in $\Gamma_1(x)$, $\Gamma_2(x)$, $\Gamma_5(x)$, $\Gamma_5(x)$, $\Gamma_7(x)$ and $\Gamma_3(x)$, respectively.*

The intersection matrix of $\Gamma(7B)$, Lemmas 8.11 (i) and 10.1 (iii) give us.

Lemma 10.10. *The orbit of length* 480 *of $E(x, y_5)$ on $\Gamma(y_5)$ is contained in $\Gamma_4(x)$.*

It follows from the proof of Lemma 10.7 that one orbit of length 840 of $E(x, y_5)$ on $\Gamma(y_5)$ is contained in $\Gamma_7(x)$. This together with Lemmas 10.9, 10.10, 10.1 (ii), (iii) and 10.2 give us the following.

Lemma 10.11. *$E(x, y_5)$ acting on $\Gamma(y_5) \cap \Gamma_4(x)$ has one orbit of length* 480, *at most two orbits of length* 840, *at most four orbits of length* 3 360 *and at most one orbit of length* 6 720.

Let σ be a subgroup of order 3 in $E(x, y_5)$. The intersection matrix of $\Gamma(3B)$ and Lemma 8.11 (iii) show that σ fixes 84 vertices in $\Gamma(y_5) \cap \Gamma_4(x)$. On the other hand Lemma 10.6 says that σ fixes 12 vertices in the orbit of length 480 and gives upper bounds on the numbers of vertices fixed by σ in orbits of lengths 840, 3360 and 6720. These bounds imply the following.

Lemma 10.12. *$p^5_{14} = 480 + 840 \cdot \gamma$ where $\gamma = 18$, 21, 22, 24, 25 or 26.*

The entry in D corresponds to $\gamma = 21$. We have checked using the computer program of D. V. Pasechnik, mentioned in the introduction that each of the other five possibilities for γ allowed by Lemma 10.12 leads to a matrix with non-integral spectrum. Since the multiplicities of the centralizer algebra are pairwise distinct, such a matrix can not possibly be the intersection matrix of Γ and we have.

Proposition 10.13. *D is the intersection matrix of Γ.*

We conclude the paper by presenting the intersection matrix B_2 corresponding to the second smallest non-trivial valency 694 980 and the character table of the centralizer algebra, both computed using the program of D. V. Pasechnik. Directly from the shapes of $B_1 = D$ and B_2 (given below) we see that the action is not distance-transitive (as was earlier proved in [CLS]).

$$B_2 = \begin{array}{rrrrrrrrr}
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
694980 & 13644 & 2592 & 1140 & 660 & 210 & 36 & 0 & 694980 \\
0 & 143360 & 62532 & 38400 & 0 & 20160 & 8064 & 0 & 38438400 \\
0 & 287280 & 174960 & 142644 & 73920 & 97860 & 68040 & 28080 & 175134960 \\
0 & 135 & 0 & 60 & 3696 & 42 & 324 & 0 & 142155 \\
0 & 241920 & 419904 & 447360 & 236544 & 478788 & 457920 & 561600 & 800616960 \\
0 & 8640 & 34992 & 64800 & 380160 & 95400 & 158580 & 93600 & 166795200 \\
0 & 0 & 0 & 576 & 0 & 2520 & 2016 & 11700 & 3592512
\end{array}$$

The character table of ${}^2E_6(2)$ acting on the cosets of Fi_{22}:

1	142155	694980	175134960	166795200	800616960	3592512	38438400	1
1	−15093	118404	5621616	−6177600	−1976832	−133056	2562560	1938
1	10827	38340	468720	1321920	−2059776	−84672	304640	48620
1	−373	12804	−139664	4800	40448	28224	53760	1828332
1	3915	3780	123120	−14400	−69120	16704	−64000	2909907
1	−1269	3204	9072	−2880	13824	−4032	−17920	29099070
1	459	324	−8208	−5184	13824	−1728	512	278555200
1	−117	−252	2160	1728	−4608	576	512	872972100

References

[BI] E. Bannai and T. Ito, "Algebraic Combinatorics I, Association Schemes", Benjamin-Cummings Lect. Notes, Benjamin, Menlo Park, 1984.

[BCN] A. Brouwer, A. Cohen, and A. Neumaier, "Distance-Regular Graphs", Springer, Berlin, 1989.

[CLS] A. M. Cohen, M. W. Liebeck and J. Saxl, The distance-transitive representations of the exceptional Lie groups. [In preparation]

[Atlas] J. H. Conway, R. T. Curtis, S. P. Norton, R. A. Parker and R. A. Wilson, "Atlas of Finite Groups", Clarendon Press, Oxford, 1985.

[Coo] B. Cooperstein, The geometry of root subgroups in exceptional groups, II, Geom. Dedicata, **15** (1983), 1–45.

[Ivn] A. A. Ivanov, A geometric characterization of Fischer's Baby Monster, J. Algebraic Combin., **1** (1992), 45–70.

[ILLSS] A. A. Ivanov, S. A. Linton, K. Lux, J. Saxl and L. Soicher, Distance-transitive representations of the sporadic groups, Comm. Algebra, **23** (1995), 3379–3427.

[JLPW] Ch. Jansen, K. Lux, R. Parker and R. Wilson, "An Atlas of Brauer Characters", Clarendon Press, Oxford, 1995.

[KL] P. B. Kleidman and M. W. Liebeck, "The Subgroup Structure of Finite Classical Groups", LMS Lecture Notes Ser. **129**, Cambridge Univ. Press 1990.

[KW] P. B. Kleidman and R. A. Wilson, The maximal subgroups of Fi_{22}, Math. Proc. Cambridge Philos. Soc., **102** (1987), 17–23.

[Nor] S. P. Norton, Constructing the Monster, in "Groups, Combinatorics and Geometry, Durham 1990" (M. Liebeck and J. Saxl eds.), pp. 63–76, London Math. Soc. Lect. Notes **165**, Cambridge Univ. Press, Cambridge 1992.

[NW] S. P. Norton and R. A. Wilson, The maximal subgroups of $F_4(2)$ and its automorphism group, Comm. Algebra, **17** (1989), 2809–2824.

[RW] P. Rowley and L. Walker, On the Fi_{22}-minimal parabolic geometry, I and II. Manchester Center for Pure Math., Preprint 1993/(7-8).

[GAP] M. Schönert, et. al., GAP: Groups, Algorithms and Programming, Lehrstuhl D für Mathematik, RWTH Aachen, 1993.

A. A. Ivanov
Department of Mathematics, Imperial College
180 Queen's Gate, London SW7 2BZ, England

Jan Saxl
Department of Pure Mathematics and Mathematical Statistics
University of Cambridge
16 Mill Lane, Cambridge CB2 1SB, England

Advanced Studies in Pure Mathematics 24, 1996
Progress in Algebraic Combinatorics
pp. 197–225

Towards a Classification of Spin Models in terms of Association Schemes

François Jaeger

§1. Introduction

The *spin models* considered here have been introduced by V. Jones [Jo] (in the symmetric case) and by Kawagoe, Munemasa, Watatani [KMW] (in the general case) as basic data for a certain construction of invariants of links in 3-space. Such a spin model consists in a pair of square matrices satisfying some constraints which we call *invariance equations*. Links are represented by plane diagrams and the matrices of the spin model are used to assign to every such diagram a number (this number is the value of the *partition function*). The invariance equations represent sufficient conditions on the matrices of the spin model which insure that the partition function (multiplied by a suitable normalization factor) is invariant under simple deformations of diagrams called *Reidemeister moves*. These moves describe in terms of diagrams the natural topological equivalence of links, and hence the partition function of every spin model defines a link invariant.

The pioneering work of Jones gave two examples of symmetric spin models and raised the question of finding new ones. It turned out that this question is intimately related with the theory of association schemes. Indeed many subsequent works (in particular [Ja1], [B3], [BB1], [BB2], [BBIK], [BBJ], [BJS], [I1], [I2], [I3], [Ja3], [N1], [N2], [N3]) confirmed the importance of the following situation: the matrices of a spin model belong to the Bose-Mesner algebra of some self-dual association scheme, and can be obtained by solving a certain *modular invariance equation* associated with the (suitably indexed) first eigenmatrix of the scheme (more details can be found in the surveys [B2], [Ja2], [Ja4]). The main purpose of the present paper is to show in the symmetric case that actually this situation is completely general.

Received May 15, 1995.
Revised July 3, 1995.

The idea of the proof is to associate with every symmetric spin model a matrix-valued partition function Z defined on plane tangle diagrams with four ends. The required Bose-Mesner algebra will be the linear span of the image of Z (which we call the tangle closure of the spin model). The fact that this is indeed a self-dual Bose-Mesner algebra is a consequence of the following properties: Z converts the "vertical" and "horizontal" products of tangles into the ordinary and Hadamard products of matrices, and converts the rotation through angle $\frac{\pi}{2}$ into a duality map. The modular invariance equation is also an algebraic translation via Z of a simple topological equivalence of tangles.

Shortly after we obtained this "topological" proof, K. Nomura found a simpler and purely algebraic proof [N4]. This proof was then extended in [JMN] to non symmetric spin models. We believe that our topological proof is still of interest, because it gives some topological insight on the relationship between spin models and association schemes. Actually we indicate in Section 5.3 how Nomura's proof could be interpreted topologically. The same thing would be possible for most of the results in [JMN]. However for the sake of simplicity we have chosen to restrict the present study to the case of symmetric spin models.

The paper is organized as follows. Section 2 gives the necessary preliminaries on association schemes and spin models. Section 3 introduces tangles and their relevant properties. Section 4 presents our main results, Theorem A (on the tangle closure of a symmetric spin model), Theorem B (on a Bose-Mesner subalgebra of the tangle closure which has an explicit algebraic description and which we call the algebraic closure of the spin model), and Theorem C which deals with the modular invariance equation. Section 5 gives some additional results: the Bose-Mesner algebras of Theorems A and B satisfy the *planar duality property* introduced in [Ja3]; the link invariants associated with spin models are invariant under *mutation* of links; the algebras of Theorems A and B are contained in Nomura's algebra. Finally we conclude in Section 6 with some remarks and open questions.

This work was done while the author had a visiting research position at RIMS at the invitation of the project "Algebraic Combinatorics 94". I would like to express my warmest gratitude to the organizers of this project and to all the people who made my stay at RIMS enjoyable and fruitful, and in particular to Eiichi Bannai, Kyoji Saito, Bill Kantor, Viakhalatur Sunder, Makoto Matsumoto, Akihiro Munemasa. I would also like to thank R. Bacher, P. de la Harpe and V. Jones for very helpful conversations; their use of diagrams and partition functions to describe some algebraic properties of spin models (see [BHJ]) had a strong influence on the present work.

§2. Association schemes and spin models for link invariants

2.1. Symmetric spin models on link diagrams

A (tame) *link* consists of a finite collection of disjoint simple closed smooth curves (the *components* of the link) embedded in 3-space. If each component has received an orientation, the link is said to be *oriented.* (Oriented) links can be represented by (*oriented*) *diagrams.* A diagram of a link is a generic plane projection (there is only a finite number of multiple points, each of which is a simple crossing), together with a "marking" at each crossing to indicate which part of the link goes over the other. For oriented links, the edges of the diagrams are oriented according to the orientations of the corresponding link components.

Two links are *ambient isotopic* if there exists a (smooth) isotopy of the ambient 3-space which carries one onto the other (for oriented links, this isotopy must preserve the orientations). This natural equivalence of links is described in terms of diagrams by Reidemeister's Theorem, which asserts that two diagrams represent ambient isotopic links if and only if one can be obtained from the other by a finite sequence of elementary local transformations, the *Reidemeister moves.* These moves belong to three basic types described for the unoriented case in Figure 1.

A move is performed by replacing a part of diagram which is one of the configurations of Figure 1 by an equivalent configuration without

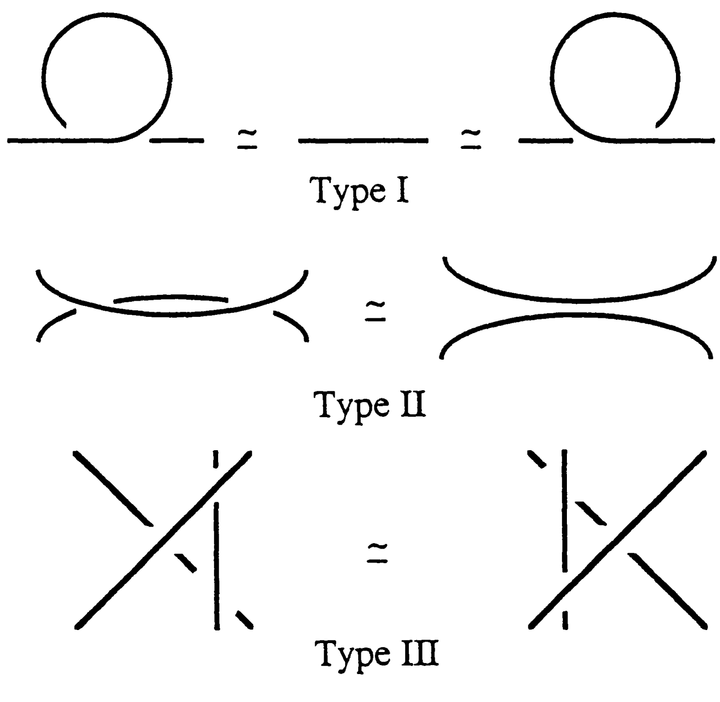
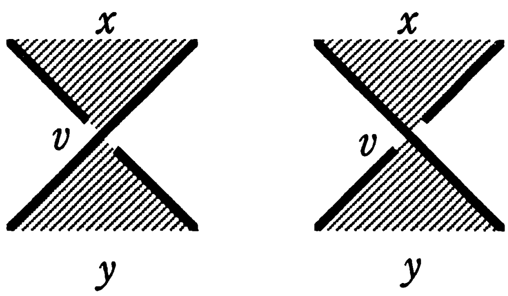


Fig. 1

modifying the remaining part of the diagram. For the oriented case, all local orientations of these pairs of equivalent configurations must be considered. More details can be found for instance in [BZ] or [K].

Reidemeister's Theorem allows the combinatorial definition of a link invariant as an assignment of values to diagrams such that the value of any diagram is preserved by Reidemeister moves performed on this diagram. It is shown in [Jo] that the concept of *spin model*, which plays an important role in statistical mechanics, can be used to obtain such invariants. This idea can be briefly described as follows.

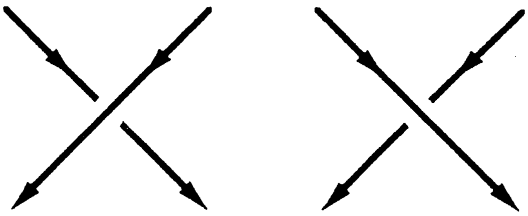
Let X be a finite set of "spins". Given a connected diagram L, we color the faces of the diagram (i.e., the connected components of the complement of L in the plane) with two colors, black and white, in such a way that adjacent faces receive different colors (we then obtain a *face-colored diagram*). Let $B(L)$ be the set of black faces. The mappings from $B(L)$ to X are called *states*. Every state σ determines at every crossing v a *local weight* $< \sigma, v >$. The local weight $< \sigma, v >$ is of the form $W_{\pm}[x,y]$, where x, y are the values of σ on the two black faces incident to v, and W_+, W_- are complex matrices with rows and columns indexed by X. The choice between these two matrices is made as shown on Figure 2. To insure that $< \sigma, v >$ is well defined we assume that W_+, W_- are symmetric (in the general case we should take into account some orientation of the diagram as in [KMW], but this case will not be considered here). The *weight* of the state σ is the product over all crossings v of the local weights $< \sigma, v >$, and the *partition function* is the sum of weights of all states.



$< \sigma, v >= W_+[x,y]$ $< \sigma, v >= W_-[x,y]$

Fig. 2

Our aim is to obtain an invariant of links from the partition function. For this it is actually necessary to introduce an orientation of L and multiply the partition function by some *normalization factor* which depends in part on this orientation. The *Tait number* (or *writhe*) $T(L)$



+1 −1

Fig. 3

of an oriented diagram L is the sum of signs of its crossings, where the sign of a crossing is defined on Figure 3. Let q be a square root of $|X|$ and let a be some non-zero complex number. The *normalized partition function* is $Z(L) = a^{-T(L)}q^{-|B(L)|}Z'(L)$, where $Z'(L)$ denotes the partition function. Then it is not difficult to derive (from a simple analysis of the effect of Reidemeister moves on the states and their weights) some conditions on W_+, W_- which guarantee that the normalized partition function defines an invariant of oriented links (see [Jo]). We shall call these conditions *invariance equations.*

Before writing down these equations we need the following notations. We denote by I the identity matrix and by J the all-one matrix of appropriate sizes. $\mathcal{M}(X)$ is the set of complex matrices with rows and columns indexed by X. The Hadamard product of two matrices A, B in $\mathcal{M}(X)$ is denoted by $A \circ B$ and given by $(A \circ B)[x,y] = A[x,y]B[x,y]$.

The invariance equations are

(1) $$I \circ W_+ = aI, \qquad I \circ W_- = a^{-1}I;$$

(2) $$JW_+ = W_+J = qa^{-1}J, \qquad JW_- = W_-J = qaJ;$$

(3) $$W_+W_- = |X|I;$$

(4) $$W_+ \circ W_- = J;$$

(5) (Star-triangle equation) for every α, β, γ in X,

$$\sum_{x \in X} W_+[\alpha,x]W_+[\beta,x]W_-[\gamma,x] = qW_+[\alpha,\beta]W_-[\beta,\gamma]W_-[\gamma,\alpha].$$

We shall call a *symmetric spin model* a 5-tuple (X, W_+, W_-, a, q), where $a \neq 0$, $q^2 = |X|$, and W_+, W_- are symmetric matrices in $\mathcal{M}(X)$ which satisfy properties (1) to (5).

We shall need in the sequel a variation of the normalization factor which does not require the link diagram L to be connected. By adding a point at infinity in the unbounded face we now view L as drawn on the sphere. We denote by $\chi_b(L)$ the Euler characteristic of the union of the black faces of L. In general, if a surface is decomposed into v vertices, e edges (each of which is homeomorphic to an open segment) which connect these vertices, and r regions (each of which is homeomorphic to an open disk), its Euler characteristic is equal to $v-e+r$. Note that if L is connected, each black face is an open disk and hence $\chi_b(L) = |B(L)|$. Otherwise some "holes" may appear in the black faces, and each such hole contributes -1 to $\chi_b(L)$. It is easy to see that, if the invariance equations hold, the expression $Z(L) = a^{-T(L)}q^{-\chi_b(L)}Z'(L)$ defines the same link invariant as before.

2.2. Symmetric association schemes

We shall need the following basic facts concerning symmetric association schemes (see [BI], [D], [BCN] for more details).

A *d-class symmetric association scheme* on the finite non-empty set X is a partition of $X \times X$ into $d+1$ non-empty symmetric relations R_i, $i = 0, \ldots, d$, where $R_0 = \{(x,x)|\ x \in X\}$, which satisfies the following property:

(6) For x, y in X, the number of elements z which satisfy given scheme relations with x and y only depends on which scheme relation is satisfied by the pair (x,y). That is, for every i, j, k in $0, \ldots, d$ there exists an integer p_{ij}^k such that

$$|\{z \in X|\ (z,x) \in R_i, (z,y) \in R_j\}| = p_{ij}^k$$

for every x, y in X with (x,y) in R_k.
Define matrices A_i, $i = 0, \ldots, d$, in $\mathcal{M}(X)$ by

(7) $A_i[x,y]$ equals 1 if $(x,y) \in R_i$, and equals 0 otherwise.

The above properties can then be reformulated as follows:

(8) $A_i^t = A_i$ (where A_i^t denotes the transpose of A_i);

(9) $$A_i \neq 0,\ A_i \circ A_j = \delta(i,j)A_i;$$

(10) $$A_0 = I;$$

(11) $$\sum_{i=0,\ldots,d} A_i = J;$$

(12) $$A_iA_j = A_jA_i = \sum_{k=0,\ldots,d} p_{ij}^k A_k.$$

Let $\mathcal{A}$ be the subspace of the complex vector space $\mathcal{M}(X)$ spanned by the matrices A_i, $i = 0, \ldots, d$. By (9) these matrices are linearly independent and hence form a basis of $\mathcal{A}$. Then (9) and (11) imply that, under Hadamard product, $\mathcal{A}$ is an associative commutative algebra with unit J, and $\{A_i,\ i=0,\dots,d\}$ is a basis of orthogonal idempotents of this algebra. Moreover by (8) $\mathcal{A}$ consists of symmetric matrices. Finally it follows from (10) and (12) that under ordinary matrix product $\mathcal{A}$ is also an associative commutative algebra with unit I. The subspace $\mathcal{A}$ of $\mathcal{M}(X)$ is called the *Bose-Mesner algebra* [BM] of the association scheme. Conversely, a $(d+1)$-dimensional subspace of $\mathcal{M}(X)$ which contains I, J, consists of symmetric matrices, and is closed under Hadamard product and ordinary matrix product, is the Bose-Mesner algebra of some symmetric d-class association scheme (see [BCN], Th. 2.6.1). Such a subspace will be called here a *symmetric Bose-Mesner algebra* on X.

A *duality* of a symmetric Bose-Mesner algebra $\mathcal{A}$ on X is a linear map Ψ from $\mathcal{A}$ to itself which satisfies the following properties:

(13) For every matrix M in $\mathcal{A}$, $\Psi(\Psi(M)) = |X|M$;

(14) For any two matrices M, N in $\mathcal{A}$, $\Psi(MN) = \Psi(M)\circ\Psi(N)$.

It easily follows that $\Psi(I) = J$ and $\Psi(J) = |X|I$.

A symmetric Bose-Mesner algebra will be called *self-dual* if it admits a duality.

§3. Tangles

Tangles and tangle diagrams are classical objects in knot theory (see for instance [C]) which also play an important role in the study of quantum invariants (see for instance [Tur]). Most of the content of this section is well known, but some of our definitions and notations are not standard.

3.1. Tangles and tangle diagrams

In what follows the Euclidean 3-dimensional space $\mathbb{R}^3$ is defined by Cartesian coordinates x, y, z.

Let S be the square $[-1,1]\times[-1,1]$ in the (x,y)-plane and let $B = S\times\mathbb{R}$. We shall call an *S-tangle* any subset T of B which is the disjoint union of a finite collection of simple smooth curves (the *components* of T) and which satisfies the following condition:

(15) $$\partial T = T\cap\partial B = \{\mathrm{NE},\mathrm{NW},\mathrm{SW},\mathrm{SE}\},$$

where $\mathrm{NE}=(1,1,0)$, $\mathrm{NW}=(-1,1,0)$, $\mathrm{SW}=(-1,-1,0)$, $\mathrm{SE}=(1,-1,0)$.

Thus exactly two components of T are homeomorphic to a closed interval, while every other component is homeomorphic to a circle.

An S-tangle is *oriented* if each of its components has received an orientation.

The notions and properties which we now introduce for S-tangles are immediate extensions of classical notions and properties for knots and links (see for instance [BZ], [K]).

Two S-tangles T_0, T_1 are *isotopic* if there exists a (smooth) ambient isotopy $H : B\times[0,1]\to B\times[0,1]$, $H(b,t) = (h_t(b),t)$, such that each h_t fixes ∂B pointwise, h_0 is the identity map, and $h_1(T_0) = T_1$. For oriented S-tangles it is also required that H preserves the orientation of each component.

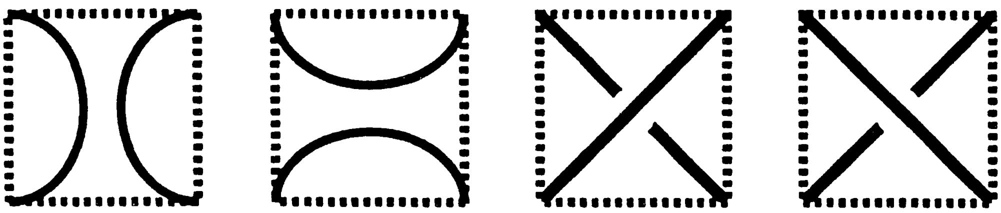
We denote by $\mathcal{T}$ (respectively: $\mathcal{OT}$) the set of isotopy classes of S-tangles (respectively: oriented S-tangles).

Isotopy classes of S-tangles can be represented by diagrams, which we shall call *S-diagrams*, as follows.

Given an element T of $\mathcal{T}$ or $\mathcal{OT}$, we may choose an S-tangle in the isotopy class T which lies in general position with respect to the (x,y)-plane, so that its projection onto that plane is generic. This means that there is only a finite number of multiple points, each of which is a simple crossing. Thus we may consider this projection as a finite graph D embedded in the square S (identified with $S\times 0$). The graph D lies in the interior of S except for four vertices of degree 1 situated at the points NE, NW, SW, and SE. The other vertices have degree 4 and will be provided with a marking indicating in the usual fashion the spatial structure of the corresponding crossings. The resulting marked graph embedded in S will be called an S-diagram of T. If T belongs to $\mathcal{OT}$, each edge of the S-diagram will receive the orientation inherited from the corresponding S-tangle component, and the S-diagram will be said to be oriented. S-diagrams will be considered up to isomorphisms of plane graphs which preserve markings, the above mentioned properties of the embeddings in S, and, in the case of oriented S-diagrams, orientations. Let $\mathcal{D}$ (respectively: $\mathcal{OD}$) be the set of S-diagrams (respectively: oriented S-diagrams). Clearly any S-diagram of T uniquely determines T. Hence we may define a map θ from $\mathcal{D}$ onto $\mathcal{T}$ and a map also denoted by θ from $\mathcal{OD}$ onto $\mathcal{OT}$ such that, for every (possibly oriented) S-diagram D, $\theta(D)$ is the isotopy class of S-tangles represented by D.

Figure 4 displays the basic S-diagrams X_0, X_∞, X_+, X_-. Here and in other pictures of S-diagrams, the x-axis is horizontal and the y-axis is vertical.

Reidemeister moves for S-diagrams are defined exactly as for link diagrams: it is enough to require that all configurations appearing in Figure 1 (or in its oriented versions) lie in the interior of S. Let us say that two S-diagrams (respectively: oriented S-diagrams) D_1, D_2 are



X_0 X_∞ X_+ X_-

Fig. 4

Reidemeister equivalent (this will be written $D_1 =_R D_2$) if it is possible to transform one into the other by a finite sequence of Reidemeister moves (respectively: oriented Reidemeister moves). Then Reidemeister's Theorem for S-tangles (see for instance Section 3.5 of [Tur]) states that two S-diagrams D_1, D_2 represent the same element of $\mathcal{T}$ (or the same element of $\mathcal{OT}$) if and only if they are Reidemeister equivalent. We record this as:

$$(16) \qquad D_1 =_R D_2 \text{ if and only if } \theta(D_1) = \theta(D_2)$$

whenever D_1 and D_2 are two elements of $\mathcal{D}$, or two elements of $\mathcal{OD}$.

In the sequel we shall mainly work with S-diagrams. However the interpretation (16) of Reidemeister equivalence in terms of S-tangles via the map θ will be crucial for the proof of the Main Lemma of Section 3.3.

3.2. Products

Let T_1, T_2 be elements of $\mathcal{T}$. Intuitively speaking, their *vertical product* $T_1 \div T_2$ will be obtained by placing T_1 above T_2 and gluing them together. More precisely, we translate an S-tangle of type T_1 by the vector $(0, 1, 0)$, we translate an S-tangle of type T_2 by the vector $(0, -1, 0)$, and we apply to the union the transformation $(x, y, z) \to (x, y/2, z)$. We may then use an ambient isotopy of B to deform the resulting object into an S-tangle (we only need to move the two points of intersection with the x-axis towards the interior of B and realize smoothness at these points). Clearly the resulting S-tangle only depends on the isotopy classes T_1, T_2 and we denote by $T_1 \div T_2$ its isotopy class.

We define similarly the *horizontal product* $T_1 \# T_2$ which is obtained by placing T_1 to the left of T_2 and gluing them together.

We also define in an analogous way the vertical product $D_1 \div D_2$ and the horizontal product $D_1 \# D_2$ of two S-diagrams D_1, D_2 in $\mathcal{D}$ (see

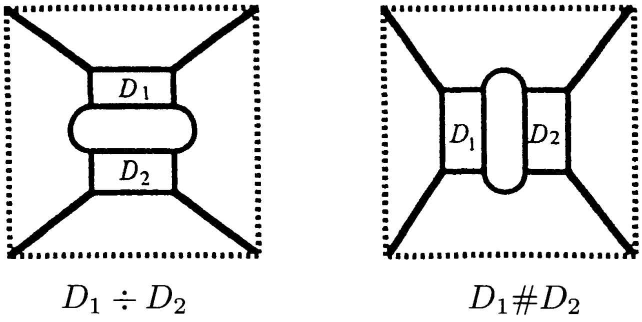


Fig. 5

Figure 5). Then

$$\text{(17)}\qquad \begin{aligned}&\theta(D_1 \div D_2) = \theta(D_1) \div \theta(D_2) \text{ and}\\ &\theta(D_1 \# D_2) = \theta(D_1) \# \theta(D_2) \text{ for all } D_1,\ D_2 \text{ in } \mathcal{D}.\end{aligned}$$

Remark. It is clear that both products $\div$ and $\#$ on $\mathcal{D}$ are associative. The vertical product has unity element X_0 and the horizontal product has unity element X_∞ (see Figure 4). Similar properties hold for $\mathcal{T}$.

3.3. Transformations and identities

We shall need the following notations for some transformations of $\mathbb{R}^3$: ρ is the rotation about the z-axis through angle $\pi/2$ such that $\rho((1,0,0)) = (0,1,0)$, α is the rotation about the x-axis through angle π, and β is the rotation about the y-axis through angle π.

These smooth transformations preserve B, ∂B and $\{\mathrm{NE}, \mathrm{NW}, \mathrm{SW}, \mathrm{SE}\}$. Hence the image of an S-tangle under one of these transformations is also an S-tangle. Since the images of two isotopic S-tangles are also isotopic, this defines actions of ρ, α, β on $\mathcal{T}$ and on $\mathcal{OT}$.

Similarly, since the transformations ρ, α, β preserve S, they act naturally on S-diagrams. Their action on the underlying (possibly oriented) graphs is clear. Their action on markings will be determined by the following requirements:

(18) for every S-diagram D, $\theta(\rho(D)) = \rho(\theta(D))$ and similarly when ρ is replaced by α or β.

Examples of the action of ρ, α, β on $\mathcal{D}$ are displayed on Figure 6.

The following identities are immediate from Figure 4:

$$\text{(19)}\qquad \rho(X_0) = X_\infty,\ \rho(X_\infty) = X_0,\ \rho(X_+) = X_-,\ \rho(X_-) = X_+.$$

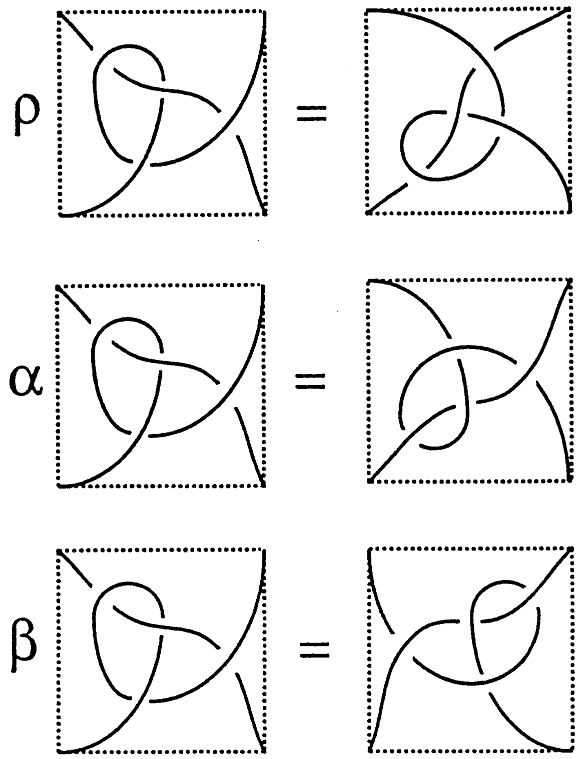


Fig. 6

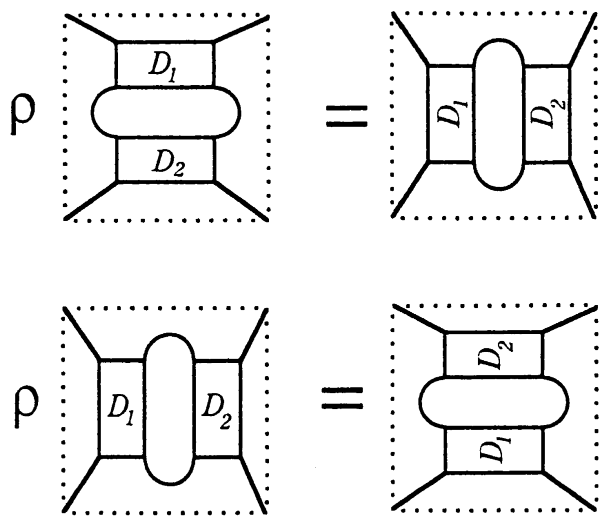


Fig. 7

The following identities are immediate from Figure 7:

(20) $\rho(D_1 \div D_2) = \rho(D_1) \# \rho(D_2)$ for all D_1, D_2 in $\mathcal{D}$,

(21) $\rho(D_1 \# D_2) = \rho(D_2) \div \rho(D_1)$ for all D_1, D_2 in $\mathcal{D}$.

Finally the following result will be essential in the sequel.

Main Lemma.

(i) $(X_+ \div (X_- \# D)) \# X_- =_R \rho(D)$ *for every* D *in* $\mathcal{D}$.
(ii) $(X_- \div (X_+ \# D)) \# X_+ =_R \rho(D)$ *for every* D *in* $\mathcal{D}$.

Proof. (i) By (16), (17), (18), this is equivalent to the equality

$$(\theta(X_+) \div (\theta(X_-) \# T)) \# \theta(X_-) = \rho(T) \qquad \text{for every } T \text{ in } \mathcal{T}.$$

This equality is proved on Figure 8.

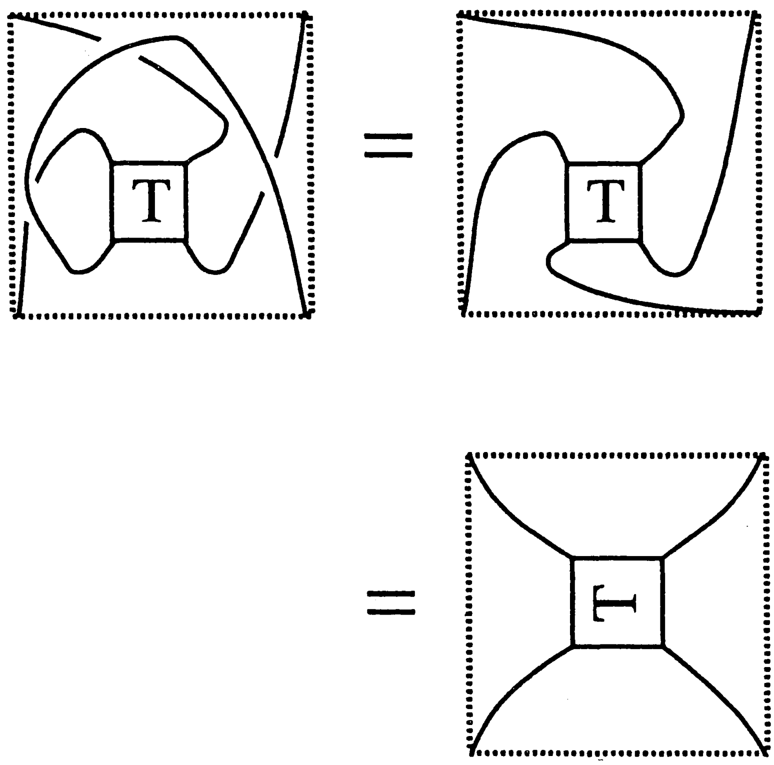


Fig. 8

The proof for (ii) is obtained from the proof for (i) by reflection in the (x, y)-plane (this is the *mirror image operation*). Q.E.D.

§4. Bose-Mesner algebras associated with symmetric spin models

4.1. The matrix-valued partition function Z

We call *faces* of an S-diagram D the connected components of $(S - \partial S) - D$. The *North face* $N(D)$ is the one whose boundary contains $[-1,1]\times\{1\}$ and the *South face* $S(D)$ is the one whose boundary contains $[-1,1]\times\{-1\}$ (these faces can be identical: for instance this is the case for X_0). The faces of D can be (uniquely) colored with two colors, black and white, in such a way that any two faces adjacent along an edge of D receive different colors, and $N(D)$, $S(D)$ are colored black. From now on the faces of every S-diagram D are colored in this way, and we denote by $B(D)$ the set of black faces of D. We call *black set* of D the union of the black faces of D and we denote by $\chi(D)$ the Euler characteristic of this set. Finally, if D is oriented, we denote by $T(D)$ the *Tait number* (or *writhe*) of D, that is, the sum of signs of its crossings, where the sign of a crossing is defined on Figure 3.

Let (X, W_+, W_-, a, q) be a symmetric spin model. We define a mapping Z from $\mathcal{D}$ to $\mathcal{M}(X)$ as follows.

Let D be an S-diagram. For i, j in X we call *state of D of type (i,j)* any mapping σ from $B(D)$ to X such that $\sigma(N(D)) = i$, $\sigma(S(D)) = j$. Then for any such mapping σ and for any vertex v of degree 4 of D, we define a *local weight* $<\sigma, v>$ as in the case of link diagrams (see Figure 2). Now the *weight* $<\sigma, D>$ of the state σ is the product of the $<\sigma, v>$ over all vertices v of degree 4 of D (this product being 1 if there are no such vertices). We define $Z'(D)$ to be the matrix in $\mathcal{M}(X)$ with (i,j) entry given by

$$Z'(D)[i,j] = \sum <\sigma, D>, \tag{22}$$

where the summation runs over all states σ of D of type (i,j) (the sum being 0 if there are no such states).

Finally, we define the mapping Z from $\mathcal{D}$ to $\mathcal{M}(X)$ by

$$Z(D) = q^{2-\chi(D)} Z'(D). \tag{23}$$

Proposition 1. *If D'_1, D'_2 are Reidemeister equivalent oriented S-diagrams, and if D_i is the S-diagram obtained from D'_i by forgetting its orientation $(i = 1, 2)$, $a^{-T(D'_1)} Z(D_1) = a^{-T(D'_2)} Z(D_2)$.*

Proof. We may proceed in essentially the same way as for link diagrams. There is only one minor difference. We need to show the invariance under Reidemeister moves of $Z(D)[i,j]$ (up to a writhe factor) for each given pair (i,j). But the fact that we restrict the summation in (22) to states of type (i,j) is not significant since, in the analysis of a given Reidemeister move, we consider (as in the case of link diagrams) the state values as fixed for all black faces except possibly one "central face" of the move. Q.E.D.

The proof of the following result is easy (see Figures 2, 4) and is left to the reader.

Proposition 2.

(i) $Z(X_0) = qI$,
(ii) $Z(X_\infty) = J$,
(iii) $Z(X_+) = W_+$, $Z(X_-) = W_-$.

We shall need the following important properties of the mapping Z.

Proposition 3.

(i) $Z(D_1 \div D_2) = q^{-1} Z(D_1) Z(D_2)$ *for all* D_1, D_2 *in* $\mathcal{D}$,
(ii) $Z(D_1 \# D_2) = Z(D_1) \circ Z(D_2)$ *for all* D_1, D_2 *in* $\mathcal{D}$,
(iii) $Z(\rho^2(D)) = (Z(D))^t$ *for all* D *in* $\mathcal{D}$.

Proof. (i) Figure 9 (i) shows that there is a bijective correspondence between the set of states of type (i,j) of $D_1 \div D_2$ and the set of triples (k, σ_1, σ_2), where k is an element of X, σ_1 is a state of type (i,k) of D_1, and σ_2 is a state of type (k,j) of D_2, such that $< \sigma, D_1 \div D_2 > = < \sigma_1, D_1 >< \sigma_2, D_2 >$ whenever σ corresponds to (k, σ_1, σ_2), for some k in X. Hence

$$Z'(D_1 \div D_2)[i,j] = \sum_{k \in X} Z'(D_1)[i,k] Z'(D_2)[k,j].$$

Moreover the black set of $D_1 \div D_2$ can be obtained from the disjoint union of the black sets of D_1 and D_2 by the attachment of a single band. Hence

$$\chi(D_1 \div D_2) = \chi(D_1) + \chi(D_2) - 1.$$

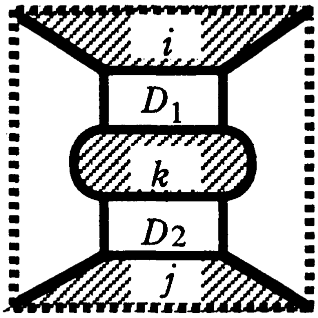
The two above equalities together yield (i).

(ii) The proof is quite similar to that of (i). Figure 9 (ii) shows that

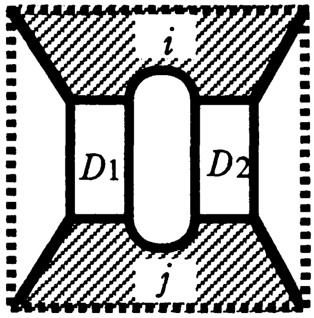
$$Z'(D_1 \# D_2)[i,j] = Z'(D_1)[i,j] Z'(D_2)[i,j].$$

Now the black set of $D_1 \# D_2$ is obtained from the disjoint union of the black sets of D_1 and D_2 by the attachment of two bands, and hence

$$\chi(D_1 \# D_2) = \chi(D_1) + \chi(D_2) - 2.$$



(i)



(ii)

Fig. 9

The result follows immediately.

(iii) Clearly ρ^2 defines a bijective correspondence between $B(D)$ and $B(\rho^2(D))$ which shows that $\chi(\rho^2(D)) = \chi(D)$, and also yields a weight-preserving bijective correspondence between states of type (i, j) of D and states of type (j, i) of $\rho^2(D)$. Q.E.D.

4.2. The tangle closure of a symmetric spin model

Let us consider a symmetric spin model (X, W_+, W_-, a, q) and the associated map Z: $\mathcal{D} \to \mathcal{M}(X)$ defined in the previous section. Let us denote by $< Z(\mathcal{D}) >$ the linear span of $Z(\mathcal{D})$, which we shall call the *tangle closure* of the spin model.

Theorem A. *$< Z(\mathcal{D}) >$ is a symmetric Bose-Mesner algebra which contains the spin model matrices W_+, W_-. Moreover the map Ψ: $\mathcal{M}(X) \to \mathcal{M}(X)$ defined by $\Psi(M) = aW_-\circ(W_+(W_-\circ M))$ induces a duality on $< Z(\mathcal{D}) >$.*

Proof. By Proposition 2, $< Z(\mathcal{D}) >$ contains W_+,W_-, I and J. By definition $< Z(\mathcal{D}) >$ is a linear subspace of $\mathcal{M}(X)$, and it follows immediately from Proposition 3 (i), (ii) that $< Z(\mathcal{D}) >$ is closed under ordinary matrix product and under Hadamard product. Hence Ψ defines a linear map from $< Z(\mathcal{D}) >$ to $< Z(\mathcal{D}) >$. By properties (3) and (4) of spin models, Ψ is invertible and $\Psi^{-1}(M) = a^{-1}|X|^{-1}W_+\circ(W_-(W_+\circ M))$ for every M in $< Z(\mathcal{D}) >$.

Let D be any element of $\mathcal{D}$. By (i) of the Main Lemma and Proposition 1, $Z((X_+ \div (X_-\#D))\#X_-) = a^k Z(\rho(D))$, where k is the Tait number of some orientation of $(X_+ \div(X_-\#D))\#X_-$ minus the Tait number of the corresponding orientation of $\rho(D)$. One easily obtains from Figure 8 that $k = -1$ for each of its six oriented versions. Hence

$$Z(\rho(D)) = aZ((X_+\div(X_-\#D))\#X_-).$$

Applying Proposition 2 (iii) and Proposition 3 (i), (ii), and using the commutativity of the Hadamard product, we obtain

$$Z(\rho(D)) = aq^{-1}W_{-}\circ(W_{+}(W_{-}\circ Z(D)))$$

or equivalently

$$\Psi(Z(D)) = qZ(\rho(D)). \tag{24}$$

Similarly, using (ii) of the Main Lemma, and considering the mirror images of the six oriented versions of Figure 8, we obtain

$$Z(\rho(D)) = a^{-1}Z((X_{-} \div (X_{+}\#D))\#X_{+})$$

and hence

$$Z(\rho(D)) = a^{-1}q^{-1}W_{+}\circ(W_{-}(W_{+}\circ Z(D))).$$

This is equivalent to

$$\Psi^{-1}(Z(D)) = q^{-1}Z(\rho(D)). \tag{25}$$

It follows from (24), (25) that $\Psi^{-1}(Z(D)) = q^{-2}\Psi(Z(D))$ and hence

$$\Psi^{2}(A) = |X|A \text{ for every } A \text{ in } < Z(\mathcal{D}) >. \tag{26}$$

Moreover, for any two elements D_1, D_2 of $\mathcal{D}$ we have:

$$\begin{aligned}
\Psi(Z(D_1)Z(D_2)) &= q\Psi(Z(D_1 \div D_2)) && \text{(by Proposition 3 (i))}\\
&= q^2 Z(\rho(D_1 \div D_2)) && \text{(by (24))}\\
&= q^2 Z(\rho(D_1)\#\rho(D_2)) && \text{(by (20))}\\
&= q^2 Z(\rho(D_1))\circ Z(\rho(D_2)) && \text{(by Proposition 3 (ii))}\\
&= \Psi(Z(D_1))\circ\Psi(Z(D_2)) && \text{(by (24))}
\end{aligned}$$

Hence $\Psi(AB) = \Psi(A)\circ\Psi(B)$ for every A, B in $< Z(\mathcal{D}) >$. This together with (26) shows that Ψ induces a duality on $< Z(\mathcal{D}) >$.

Finally, by applying (24) twice to the equality $(Z(D))^t = Z(\rho^2(D))$ of Proposition 3(iii), and using (26), we obtain $(Z(D))^t = Z(D)$. Hence all matrices in $< Z(\mathcal{D}) >$ are symmetric. Q.E.D.

Remark. It follows from (24) and Proposition 2 (iii), or from properties (2), (4) of spin models and the expression of Ψ given in Theorem A, that $\Psi(W_{+}) = qW_{-}$.

4.3. The algebraic closure of a symmetric spin model

So far the symmetric Bose-Mesner algebra $< Z(\mathcal{D}) >$ is related to the corresponding spin model in an abstract way via S-tangles or S-diagrams, and we are not able to describe this Bose-Mesner algebra explicitly. We now propose a way to overcome this difficulty.

Let us say that a subset of $\mathcal{M}(X)$ is *weakly Bose-Mesner* (WBM for short) if it is closed under complex linear combinations, ordinary matrix product, Hadamard product, and contains I, J. Thus symmetric Bose-Mesner algebras on X are exactly the WBM subsets of $\mathcal{M}(X)$ consisting only of symmetric matrices.

Clearly $\mathcal{M}(X)$ is WBM, and the intersection of a family of WBM subsets is again WBM. Thus every subset F of $\mathcal{M}(X)$ has a unique *WBM-closure* $\mathcal{C}(F)$ which is the smallest WBM subset of $\mathcal{M}(X)$ containing it. More precisely we may define $\mathcal{C}(F)$ in the following two equivalent ways:

(i) abstract definition: $\mathcal{C}(F)$ is the intersection of all WBM subsets of $\mathcal{M}(X)$ containing F.

(ii) algorithmic definition: $\mathcal{C}(F)$ is obtained by iterating the following process involving a finite current set F', and taking as the initial instance of F' a maximal linearly independent subset of $F \cup \{I, J\}$. Compute ordinary or Hadamard products of two elements of F', checking for each resulting matrix if it belongs to the linear span of F'. If not, the matrix is incorporated into F'. If no such incorporation is possible, the linear span of F' is $\mathcal{C}(F)$.

Theorem B. *Let (X, W_+, W_-, a, q) be a symmetric spin model. $\mathcal{C}(\{W_+\}) = \mathcal{C}(\{W_-\})$ is a symmetric Bose-Mesner algebra. Moreover the map $\Psi : \mathcal{M}(X) \to \mathcal{M}(X)$ defined by $\Psi(M) = aW_- \circ (W_+(W_- \circ M))$ induces a duality on this Bose-Mesner algebra.*

Proof. By Theorem A, W_+ belongs to the symmetric Bose-Mesner algebra $< Z(\mathcal{D}) >$. Since $< Z(\mathcal{D}) >$ is WBM, it contains $\mathcal{C}(\{W_+\})$ and hence all matrices in $\mathcal{C}(\{W_+\})$ are symmetric. The equality $\mathcal{C}(\{W_+\}) = \mathcal{C}(\{W_-\})$ easily follows from property (3) of spin models. Indeed the inverse of W_+ is a linear combination of powers of W_+ and hence W_- belongs to $\mathcal{C}(\{W_+\})$. This shows that $\mathcal{C}(\{W_+\}) \supset \mathcal{C}(\{W_-\})$ and the reverse inclusion is proved similarly. Finally the definition of Ψ shows immediately that $\Psi(\mathcal{C}(\{W_+\})) = \mathcal{C}(\{W_+\})$ and the result follows from Theorem A. Q.E.D.

The Bose-Mesner algebra $\mathcal{C}(\{W_+\})$ will be called the *algebraic closure* of the spin model (X, W_+, W_-, a, q). This Bose-Mesner algebra can also be understood in terms of S-diagrams as follows. Call an S-diagram *algebraic* if it can be generated using only vertical and horizontal products from the basic S-diagrams X_0, X_∞, X_+, X_- (the corresponding S-tangles are called algebraic in [C]). Let $\mathcal{AD}$ be the set of algebraic S-diagrams. It is clear from Propositions 2 and 3 (i), (ii) that $Z(\mathcal{AD})$ is contained in $\mathcal{C}(\{W_+, W_-\}) = \mathcal{C}(\{W_+\})$. Conversely, it is easy to see that $\mathcal{C}(\{W_+\})$ is linearly spanned by elements in $Z(\mathcal{AD})$. Thus $\mathcal{C}(\{W_+\})$ is the linear span of $Z(\mathcal{AD})$, which we denote by $< Z(\mathcal{AD}) >$, and clearly we could give a "topological" proof of Theorem B exactly similar to that of Theorem A.

Theorem B generalizes previous results of [Ja1]. In that paper two algebras $\mathcal{M}$ and $\mathcal{H}$ are associated with any symmetric spin model (X, W_+, W_-, a, q): $\mathcal{M}$ is the algebra under matrix product generated by J, W_+, and $\mathcal{H}$ is the algebra under Hadamard product generated by I, W_-. Clearly $\mathcal{M}$ and $\mathcal{H}$ are contained in $\mathcal{C}(\{W_+\})$, and, since $\Psi(W_+) = qW_-$ and $\Psi(J) = |X|I$, $\Psi(\mathcal{M}) = \mathcal{H}$. This equality is the content of Proposition 3 of [Ja1], with however different expressions for Ψ (which are easily obtained from the expression of Theorem B using the star-triangle equation (5)). Assume now that $\mathcal{M}$ is closed under Hadamard product. Then $\mathcal{M}$ is WBM (because it contains I by (3)) and hence equals $\mathcal{C}(\{W_+\})$. This implies Proposition 4 of [Ja1].

We do not know any example of a symmetric spin model for which the tangle closure $< Z(\mathcal{D}) >$ and the algebraic closure $< Z(\mathcal{AD}) >$ are different, but we believe that such an example should exist.

4.4. Modular invariance

Let (X, W_+, W_-, a, q) be a symmetric spin model. We have introduced two symmetric Bose-Mesner algebras $< Z(\mathcal{D}) >$ and $< Z(\mathcal{AD}) >$, each of which contains W_+, W_- and is provided with a duality Ψ given by the expression $\Psi(M) = aW_- \circ (W_+(W_- \circ M))$. Now, given a symmetric Bose-Mesner algebra with duality Ψ, we would like to know if it is associated in this way with some symmetric spin model, and classify such spin models. A key concept here is that of *modular invariance*, formulated by Eiichi Bannai and his coworkers in their study of relations between fusion algebras of conformal field theories and association schemes (see [B1]), and used by them to construct new spin models ([BB2], [BBIK])

We have seen that the Bose-Mesner algebra $\mathcal{A}$ of a d-class symmetric association scheme on a set X has a basis $\{A_i, i = 0, \dots, d\}$ of orthogonal idempotents for the Hadamard product. It is well known that it has also a basis $\{E_i, i = 0, \dots, d\}$ of orthogonal idempotents for the ordinary matrix product, where $E_0 = |X|^{-1}J$ (see [BI], Section II.3). The *first* *eigenmatrix* P is defined by

$$A_j = \sum_{i=0,\dots,d} P[i,j]E_i. \tag{27}$$

Thus P is the transition matrix from the basis $\{A_i, i = 0, \dots, d\}$ to the basis $\{E_i, i = 0, \dots, d\}$, and only depends on the indexing of the A_i and the E_i. If Ψ is a duality, the matrices $\Psi(E_i)$ are the A_i in some order. In particular, by (10), $\Psi(E_0) = |X|^{-1}\Psi(J) = I = A_0$. Thus we may choose the indices in such a way that $\Psi(E_i) = A_i$ for $i = 0, \dots, d$. Then by (27) P is the matrix of Ψ with respect to the basis $\{E_i, i = 0, \dots, d\}$, and $P^2 = |X|I$ by (13).

The modular invariance property asserts that there exists a diagonal matrix Δ such that $(P\Delta)^3$ is a non-zero multiple of the identity. The following relationship between this property and spin models is presented in [BBJ].

Let us write $W_- = \sum_{i=0,\dots,d} t_i A_i$ and define W_+ by $\Psi(W_+) = qW_-$. Thus $W_+ = q^{-1}\Psi(W_-) = q^{-1}\sum_{i=0,\dots,d} t_i \Psi(A_i) = q\sum_{i=0,\dots,d} t_i E_i$.

Assume that $\Psi(M) = aW_- \circ (W_+(W_- \circ M))$ for every M in the Bose-Mesner algebra $\mathcal{A}$. We write the corresponding equality in terms of matrices with respect to the basis $\{A_i, i = 0, \dots, d\}$.

The matrix of the linear map $M \to W_- \circ M$ is the diagonal matrix Δ with $\Delta(i,i) = t_i$ for $i = 0, \dots, d$.

The matrix of the linear map $M \to W_+M$ is $P^{-1}(q\Delta)P = q^{-1}P\Delta P$. The matrix of Ψ is $P^{-1}(P)P = P$. We obtain the equality $P = q^{-1}a\Delta P\Delta P\Delta$, or equivalently $(P\Delta)^3 = q^3a^{-1}I$.

We have proved the following result.

Theorem C. *Every symmetric spin model (X, W_+, W_-, a, q) is characterized by some solution Δ to the modular invariance equation $(P\Delta)^3 = q^3a^{-1}I$, where P is the first eigenmatrix of some self-dual symmetric Bose-Mesner algebra (with indices chosen such that $P^2 = |X|I$).*

In general, a solution to the above modular invariance equation needs not correspond to some spin model (see [BBJ]). However, Theorem C should be a very useful tool for the classification of spin models (see for instance [B3], [CS]).

§5. Complements

5.1. Planar duality

We have just seen that the modular invariance property gives a necessary condition for a symmetric Bose-Mesner algebra to correspond to a symmetric spin model as in Theorems A or B. Another necessary condition is given by the *planar duality property* introduced in [Ja3]. This property is defined in terms of spin models on plane graphs, and some special cases are well known in physics (see also [Bi1], [Bi2]).

Let G be a finite undirected graph (loops and multiple edges are allowed). Its vertex-set and edge-set will be denoted by $V(G)$ and $E(G)$ respectively. Let X be a non-empty finite set and let w be a mapping from $E(G)$ to $\mathcal{M}(X)$ whose values are symmetric matrices. This mapping defines a *spin model on the graph* G. Let us call *state* of G any mapping σ from $V(G)$ to X. If the edge e has ends v_1, v_2, let $w(e|\sigma)$ be the $(\sigma(v_1), \sigma(v_2))$-entry of the matrix $w(e)$. The *weight* $w(\sigma)$ *of the state* σ is then the product of the $w(e|\sigma)$ over all edges e (this will be set to 1 if G has no edge). Finally, the *partition function* $Z(G, w)$ is the sum of weights of all states.

Let $\mathcal{A}$ be a symmetric Bose-Mesner algebra on X and let Ψ be a duality of $\mathcal{A}$. The pair $(\mathcal{A}, \Psi)$ is said to satisfy the planar duality property if for every connected plane graph G and mapping w from $E(G)$ to $\mathcal{A}$

$$Z(G, w) = |X|^{1-|V(G^*)|} Z(G^*, \Psi w \phi), \tag{28}$$

where G^* is the dual plane graph of G and $\phi : E(G^*) \to E(G)$ associates to every edge of G^* its dual edge.

Proposition 4. *Let (X, W_+, W_-, a, q) be a symmetric spin model. Each of the pairs $(< Z(\mathcal{D}) >, \Psi)$, defined in Theorem A, and $(\mathcal{C}(\{W_+\}), \Psi)$, defined in Theorem B, satisfies the planar duality property.*

Proof. The following argument is essentially a symmetric version of the one used for Proposition 12 of [Ja3] (see also the associated remark).

We consider a connected plane graph G and we want to check (28) for every mapping w from $E(G)$ to $\mathcal{A}$, where $\mathcal{A}$ is either $< Z(\mathcal{D}) >$ or $< Z(\mathcal{AD}) > = \mathcal{C}(\{W_+\})$. The set of these mappings can be identified with a product of m copies of the vector space $\mathcal{A}$, where $m = |E(G)|$. Then each side of (28), considered as a function of w, is clearly a multilinear form on this product of vector spaces. It follows that it is enough to check (28) for mappings w which take their values in some fixed basis of $\mathcal{A}$. Since $\mathcal{A}$ is either $< Z(\mathcal{D}) >$ or $< Z(\mathcal{AD}) >$, we may choose a basis contained in $Z'(\mathcal{D})$ (see (22), (23)). Hence it is enough to check (28) when, for every edge e of G, $w(e) = Z'(D_e)$ for some S-diagram D_e.

We shall need the following classical description of the pair (G, G^*). Let H be a connected 4-regular plane graph. We color its faces in black and white in such a way that adjacent faces receive different colors. The *graph of black faces* of H is a connected plane graph defined as follows. It has one vertex f° for each black face f of H, placed in the interior of that face, and one edge v° for each vertex v of H. If the vertex v of H is incident to the black faces f, g, the edge v° is embedded as a simple curve joining the vertices f°, g° across v. The *graph of white faces* of H is defined similarly, and it is easily seen that it is the dual of the graph of black faces. Moreover, every pair of dual connected plane graphs arises in this way (see [O] p.47).

Thus we may assume that G is the graph of black faces and G^* is the graph of white faces of some face-colored connected 4-regular plane graph H. We now associate with w (satisfying $w(e) = Z'(D_e)$ for every edge e of G) a face-colored connected link diagram L. To construct L, for every edge e of G we locally replace as shown on Figure 10 the corresponding vertex of H by the S-diagram D_e. There is some ambiguity in this construction since we do not specify which black face of H incident to this vertex corresponds to the North face of D_e, but we shall see that this will not matter, due to the fact that $Z'(D_e)$ is a symmetric matrix.

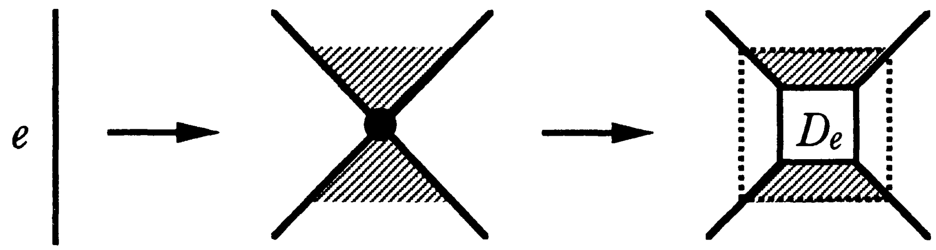


Fig. 10

It is clear from the definitions of the partition functions $Z'(L)$ (see Section 2.1), $Z'(D_e)$ (Section 4.1), and $Z(G, w)$, that

$$Z'(L) = Z(G, w). \tag{29}$$

Let us now exchange black and white in the face-coloring of L. The same argument (and examination of Figure 10) shows that the partition function $Z''(L)$ computed with respect to the new face-coloring satisfies

$$Z''(L) = Z(G^*, w^*), \tag{30}$$

where w^*: $E(G^*) \to \mathcal{A}$ is defined by $w^*(e^*) = Z'(\rho(D_e))$ for all e^* in $E(G^*)$ with dual edge e in $E(G)$. Using (23), (24) we see that

$$\begin{aligned} w^*(e^*) &= q^{\chi(\rho(D_e))-2} Z(\rho(D_e)) \\ &= q^{\chi(\rho(D_e))-3} \Psi(Z(D_e)) \\ &= q^{\chi(\rho(D_e))-3} \Psi(q^{2-\chi(D_e)} Z'(D_e)) \\ &= q^{\chi(\rho(D_e))-\chi(D_e)-1} \Psi(w(e)) \\ &= q^{\chi(\rho(D_e))-\chi(D_e)-1} (\Psi w \phi)(e^*). \end{aligned}$$

Hence (28) is equivalent to

$$Z'(L) = |X|^{1-|V(G^*)|} q^{\mu} Z''(L), \tag{31}$$

where $\mu = \sum_{e \in E(G)} (-\chi(\rho(D_e)) + \chi(D_e) + 1)$.

Recall from Section 2.1 that $Z(L) = a^{-T(L)} q^{-\chi_b(L)} Z'(L)$ defines a link invariant. It is easy to show that the exchange of black and white in the face-coloring of L does not modify the value of this invariant (see [Jo], Proposition 2.14). Hence $q^{-\chi_b(L)} Z'(L) = q^{-\chi_w(L)} Z''(L)$, where $\chi_w(L)$ is the Euler characteristic of the union of the white faces of L. It follows that (31) reduces to the equality

$$\chi_b(L) - \chi_w(L) = 2(1 - |V(G^*)|) + \mu.$$

Clearly $\chi_b(L) = |V(G)| + \sum_{e \in E(G)} (\chi(D_e) - 2)$, and similarly

$$\chi_w(L) = |V(G^*)| + \sum_{e \in E(G)} (\chi(\rho(D_e)) - 2).$$

Hence $\chi_b(L) - \chi_w(L) = |V(G)| - |V(G^*)| + \mu - |E(G)|$ and the result follows from Euler's formula applied to the connected plane graph G. Q.E.D.

The efficiency of Proposition 4 for the classification of spin models is still unclear, mainly because it is computationally difficult to verify the planar duality property (or even its restriction to the case where the plane graph G is the 4-clique, which is of special significance as shown in [Ja3]) .

5.2. Mutation and link invariants described by spin models

Let D be an S-diagram. We denote by $K(D)$ the link diagram obtained from D by joining together the ends of D as shown on Figure 11 (with a face-coloring extending the face-coloring of D). Clearly, every

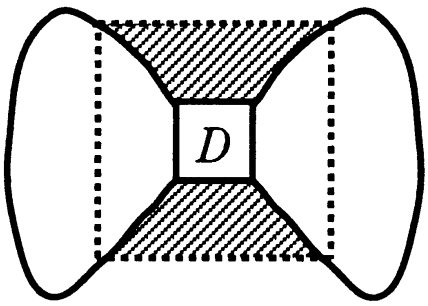


Fig. 11. $K(D)$

link diagram is of the form $K(D)$ for some S-diagram D. Moreover, given any symmetric spin model,

$$Z'(K(D)) = \mathrm{Trace}(Z'(D)). \tag{32}$$

A link diagram L' is obtained by *mutation* of the link diagram L if L is of the form $K(D_1 \div D_2)$ and L' is of the form $K(D'_1 \div D_2)$, where D'_1 is obtained from D_1 by a rotation about one of the three coordinate axes through angle π, that is, by one of the transformations ρ^2, α, β (see Section 3.3 and Figure 6). For oriented link diagrams, the orientation on D'_1 is fully reversed if necessary to match the orientation of D_2. A link is said to be obtained from another by mutation if the same statement holds for some diagrams of these links, and this property can also be defined in a 3-dimensional setting (see for instance [LM]).

If $D'_1 = \rho^2(D_1)$, $Z(D'_1) = Z(D_1)$ by Proposition 3(iii) and Theorem A.

If $D'_1 = \beta(D_1)$, there is an easy weight-preserving bijection between the sets of states of a given type (i,j) of D'_1 and D_1. Hence $Z'(D'_1) = Z'(D_1)$, and also $Z(D'_1) = Z(D_1)$ since D'_1 and D_1 have homeomorphic black sets.

Since $\alpha = \beta\rho^2$, we have also $Z(D'_1) = Z(D_1)$ if $D'_1 = \alpha(D_1)$. Then $Z(D'_1 \div D_2) = Z(D_1 \div D_2)$ by Proposition 3(i), and consequently $Z'(D'_1 \div D_2) = Z'(D_1 \div D_2)$. It now follows from (32) that $Z'(L') = Z'(L)$. Noting that the normalization factors for these two link diagrams are the same (if some orientation of D'_1 is fully reversed the corresponding sum of signs of crossings is not modified) we obtain the following result.

Proposition 5. *If an invariant of oriented links is defined by a symmetric spin model it is invariant under mutation.*

5.3. Nomura's algebra

Immediately after the announcement of Theorem A above, K. Nomura found a simpler and purely algebraic proof of the fact that the matrices of a symmetric spin model belong to some symmetric Bose-Mesner algebra [N4].

Let W_+, W_- be two symmetric matrices in $\mathcal{M}(X)$ which satisfy (3), (4). For every b, c in X, let $\mathbf{u}_{b,c}$ be the column vector indexed by X with x-entry $\mathbf{u}_{b,c}(x) = W_+[b,x]W_-[c,x]$. Let $\mathcal{N}(W_+)$ be the set of symmetric matrices A in $\mathcal{M}(X)$ such that $\mathbf{u}_{b,c}$ is an eigenvector of A for every b, c in X. It is shown in [N4] that $\mathcal{N}(W_+)$ is a symmetric Bose-Mesner algebra. Moreover if (5) holds for some square root q of $|X|$, W_+ and W_- belong to $\mathcal{N}(W_+)$. This Bose-Mesner algebra is related as follows with the Bose-Mesner algebra of Theorem A.

Proposition 6. *Let (X, W_+, W_-, a, q) be a symmetric spin model. The corresponding algebra $< Z(\mathcal{D}) >$ is contained in the Nomura algebra $\mathcal{N}(W_+)$.*

Proof. (sketch): We introduce "hexagonal" tangles with six ends, which we call H-tangles, and corresponding H-diagrams such as the one depicted on Figure 12. It is easy to formalize these notions as we did for S-tangles and S-diagrams. We introduce a standard face-coloring (with the "North", "Southwest" and "Southeast" faces black) and define a partition function Z' and normalized partition function Z for H-diagrams as we did for S-diagrams (but now Z' takes its values in a space of tensors with 3 indices).

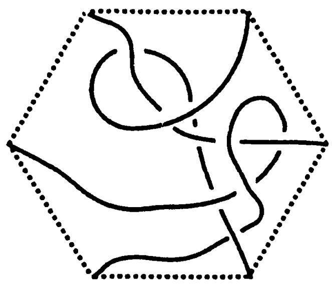


Fig. 12

Evaluation of Z on the two Reidemeister equivalent H-diagrams of Figure 13, where D is any S-diagram, shows that, for every i, b, c in X:

$$\sum_{x \in X} Z(D)[i,x]W_+[b,x]W_-[c,x] = qZ(\rho(D))[c,b]W_+[b,i]W_-[c,i]$$

or equivalently, for every b, c in X:

$$Z(D)\mathbf{u}_{b,c} = qZ(\rho(D))[c,b]\mathbf{u}_{b,c}. \tag{33}$$

Since $Z(D)$ is symmetric by Theorem A, it belongs to $\mathcal{N}(W_+)$. Q.E.D.

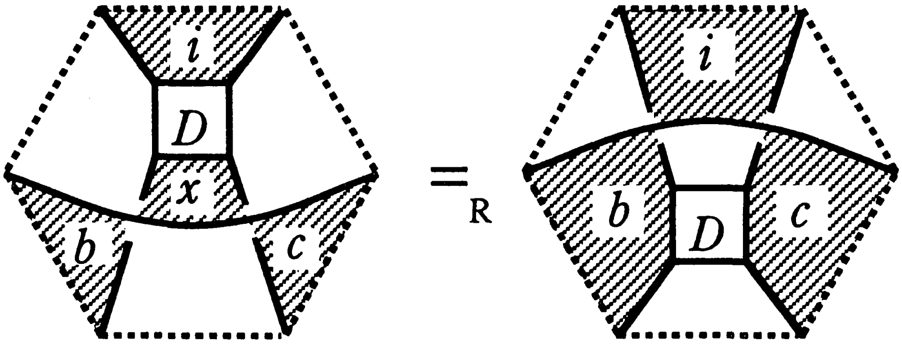


Fig. 13

Using (24), (33) becomes $Z(D)\mathbf{u}_{b,c} = \Psi(Z(D))[c,b]\mathbf{u}_{b,c}$, and hence $A\mathbf{u}_{b,c} = \Psi(A)[c,b]\mathbf{u}_{b,c}$ for every b, c in X and A in $< Z(\mathcal{D}) >$.

We may use this formula for every A in $\mathcal{N}(W_+)$ to define a linear map Ψ: $\mathcal{N}(W_+) \to \mathcal{M}(X)$. It is shown in [JMN] that Ψ is a duality of $\mathcal{N}(W_+)$ which is given by the expression $\Psi(M) = aW_-\circ(W_+(W_-\circ M))$ (i.e., the modular invariance property holds). Thus Nomura's algebra $\mathcal{N}(W_+)$ can play the same role for symmetric spin models as the algebras $< Z(\mathcal{D}) >$ of Theorem A or $< Z(\mathcal{AD}) >$ of Theorem B.

It would not be difficult to develop a "topological" approach to the results on the algebra $\mathcal{N}(W_+)$ given in [N4] and [JMN]. This could be a way to obtain some geometric intuition on these algebraic results (the starting point would be a "picture" of the definition of $N(W_+)$ similar to Figure 13), or even more, a way to actually prove the results (this would require of course an extension of the notion of S-diagram). As an example, we give in Figure 14 a diagrammatic proof (which the reader is invited to convert into an algebraic proof) of the fact that $\mathcal{N}(W_+)$ is closed under Hadamard product.

It is natural to ask whether or not $< Z(\mathcal{D}) >$ and $\mathcal{N}(W_+)$ can be different. Consider a symmetric spin model which satisfies $W_+ = W_-$ (hence W_+ is a Hadamard matrix). Then, for any S-diagram D, $Z(D)$ will be invariant under modification of the spatial structure of crossings. This, together with Proposition 1, shows that for every S-diagram D, $Z(D)$ is a scalar multiple of $Z(X_0)$, $Z(X_\infty)$, or $Z(X_+) = Z(X_-)$. Hence $< Z(\mathcal{D}) >$ has dimension at most 3 (actually, this dimension is 3 except when $|X| = 4$ and $W_+ = 2I - J$). On the other hand, it is shown in [JMN] that there exist symmetric spin models with $W_+ = W_-$ and dim $\mathcal{N}(W_+) = |X|$ whenever $|X|$ is an even power of 2. Thus the gap between the dimensions of $< Z(\mathcal{D}) >$ and $\mathcal{N}(W_+)$ can be arbitrarily large.

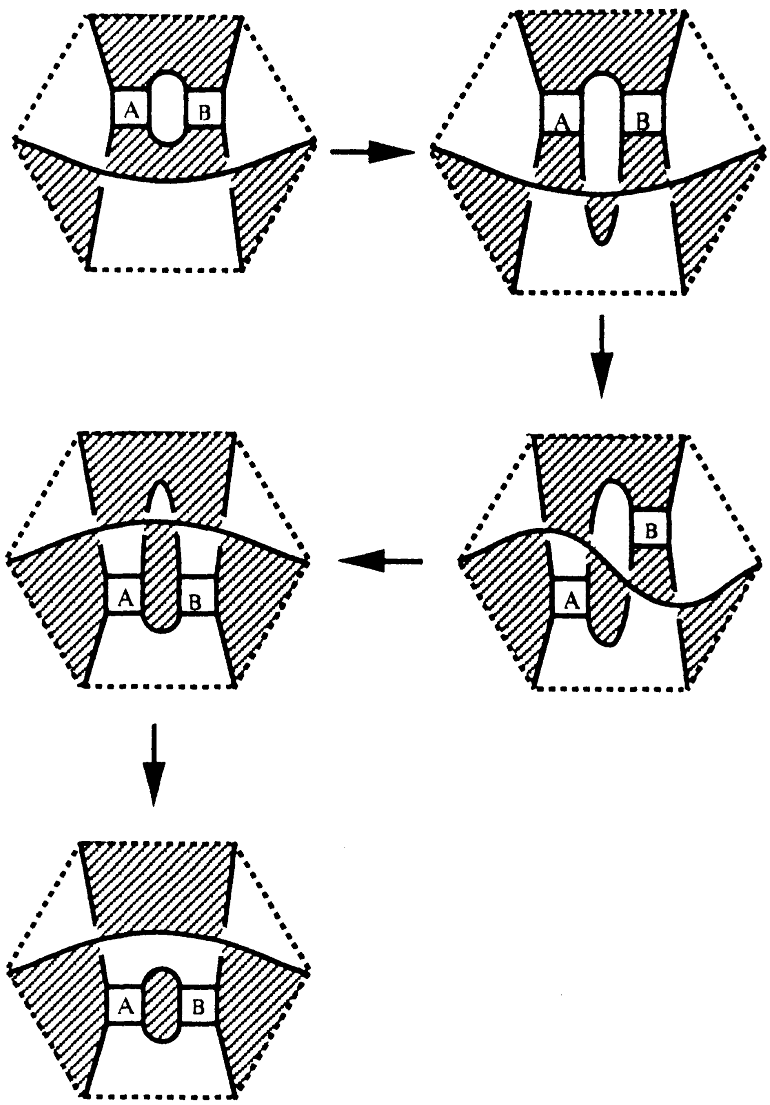


Fig. 14

§6. Conclusion

We believe that our results represent a significant step towards the classification of symmetric spin models: we can restrict our attention to the solutions of the modular invariance equations for self-dual symmetric Bose-Mesner algebras, and then check these solutions to retain those which actually yield spin models. We have proposed the planar duality property as an additional criterion, but it would be very interesting to find other criteria. Ideally we would like to have an intrinsic characterization of symmetric Bose-Mesner algebras which contain some spin model matrices. This question is solved for the 3-dimensional case in [Ja1]. This leads us to concentrate our study on the next and simplest open case of 4-dimensional symmetric Bose-Mesner algebras.

Theorem C is generalized in [JMN] to the non-symmetric spin models of [KMW] using the algebraic approach initiated in [N4]. Here the topological approach does not work well, because to define products for oriented S-tangles one must introduce compatibility conditions on the orientations. However, using the topological approach, we have obtained some preliminary results on the 4-weight spin models of [BB3], which generalize some results in [BB1], but the overall picture is still very unclear.

References

[B1] Ei. Bannai, Association schemes and fusion algebras (an introduction) J. Algebraic Combin., **2** (1993), 327–344.

[B2] Ei. Bannai, Algebraic Combinatorics —recent topics on association schemes, Sugaku **45** (1993), 55–75 (in Japanese). English translation in Sugaku Expositions, **7**, no. 2 (1994), 181–207.

[B3] Et. Bannai, Modular invariance property and spin models attached to cyclic groups association schemes, J. Statist. Plann. Inference, to appear.

[BB1] E. Bannai and E. Bannai, Generalized spin models and association schemes, Mem. Fac. Sci. Kyushu Univ. Ser. A, **47**, no. 2 (1993), 397–409.

[BB2] E. Bannai and E. Bannai, Spin models on finite cyclic groups, J. Algebraic Combin., **3** (1994), 243–259.

[BB3] E. Bannai and E. Bannai, Generalized generalized spin models (four-weight spin models), Pacific J. Math., **170** (1995), 1–16.

[BBIK] E. Bannai, E. Bannai, T. Ikuta and K. Kawagoe, Spin models constructed from the Hamming association schemes, in "Proceedings 10th Algebraic Combinatorics Symposium", Gifu Univ. 1992, pp. 91–106.

[BBJ] E. Bannai, E. Bannai and F. Jaeger, On spin models, modular invariance, and duality, J. Algebraic Combin., to appear.

[BCN] A. E. Brouwer, A. M. Cohen, A. Neumaier, "Distance-Regular Graphs", Springer-Verlag, Ergebnisse der Mathematik und ihrer Grenzgebiete, 3. Folge, Band 18, 1989.

[BHJ] R. Bacher, P. de la Harpe, V. F. R. Jones, Tours de centralisateurs pour les paires d'algèbres, modèles à spins et modèles à vertex, in preparation.

[BI] Ei. Bannai and T. Ito, "Algebraic Combinatorics I, Association Schemes", Benjamin/Cummings, Menlo Park, 1984.

[BJS] Ei. Bannai, F. Jaeger and A. Sali, Classification of small spin models, Kyushu J. Math., **48** (1994), 185–200.

[BM] R. C. Bose, D. M. Mesner, On linear associative algebras corresponding to association schemes of partially balanced designs, Ann. Math. Statist., **30** (1959), 21–38.

[BZ] G. Burde, H. Zieschang, "Knots", de Gruyter, Berlin, New York, 1985.

[Bi1] N. L. Biggs, On the duality of interaction models, Math. Proc. Cambridge Philos. Soc., **80** (1976), 429–436.

[Bi2] N. L. Biggs, "Interaction Models", London Math. Soc. Lecture Notes 30, Cambridge University Press, 1977.

[C] J. H. Conway, An enumeration of knots and links, and some of their algebraic properties, in "Computational Problems in Abstract Algebra", (J. Leech ed.), Pergamon Press, 1969, pp. 329–358.

[CS] L. Chihara, D. Stanton, A matrix equation for association schemes, Graphs Combin., **11** (1995), 103–108.

[D] P. Delsarte, An algebraic approach to the association schemes of coding theory, Philips Research Reports Supplements, **10** (1973).

[I1] T. Ikuta, Spin models corresponding to nonsymmetric association schemes of class 2 , Mem. Fac. Sci. Kyushu Univ. Ser. A, **47**, no. 2 (1993), 383–390.

[I2] T. Ikuta, Spin models corresponding to nonsymmetric association schemes of class three, preprint, 1993.

[I3] T. Ikuta, On spin models attached to association schemes, Ph.D thesis, Kyushu University, 1994.

[JMN] F. Jaeger, M. Matsumoto, K. Nomura, Bose-Mesner algebras related with type II matrices and spin models, preprint.

[Ja1] F. Jaeger, Strongly regular graphs and spin models for the Kauffman polynomial, Geom. Dedicata, **44** (1992), 23–52.

[Ja2] F. Jaeger, Modèles à spins, invariants d'entrelacs, et schémas d'association, Actes du Séminaire Lotharingien de Combinatoire, $30^{\text{ième}}$ session, (R. Konig et V. Strehl, éd), IRMA, Strasbourg, 1993, 43–60.

[Ja3] F. Jaeger, On spin models, triply regular association schemes, and duality, J. Algebraic Combin., **4** (1995), 103–144.

[Ja4] F. Jaeger, Spin models for link invariants, invited talk at the 15th British Combinatorial Conference, London Math. Soc. Lecture Notes Series, **218** (1995), 71–101.

[Jo] V. F. R. Jones, On knot invariants related to some statistical mechanical models, Pacific J. Math., **137**, no. 2 (1989), 311–334.

[K] L. H. Kauffman, "On Knots", Annals of Mathematical Studies 115, Princeton University Press, Princeton, New Jersey, 1987.

[KMW] K.Kawagoe, A. Munemasa, Y. Watatani, Generalized spin models, J. Knot Theory Ramifications, **3**, no. 4 (1994), 465–475.

[LM] W. B. R. Lickorish, K. C. Millett, A polynomial invariant of oriented links, Topology, **26** (1987), 107–141.

[N1] K. Nomura, Spin models constructed from Hadamard matrices, J. Combin. Theory Ser. A, **68** (1994), 251–261.

[N2] K. Nomura, Spin models on bipartite distance regular graphs, J. Combin. Theory Ser. B, **64** (1995), 300–313.

[N3] K. Nomura, Spin models on triangle-free connected graphs, submitted, 1994.

[N4] K. Nomura, An algebra associated with a spin model, J. Algebraic Combin., to appear.

[O] O. Ore, "The Four-Color Problem", Academic Press, New York, 1967.

[Tur] V. G. Turaev, Operator invariants of tangles and R-matrices, Math. USSR Izvestiya, **35**, no. 2 (1990), 411–444.

Laboratoire de Structures Discrètes et de Didactique
IMAG (CNRS)
BP 53X, 38041 Grenoble
FRANCE

Advanced Studies in Pure Mathematics 24, 1996
Progress in Algebraic Combinatorics
pp. 227–242

Orthogonal Spreads and Translation Planes

William M. Kantor

Abstract.

There have been a number of striking new results concerning translation planes of characteristic 2, obtained using orthogonal and symplectic spreads. The impetus for this came from coding theory. This paper surveys the geometric advances, while providing a hint of their coding–theoretic connections.

§1. Introduction

Spreads are familiar in finite geometry since they produce translation planes (cf. Section 3.1 below). Orthogonal and symplectic spreads are less familiar. They have an underlying additional structure, produced by a quadratic form or an alternating bilinear form on the vector space. When the field is $\mathbb{Z}_2$ they also produce Kerdock codes over $\mathbb{Z}_2$ and $\mathbb{Z}_4$. This note summarizes results that are more than 10 years old, while setting the stage for a discussion of new advances.

§2. Orthogonal spreads

Let $V = \mathrm{GF}(q)^{2n} = X \oplus Y$ for subspaces X and Y both of which are identified with $\mathrm{GF}(q)^n$. Equip V with the quadratic form Q defined by $Q(x, y) = x \cdot y$ (using the usual dot product on $\mathrm{GF}(q)^n$); this form is nonsingular, with isometry group $\mathrm{O}^+(2n, q)$ and associated symmetric bilinear form $(\, , \,)$. Then V has $(q^n - 1)(q^{n-1} + 1)$ nonzero singular vectors and each totally singular n–space (such as X and Y) contains $q^n - 1$ nonzero singular vectors. This suggests that there might be families of $q^{n-1} + 1$ totally singular n–spaces that partition the set of all nonzero singular vectors; such a family is called an *orthogonal spread. We will*

Received February 5, 1995.
Revised May 11, 1995.
This paper was written during a visit to RIMS in Kyoto University, whose hospitality is gratefully acknowledged.

assume that q is even and see that such a family cannot exist unless n is even, in which case there is always at least one orthogonal spread.

2.1. Matrices

Fix a basis $x_1, \dots, x_n$ of X and let $y_1, \dots, y_n$ be the dual basis of Y: $(x_i, y_j) = \delta_{ij}$. Write matrices with respect to the basis $x_1, \dots, x_n, y_1, \dots, y_n$. It is easy to check that the group $\mathrm{O}^+(2n,q)_{(Y)}$ of isometries of V that fix every vector of Y consists of those linear transformations whose matrices are $\left(\begin{smallmatrix} I & M \\ 0 & I \end{smallmatrix}\right)$ for some skew–symmetric $n \times n$ matrix M over $\mathrm{GF}(q)$ (in characteristic 2, "skew–symmetric" means "symmetric with 0 diagonal"); $\mathrm{O}^+(2n,q)_{(Y)}$ is isomorphic to the vector space of all skew–symmetric $n \times n$ matrices over $\mathrm{GF}(q)$, and is regular on the set of totally singular n–spaces Z such that $Y \cap Z = 0$.

Note that $\dim X\left(\begin{smallmatrix} I & M \\ 0 & I \end{smallmatrix}\right) \cap X\left(\begin{smallmatrix} I & N \\ 0 & I \end{smallmatrix}\right) = n - \mathrm{rank}(M-N)$. In particular, if two such totally singular n–spaces meet only at 0, then n must be even (since $M - N$ is skew–symmetric). There is another view of this parity remark: the totally singular n–spaces fall into two families such that two such subspaces are in the same family if and only if the dimension of their intersection has the same parity as n, so that there can be three such subspaces pairwise having only 0 in common only if n is even.

It is now straightforward to prove

Proposition 2.1. (i) *If Σ is an orthogonal spread of V that contains both X and Y, then*

$$\Sigma := \{Y\} \bigcup \left\{ X\begin{pmatrix} I & M \\ O & I \end{pmatrix} \;\middle|\; M \in \mathcal{K} \right\}$$

for a set $\mathcal{K}$ of $n \times n$ skew–symmetric matrices, containing O, and such that the difference of any two is nonsingular (*a* **Kerdock set** *of matrices*).

(ii) *Conversely, if $\mathcal{K}$ is a Kerdock set of $n \times n$ skew–symmetric matrices, then the set Σ defined in* (i) *is an orthogonal spread of V that contains both X and Y.*

Of course, since $\mathrm{O}^+(2n,q)$ is transitive on the ordered pairs of totally singular n–spaces having intersection 0, the restriction in (i) is insignificant.

Definition 2.2. Kerdock sets $\mathcal{K}_1$ and $\mathcal{K}_2$ are *equivalent* if $A^t\mathcal{K}_1^\tau A + M = \mathcal{K}_2$ for some $A \in \mathrm{GL}(n,q)$, some skew–symmetric matrix M, and some field automorphism τ. Orthogonal spreads Σ_1 and Σ_2 are *equivalent* if there is an element of $\Gamma\mathrm{O}(V)$ sending Σ_1 to Σ_2. (Here, $\Gamma\mathrm{O}(V)$ is the set of semilinear maps g on V that preserve Q projectively:

$Q(vg) = aQ(v)^\tau$ for some nonzero scalar a, some field automorphism τ and all $v \in V$.)

Evidently, in Proposition 2.1(ii), Σ depends on $\mathcal{K}$. It turns out that it is straightforward to determine more about the interdependence of Σ and $\mathcal{K}$:

Proposition 2.3. *Let $\mathcal{K}_1$ and $\mathcal{K}_2$ be Kerdock sets of $n \times n$ matrices over* GF(q). *Then the following are equivalent*:

(i) *$\mathcal{K}_1$ and $\mathcal{K}_2$ are equivalent*;
(ii) *The orthogonal spreads Σ_1 and Σ_2 of V, determined by $\mathcal{K}_1$ and $\mathcal{K}_2$ via* Proposition 2.1, *are equivalent by an element of* $\Gamma\mathrm{O}(V)$ *sending Y to itself.*

It is easy to deduce that there are many choices of inequivalent Kerdock sets that produce equivalent orthogonal spreads.

2.2. To symplectic spreads

Let z denote any nonsingular point (1–space) of V: $Q(z) \neq 0$. If Σ is any orthogonal spread of V, then n is even and

$$\{\, Z \cap z^\perp \mid Z \in \Sigma \,\}$$

is a family of totally singular $n-1$–spaces of $z^\perp$ that partitions the set of nonzero singular vectors.

Since the characteristic is 2, z is contained in the hyperplane $z^\perp$. The $2n-2$–space $z^\perp/z$ is turned into a symplectic space using the inherited alternating bilinear form $(u+z, v+z) := (u,v)$ (for $u, v \in z^\perp$). Then

$$\Sigma_z := \{\langle Z \cap z^\perp, z\rangle/z \mid Z \in \Sigma\}$$

is a family of $|\Sigma| = q^{n-1}+1$ totally isotropic $n-1$–spaces of $z^\perp/z$ that partitions the set of nonzero singular vectors. Such a family is called a *symplectic spread* of the symplectic space $z^\perp/z$. (N.B.—There is no quadratic form inherited by $z^\perp/z$.)

2.3. From symplectic spreads

The preceding construction can be reversed, proceeding from symplectic spreads to orthogonal ones.

Namely, let m be odd, and start with a symplectic space V' of dimension $2m$ over GF(q) together with a symplectic spread Σ' in it. If $m = n-1$ then we can identify V' with the symplectic space $z^\perp/z$ arising, as above, from the orthogonal space V and one of its nonsingular points z. Each totally isotropic subspace of V' is the projection, mod

z, of a unique totally singular subspace of $z^\perp/z$. In particular, Σ' arises from a family

$$\Sigma'^{\uparrow} := \{U \mid \langle U, z\rangle/z \in \Sigma'\}$$

of totally singular $n-1$–spaces of $z^\perp$ such that each nonzero vector of $z^\perp/z$ lies in just one of its members.

Finally, each totally singular $n-1$–space of $z^\perp$ lies in exactly two totally singular n–spaces of V, one from each family. Pick a family $\mathcal{M}$ of such n–spaces, and let

$$\Sigma := \{Z \mid Z \in \mathcal{M} \text{ and } Z \text{ contains a member of } \Sigma'^{\uparrow}\}.$$

Then Σ *is an orthogonal spread of* V, *and* $\Sigma_z = \Sigma'$. Note that this passage from symplectic to orthogonal spreads is essentially unique: it only depends on the choice of the family $\mathcal{M}$. (Moreover, the nontrivial orthogonal transvection with center z interchanges $\mathcal{M}$ with the other family while leaving Σ' unchanged.)

2.4. Back and forth

Starting with a symplectic spread Σ' in a $2m$–dimensional symplectic space over $\mathrm{GF}(q)$ with m odd, Section 2.3 produces an orthogonal spread in a $2m+2$–dimensional orthogonal space, in such a way that there is a nonsingular point z for which Σ_z is Σ'. Once we have Σ, Section 2.2 can be used to form a *different* symplectic spread Σ_{z^*} using a *different* nonsingular point z^*.

2.5. Changing fields: up and down

Iterating the procedure in Section 2.4 never produces a "new" orthogonal spread. There is a simple way to modify that procedure in order to get large numbers of new orthogonal and symplectic spreads.

Start with a symplectic spread Σ' in a $2m$–dimensional vector space V' over $K = \mathrm{GF}(q)$. Let L be any proper subfield of K over which K has odd degree, and let $T\colon K \to L$ be the trace map. Then $T(u, v)$ defines a nonsingular alternating L-bilinear form on the L–space V'. We can view Σ' as a family of subspaces of this L–space. It is still a spread, and each of its members is still totally isotropic with respect to the new form. Thus, Σ' *is a symplectic spread of the* L*–space* V'. Here, $\dim_L V' = m[K:L]$.

Now Section 2.3 can be applied, producing an orthogonal spread of a $(2m[K:L]+2)$–dimensional orthogonal L–space. In fact, Section 2.4 now gives us "new" symplectic spreads. It is a difficult problem to decide, in general, whether these spreads are actually new: conceivably some are equivalent to ones already obtained.

Up and down process. This process of *repeatedly* going from a symplectic spread over some field, changing fields, going up to an orthogonal spread and then back down to a symplectic spread, is called the *up and down process.* It is difficult to keep control over properties of these spreads. However, in important special cases control can be maintained, a surprising discovery of Williams [Wi] that will be discussed shortly (Sections 3.5–3.9).

§3. Projective planes

An entirely different type of geometric view of symplectic spreads is provided by projective planes, and provides one of the principal motivations for their study. For this purpose we begin by ignoring the symplectic structure.

3.1. From spreads to projective planes

Let V' be a $2m$–dimensional vector space over $\mathrm{GF}(q)$ (no restriction is placed even on the parity of q and m).

Spreads A *spread* of V' is a family Σ' of q^m+1 subspaces of dimension m whose union is all of V'. This means that every nonzero vector is in a unique member of Σ'. *Any family of* q^m+1 m*–spaces in a* $2m$*–space, any two of which have only* 0 *in common, is a spread.* (N.B.—An orthogonal spread is not a spread in this sense, but a symplectic spread is.)

Example 3.1. If V' is a 2–dimensional vector space over a finite field E, its set Σ' of 1–spaces is a *desarguesian* (*or "regular"*) *spread.* Note that this spread is symplectic with respect to any alternating bilinear form on V'. It is also symplectic when V' is viewed as a vector space over any subfield of E (cf. Section 2.5).

Translation planes Any spread of V' determines a *translation plane* $\mathbf{A}(\Sigma')$, an affine plane of order q^m whose points are vectors and whose lines are the cosets $W+v$ with $W \in \Sigma', v \in V'$. The plane $\mathbf{A}(\Sigma')$ corresponding to a desarguesian spread Σ' is a desarguesian plane.

Any isomorphism between two translation planes is induced by a semilinear transformation of the underlying vector spaces. See [De] for more background concerning translation planes. The transition to *projective planes* is standard: introduce a line at infinity whose points are all parallel classes of lines, in order to obtain a projective plane of order q^m.

3.2. Symplectic translation planes

Example 3.2. (Example 3.1 *continued.*) Starting with a desarguesian spread Σ' in $\mathrm{GF}(q)^{2m}$, where m is odd, by Section 2.3 we obtain an orthogonal spread Σ in $\mathrm{GF}(q)^{2m+2}$, and hence a Kerdock set. This latter Kerdock set is the one first discovered by Kerdock [Ke] when $q = 2$ (cf. [Di] and [MS, Ch. 15 §5], among many other references).

Each orthogonal spread appears to produce large numbers of symplectic spreads Σ_z. This leads us to the isomorphism question: when are two planes $\mathbf{A}(\Sigma_z)$ obtained in this manner isomorphic? If there is a symplectic transformation sending one spread to the other, the planes are certainly isomorphic. It seems surprising that the converse is (essentially) true:

Theorem 3.3. *For $i = 1, 2$, let Σ_i be a symplectic spread in a $2m$-dimensional symplectic space V_i over* $\mathrm{GF}(q)$. *Let $g\colon \mathbf{A}(\Sigma_1) \to \mathbf{A}(\Sigma_2)$ be an isomorphism that sends the point 0 to the point 0. Then there is an invertible semilinear transformation $h\colon V_1 \to V_2$ such that the following hold:*

(i) $(\Sigma_1)h = \Sigma_2$,
(ii) *There is a field automorphism τ, and a nonzero scalar a, such that $(uh, vh) = a(u,v)^\tau \ \ \forall u, v \in V_1$, and*
(iii) *$g^{-1}h$ fixes every member of Σ_2.*

The elementary proof is in [Ka1, I (3.5)]. The set of all nonsingular linear transformations fixing every member of Σ_2 (as in (iii)), together with 0, is a field, the *kernel* of the translation plane. It is the largest field over which the spread consists of subspaces.

The preceding theorem implies that isomorphic planes can only arise from equivalent orthogonal spreads (Definition 2.2). Moreover:

Corollary 3.4. *Two translation planes $\mathbf{A}(\Sigma_{z_1})$ and $\mathbf{A}(\Sigma_{z_2})$ arising from the same orthogonal spread Σ are isomorphic if and only if z_1 and z_2 are in the same orbit of the group $G(\Sigma)$ of all elements of $\Gamma\mathrm{O}(V)$ that preserve Σ.*

Theorem 3.3 also permits the determination of the full automorphism groups of many of these planes. The construction techniques for planes, using Kerdock sets and orthogonal and symplectic spreads, are very flexible. They have produced planes with relatively large collineation groups (Sections 3.5–3.7) as well as planes with unexpectedly small collineation groups (Section 3.8).

3.3. Prequasifields

A translation plane is usually coordinatized by an algebraic system called a *quasifield* [De]. Here it will be convenient to consider a weaker, but geometrically equivalent system, called a *prequasifield.*

Definition 3.5. Consider a binary operation $*$ on $F = \mathrm{GF}(q^m)$ related to field addition by the following conditions (for all $x, y, z \in F$):

- $(x + y) * z = x * z + y * z.$
- $x * y = x * z \Longrightarrow x = 0$ or $y = z.$
- $x * y = 0 \iff x = 0$ or $y = 0.$

Then $(F, +, *)$ will be called a *prequasifield.* It is a *quasifield* if it has an identity element; in order to be able to state (3.6), it is preferable to delete this condition even though an identity element is readily introduced. $(F, +, *)$ is a *presemifield* if both distributive laws hold, and a *semifield* if, in addition, there is an identity element.

A translation plane is obtained by using $F \oplus F$ as point–set and letting the lines have the familiar appearance

$$\text{“}x = c\text{” and “}y = x * s + b\text{”} \ \forall b, c, s \in F.$$

If we view F and $F \oplus F$ as vector spaces over $K = \mathrm{GF}(q)$, then the spread $\Sigma(*)$ of $F \oplus F$ associated with $(F, +, *)$ consists of the lines “$y = x * s$” through 0. We will always assume that our quasifield associates with K in the following manner:

$$(kx) * s = k(x * s) \quad \forall k \in K; \ x, s \in F,$$

so that $x \mapsto x * s$ is a K–linear map for each $s \in F$. Thus, K is contained in the kernel of the plane, since $(x, y) \mapsto (kx, ky)$ fixes each member “$y = x * s$” of $\Sigma(*)$ whenever $k \in K^*$.

In order to consider symplectic translation planes, we use a substitute for the dot product. The trace map $T\colon F \to K$ determines an inner product $T(xy)$ on F having an orthonormal basis that lets us identify F, equipped with this inner product, and K^m, equipped with its usual dot product.

Finally, we assume in addition that m is *odd* and that $*$ satisfies the following condition:

$$T(x(x * y)) = T(xy)^2 \quad \forall x, y \in F. \tag{3.6}$$

One example of such a binary operation is $x*y = xy^2$; the corresponding plane is desarguesian. Soon we will present many more examples. Note that, if we had required that our prequasifield have an identity element, then we would have had to use a more complicated version of the inner product. Thus, for example, it is more convenient in the present context to use the preceding inconvenient–looking modification xy^2 of ordinary multiplication in F.

Replacing x in turn by x, z, $x+z$ in (3.6) and subtracting, we find that

$$T(x(z*y)) = T(z(x*y)) \ \ \forall x,y,z \in F\,. \tag{3.7}$$

By a simple calculation:

Proposition 3.8. *Equip $F \oplus F$ with the alternating bilinear form*

$$((x_1,y_1),(x_2,y_2)) := T(x_1y_2 - x_2y_1). \tag{3.9}$$

Then the spread $\Sigma(*)$ of $F \oplus F$ associated with a prequasifield $(F,+,*)$ is symplectic if and only if $(F,+,*)$ satisfies (3.7).

In view of this result, it may seem as if condition (3.6) is unnecessarily restrictive. When searching for examples, this may be so, but in fact it is no serious restriction at all:

Proposition 3.10. *If $(F,+,*)$ is a symplectic prequasifield, then there is a permutation μ of F such that $x \circ y := x * y^\mu$ defines a prequasifield $(F,+,\circ)$ that is symplectic with respect to the same form* (3.9), *defines the same plane, and behaves as in* (3.6).

Namely, by (3.7) for each $y \in F$ the map $x \mapsto T(x(x*y))^{1/2}$ is K–linear, so $T(x(x*y)) = T(xy^\pi)^2$ for some map $\pi\colon F \to F$ with $0^\pi = 0$. If π is bijective, let μ denote its inverse and note that $T(x(x*y^\mu)) = T(xy)^2$ for all x, y, in which case $x \circ y := x * y^\mu$ behaves as required.

Suppose that π is not bijective, and let $y, z \in F$, $y \neq z$, with $y^\pi = z^\pi$. If g denotes the K–linear map $x \mapsto x*y - x*z$, then $T(xx^g) = T(x(x*y)) - T(x(x*z)) = 0\ \forall x \in F$. Then the K-bilinear map $(u,v) := T(uv^g)$ on F satisfies $(u,u) = 0\ \forall u \in F$, and hence is an alternating bilinear form. It is nonsingular since g is (i.e., $T(Fy^g) = 0 \Longrightarrow y^g = 0 \Longrightarrow y = 0$). Since $[F:K] = m$ is odd, this is impossible.

3.4. Up to Kerdock sets and orthogonal spreads, and down again

Now equip $F \oplus K$ with the inner product

$$\big((x,a),(y,b)\big) := T(xy) + ab.$$

Proposition 3.11. *The linear maps*

$$M_s\colon (x,a) \mapsto (x * s + sT(sx) + as, T(sx)),\ \ s \in F\,,$$

form a Kerdock set of $(m+1)\times(m+1)$ skew–symmetric matrices over K. If the above inner product is used, then every Kerdock set is equivalent to one arising in this manner.

The proof is straightforward. Corresponding to this Kerdock set is the orthogonal spread $\Sigma[*]$ in $F\oplus K\oplus F\oplus K$ consisting of $0\oplus 0\oplus F\oplus K$ together with the subspaces

$$(3.12)\quad \Big\{(x,k,x*s+T(xs)s+ks,T(xs)\ \Big|\ x\in F,\ k\in K\Big\}\ \text{ for } s\in F;$$

here, the quadratic form is $Q(x,a,y,b) = T(xy) + ab$.

For some choices of a nonsingular point z it is easy to write down the symplectic spread $\Sigma[*]_z$. Namely, if $z = \langle(0,a,\zeta,1)\rangle$ with $a\in K^*, \zeta\in F$, then a straightforward calculation shows that the following symplectic prequasifield multiplication $\circ$ gives rise to an equivalent copy of $\Sigma[*]_z$ lying inside $F\oplus F$, where the alternating form is (3.9):

$$(3.13)\qquad x\circ s = \big[x*s + (1+a)T(xs)s + T(xs)\zeta + T(x\zeta)s\big]/a.$$

(Division by a in (3.13) is only included so that (3.6) will hold for $\circ$. Namely, $T(x[x*s+(1+a)T(xs)s+T(xs)\zeta+T(x\zeta)s]) = T(xs)^2 + (1+a)T(xs)T(xs) + T(xs)T(x\zeta) + T(x\zeta)T(xs) = aT(xs)^2$.)

3.5. Semifield planes

Let F and K be as before, with $K\supset \mathrm{GF}(2)$, and let T_1 denote the trace map $F\to K$. The presemifield

$$x*y := xy^2 + T_1(x)y + T_1(xy)$$

was introduced and studied in [Ka1, II]. The corresponding spread arises by starting with the desarguesian spread, going up and down once (cf. Section 2.5) while preserving the group of $|F|$ elations with axis $0\oplus F$. This produces a nondesarguesian semifield plane.

This approach was greatly generalized in [Wi]. The presemifields studied there arise by the up and down process (Section 3.4), carefully retaining elations having a finite axis. In fact, by iterating (3.13) but always using $a = 1$, these presemifields can be described explicitly as follows. Let $F = F_0\supset F_1\supset\cdots\supset F_n = K$ be a sequence of fields with

$n \geq 3$, let $T_i\colon F \to F_i$ be the trace map, and choose $\zeta_i \in F$ for each $i \geq 1$. Then

$$x * y = xy^2 + \sum_1^n \Big(T_i(\zeta_i x)y + \zeta_i T_i(xy)\Big) \tag{3.14}$$

defines a 2–sided distributive binary operation on F that produces a symplectic semifield plane.

Theorem 3.15 ([Wi]). *Assume that a sequence $F = F_0 \supset F_1 \supset \cdots \supset F_n = K$ is given as above with $[F_0 : F_1] \geq 7$ and $[F\colon K]$ odd. If (ζ_i) and (ζ_i') are sequences as above, then they define isomorphic planes if and only if $\zeta_i' = a\zeta_i^\tau$ for some $a \in F^*$, some $\tau \in \operatorname{Aut} F$, and all i.*

When all ζ_i are 0, the plane is desarguesian. The theorem implies, for example, that *there are at least $|F|^{n-2}/(|F|-1)\log_2|F|$ pairwise nonisomorphic symplectic semifield planes defined by* (3.14)—provided that $m = [F : K]$ has at least $n \geq 3$ (not necessarily distinct) prime factors, at least one of which is ≥ 7. Stronger versions of this result appear in [Wi].

3.6. Nearly flag–transitive planes

If F and K are as before, and if $a \in K - \mathrm{GF}(2)$, then the prequasifield

$$x * y := xy^2 + aT(xy)y$$

was introduced and studied in [Ka1, II]. As in Section 3.5, this spread (and those in Section 3.7 below) arises by starting with the desarguesian spread, going up to the orthogonal spread in Example 3.2, and then coming down in a different manner (cf. Section 2.2). This time the group preserved is isomorphic to F^*: it has the form $(x, y) \mapsto (x\alpha, y/\alpha)$ with $\alpha \in F^*$, fixes two members of the symplectic spread, and cyclically permutes the remaining ones.

This approach was again generalized in [Wi] by iterating (3.13) but this time always using $\zeta = 0$. Let $F = F_0 \supset F_1 \supset \cdots \supset F_n = K \supset \mathrm{GF}(2)$ be a sequence of fields, where $[F : K]$ is odd. For each i, let $T_i\colon F \to F_i$ be the trace map, where $T_n = T$ in our earlier notation; and choose $c_i \in F^*$ such that $c_0 = 1$ and $c_i/c_{i-1} \in F_i$ for each i. Then

$$x * y := xy^2 + \sum_1^n \Big(c_{i-1}yT_i(c_{i-1}xy) + c_i yT_i(c_i xy)\Big)$$

defines a prequasifield, the corresponding plane is nondesarguesian, and the maps $(x, y) \mapsto (x\alpha, y/\alpha)$, where $\alpha \in F^*$, form a cyclic collineation *group of order* $|F|-1$. Isomorphisms among these planes are determined in [Wi].

This construction suggests the following general approach, for any characteristic and unrelated to symplectic spreads. Suppose that F is a finite field and $g\colon F \to F$ *is an additive map such that* $x \mapsto xg(x)$ *is bijective.* Then $(F,+,*)$ *is a prequasifield, where*

$$x * y := g(xy)y.$$

(Namely, left distributivity is clear, and $x*y = x*z \Rightarrow xg(xy)y = xg(xz)z \Rightarrow xy = xz$, as required.) Once again *the maps* $(x,y) \mapsto (\alpha x, y/\alpha)$ *with* $\alpha \in F^*$ *form a cyclic collineation group fixing the* $x-$ *and* $y-$*axes and transitively permuting the remaining lines through the origin.* (Namely, $(x, x*s)$ is sent to $(\alpha x, g(\alpha x s/\alpha)s/\alpha) = (\alpha x, (\alpha x) * (s/\alpha))$.)

Soon after I mentioned to Y. Hiramine this condition on a map g, he produced the following example: if $\omega \in \mathrm{GF}(2^6)$ and $\omega^6 = \omega + 1$, let $g(x) = x + \omega x^4 + \omega^{47}x^{16}$. However, the proof that this satisfies the required condition, and hence produces a plane of order 64, involves a long and ingenious case argument.

In the examples given earlier, $g(x) = x + \sum_1^n \Big(c_{i-1}T_i(c_{i-1}x) + c_iT_i(c_ix)\Big)$ (or, somewhat more precisely, $g(x/c_n)$ is the preceding right hand side in order to make (3.6) hold). That the resulting spread is symplectic comes from fact that g has following additional property:

$$T(xg(z)) = T(zg(x))\ \forall x, z \in F$$

(cf. (3.7)).

3.7. Flag–transitive planes

There is one further way to obtain planes from a desarguesian spread, while retaining a large collineation group [KW]. In the previous sections we preserved a group of order $|F|$ or $|F|-1$, this time the group will have order $|F|+1$. Once again, the planes are obtained by starting with the desarguesian spread and using the up and down process (Section 3.4). This time, in order to describe these planes we need to use the field $E = \mathrm{GF}(q^{2m})$ (where m is odd), and its multiplicative subgroup of order $q^m + 1$.

Let $E = E_0 \supset \cdots \supset E_n$ be a sequence of fields, where $[E : E_n]$ is odd and $|E|$ is a square; let "overbar" denote the involutory field automorphism of E. For each i let F_i be the subfield of E_i over which E_i has degree 2, let $T_i\colon F_0 \to F_i$ be the trace map, and write $W_i := \ker T_{i+1}\big|_{F_i}$. Pick any $\zeta_i \in E_i$, where $\zeta_i\overline{\zeta}_i = 1$ and $\zeta_0 = 1$, and write

$\gamma_i := \Pi_0^i \zeta_l$. Then

$$(3.16)\qquad \Big\{\theta\big(\sum_0^{n-1} W_i\gamma_i + F_n\gamma_n\big) \;\Big|\; \theta\bar\theta = 1\Big\}$$

is a symplectic spread in E, and $\{z \mapsto \theta z + w \mid \theta, w \in F,\ \theta\bar\theta = 1\}$ *is a sharply flag–transitive collineation group.*

This produces exponential numbers of flag–transitive affine planes of order q^m. In [KW] there is a complete determination of when two of them are isomorphic, as well as a discussion of the iteration involved in the construction. Once again, the simplest of these planes were first studied in [Ka1, II].

3.8. Orthogonal spreads and boring planes

The group $G(\Sigma)$ has been determined for various orthogonal spreads Σ [Ka1; Ka2; Ka4; KW; Wi]. For many of the ones in Sections 3.5–3.7, $G(\Sigma)$ is generated by the group preserved in the specific section (of order $|F|$ or $|F|\pm 1$) together with scalar transformations and some elements of $\mathrm{Aut}\, F$. It is then possible to find nonsingular points z such that $G(\Sigma)_z$ consists only of scalars. In view of Theorem 3.3, this means that the collineation group of $\mathbf{A}(\Sigma_z)$ consists entirely of perspectivities. Showing that the stabilizer of 0 is (isomorphic to) K^* can be messy (as in [Ka4]) or partly pleasant (as in [Wi]), depending on the specific circumstances.

The most interesting case is that arising in Section 3.5. There, the orthogonal spread Σ occurs at the end of an iterative process. The last step of the iteration starts with an orthogonal spread $\widehat{\Sigma}$ in a smaller–dimensional space over a field properly between F and K, forms a symplectic semifield spread $\widehat{\Sigma}_{\widehat{z}}$, and identifies this with a symplectic semifield spread Σ_z arising from our orthogonal spread Σ over the smaller field K (so $\widehat{\Sigma}_{\widehat{z}} = \Sigma_z$. In [Wi], Williams proceeds as follows: he identifies all of the nonsingular points z' such that $\Sigma_{z'}$ is a semifield spread, and then shows that z is the only such z' for which the kernel of $\mathbf{A}(\Sigma_{z'})$ is larger than K. It follows that $G(\Sigma)$ must fix z, and hence is determined by $\mathrm{Aut}\,\mathbf{A}(\Sigma_z) = \mathrm{Aut}\,\mathbf{A}(\widehat{\Sigma}_{\widehat{z}})$. Then $G(\Sigma)$ is determined by $G(\widehat{\Sigma})$ (cf. Theorem 3.3), and induction can be used. This outline is the pleasant part of the argument. The difficult part is in the implementation: calculating the kernels of planes defined using the formula (3.14).

Boring planes. A *boring plane* is a translation plane $\mathbf{A}$ of order q^m with kernel $\mathrm{GF}(q)$ such that $|\,\mathrm{Aut}\,\mathbf{A}| = q^{2m}(q-1)$ is as small as possible. The reason for this name "boring" is that such planes are contrary to those usually studied in finite geometry, in which collineation groups are assumed to be in some sense "big". The only examples known in odd characteristic are two planes of order 17^2 [Ch]. By contrast, there are too many boring planes when the characteristic is 2 [Ka4; Wi]. These planes arise as follows: as already indicated, $G(\Sigma)$ is known for many orthogonal spreads Σ. For most of these there are nonsingular points z such that $G(\Sigma)_z = 1$. In view of Theorem 3.3, this means that $\mathrm{Aut}\,\mathbf{A}(\Sigma_z)$ consists of perspectivities. All that then remains is to show that the kernel of this plane is just K. This step involves calculations that are very different in the proofs for $q = 2$ [Ka4] or $q > 2$ [Wi]. (Neither proof extends to the situation in the other part of the theorem.) Clearly this theorem still leaves open the case of other values of m, as well as the entirely different case in which m is even—and of course, the case q odd also needs to be investigated. It is very likely that there large numbers of boring planes in all of these cases as well.

Similarly, a *boring semifield plane* is a semifield plane of order q^m with kernel $\mathrm{GF}(q)$ such that $|\,\mathrm{Aut}\,\mathbf{A}| = q^{3m}(q-1)$, which again is as small as possible. Once again large numbers of these are obtained in [Wi] using (3.14) and $G(\Sigma)$ for the corresponding orthogonal spread Σ.

3.9. The number of Kerdock sets and orthogonal spreads

In view of Proposition 2.3 and Theorem 3.3, the planes we have been discussing produce exponential numbers of inequivalent Kerdock sets and orthogonal spreads. We refer to [Wi] for estimates of the numbers of these, which significantly improve previous estimates in [Ka1; Ka2].

§4. Additional uses of Kerdock sets

Symplectic and orthogonal spreads are also important for reasons quite different than the construction of planes. The basic constructions of the objects discussed presently depend on planes, which are involved in all present descriptions — with the exception of the original approach used in the construction of Kerdock codes [Ke] (and we have seen that this can be viewed as dealing with desarguesian spreads, albeit in a somewhat indirect manner).

The recent resurgence of interest in Kerdock codes (and hence of orthogonal spreads) stems from their versions over $\mathbb{Z}_4$ (Section 4.2) [CCKS; Wi].

4.1. Kerdock codes

Assume that the underlying field is $\mathbb{Z}_2$. Fix an ordering of the vectors in $\mathbb{Z}_2^n$, where n is even.

Each Kerdock set $\mathcal{K}$ determines a *Kerdock code*

$$(4.1)\qquad C(\mathcal{K}) := \{\, (Q_B(v) + sv^t + \varepsilon)_{v\in\mathbb{Z}_2^n} \mid B \in \mathcal{K}, s \in \mathbb{Z}_2^n, \varepsilon \in \mathbb{Z}_2 \,\},$$

where Q_B denotes any quadratic form whose associated bilinear form is uBv^t. The code $C(\mathcal{K})$ has length 2^n, consists of $2^{n-1}2^n2 = 2^{2n}$ codewords (i.e., vectors), and has minimum distance $2^{n-1} - 2^{(n-2)/2}$.

The resulting codes have interesting combinatorial properties, and were investigated starting in [Ke] and continuing in [Di; Ka1; Ka2; Wi]. See [MS, Ch. 15 §5; Li; CL] for further background concerning these codes.

Quasi-equivalence of codes Two binary codes of the same length will be called *quasi-equivalent* if there is an isometry of the underlying Hamming space sending one to the other. This means: permute the coordinates of the first code and then add a constant vector to all codewords in order to get the second code. The codes are called *equivalent* if only a permutation of coordinates is used. The latter is the more standard notion. However, we need the broader notion of quasi-equivalence in view of the following elementary fact: *Two Kerdock codes are quasi-equivalent if and only if they arise from equivalent Kerdock sets.*

Since we already know that there are large numbers of inequivalent Kerdock sets, it follows that the same is true for Kerdock codes.

4.2. $\mathbb{Z}_4$-codes

Each code $C(\mathcal{K})$ is nonlinear. In [CHKSS], unexpected relationships were discovered between codes over $\mathbb{Z}_4$ and binary codes, allowing the original Kerdock codes [Ke] to be viewed as codes over $\mathbb{Z}_4$ that are $\mathbb{Z}_4$–linear. This was generalized in [CCKS]: with each (binary) Kerdock code $C(\mathcal{K})$ of length 2^{m+1} is associated a $\mathbb{Z}_4$–code $C_4(\mathcal{K})$ of length 2^m that is *isometric* to $C(\mathcal{K})$, where a suitable natural metric is used on $\mathbb{Z}_4^{2^m}$: the *Lee metric* d_L. (This is defined by $d_L((a_i),(b_i)) = \sum |a_i - b_i|$, where $|a_i - b_i|$ is reduced mod 4 so as to be in $\{0,1,2\}$ and the sum is taken in $\mathbb{Z}$.)

We will define $C_4(\mathcal{K})$ using a binary operation as in Sections 3.3–3.4. By (3.7), for each $r \in F$ the map $P_r\colon x \mapsto x * r$ is self-adjoint with respect to the inner product $T(xy)$ on F. We fix an orthonormal basis for F, and view P_r as a matrix $\widehat{P_r}$ with entries 0 and 1 *in* $\mathbb{Z}_4$ *rather than* $\mathbb{Z}_2$. Similarly, we view each $x \in F$ as a row vector $\widehat{x}$ with entries $0, 1 \in \mathbb{Z}_4$. If $\mathcal{K}$ denotes the Kerdock set given in Proposition 3.11, then

$$C_4(\mathcal{K}) := \{\widehat{x}\widehat{P_r}\widehat{x}^t + 2\widehat{s}\cdot\widehat{x} + \varepsilon)_{x\in F} \mid r \in F, s \in F, \varepsilon \in \mathbb{Z}_4\}$$

is a $\mathbb{Z}_4$–*Kerdock code.* The similarity of this definition to (4.1) is evident.

Moreover, $C_4(\mathcal{K})$ *is* $\mathbb{Z}_4$*–linear if and only if* $*$ *is* 2*–sided distributive.* Part of this is easy to see: suppose that $*$ is 2–sided distributive. Then, for any $s, s' \in F$, $\widehat{P}_{s+s'} - \widehat{P}_s - \widehat{P}_{s'}$ is twice a symmetric matrix, and hence $x \mapsto \widehat{x}[\widehat{P}_{s+s'} - \widehat{P}_s - \widehat{P}_{s'}]\widehat{x}^t$ is additive from F to $2\mathbb{Z}_4$: it looks like $x \mapsto 2\widehat{r} \cdot \widehat{x}$ for some $r \in F$. Thus, semifields enter coding theory. These results, and a thorough discussion of equivalences among these $\mathbb{Z}_4$–codes, can be found in [CCKS; Wi].

If P is a symmetric binary $m \times m$ matrix then the map $x \mapsto \widehat{x}\widehat{P}\widehat{x}^t$ is called a $\mathbb{Z}_4$*–valued quadratic form* [Br]. In view of the above connection, it appears that the $\mathbb{Z}_4$–module of all of these needs to be investigated from a combinatorial point of view (cf. [Wo]).

4.3. Further topics

[CCKS] and [Wi] discuss relationships between Kerdock sets, extraspecial 2–groups, and extremal line–sets in real and complex vector spaces.

Symplectic and orthogonal spreads produce other types of combinatorial objects: partial geometries [DDT] or strongly regular graphs [Ka3].

Relationships of symplectic and orthogonal spreads with Lie algebras are surveyed in [Ka5].

Finally, Kerdock codes over $\mathbb{Z}_2$ and $\mathbb{Z}_4$ have suggested natural variations: codes over the quaternion group of order 8 [Ka6].

References

[Br] E. H. Brown, Generalizations of Kervaire's invariant, Ann. of Math., **95** (1972), 368–383.

[CCKS] A. R. Calderbank, P. J. Cameron, W. M. Kantor and J. J. Seidel, $\mathbb{Z}_4$–Kerdock codes, orthogonal spreads, and extremal Euclidean line–sets (submitted).

[CHKSS] A. R. Calderbank, A. R. Hammons, Jr., P. V. Kumar, N. J. A. Sloane, and P. Solé, The $\mathbb{Z}_4$–linearity of Kerdock, Preparata, Goethals and related codes, IEEE Trans. Inform. Theory, **40** (1994), 301–319.

[CL] P. J. Cameron and J. H. van Lint, "Designs, Graphs, Codes and their Links", London Math. Soc. Student Texts 22, Cambridge Univ. Press, Cambridge, 1991.

[Ch] C. Charnes, A non–symmetric translation plane of order 17^2. J. Geom., **37** (1990), 77–83.

[DDT] F. DeClerck, R. H. Dye and J. A. Thas, An infinite class of partial geometries associated with the hyperbolic quadric in $PG(4n-1,2)$, Europ. J. Combin.,**1** (1980), 323–326.

[De] P. Dembowski, "Finite Geometries", Springer, Berlin-Heidelberg-New York, 1968.

[Di] J. F. Dillon, Elementary Hadamard difference sets, Ph.D. thesis, U. of Maryland, 1974.

[Ka1] W. M. Kantor, Spreads, translation planes and Kerdock sets, I,II, SIAM J. Alg. Discr. Math., **3** (1982), 151–165 and 308–318.

[Ka2] W. M. Kantor, An exponential number of generalized Kerdock codes, Inform. and Control,**53** (1982), 74–80.

[Ka3] W. M. Kantor, Strongly regular graphs defined by spreads, Israel J. Math., **41** (1982), 298–312.

[Ka4] W. M. Kantor, Projective planes of order q whose collineation group has order q^2, J. Algebraic Combin., **3** (1994) 405–425.

[Ka5] W. M. Kantor, Note on Lie algebras, finite groups and finite geometries (to appear in: Proc. Ohio State U. Groups and Geometries Conference).

[Ka6] W. M. Kantor, Quaternionic line-sets and quaternionic Kerdock codes, Linear Algebra Appl., **226–228** (1995), 749–779.

[KW] W. M. Kantor and M. E. Williams, New flag–transitive affine planes of even order (to appear in J. Combin. Theory Ser. A).

[Ke] A. M. Kerdock, A class of low–rate nonlinear binary codes, Inform. and Control, **20** (1972), 182–187.

[Li] J. H. van Lint, Kerdock and Preparata codes, Congr. Numer., **39** (1983), 25–41.

[MS] F. J. MacWilliams and N. J. A. Sloane, "The theory of error-correcting codes", North–Holland, Amsterdam 1977.

[Wi] M. E. Williams, $\mathbb{Z}_4$–linear Kerdock codes, orthogonal geometries, and non–associative division algebras, Ph.D. thesis, University of Oregon, 1995.

[Wo] J. Wood, Witt's extension theorem for mod four valued quadratic forms, Trans. Amer. Math. Soc., **336** (1993), 445–461.

Department of Mathematics
University of Oregon
Eugene, OR 97403
U.S.A.

Advanced Studies in Pure Mathematics 24, 1996
Progress in Algebraic Combinatorics
pp. 243–264

Level-Rank Duality of Witten's 3-Manifold Invariants

Toshitake Kohno and Toshie Takata

§1. Introduction

The main object of this paper is to establish a duality satisfied by Witten's 3-manifold invariants for $sl(n, \mathbf{C})$ at level k and those for $sl(k, \mathbf{C})$ at level n. This type of duality, which is called the level-rank duality, has been encountered in several contexts in solvable lattice models and conformal field theory — the Boltzmann weights of solvable lattice models [JMO], quantum groups at roots of unity [SA], link invariants related to Chern-Simons gauge theory [NRS], fusion algebras [KN], and the space of conformal blocks in Wess-Zumino-Witten conformal field theory [NT]. This subject has been also treated by many other authors from different viewpoints. More recently, Witten [W2] described the relationship between the fusion algebra and the quantum cohomology of the Grassmann manifold and explained the level-rank duality from this point of view. However, a precise formulation for the level-rank duality of Witten's 3-manifold invariants has not appeared in the literature, as far as the authors know.

Let M be a closed oriented 3-manifold and we denote by $Z_k(M, SU(n))$ Witten's 3-manifold invariant for $sl(n, \mathbf{C})$ at level k discovered in the seminal article [W1]. Subsequently these invariants were studied in detail in [RT] and [KM]. Our notation corresponds to $\tau_r(M)$ in [KM] with $r = k + 2$ in the case $n = 2$. To describe the duality between $Z_k(M, SU(n))$ and $Z_n(M, SU(k))$ we first factorize the invariants by the Dynkin diagram automorphism. This is the $sl(n, \mathbf{C})$ counterpart of the symmetry principle discovered in [KM] for $sl(2, \mathbf{C})$. In this paper we assume that the integers n and k are relatively prime. Let us suppose that M is obtained by the Dehn surgery on a framed link L in S^3. We recall that the invariant $Z_k(M, SU(n))$ is written as a weighted sum of link invariants obtained by associating with each component of the link

Received September 25, 1995.
Revised December 22, 1995.

a dominant integral weight for $sl(n, \mathbf{C})$ at level k (see 4.1.1). By means of the $\mathbf{Z}_n$ action on the above set of weights, we show that the invariant $Z_k(M, SU(n))$ can be written in the form $\xi_{n,k}(M) Z_k(M, PSU(n))$, where $\xi_{n,k}(M)$ is an invariant of M defined by the linking matrix.

The level-rank duality holds for the above $PSU(n)$ invariant. In Theorem 4.2.7, we prove the duality relation

$$Z_k(M, PSU(n)) = \overline{Z_n(M, PSU(k))}.$$

The argument of our proof also permits us to describe the level-rank duality of representations of the mapping class groups on the space of conformal blocks. We show that the two representations of the mapping class groups, one for $sl(n, \mathbf{C})$ at level k and one for $sl(k, \mathbf{C})$ at level n are contragredient to each other.

The paper is organized in the following way. In Section 2, we start from recalling basic properties of the fusion algebra of type A. We introduce the fusion algebra as a truncated representation ring of the Lie algebra $sl(n, \mathbf{C})$. Then, we define the cyclic group action on the set of dominant integral weights at level k, which is induced from the Dynkin diagram automorphism of the corresponding affine Lie algebra. By considering the orbit of this action we describe the level-rank duality for the fusion algebra. Although it is not directly used in the article, for the reader's convenience, we explain briefly the appearance of the fusion algebra in conformal field theory at the end of Section 2. In Section 3, we explain the level-rank duality of link invariants. Given a framed link in S^3 we associate a dominant integral weight at level k for each component. We define an invariant of the above colored framed link, which is a generalization of the Jones polynomial at roots of unity. We describe the level-rank duality for the invariants of colored framed links. Section 4 is devoted to the proof of the level-rank duality of $PSU(n)$ invariants of 3-manifolds. In the case n and k are relatively prime we can factorize the $SU(n)$ invariant at level k by the action of the Dynkin diagram automorphism to get the $PSU(n)$ invariant. We prove the duality for the $PSU(n)$ invariant at level k and the $PSU(k)$ invariant at level n.

Acknowledgement. The authors would like to thank Hitoshi Murakami for a stimulating correspondence. This work was partially supported by Grant-in-Aid for Scientific Research on Priority Areas 231 "Infinite Analysis".

§2. Duality of fusion algebras

2.1. Definition of the fusion algebra of type A

In this section we summarize basic facts about the fusion algebra of type A. First, let us fix some notation. Let $\mathbf{h}$ denote the Cartan subalgebra of $sl(n, \mathbf{C})$ fixed as the set of diagonal matrices in $sl(n, \mathbf{C})$. The linear form $\varepsilon_i : \mathbf{h} \to \mathbf{C}$, $1 \le i \le n$, is defined by associating to $X \in \mathbf{h}$ its (i,i) component X_{ii}. The set of roots of $sl(n, \mathbf{C})$ is given by

$$\Delta = \{\varepsilon_i - \varepsilon_j \ ; \ 1 \le i \neq j \le n\}.$$

We fix the set of fundamental roots as

$$\alpha_i = \varepsilon_i - \varepsilon_{i+1}, \quad 1 \le i \neq j \le n-1.$$

The Cartan-Killing form induces an inner product on $\mathbf{h}^*$ defined by

$$\langle \varepsilon_i, \varepsilon_j \rangle = \delta_{ij}.$$

The fundamental system of weights is

$$\Lambda_i = (\varepsilon_1 + \cdots + \varepsilon_i) - \frac{i}{n} \sum_{j=1}^{n} \varepsilon_j, \quad 1 \le i \le n,$$

which is characterized by

$$\langle \Lambda_i, \alpha_j \rangle = \delta_{ij}.$$

We denote by ρ the half sum of the positive roots $\frac{1}{2} \sum_{i<j} (\varepsilon_i - \varepsilon_j)$, which is equal to $\sum_{i=1}^{n-1} \Lambda_i$.

By definition the set of dominant integral weights of $sl(n, \mathbf{C})$ is given by

$$P_+(n) = \{\sum_{i=1}^{n-1} a_i \Lambda_i \ ; \ a_i \in \mathbf{Z}, a_i \ge 0\}.$$

For $\lambda = \sum_{i=1}^{n-1} a_i \Lambda_i \in P_+(n)$ we put $|\lambda| = \sum_{i=1}^{n-1} i a_i$, which is the number of nodes in the corresponding Young diagram. Here $|\lambda|$ is considered modulo n. We set

$$P = \mathbf{Z}\Lambda_1 \oplus \cdots \oplus \mathbf{Z}\Lambda_{n-1},$$

which is called the weight lattice. The Weyl group W acts on P by means of the reflections s_{α_i}, $1 \le i \le n-1$, defined by

$$s_{\alpha_i}(v) = v - \langle v, \alpha_i \rangle \alpha_i.$$

A dominant integral weight λ is characterized by the property

$$\langle \lambda, \alpha_i \rangle \in \mathbf{Z}, \quad \langle \lambda, \alpha_i \rangle \geq 0, \quad 1 \leq i \leq n-1$$

and the set $P_+(n)$ is in one-to-one correspondence with the set of finite dimensional irreducible representations of $sl(n, \mathbf{C})$.

We denote by V_λ the irreducible representation of $sl(n, \mathbf{C})$ with highest weight $\lambda \in P_+(n)$. As a representation space of $sl(n, \mathbf{C})$ the tensor product $V_\lambda \otimes V_\mu$ admits a decomposition

$$V_\lambda \otimes V_\mu = \sum_\nu M_{\lambda\mu}^\nu \otimes V_\nu,$$

where the multiplicity $n_{\lambda\mu}^\nu = \dim M_{\lambda\mu}^\nu$ is called the Littlewood-Richardson coefficient. Let us recall that the representation ring R_n is a free $\mathbf{Z}$ module with basis $\lambda \in P_+(n)$ equipped with the product structure defined by

$$\lambda \cdot \mu = \sum_\nu n_{\lambda\mu}^\nu \, \nu \quad \text{in} \quad R_n.$$

It is well known that R_n is isomorphic to the polynomial ring

$$\mathbf{Z}[\Lambda_1, \ldots, \Lambda_{n-1}].$$

Let us now introduce the fusion algebra of type A_{n-1} in a combinatorial manner. The set of dominant integral weights at level k is by definition

$$P_+(n,k) = \{\sum_{i=1}^{n-1} a_i \Lambda_i \,;\, a_i \in \mathbf{Z},\, a_i \geq 0,\, \sum_{i=1}^{n-1} a_i \leq k\}.$$

Let us consider the natural inclusion $j : P_+(n,k) \to P_+(n,k+1)$ and we put $\partial P(n,k) = P_+(n,k+1) \setminus P_+(n,k)$. Let $I_{n,k}$ be the ideal of R_n generated by the elements of $\partial P(n,k)$. We define the fusion algebra $R_{n,k}$ as the truncated representation ring $R_n / I_{n,k}$. It is known by [GW] that the fusion algebra $R_{n,k}$ is a free $\mathbf{Z}$ module whose basis is in one-to-one correspondence with the set $P_+(n,k)$. Let us denote the basis by the same letter $\lambda \in P_+(n,k)$. We define the fusion rule $N_{\lambda\mu}^\nu$ to be the structure constant defined by

$$\lambda \cdot \mu = \sum N_{\lambda\mu}^\nu \nu \quad \text{in} \quad R_{n,k}.$$

Example. In the case of $sl(2, \mathbf{C})$, the representation ring is the polynomial ring $\mathbf{Z}[\Lambda_1]$. The representation with highest weight $m\Lambda_1$ can be written as the polynomial

$$P_m(\Lambda_1) = \sum_{i=0}^{[m/2]} (-1)^i \binom{m-1}{i} \Lambda_1^{m-2i}.$$

The fusion algebra $R_{2,k}$ is by definition $\mathbf{Z}[\Lambda_1]/(P_{k+1}(\Lambda_1))$. Let us denote by v_j the element corresponding to the representation with highest weight $2j\Lambda_1$. Then the structure constant $N_{j_1 j_2}^{j_3}$ is 1 if the condition

$$|j_1 - j_2| \leq j_3 \leq j_1 + j_2,\ j_1 + j_2 + j_3 \in \mathbf{Z},\ j_1 + j_2 + j_3 \leq k$$

is satisfied and is 0 otherwise.

We have an involution on $P_+(n,k)$ defined by $\lambda^* = -w(\lambda)$ where w denotes the longest element in the Weyl group. We put $N_{\lambda\mu\nu} = N_{\lambda\mu}^{\nu^*}$. We have the following basic properties of the structure constant $N_{\lambda\mu}^{\nu}$.

(1) $0 \leq N_{\lambda\mu}^{\nu} \leq n_{\lambda\mu}^{\nu}$
(2) $N_{\lambda\mu}^{\nu}$ is symmetric with respect to λ, μ and ν.
(3) $N_{0\mu}^{\nu} = \delta_{\mu\nu}$

2.2. Dynkin diagram automorphism

Let $\widehat{\Lambda}_i, 0 \leq i \leq n-1$, denote the fundamental weights of the affine Lie algebra $\widehat{sl(n, \mathbf{C})}$. We define the set of dominant integral weights at level k by

$$\widehat{P}_+(n,k) = \{\sum_{i=0}^{n-1} a_i \widehat{\Lambda}_i\,;\ a_i \in \mathbf{Z},\ a_i \geq 0,\ \sum_{i=0}^{n-1} a_i = k\}.$$

We have a natural injection $j : \widehat{P}_+(n,k) \to P_+(n)$ defined by $j(\lambda) = \sum_{i=1}^{n-1} a_i \Lambda_i$ for $\lambda = \sum_{i=0}^{n-1} a_i \widehat{\Lambda}_i$ and its image is equal to $P_+(n,k)$. It is known that $P_+(n,k)$ is in one to one correspondence with the set of integrable highest weight representations of the affine Lie algebra $\widehat{sl(n, \mathbf{C})}$ (see [K]).

The cyclic group $\mathbf{Z}_n$ acts on the set $\widehat{P}_+(n,k)$ by the Dynkin diagram automorphism

$$\sigma(\widehat{\Lambda}_i) = \widehat{\Lambda}_{i+1},$$

where the suffix is taken modulo n. By means of the identification using j, the Dynkin diagram automorphism induces a $\mathbf{Z}_n$ action on $P_+(n,k)$. More explicitly, this action is defined by

$$\sigma(\sum_{i=1}^{n-1} a_i\Lambda_i) = (k - \sum_{i=1}^{n-1} a_i)\Lambda_1 + \sum_{j=2}^{n-1} a_{j-1}\Lambda_j.$$

Lemma 2.2.1. *We have a bijection*

$$P_+(n,k)/\mathbf{Z}_n \to P_+(k,n)/\mathbf{Z}_k.$$

Proof. To see this correspondence it is convenient to consider for $\hat{\lambda} \in \widehat{P}_+(n,k)$ the sum $\hat{\lambda} + \sum_{i=0}^{n-1} \hat{\Lambda}_i$, which is expressed as $\alpha_0\hat{\Lambda}_0 + \cdots + \alpha_{n-1}\hat{\Lambda}_{n-1}$ with $\sum_{i=0}^{n-1} \alpha_i = k+n$. Let us consider a circle of the circumference $n+k$ and divide the circle with $n+k$ points in such a way that the length of each divided arc is equal to 1. We take n of the above $n+k$ points so that the lengths of the arcs are $\alpha_0, \ldots, \alpha_{n-1}$. Considering up to rotation, the set of ways of dividing circles in the above manner are in one to one correspondence with $P_+(n,k)/\mathbf{Z}_n$. Now let us take the complementary k points on the circle, which defines an element of $P_+(k,n)/\mathbf{Z}_k$. It is clear that this gives a bijection between $P_+(n,k)/\mathbf{Z}_n$ and $P_+(k,n)/\mathbf{Z}_k$.

Let us suppose that n and k are relatively prime. Since we have $|\sigma(\lambda)| \equiv |\lambda| \bmod n$, each orbit of the $\mathbf{Z}_n$ action on $P_+(n,k)$ contains a unique dominant integral weight λ of level k with $|\lambda| \equiv 0 \bmod n$. The orbit space $P_+(n,k)/\mathbf{Z}_n$ is identified with the set

$$\Omega_{n,k} = \{\lambda \in P_+(n,k) \ ; \ |\lambda| \equiv 0 \bmod n\}.$$

For $\lambda = \sum_{i=1}^{n-1} a_i\Lambda_i \in P_+(n,k)$ we associate the Young diagram of type $(a_1 + \cdots + a_{n-1}, a_2 + \cdots + a_{n-1}, \ldots, a_{n-1})$. We express λ as

$$\lambda = a_{i_1}\Lambda_{i_1} + \cdots + a_{i_s}\Lambda_{i_s}, \quad a_{i_1}, \ldots, a_{i_s} \neq 0.$$

Let us consider the transposed Young diagram and the associated weight ${}^t\lambda \in P_+(k,n)$ given by

$${}^t\lambda = (i_s - i_{s-1})\Lambda_{\alpha_{i_s}} + (i_{s-1} - i_{s-2})\Lambda_{\alpha_{i_s}+\alpha_{i_{s-1}}} + \cdots + i_1\Lambda_{\alpha_{i_1}+\cdots+\alpha_{i_s}}.$$

We can easily check the following lemma.

Lemma 2.2.2. *In the case n and k are relatively prime, we have a bijection*

$$\tau : \Omega_{n,k} \to \Omega_{k,n}$$

given by

$$\tau(\lambda) = \sigma^{k-|\lambda|/n}({}^t\lambda).$$

2.3. Verlinde formula

In this paragraph, we recall the Verlinde formula which relates the fusion rule and the modular transformation S matrix. For λ, $\mu \in P_+(n,k)$, we set

$$S_{\lambda\mu} = \frac{(\sqrt{-1})^{n(n-1)/2}}{\sqrt{n(k+n)^{n-1}}} \sum_{w \in W} \det w \exp\left(-\frac{2\pi\sqrt{-1}}{k+n} \langle w(\lambda+\rho), \mu+\rho \rangle \right),$$

$$T_{\lambda\mu} = \delta_{\lambda\mu} \exp 2\pi\sqrt{-1} \left(\Delta_\lambda - \frac{c_{n,k}}{24} \right),$$

where $c_{n,k}$ is the so-called central charge

$$c_{n,k} = \frac{k \dim sl(n, \mathbf{C})}{k+n}.$$

For $\lambda \in P_+(n,k)$ we denote by $\mathcal{H}_\lambda$ the integrable highest weight module of the affine Lie algebra $\widehat{sl(n,\mathbf{C})}$ at level k. The character χ_λ is by definition

$$\chi_\lambda(\tau) = Tr_{\mathcal{H}_\lambda} q^{L_0 - \frac{c_{n,k}}{24}}$$

where L_0 is the 0-th Sugawara operator and $q = e^{2\pi\sqrt{-1}\tau}$ with Im $q > 0$. Let us recall the following fundamental result due to Kac and Peterson.

Theorem 2.3.1 [KP]. *The set of characters χ_λ, $\lambda \in P_+(n,k)$, are invariant under the modular transformation and they satisfy*

$$\chi_\lambda(-1/\tau) = \sum_{\mu \in P_+(n,k)} S_{\lambda\mu}\, \chi_\mu(\tau),$$

$$\chi_\lambda(\tau+1) = \exp 2\pi\sqrt{-1} \left(\Delta_\lambda - \frac{c_{n,k}}{24} \right) \chi_\lambda(\tau).$$

The matrices $S = (S_{\lambda\mu})$ and $T = (T_{\lambda\mu})$ are unitary and symmetric and satisfy the relation

$$(ST)^3 = S^2 = (\delta_{\lambda\mu^*}),$$

where $\lambda^* = -w(\lambda)$ with the longest element $w \in W$ and is the highest weight for the dual representation V_λ^*.

Let us observe that $S_{0\lambda}$ is a real number and is written as

$$S_{0\lambda} = \frac{1}{\sqrt{n(k+n)^{n-1}}} \prod_{\alpha \in \Delta_+} 2 \sin \frac{\pi \langle \lambda + \rho, \alpha \rangle}{k+n}.$$

For our purpose the following expression for the modular transformation S matrix is also useful. We put

$$\lambda + \rho = \sum_{i=1}^{n} x_i \varepsilon_i, \quad \mu + \rho = \sum_{i=1}^{n} y_i \varepsilon_i.$$

Then we have

$$S_{\lambda\mu} = \frac{(\sqrt{-1})^{n(n-1)/2}}{\sqrt{n(k+n)^{n-1}}} \sum_{\sigma \in S_n} \det \sigma \; \zeta^{x_{\sigma(1)} y_1} \cdots \zeta^{x_{\sigma(n)} y_n}$$

with $\zeta = \exp \frac{2\pi\sqrt{-1}}{k+n}$, which is a minor of the Vandermonde determinant.

The Verlinde formula relates the fusion rule and the modular transformation S matrix as

$$N_{\lambda\mu}^{\nu} = \sum_{\alpha \in P_+(n,k)} \frac{S_{\lambda\alpha} S_{\mu\alpha} \overline{S_{\nu\alpha}}}{S_{0\alpha}} \tag{2.3.2}$$

(see [K] for details).

2.4. Behaviour of S, N and Δ under σ and τ

We put

$$\Delta_\lambda = \frac{\langle \lambda, \lambda + 2\rho \rangle}{2(k+n)},$$

which is the value of the operator L_0 on the highest weight representation V_λ.

Lemma 2.4.1 [KT]. *With respect to the action σ, the conformal weight Δ and the modular transformation matrix S satisfy*

$$(1) \quad \Delta_{\sigma(\lambda)} - \Delta_\lambda = \frac{1}{n} \left(\frac{(n-1)k}{2} - |\lambda| \right),$$

$$(2) \quad S_{\sigma(\lambda)\mu} = \exp \left(\frac{2\pi\sqrt{-1}|\mu|}{n} \right) S_{\lambda\mu}.$$

Now we describe the action of τ on the modular transformation matrix $S_{\lambda\mu}$, the structure constant of the fusion algebra $N_{\lambda\mu}^{\nu}$ and the conformal weight Δ_{λ}. To specify the rank and the level, we denote them by $S_{\lambda\mu}[n,k]$, $N_{\lambda\mu}^{\nu}[n,k]$ and $\Delta_{\lambda}[n,k]$ for the highest weights $\lambda,\mu,\nu \in P_{+}(n,k)$.

Lemma 2.4.2. *For $\tau : \Omega_{n,k} \to \Omega_{k,n}$, we have the following properties.*

$$\begin{aligned}
&(1) \quad S_{\lambda\mu}[n,k] = \sqrt{\frac{k}{n}}\,\overline{S_{\tau(\lambda)\tau(\mu)}[k,n]} \\
&(2) \quad N_{\lambda\mu}^{\nu}[n,k] = N_{\tau(\lambda)\tau(\mu)}^{\tau(\nu)}[k,n] \\
&(3) \quad \Delta_{\lambda}[n,k] + \Delta_{\tau(\lambda)}[k,n] = 0 \mod \mathbf{Z}
\end{aligned}$$

Proof. (1) By Lemma 4 in [KN], we have

$$S_{\lambda\mu}[n,k] = \sqrt{\frac{k}{n}} \exp \frac{2\pi\sqrt{-1}}{nk}|\lambda||\mu|\overline{S_{{}^{t}\lambda{}^{t}\mu}[k,n]}$$

for $\lambda,\mu \in P_{+}(n,k)$. On the other hand, it follows from the property of $S_{\lambda\mu}$ in Lemma 2.4.1 that

$$\begin{aligned}
S_{\tau(\lambda)\tau(\mu)}[k,n] &= \exp 2\pi\sqrt{-1}\left(k - \frac{|\mu|}{n}\right)\frac{|\lambda|}{k} S_{{}^{t}\lambda{}^{t}\mu}[k,n] \\
&= \exp 2\pi\sqrt{-1}\left(-\frac{|\lambda||\mu|}{nk}\right) S_{{}^{t}\lambda{}^{t}\mu}[k,n].
\end{aligned}$$

Thus we obtain the assertion (1).

(2) Using the assertion (1), we obtain

$$\begin{aligned}
&n\,\frac{S_{\lambda\varepsilon}[n,k]S_{\mu\varepsilon}[n,k]\overline{S_{\nu\varepsilon}[n,k]}}{S_{0\varepsilon}[n,k]} \\
&\qquad = k\,\frac{\overline{S_{\tau(\lambda)\tau(\varepsilon)}[k,n]}\ \ \overline{S_{\tau(\mu)\tau(\varepsilon)}[k,n]}S_{\tau(\nu)\tau(\varepsilon)}[k,n]}{\overline{S_{\tau(0)\tau(\varepsilon)}[k,n]}}
\end{aligned}$$

for $\lambda,\mu,\nu,\varepsilon \in \Omega_{n,k}$. By the Verlinde formula we have

$$\begin{aligned}
N_{\lambda\mu}^{\nu}[n,k] &= \sum_{\alpha\in P_{+}(n,k)} \frac{S_{\lambda\alpha}[n,k]S_{\mu\alpha}[n,k]\overline{S_{\nu\alpha}[n,k]}}{S_{0\alpha}[n,k]} \\
&= \sum_{\varepsilon\in\Omega_{n,k}}\sum_{0\le x\le n-1} \frac{S_{\lambda\sigma^{x}(\varepsilon)}[n,k]S_{\mu\sigma^{x}(\varepsilon)}[n,k]\overline{S_{\nu\sigma^{x}(\varepsilon)}[n,k]}}{S_{0\sigma^{x}(\varepsilon)}[n,k]}.
\end{aligned}$$

Since $|\lambda|, |\mu|, |\nu| \equiv 0 \bmod n$, we have

$$S_{\lambda\sigma^x(\varepsilon)} = S_{\lambda\epsilon}, \quad S_{\mu\sigma^x(\varepsilon)} = S_{\mu\varepsilon}, \quad S_{\nu\sigma^x(\varepsilon)} = S_{\nu\varepsilon}.$$

This implies

$$N_{\lambda\mu}^{\nu}[n,k] = \sum_{\varepsilon\in\Omega_{n,k}} n\, \frac{S_{\lambda\varepsilon}[n,k] S_{\mu\varepsilon}[n,k] \overline{S_{\nu\varepsilon}[n,k]}}{S_{0\varepsilon}[n,k]}.$$

Combining with a similar formula for $N_{\tau(\lambda)\tau(\mu)}^{\tau(\nu)}[k,n]$ we obtain the assertion.

(3) For $\lambda = \sum_{i=1}^{n-1} a_i\Lambda_i \in P_+(n,k)$, we associate the Young diagram of type $(x_1,\ldots,x_n)$ with $x_i = \sum_{j=i}^{n-1} a_j, 1 \le i \le n-1$, and $x_n = 0$. We put $|\lambda| = \sum_{i=1}^{n} x_i$. In a similar way, we denote by $(y_1,\ldots,y_k)$ the Young diagram for ${}^t\lambda$. We have

$$\begin{aligned}
\langle \lambda, \lambda + 2\rho\rangle &= \langle \lambda+\rho, \lambda+\rho\rangle - \langle \rho,\rho\rangle \\
&= \sum_{i=1}^{n} \left(x_i - i - \frac{|\lambda|}{n} + \frac{n+1}{2} \right)^2 - \frac{n(n^2-1)}{12} \\
&= \sum_{i=1}^{n} x_i^2 - 2\sum_{i=1}^{n} i x_i - \frac{|\lambda|^2}{n} + (n+1)|\lambda|.
\end{aligned}$$

This implies

$$\begin{aligned}
\langle \lambda, \lambda+2\rho\rangle + \langle {}^t\lambda, {}^t\lambda + 2\rho\rangle =& \sum_{i=1}^{n} x_i^2 - 2\sum_{i=1}^{n} i x_i + \sum_{j=1}^{k} y_j^2 - 2\sum_{j=1}^{k} j y_j \\
&- \frac{|\lambda|^2}{nk}(n+k) + (n+k)|\lambda| + 2|\lambda|.
\end{aligned}$$

According to Lemma 2 in [KN], the $n+k$ numbers $x_i + n - i$ $(1 \le i \le n)$, and $n-1+j-y_j$ $(1 \le j \le k)$ are both obtained as a permutation of $\{0,1,\ldots,n+k-1\}$. Since

$$\sum_{i=1}^{n} x_i^2 - 2\sum_{i=1}^{n} i x_i + \sum_{j=1}^{k} y_j^2 - 2\sum_{j=1}^{k} j y_j = -2|\lambda|,$$

we obtain the equation

$$\Delta_\lambda + \Delta_{{}^t\lambda} = \frac{|\lambda|}{2}\left(1 - \frac{|\lambda|}{nk}\right).$$

Applying Lemma 2.4.1, one has

$$\Delta_{\tau(\lambda)} - \Delta_{{}^t\lambda} = \frac{|\lambda|}{2}\left(\frac{|\lambda|}{nk} - 1\right) \mod \mathbf{Z}.$$

Thus we obtain the assertion $\Delta_\lambda + \Delta_{\tau(\lambda)} = 0 \mod \mathbf{Z}$.

2.5. Fusion algebras in conformal field theory

The fusion algebra was originally defined in the context of conformal field theory (see [V], [TUY] and [B]). Let us describe briefly the fusion algebra associated with the $SU(n)$ Wess-Zumino-Witten model at level k on the Riemann sphere. We fix a coordinate function t on the Riemann sphere $\mathbf{C}P^1$. Let $P_1, \ldots, P_m$ be distinct points of $\mathbf{C}P^1$ with $t(P_j) = \xi_j, \xi_j \neq 0, 1 \leq j \leq m$ and to each point we associate $\lambda_1, \ldots, \lambda_m \in P_+(n,k)$.

Let T denote the endomorphism on $V_{\lambda_1} \otimes \cdots \otimes V_{\lambda_m}$ defined by

$$T(v_1 \otimes \cdots \otimes v_m) = \sum_{i=1}^m \xi_i v_1 \otimes \cdots \otimes X_\theta v_i \otimes \cdots \otimes v_m$$

where X_θ is associated with the longest root θ. More explicitly $X_\theta \in sl(n, \mathbf{C})$ is written as E_{1n} where E_{ij} is the matrix unit such that ij component is 1 and the other components are 0. We define

$$V^\dagger_{\mathbf{C}P^1}(P_1, \ldots, P_m; \lambda_1, \ldots, \lambda_m)$$

to be the space of $sl(n, \mathbf{C})$ invariant m-linear form

$$\varphi : V_{\lambda_1} \otimes \cdots \otimes V_{\lambda_m} \to \mathbf{C}$$

satisfying $\varphi \circ T^{k+1} = 0$. It turns out that the above vector space is isomorphic to the dual of the space of conformal blocks (see [TUY] and [B]). We put

$$N(\lambda_1, \ldots, \lambda_m) = \dim V^\dagger_{\mathbf{C}P^1}(P_1, \ldots, P_m; \lambda_1, \ldots, \lambda_m),$$

which is related to the structure constant of the fusion algebra by

$$N^\nu_{\lambda\mu} = N(\lambda, \mu, \nu^*).$$

§3. Level-rank duality of link invariants

3.1. Link invariants via face Boltzmann weights

Let us recall some basic properties of the link invariants for oriented framed links associated with $sl(n, \mathbf{C})$. For a more detailed description we refer the readers to [TW] and [KT]. Let L be an oriented framed link in S^3 with m components. To each component L_i, $1 \leq i \leq m$, we assign a highest weight $\lambda_i \in P_+(n,k)$, and we denote by $J(L, \lambda_1, \ldots, \lambda_m)$ the associated invariant. We put

$$q = \exp\left(\frac{2\pi\sqrt{-1}}{k+n}\right), \quad t = \exp\left(\frac{\pi\sqrt{-1}}{n(k+n)}\right).$$

In the case $\lambda_i = \Lambda_1$, $1 \leq i \leq m$, the invariant $J_L = J(L, \Lambda_1, \ldots, \Lambda_1)$ is characterized by the skein relation

$$tJ_{L_+} - t^{-1}J_{L_-} = (q^{1/2} - q^{-1/2})J_{L_0}$$

and the condition

$$J_{\bigcirc} = \frac{q^{n/2} - q^{-n/2}}{q^{1/2} - q^{-1/2}}.$$

With respect to the following local modifications of the framed link we have

$$J \;\lambda = \frac{S_{0\lambda}}{S_{00}} J$$

$$J \;\lambda = \exp 2\pi\sqrt{-1}\Delta_\lambda \, J \;\lambda,$$

$$J \;{}^{\lambda}\;\mu = \frac{S_{\lambda\mu}}{S_{0\lambda}} J \;\lambda.$$

The invariant $J(L, \lambda_1, \ldots, \lambda_m)$ is related to the fusion algebra $R_{n,k}$ in the following way. We have a multi-linear map

$$J(L, \cdot) : R_{n,k}^{\otimes m} \to \mathbf{C}$$

by associating to $\lambda_1 \otimes \cdots \otimes \lambda_m$ the invariant $J(L, \lambda_1, \ldots, \lambda_m)$. This map is compatible with the product structure of the fusion algebra (see [Ko2]).

In view of the above multi-linear map, we first deal with the case when λ_i, $1 \leq i \leq m$, is one of the fundamental weights $\Lambda_1, \ldots, \Lambda_{n-1}$. For the purpose of describing the symmetry, it is useful to introduce an expression of the link invariants using the Boltzmann weights for face models.

Let us denote by Φ_n the set of fundamental weights $\{\Lambda_1, \ldots, \Lambda_{n-1}\}$. We consider a projection diagram of the oriented framed link L with the blackboard framing. We fix $\lambda_1, \ldots, \lambda_m \in \Phi_n$ and to each component L_i we associate λ_i. A state is a map s from the set of regions of the projection diagram to $P_+(n,k)$ such that the following admissibility conditions (1) and (2) are satisfied.

(1) If D_i and D_j are adjacent regions as shown in Figure 1, then for $\mu = s(D_i)$ and $\nu = s(D_j)$ we have $N^{\nu}_{\lambda_i \mu} \neq 0$ where λ_i is the highest weight associated to the edge incident to both D_i and D_j.

(2) For the non-compact region D_0 one has $s(D_0) = 0$.

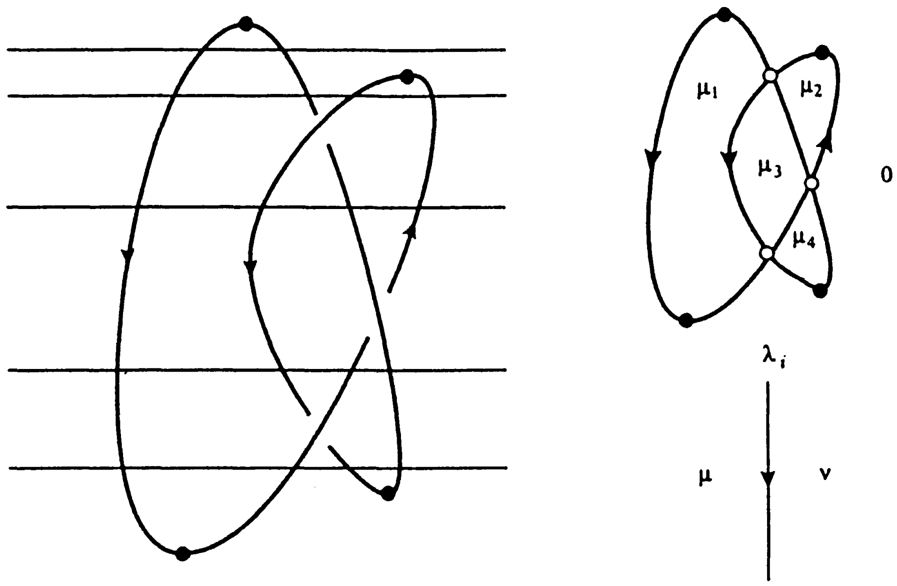


Fig. 1.

We observe that in the above situation the structure constant $N^{\nu}_{\lambda_i \mu}$ is equal to 1. This follows from the Littlewood-Richardson rule and the inequality (1) in 2.1. We have 4 kinds of vertices corresponding to overcrossing, undercrossing, creation and annihilation. For each vertex v and a state s we have a way to associate the face Boltzmann weight $W_v(s)$ such that the link invariant $J(V, \lambda_1, \ldots, \lambda_m)$ is expressed as

$$J(L, \lambda_1, \ldots, \lambda_m) = \sum_{s:\text{state}} \prod_{v:\text{vertex}} W_v(s).$$

An explicit form of such Boltzmann weights might be found in [JMO] up to some normalization.

3.2. $\mathbf{Z}_n$ symmetry for link invariants

Let us review the behaviour of the invariant $J(L, \lambda_1, \ldots, \lambda_m)$ under the action of the Dynkin diagram automorphism. A detailed account of the subject is given in [KT]. For $\lambda = (\lambda_1, \dots, \lambda_m)$ we define a map $\phi_\lambda : (\mathbf{Z}_n)^{\oplus m} \to \mathbf{Z}_{2n}$ by

$$\phi_\lambda(x_1, \dots, x_m) = nk \sum_{i=1}^{m} l_{ii} x_i - k \sum_{i,j} l_{ij} x_i x_j - 2 \sum_{i,j} |\lambda(j)| l_{ij} x_i.$$

Proposition 3.2.1 [KT]. *Using the above notation, we have*

$$\frac{J(L, \sigma^{x_1}(\lambda_1), \dots, \sigma^{x_m}(\lambda_m))}{J(L, \lambda_1, \dots, \lambda_m)} = \exp\left(\frac{\pi\sqrt{-1}}{n} \phi_\lambda(x_1, \dots, x_m)\right).$$

If $\lambda_1, \dots, \lambda_m \in \Omega_{n,k}$, then the above formula is simplified as

(3.2.2)

$$\frac{J(L, \sigma^{x_1}(\lambda_1), \dots, \sigma^{x_m}(\lambda_m))}{J(L, \lambda_1, \dots, \lambda_m)} = \exp\left(\frac{\pi\sqrt{-1}}{n}((n-1)k \sum_{i,j} l_{ij} x_i x_j)\right).$$

3.3. Level-rank duality of link invariants

As in the previous section we denote by Φ_n the set of fundamental weights for $sl(n, \mathbf{C})$. We define $\Phi_{n,k}$ to be the set of weights $\lambda \in P_+(n,k)$ such that there exists $g \in \mathbf{Z}_n$ with $g(\lambda) \in \Phi_n$. Namely, $\Phi_{n,k}$ consists of the elements which are $\mathbf{Z}_n$ equivalent to fundamental weights with respect to the Dynkin diagram automorphism.

For an oriented framed link L with m components, we consider a similar construction as in 3.1 with $\lambda_1, \dots, \lambda_m \in \Omega_{n,k}$. Let us denote by $\mathcal{S}_{n,k}$ the set of admissible states satisfying (1) and (2) in 3.1. Since $N_{\lambda\mu}^{\nu} = N_{\sigma^p(\lambda)\sigma^q(\mu)}^{\sigma^{p+q}(\nu)}$ the multiplicity appearing in this construction is also at most 1. The local Boltzmann weight behaves as follows with respect to the map $\tau : \Omega_{n,k} \to \Omega_{k,n}$.

We have a one-to-one correspondence $T : \mathcal{S}_{n,k} \to \mathcal{S}_{k,n}$ defined by $(T(s))(D) = \tau(s(D))$. The following lemma is essentially due to [JMO] (see also [SA]), where the corresponding statement is shown for the fundamental representation up to some phase factor. Computing the phase factor by means of Lemma 2.4.1 and Lemma 2.4.2, we obtain the following lemma.

Lemma 3.3.1. *For any type of vertex v in the graph associated with the link diagram we have we have*

$$W_v^{\tau(\lambda)}(s) = \overline{W_v^{\lambda}(T(s))}$$

for any admissible state $s \in \mathcal{S}_{n,k}$.

Proposition 3.3.2. *Let L be an oriented framed link with m components. For $\lambda_1, \ldots, \lambda_m \in \Omega_{n,k}$ we have*

$$J(L, \tau(\lambda_1), \ldots, \tau(\lambda_m)) = \overline{J(L, \lambda_1, \ldots, \lambda_m)}.$$

Proof. By the behaviour of the face Boltzmann weights with respect to τ it can be shown that the assertion holds if $\lambda_1, \ldots, \lambda_m \in \Phi_{n,k}$. For $\lambda, \mu \in \Omega_{n,k}$, we observe by means of the Littlewood-Richardson rule and the inequality (1) in 2.1 that $N_{\lambda\mu}^{\nu} \neq 0$ only if $\nu \in \Omega_{n,k}$. Moreover, it is clear that $\Omega_{n,k}$ is generated by $\Phi_{n,k}$ in the fusion algebra. Let us suppose that $\lambda_j \in \Omega_{n,k}$ is written as a polynomial

$$\lambda_j = P_j(\mu_1, \mu_2, \ldots), \quad \mu_1, \mu_2, \ldots \in \Phi_{n,k}.$$

We have

$$\tau(\lambda_j) = P_j(\tau(\mu_1), \tau(\mu_2), \ldots)$$

and it follows from the expression of the link invariant by the face Boltzmann weights and Lemma 3.3.1 together with the compatibility of $R_{n,k}^{\otimes m} \to \mathbf{C}$ with the product structure of the fusion algebra explained in 3.1 that the assertion holds for any $\lambda_1, \ldots, \lambda_m \in \Omega_{n,k}$. This completes the proof.

§4. $PSU(n)$ invariants of 3-manifolds

4.1. Review of the definition of $SU(n)$ invariants

Let M be a closed oriented 3-manifold. We suppose that M is obtained by the Dehn surgery on a framed link L with m components in S^3. As in [KT] we put $C_{n,k} = \exp\left(-\frac{2\pi\sqrt{-1}}{8} c_{n,k}\right)$ with the central charge $c_{n,k} = \frac{k \dim sl(n,\mathbf{C})}{k+n}$. Let m be the number of components of the link L. We consider the sum

$$\begin{aligned} &Z_k(M, SU(n)) \\ &= C_{n,k}^{sign(L)} \sum_{\lambda_1, \ldots, \lambda_m \in P_+(n,k)} S_{0\lambda_1} \cdots S_{0\lambda_m} J(L, \lambda_1, \ldots, \lambda_m) \end{aligned} \tag{4.1.1}$$

where $sign(L)$ denotes the signature of the linking matrix of L. It was shown in [KT] that $Z_k(M, SU(n))$ is a topological invariant of M (see also [TW]).

4.2. $PSU(n)$ invariants and level-rank duality

As is shown in 2.2, the set of weights $P_+(n,k)$ admits a $\mathbf{Z}_n$ action derived from the Dynkin diagram automorphism. Let us suppose that n and k are relatively prime. In this case the $\mathbf{Z}_n$ action is fixed point free and moreover, the invariant $Z_k(M, SU(n))$ admits the following factorization.

We consider the Gauss sum

$$G_{n,k} = \sum_{j=0}^{n-1} e((n-1)kj^2)$$

where $e(x)$ denotes $\exp(\pi\sqrt{-1}x/n)$. For $\lambda \in \Omega_{n,k}$ we have

$$\sum_{j=0}^{n-1} \exp 2\pi\sqrt{-1}\Delta_{\sigma^j(\lambda)} = G_{n,k} \exp 2\pi\sqrt{-1}\Delta_\lambda.$$

As in the previous paragraph, we suppose that M is a closed oriented 3-manifold obtained as the Dehn surgery on a framed link L in S^3. Let $(l_{ij})_{1\le i,j\le m}$ be the linking matrix of L. It is known by [MOO] (see also [Ko1]) that

(4.2.1)
$$\xi_{n,k}(M) = \left(\frac{\sqrt{n}}{G_{n,k}}\right)^{sign(L)} \left(\frac{1}{\sqrt{n}}\right)^m \sum_{x_1,\ldots,x_m\in\mathbf{Z}_m} e((n-1)k\sum_{i,j} l_{ij}x_ix_j)$$

is a topological invariant of M. We put

(4.2.2)
$$Z_k(M, PSU(n)) = \left(\frac{C_{n,k}G_{n,k}}{\sqrt{n}}\right)^{sign(L)} (\sqrt{n})^m \times \sum_{\lambda_1,\ldots\lambda_m\in\Omega_{n,k}} S_{0\lambda_1}\ldots S_{0\lambda_m} J(L,\lambda_1,\ldots,\lambda_m).$$

Proposition 4.2.3. *Let us suppose that n and k are relatively prime. Then we have*

$$Z_k(M, SU(n)) = \xi_{n,k}(M) Z_k(M, PSU(n)).$$

Proof. By using 3.2.2, we have

$$\begin{aligned}
&\sum_{\lambda_1,\dots,\lambda_m\in P_+(n,k)} S_{0\lambda_1}\cdots S_{0\lambda_m} J(L,\lambda_1,\dots,\lambda_m)\\
&= \sum_{\lambda_1,\dots,\lambda_m\in \Omega_{n,k}} \sum_{x_1,\dots,x_m\in \mathbf{Z}_n} S_{0\lambda_1}\cdots S_{0\lambda_m} J(L,\sigma^{x_1}(\lambda_1),\dots,\sigma^{x_m}(\lambda_m))\\
&= e((n-1)k \sum_{x_1,\dots,x_m\in \mathbf{Z}_n} l_{ij}x_ix_j)\\
&\qquad \times \sum_{\lambda_1,\dots,\lambda_m\in \Omega_{n,k}} S_{0\lambda_1}\cdots S_{0\lambda_m} J(L,\lambda_1,\dots,\lambda_m),
\end{aligned}$$

which shows the assertion.

We are going to show that $Z_k(M, PSU(n))$ is actually a topological invariant of M. This is not obvious from the above factorization, since the invariant $\xi_{n,k}(M)$ defined by the linking matrix might be equal to zero.

Lemma 4.2.4. *For $\lambda, \mu, \nu \in \Omega_{n,k}$ we have*

$$C_{n,k}G_{n,k} \sum_{\mu\in\Omega_{n,k}} S_{\lambda\mu}S_{\mu\nu} \exp 2\pi\sqrt{-1}(\Delta_\lambda+\Delta_\mu+\Delta_\nu) = S_{\lambda\nu}.$$

Proof. Let us start with the equality

$$C_{n,k} \sum_{\mu\in P_+(n,k)} S_{\lambda\mu}S_{\mu\nu} \exp 2\pi\sqrt{-1}(\Delta_\lambda+\Delta_\mu+\Delta_\nu) = S_{\lambda\nu} \tag{4.2.5}$$

which follows from the explicit form of the matrix S given in 2.3. We see that the equality (4.2.5) reveals the fact that the modular group acts on the space of characters $\chi_\lambda, \lambda \in P_+(n,k)$ as stated in Theorem 2.3.1. Since $\lambda, \nu \in \Omega_{n,k}$ we have $S_{\lambda\sigma(\mu)} = S_{\lambda\mu}$ by Lemma 2.4.1 (2). We decompose the left hand side of the equality 4.2.5 with respect to the $\mathbf{Z}_n$ action on μ. Combining with Lemma 2.4.1 (1), we obtain the desired equality.

The above lemma shows that the matrices

$$\tilde{S} = (\sqrt{n}S_{\lambda\mu})_{\lambda,\mu\in\Omega_{n,k}}$$

and

$$\tilde{T} = diag\left(\exp 2\pi\sqrt{-1}\left(\Delta_\lambda - \frac{c_{n,k}}{24}\right)G_{n,k}^{-1/3}\right)_{\lambda\in\Omega_{n,k}}$$

satisfy

$$(\tilde{S}\tilde{T})^3 = (\tilde{S})^2$$

and define a linear representation of $SL(2, \mathbf{Z})$.

We are in a position to show the following proposition.

Proposition 4.2.6. *$Z_k(M, PSU(n))$ is a topological invariant of a closed oriented 3-manifold M.*

Proof. As is explained in [RT], we show the invariance under the Kirby moves in the sense of Fenn and Rourke [FR]. It is enough to check the equality

$$C_{n,k}G_{n,k} \sum_{\mu \in \Omega_{n,k}} S_{0\mu}\, J(L', \lambda_1, \ldots, \lambda_m, \mu) = J(L, \lambda_1, \ldots, \lambda_m)$$

for the link diagram L' obtained by adding one component with framing 1 as shown in Figure 2. Here μ stands for the representation associated to the new component. Let l denote the number of strands passing through the new component in Figure 2.

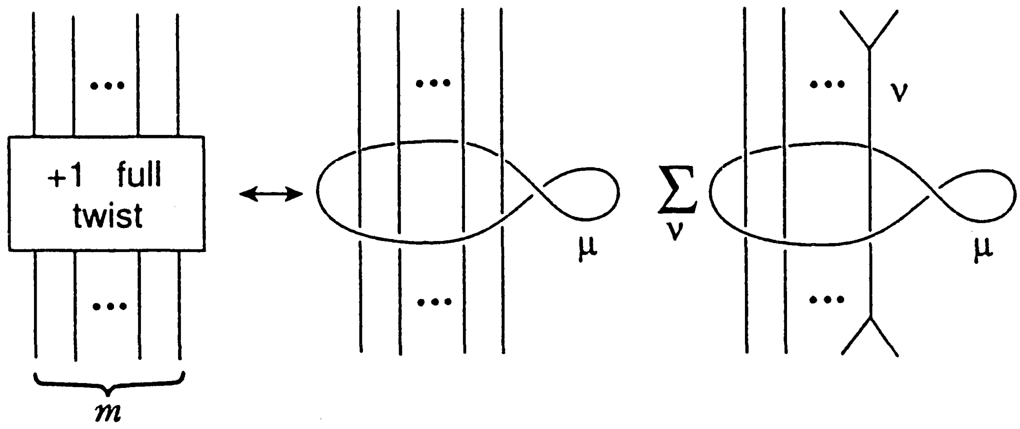


Fig. 2.

We show the assertion by induction with respect to l. As a special case of Lemma 4.2.4, we have

$$C_{n,k}G_{n,k} \sum_{\mu \in \Omega_{n,k}} S_{0\mu} S_{\lambda\mu} \exp 2\pi\sqrt{-1}(\Delta_\lambda + \Delta_\mu) = S_{0\lambda}.$$

This settles the case $l = 1$ and in particular, putting $\lambda = 0$, we obtain the case $l = 0$. Let us recall that $\Omega_{n,k}$ is closed under the product structure of the fusion algebra. Now we apply the fusion rule as shown in Figure 2, which implies that the linear map associated with the two tangles in Figure 2 are identical. This procedure permits us to reduce to the case of $l-1$ strands and we are done by induction.

Now we are in a position to show the following level-rank duality of $PSU(n)$ invariants.

Theorem 4.2.7. *Let us suppose that n and k are relatively prime. Then we have*

$$Z_k(M, PSU(n)) = \overline{Z_n(M, PSU(k))}.$$

Proof. First, we see

$$C_{n,k} = \exp(-\frac{nk-1}{4}\pi\sqrt{-1})\,\overline{C_{k,n}}$$

and

$$G_{n,k} = \exp(\frac{nk-1}{4}\pi\sqrt{-1})\sqrt{\frac{n}{k}}\,\overline{G_{k,n}}.$$

By Lemma 2.4.2 we have

$$\sqrt{n}S_{0\lambda}[n,k] = \sqrt{k}S_{0\tau(\lambda)}[k,n].$$

Here we recall that they are real numbers. Using Proposition 3.3.2, we obtain the equality

$$\begin{aligned}
&(\sqrt{n})^m \sum_{\lambda_1,\ldots,\lambda_m\in\Omega_{n,k}} S_{0\lambda_1}\cdots S_{0\lambda_m}J(L,\lambda_1,\ldots,\lambda_m)\\
=&(\sqrt{k})^m \sum_{\tau(\lambda_1),\ldots,\tau(\lambda_m)\in\Omega_{k,n}} S_{0\tau(\lambda_1)}\cdots S_{0\tau(\lambda_m)}\overline{J(L,\tau(\lambda_1),\ldots,\tau(\lambda_m))}.
\end{aligned}$$

This completes the proof.

Let us describe briefly the duality relation of the representations of the mapping class groups associated with $SU(n)$ Wess-Zumino-Witten model at level k and $SU(k)$ model at level n. Let Σ denote a closed oriented surface of genus g and $\mathcal{M}_g$ its mapping class group. Using the notation of 2.5, we put

$$\mathcal{H}_\Sigma[n,k] = \oplus_{\lambda_1,\ldots,\lambda_g\in P_+(n,k)} V^{\dagger}_{\mathbf{C}P^1}(\lambda_1,\lambda_1^*,\ldots,\lambda_g,\lambda_g^*)$$

which is isomorphic to the space of conformal blocks of $SU(n)$ Wess-Zumino-Witten model at level k in the sense of [TUY]. As in [RT] and

[Ko2], we have a projectively linear unitary representation of the mapping class group

$$\mathcal{M}_g \to GL(\mathcal{H}_\Sigma[n,k]).$$

Let us recall that this representation is expressed in terms of the Boltzmann weight of the face model by considering a $(2g,2g)$ tangle associated with an element of the mapping class group.

Let us now suppose that n and k are relatively prime. We put

$$\mathcal{H}_\Sigma[n,k]/\mathbf{Z}_n = \oplus_{\lambda_1,\ldots,\lambda_g \in \Omega_{n,k}} V^\dagger_{\mathbf{C}P^1}(\lambda_1,\lambda_1^*,\ldots,\lambda_g,\lambda_g^*).$$

Our previous construction permits us to construct a projectively linear unitary representation

$$\rho_g[n,k] : \mathcal{M}_g \to GL(\mathcal{H}_\Sigma[n,k]/\mathbf{Z}_n).$$

In particular, in the case $g = 1$ this representation is identical to the representation considered in the remark after Lemma 4.2.4. By the Verlinde formula, we see that

$$\dim \mathcal{H}_\Sigma[n,k]/\mathbf{Z}_n = \sum_{\lambda \in \Omega_{n,k}} \left(\frac{1}{\sqrt{n} S_{0\lambda}[n,k]} \right)^{2g-2}.$$

Combining with our previous discussion, we have the following proposition.

Proposition 4.2.8. *We have a non degenerate bilinear pairing of the vector spaces $\mathcal{H}_\Sigma[n,k]/\mathbf{Z}_n \times \mathcal{H}_\Sigma[k,n]/\mathbf{Z}_k \to \mathbf{C}$ and the representation of the mapping class group $\rho_g[n,k]$ is the contragredient representation of $\rho_g[k,n]$.*

References

[B] A. Beauville, Conformal blocks, fusion rules and the Verlinde formula, preprint.

[FR] R. Fenn and C. Rourke, On Kirby's calculus of links, Topology, **18** (1979), 1–15.

[GW] F. M. Goodman and H. Wenzl, Littlewood-Richardson coefficients for Hecke algebras at roots of unity, Advances in Math., **82** (1990), 244–265.

[JMO] M. Jimbo, T. Miwa and M. Okado, Solvable lattice models whose states are dominant integral weights of $A^{(1)}_{n-1}$, Lett. Math. Phys., **14** (1987), 123–131.

[K] V. G. Kac, "Infinite Dimensional Lie Algebras", 3rd edition, Cambridge University Press.

[KP] V. G. Kac and D. H. Peterson, Infinite dimensional Lie algebras, theta functions and modular forms, Advances in Math., **53** (1984), 125–264.

[KM] R. Kirby and P. Melvin, The 3-manifold invariants of Witten and Reshetikhin-Turaev for $sl(2, \mathbf{C})$, Invent. Math., **105** (1991), 473–545.

[Ko1] T. Kohno, Topological invariants for 3-manifolds using representations of mapping class groups I, Topology, **31** (1992), 203–230.

[Ko2] T. Kohno, Topological invariants for 3-manifolds using representations of mapping class groups II, Contemp. Math., **175** (1994), 193–217.

[KT] T. Kohno and T. Takata, Symmetry of Witten's 3-manifold invariants for for $sl(n, \mathbf{C})$, J. Knot Theory Ramifications, **2-2** (1993), 149–169.

[KN] A. Kuniba and T. Nakanishi, Level-rank duality in fusion RSOS models, in "Modern Quantum Field Theory", (A. Das et al., eds.), World Scientific, 1991, pp. 344–374.

[MOO] H. Murakami, T. Ohtsuki and M. Okada, Invariants of three-manifolds derived from linking matrices of framed links, Osaka. J. Math., **29** (1992), 545–572.

[NRS] S. G. Naculich, H. A. Riggs and H. J. Schnitzer, Simple current symmetries, rank-level duality, and linear skein relations for Chern-Simons graphs, Nuclear Phys. B, **394** (1993), 445–506.

[NT] T. Nakanishi and A. Tsuchiya, Level rank duality of WZW models in conformal field theory, Comm. Math. Phys., **144** (1992), 351–372.

[RT] N. Y. Reshetikhin and V. G. Turaev, Invariants of 3-manifolds via link polynomials and quantum groups, Invent. Math., **103** (1991), 547–597.

[SA] H. Saleur and D. Altschüler, Level-rank duality in quantum groups, Nuclear Phys. B, **354** (1991), 579–613.

[TUY] A. Tsuchiya, K. Ueno and Y. Yamada, Conformal field theory on universal family of stable curves with gauge symmetries, Adv. Stud. Pure Math., **19** (1989), 459–566.

[TW] V. G. Turaev and H. Wenzl, Quantum invariants of 3-manifolds associated with classical simple Lie algebras, Internat. J. Math., **4-2** (1993), 323–358.

[V] E. Verlinde, Fusion rules and modular transformations in 2-D conformal field theory, Nuclear Phys., **B300** (1988), 360–376.

[W1] E. Witten, Quantum field theory and the Jones polynomial, Comm. Math. Phys., **121** (1989), 351–399.

[W2] E. Witten, The Verlinde algebra and the cohomology of the Grassmannian, preprint, IASSNS-HEP-93/41.

Toshitake Kohno
Department of Mathematical Sciences
University of Tokyo
Meguro-ku, Tokyo 153
Japan

Toshie Takata
Graduate School of Mathematics
Kyushu University
Fukuoka 812
Japan

Advanced Studies in Pure Mathematics 24, 1996
Progress in Algebraic Combinatorics
pp. 265–283

Incidence Matrix Diagonal Forms and Integral Hecke Algebras

Robert A. Liebler

§1. Introduction

Let $\mathcal{P}_k$ denote the set of flats in the projective geometry $PG_{n-1}(q)$ arising from the k-dimensional subspaces of an n-dimensional vector space over the Galois field $GF(q)$, q a power of the prime p. Let $\mathcal{M}_k(q)$ be incidence matrix of points $\mathcal{P}_1$ versus k-flats $\mathcal{P}_k$.

The study of $\mathbb{Z}$-span of the columns of $\mathcal{M}_k(q)$ as a submodule of $\mathbb{Z}\mathcal{P}_1$ is of interest in its own right, as a source of easily implemented codes and because it may provide a means for representing, and ultimately characterizing, associated combinatorial structures like the q-analog Johnson schemes. I also find this study particularly attractive because it provides an explicit setting to further develop the application of integral representation theory and number theory to combinatorics along the lines that have been so successful with non-abelian difference sets.

It is easy to see that the Smith normal form of $\mathcal{M}_k(q)$ has all but one of its diagonal entries a divisor of the difference : $\begin{bmatrix} n-1 \\ k-1 \end{bmatrix}_q - \begin{bmatrix} n-2 \\ k-2 \end{bmatrix}_q$ of the q-binomial coefficients. Even for $n = 3$, every conceivable p-elementary divisor arises. Indeed, the set of lines of an affine net of degree p^k is in the kernel of the incidence map mod p^k but not mod p^{k+1}.

The first formulas for the p-rank of $\mathcal{M}_k(q)$ are due to Hamada [6] and to Smith [12] for $k = n - 1$. More recently Black and List [1] gave a generating function for the entries in a diagonal form of $\mathcal{M}_{n-1}(p)$ over $\mathbb{Z}$. Also, Lander [10, p 77] has apparently given information equivalent to a diagonal form of $\mathcal{M}_{n-1}(q)$ over $\mathbb{Z}$ in his (unpublished) Ph.D. thesis.

We show how diagonal forms for these incidence matrices arise naturally in the study of integral Hecke algebras and their geometrically significant eigenpotents. Within the $\mathbb{Z}$-module based on the chambers

Received February 7, 1995.
Revised January 15, 1996.
This research was partially supported by NSA grant 904-91-H-0048.

of the associated building we construct a certain lattice $\mathcal{L}$ that is invariant under the integral Hecke algebra $\mathcal{H}$. Theorem 3.7 asserts that $\mathcal{L}$ has a $\mathcal{H}$-completely reducible sublattice $\mathcal{N}$ of finite volume. It follows from Proposition 3.8 that the essential part of these incidence maps can be simultaneously diagonalized within $\mathcal{N}$ and that their Smith normal forms are computable from the multiplicites of the constituent irreducible $\mathcal{H}$-modules appearing in $\mathcal{N}$.

Unfortunately, I have no effective way of directly computing the constituent multiplicites of the $\mathbb{Z}$-forms appearing in $\mathcal{N}$. But, it turns out there is only one $\mathbb{Z}[1/p]$-form arising in $\mathcal{L} \otimes \mathbb{Z}[1/p]$. This allows the easy computation of the elementary divisors of $\mathcal{M}_k(q)$ that are not powers of p in Theorem 3.10.

We also give an independently obtained formula for a diagonal form for $\mathcal{M}_k(q)$ whenever k and n are relatively prime, Corollary 4.5. The proof of this formula involves and adaptation of an argument of Brouwer [2] to finite local rings. Some data from a Maple implementation of the formula is also given.

§2. Pure preliminaries

Let K be the quotient field of a Dedekind domain R. An *R-lattice* is a finitely generated R-torsionfree R-module. Each R-lattice M is naturally embedded in a finite dimensional K-vector space $KM := K \otimes_R M$. An R-sublattice $L \leq M$ is *pure* if M/L is R-torsionfree, or equivalently if L is a direct summand of M. Any submodule L of M has as its *pure closure* the inverse image of the torsion part of M/L. The importance of purity arises from:

Theorem 2.1 ([4, 4.12]). *Let M be an R-lattice. There is a bijective inclusion preserving correspondence between the R-pure sublattice of M and the K-subspaces of KM given by $L \to KL$ and $W \to W \cap M$.*

The theory of pure submodules is more subtle than that of vector spaces in part because a sum of two pure submodules that intersect trivially need not be pure. For example, $\mathcal{L}_\varepsilon = \{(a, \varepsilon a) | \ a \in \mathbb{Z}\}$ is a pure submodule of $\mathbb{Z}^2$ for any $\varepsilon \in \mathbb{Z}$. But $(2,0) \in \mathcal{L}_1 + \mathcal{L}_{-1}$ and $(1,0) \notin \mathcal{L}_1 + \mathcal{L}_{-1}$, so $\mathcal{L}_1 + \mathcal{L}_{-1}$ is not pure.

An *R-order* is a ring Λ whose center contains R, is finitely generated as an R-module and such that $A := K\Lambda$ is a K-algebra. Two Λ-modules M_1, M_2 that are R-lattices and with the property that KM_1 is equivalent (over K) to KM_2 are said to be *R-forms* for the associated A-module.

It is tempting to look for a Krull-Schmidt-Azumaya Theorem [4, Theorem 6.12] or even a Jordan-Hölder Theorem [4, Theorem 3.11] for

R-forms, but in Section 3.2 we see that the central object of our study $\mathcal{L}$ has a more subtle structure, not being completely reducible and failing the conclusion of the Jordan-Hölder Theorem. Indeed, (in the language of that Section) $\mathcal{L}$ contains two proper submodules $\mathcal{N}$ and $\mathcal{N}'$ each of which is a direct sum of irreducible $\mathbb{Z}$-forms for the natural representation but $\mathcal{L}/\mathcal{N}$ affords a finite version of the index representation while $\mathcal{L}/\mathcal{N}'$ affords a finite version of the Steinberg representation. Fortunately, Proposition 3.8 observes that this anomaly does not impact our incidence map study.

§3. Integral Hecke Algebras

Because this section is written in somewhat greater generality than the rest of this paper, examples are included to provide more explicit information about the most important case namely $PG_{n-1}(q)$.

Let $\mathcal{C}h$ be the set of chambers of a finite building $\mathcal{B}$ of type I and Weyl group $W = W(I) = \langle r_i | i \in I\rangle$. The *Standard module* $\mathbb{Z}\mathcal{C}h$ of $\mathcal{B}$ is the free $\mathbb{Z}$-module with $\mathcal{C}h$ as an orthonormal basis. The *integral Hecke algebra* $\mathcal{H}$ is the $\mathbb{Z}$-algebra generated by $\{\sigma_i | i \in I\} \subseteq \mathrm{End}_{\mathbb{Z}}(\mathbb{Z}\mathcal{C}h)$ where σ_i is (the adjacency matrix of) the relation "differ only in type i" on $\mathcal{C}h$. Because each such relation determines an equivalence relation with equipotent classes, each σ_i satisfies

$$(\sigma_i + 1)(\sigma_i - q_i) = 0, \tag{1}$$

for some integer q_i called the *index parameter* of σ_i.

A fundamental result of Iwahori [9],[4, Section 67] implies that the rational Hecke algebra $\mathcal{H} \otimes \mathbb{Q}$ has a $\mathbb{Q}$-basis

$$\mathcal{F} := \{\sigma_w := \sigma_{i_1}\sigma_{i_2}\cdots\sigma_{i_n} | \begin{array}{c} w \in W \text{ can be expressed} \\ \text{in reduced form } r_{i_1}r_{i_2}\ldots r_{i_n} \end{array}\}.$$

Notice that, elements of $\mathcal{F}$ are well defined and depend only the the associated element $w \in W$, not on the reduced form used to present w, because any two reduced forms for w can be transformed into each other using only rules of the form $r_ir_jr_i... = r_jr_ir_j...$, without any use of the relation $r_i^2 = 1$.

Because $\mathcal{F}$ is contained in $\mathcal{H}$, $\mathcal{H}$ is a $\mathbb{Z}$-order in the rational Hecke algebra. More than this, the result of Iwahori asserts that the rational Hecke algebra has a presentation (as an algebra) with generators $\{\sigma_i\}$ and relations: Equation (1) together with a relation equivalent to:

$$\begin{array}{c} \text{If } \sigma_w \in \mathcal{F} \text{ but } \sigma_w\sigma_j \notin \mathcal{F}, \text{ then } \sigma_w \\ \text{can be written in the form } \sigma_w = \cdots\sigma_j. \end{array} \tag{2}$$

Theorem 2.1 [4, 23.7] implies that any pure $\mathcal{H}$-submodule L of $\mathbb{Z}\mathcal{C}h$ is a $\mathbb{Z}$-form for a $\mathbb{Q}\mathcal{H}$-module.

3.1. Geometrically significant eigenpotents

For each $J \subseteq I$ and $c \in \mathcal{C}h$, define $c_J \in S$ to be the sum of all chambers having the same J residue as c. Explicitly the map $c \to c_J$ is realized by the *eigenpotent* $P_J \in \mathcal{H}$ where

$$P_J = \sum_{w \in W_J} \sigma_w$$

and W_J is the set of reduced words in the Weyl group W that involve only elements of $I \setminus J$. The elements P_J appear in [8] and [11]. The *Poincaré polynomial* [4, sec 67], p_J, of type J is obtained from P_J by replacing each σ_i with its index parameter.

In practice, one never really computes a geometrically significant eigenpotent as such a large sum. Instead, there are factored forms obtainable from chains of subgroups of the Weyl group.

Example 3.1. The building of type A_{n-1} with index parameters $q_i = q$ arises from the projective geometry $PG_{n-1}(q)$. It has as chambers the set of maximal chains (ordered by inclusion) of subspaces of the underlying vector space $GF(q)^n$. The relation giving rise to σ_i associates such a chain $0 < V_1 < \cdots < GF(q)^n$ to all other chains $0 < W_1 < \cdots < GF(q)^n$, where $W_j = V_j$ if and only if $i \neq j$. The associated Weyl group $W(A_{n-1})$ is the symmetric group S_n with the distinguished generators the transpositions $r_i = (i, i+1)$. The Iwahori relations for this Hecke algebra are Equation (1) and

$$\sigma_i \sigma_{i-1} \sigma_i = \sigma_{i-1} \sigma_i \sigma_{i-1} \text{ and } \sigma_i \sigma_j = \sigma_j \sigma_i \text{ whenever } |i-j| > 1.$$

Because S_n is doubly transitive and r_{n-1} commutes with S_{n-2}, S_n/S_{n-1} can be written $S_n = S_{n-1} \cup (S_{n-1}/S_{n-2}) r_{n-1} S_{n-1}$. It follows that

$$\sum_{w \in W(A_{n-1})} w = \lambda_{n-1} \cdots \lambda_2 \lambda_1, \text{ where } \lambda_k = (1 + \lambda_{k-1} r_k), \lambda_0 = 1.$$

Notice that the left most factor on the right hand side is the sum of left-coset representatives of S_n/S_{n-1} and, when multiplied out, each of the expressions is already in reduced form! It follows that the eigenpotent element associated with the empty set for a building of type A_{n-1} is $P_\phi(A_{n-1}) = L_{n-1} P_\phi(A_{n-2})$, where

$$L_{n-1} := 1 + L_{n-2} \sigma_{n-1} \qquad \text{and} \qquad L_0 = 1.$$

If instead we worked with right-cosets, we would obtain:

$$\sum_{w \in W(A_{n-1})} w = \rho_1 \cdots \rho_{n-1}, \text{ where } \rho_k = (1 + r_k \rho_{k-1}), \rho_0 = 1$$

and a different factorization: $P_\phi(A_{n-1}) = P_\phi(A_{n-2})R_{n-1}$, where

$$R_{n-1} := 1 + \sigma_{n-1} R_{n-2} \qquad \text{and} \qquad R_0 = 1.$$

The Weyl group $W(B_3) = \langle r_1, r_2, r_3 | (r_i r_j)^{m_{ij}} = 1 \rangle$; where $m_{ij} = m_{ji}$ and $m_{ii} = 1, m_{12} = 3, m_{13} = 2, m_{23} = 4$. The (formal) sum $T = [(1+r_1+r_2r_1)+(1+r_2+r_1r_2)r_3r_2r_1] \in \mathbb{Z}W$ of left-coset representatives for $W/\langle r_2, r_3 \rangle$ satisfies

$$\sum_{w \in W} w = T \sum_{h \in \langle r_2, r_3 \rangle} h = T(1 + r_3)(1 + r_2r_3)(1 + r_2).$$

Again, each factor is the sum of left-coset representatives and, when multiplied out, each expression is already in reduced form. For this reason the eigenpotent P_ϕ associated with the empty set in a building of type B_3 can be factored:

$$P_\phi = [(1+\sigma_1+\sigma_2\sigma_1)+(1+\sigma_2+\sigma_1\sigma_2)\sigma_3\sigma_2\sigma_1](1+\sigma_3)(1+\sigma_2\sigma_3)(1+\sigma_2).$$

The reader may check that a different sequence of subgroups leads to the more pleasant factorization:

$$P_\phi = (1 + \sigma_3)(1 + \sigma_2\sigma_3)(1 + \sigma_1\sigma_2\sigma_3)(1 + \sigma_2 + \sigma_1\sigma_2)(1 + \sigma_1).$$

The name given P_J is perhaps motivated by the fact that P_J is almost idempotent.

Lemma 3.2. $P_J^2 = p_J P_J$

Proof. It is sufficient to show that $P_J \sigma_i = P_J q_i$ for each $i \in I \setminus J$. This follows from the Iwahori relations Equations (1) and (2). This last computation can be vastly simplified by using a factored form for $P_J = \cdots (1 + \sigma_i)$ developed from a sequence of subgroups of the associated Weyl group starting from $\langle r_i \rangle$ as described in Example 3.1. Q.E.D.

Let $\mathbb{Z}\,\mathrm{typ}^{-1}(J)$ denote the image of $P_J \in \mathrm{End}(\mathbb{Z}\mathcal{C}h)$. Since no chamber is incident with two different elements of type J, $\mathbb{Z}\,\mathrm{typ}^{-1}(J)$ is a pure $\mathbb{Z}$-submodule of $\mathbb{Z}\mathcal{C}h$ having a distinguished basis $\{P_J(x) : \mathrm{typ}(x) = J\}$. It is therefore natural to identify an element x_J of type J in $\mathcal{B}$ with its image in $\mathbb{Z}\,\mathrm{typ}^{-1}(J)$.

The importance of the geometrically significant eigenpotent elements for our purposes is made clear by the following:

Lemma 3.3 ([11]). *Identify an element x_J of type J in $\mathcal{B}$ with its image in $\mathbb{Z}\,\mathrm{typ}^{-1}(J)$ as above. Then the natural incidence map from* $\mathrm{typ}^{-1}(J)$ *to* $\mathrm{typ}^{-1}(K)$ *is afforded by:*

$$P_K/p_{J\cap K} : \mathbb{Z}\,\mathrm{typ}^{-1}(J) \to \mathbb{Z}\,\mathrm{typ}^{-1}(K),$$

where $p_{J\cap K}$ is the Poincaré polynomial defined above.

Proof. The (a,c) entry of $P_K P_J$ counts the number of $b \in Ch$ where a and b share a face of type J and where b and c share a face of type K. There is no such b unless a and c share a face of type $J \cap K$ and then the number of choices for b is $p_{J\cap K}$. Q.E.D.

In the notation of the introduction, it is well known that the composition of the incidence maps $\mathcal{P}_{i-1} \to \mathcal{P}_i \to \mathcal{P}_{i+1}$ is $(q+1)$ times the incidence map $\mathcal{P}_{i-1} \to \mathcal{P}_{i+1}$. We illustrate the present language by recording this familiar fact. There follows a related factorization of the point hyperplane (non)-incidence map: $\mathcal{P}_1 \to \mathcal{P}_{n-1}$ that plays a key role in the proof of Theorem 3.7.

Corollary 3.4. *Suppose $\mathcal{B}$ is of type A_{n-1} with index parameter q. Then $\mathrm{typ}^{-1}(\{i\}) = \mathcal{P}_i$ — the set of $i-1$-flats in $PG_{n-1}(q)$ (the set of i-dimensional subspaces of the underlying vector space). Abuse notation by writing i for $\{i\}$.*

1. *The Poincaré polynomial of incident $i-1$- and i-flats is*

$$p_{i\,i+1} = \prod_{j=1}^{i-1} \frac{(q^j-1)}{(q-1)} \prod_{j=1}^{n-i-2} \frac{(q^j-1)}{(q-1)}.$$

2. *For any chamber $c \in Ch$, $P_{i-1}(c)$ is (under the above identification) the $i-2$-flat of c and $P_{i+1}P_i(P_{i-1}(c))) = p_i P_{i+1}(P_{i-1}(c))$.*
3. *For $c \in Ch$, $P_{n-1}\sigma_{n-1}\cdots\sigma_1(c)$ is the sum of all hyperplanes off the point of c.*
4. *For $c \in Ch$,*

$$(\sigma_{n-1}-q)\cdots(\sigma_1-q)(P_1(c)) = -(\sigma_{n-1}-q)P_{n-1}\sigma_{n-1}\cdots\sigma_1(c).$$

Proof. The first two parts follow from Lemma 3.3. The chambers d in (the support of) $\sigma_{n-1}\cdots\sigma_1(c)$ are obtained from c by first changing the point (to another one on the line of c), then changing the line, etc. The result is chambers d having no i-flat in common with c but for whom each i-flat meets the i-flat of c in the $i-1$-flat of d. If two such chambers shared a k-flat, then the $k-1$-flat of each must be the intersection of their common k-flat with the k-flat of c. Consequently, no two distinct such chambers share a hyperplane. For any such chamber d, $P_{n-1}(d)$ is just the hyperplane of d, by Part 1, so the indicated expresssion counts each hyperplane off the point of c exactly once. This proves Part 3.

In order to prove Part 4, we embellish the notation introduced in Example 3.1 and begin with a sublemma. Suppose $a_1, a_2, \ldots, a_k$ is a monotonic sequence of consective integers between 1 and $n-1$. Define $R_i(a_1 \ldots a_k)$ recursively by $R_i(a_1 \ldots a_k) = 1 + \sigma_{a_i} R_{i-1}(a_1 \ldots a_k)$ and $R_0(a_1 \ldots a_k) = 1$. Note that $R_i(a_1 \ldots a_k)$ depends only on $a_1, \ldots, a_i$ and so $R_{k-1}(a_1 \ldots a_k) = R_{k-1}(a_1 \ldots a_{k-1})$. Argue inductively that:

$$\sigma_{a_1} \ldots \sigma_{a_k} R_{k-1}(a_1 \ldots a_{k-1}) = R_{k-1}(a_2 \ldots a_k) \sigma_{a_1} \ldots \sigma_{a_k}. \tag{3}$$

By the recursive nature of R's and the Iwahori relations, we have:

$$\begin{aligned}
&\sigma_{a_1} \ldots \sigma_{a_k}(R_{k-1}(a_1 \ldots a_{k-1}) - 1) \\
&\quad = \sigma_{a_1} \ldots \sigma_{a_{k-2}}(\sigma_{a_{k-1}} \sigma_{a_k} \sigma_{a_{k-1}}) R_{k-2}(a_1 \ldots a_k) \\
&\quad = \sigma_{a_1} \ldots \sigma_{a_{k-2}}(\sigma_{a_k} \sigma_{a_{k-1}} \sigma_{a_k}) R_{k-2}(a_1 \ldots a_{k-1}) \\
&\quad = \sigma_{a_k}[\sigma_{a_1} \ldots \sigma_{a_{k-2}} \sigma_{a_{k-1}} R_{k-2}(a_1 \ldots a_{k-1})] \sigma_{a_k} \\
&\quad = \sigma_{a_k}[R_{k-2}(a_2 \ldots a_k) \sigma_{a_1} \ldots \sigma_{a_{k-2}} \sigma_{a_{k-1}}] \sigma_{a_k},
\end{aligned}$$

by the induction hypothesis. Therefore,

$$\sigma_{a_1} \ldots \sigma_{a_k} R_{k-1}(a_1 \ldots a_{k-1}) = (1 + \sigma_{a_k} R_{k-2}(a_2 \ldots a_k)) \sigma_{a_1} \ldots \sigma_{a_{k-1}} \sigma_{a_k}$$

as desired.

Now argue Part 4 by induction on n. In case $n = 2$:

$$\begin{aligned}
(\sigma_2 - q)(\sigma_1 - q)(\sigma_2 + 1) \quad &= \quad (\sigma_2 - q)\sigma_1(\sigma_2 + 1) - q(\sigma_2 - q)(\sigma_2 + 1) \\
= \quad (\sigma_2 - q)(-\sigma_2)\sigma_1(\sigma_2 + 1) - 0 &= -(\sigma_2 - q)(\sigma_2\sigma_1\sigma_2 + \sigma_2\sigma_1) \\
= \quad -(\sigma_2 - q)(\sigma_1\sigma_2\sigma_1 + \sigma_2\sigma_1) \quad &= -(\sigma_2 - q)(\sigma_1 + 1)\sigma_2\sigma_1
\end{aligned}$$

by the distributive law and two applications of Equation (1) for σ_2 followed by the relation $\sigma_2\sigma_1\sigma_2 = \sigma_1\sigma_2\sigma_1$.

For the general case, begin with P_1 in the factored form with *reverse* lexicographical order and right cosets: $P_1 = R_1(n-1 \ldots 2) \cdots R_{n-2}(n-1 \ldots 2)$, and note that σ_2 appears in only the last factor. Since σ_1 commutes with each σ_i, $i \neq 2$, we have by induction:

$$\begin{aligned}
&(\sigma_{n-1} - q) \cdots (\sigma_1 - q) P_1 \\
&\quad = [(\sigma_{n-1} - q) \cdots (\sigma_2 - q) R_1(n-1 \ldots 2) \cdots R_{n-3}(n-1 \ldots 2)] \\
&\qquad \times (\sigma_1 - q) R_{n-2}(n-1 \ldots 2) \\
&\quad = -[(\sigma_{n-1} - q) R_1(n-2 \ldots 2) \cdots R_{n-3}(n-2 \ldots 2) \sigma_{n-1} \cdots \sigma_2] \\
&\qquad \times (\sigma_1 - q) R_{n-2}(n-1 \ldots 2).
\end{aligned}$$

Next distribute with respect to the $(\sigma_1 - q)$ term and apply Equation (3) repeatedly to the expression:

$$\begin{aligned} &q(\sigma_{n-1} - q)[R_1(n-2\dots 2)\cdots R_{n-3}(n-2\dots 2)\sigma_{n-1}\cdots\sigma_2] \\ &\qquad \times R_{n-2}(n-1\dots 2) \\ &\quad = q(\sigma_{n-1} - q)[\cdots R_{n-4}(n-2\dots 2)\sigma_{n-1}\cdots\sigma_2] \\ &\qquad \times R_{n-3}(n-1\dots 2)R_{n-2}(n-1\dots 2) \\ &\quad \cdots = q(\sigma_{n-1} - q)\sigma_{n-1}\cdots\sigma_2[R_1(n-1\dots 2)\cdots R_{n-2}(n-1\dots 2)]. \end{aligned}$$

But the last expression P within brackets [ ] is a factored form of the trivial eigenpotent on $\{2\dots n-1\}$ and so can be rewritten in a factored form begining with $(\sigma_i + 1)$ for any $i = 2, \dots, n-1$. It follows that this term equals $q^{n-1}(\sigma_{n-1} - q)P = 0$. This shows that only the other expression survives and by one more application of Equation (3) it can be rewritten:

$$\begin{aligned} &-(\sigma_{n-1} - q)R_1(n-2\dots 2)\cdots R_{n-3}(n-2\dots 2) \\ &\qquad \times[\sigma_{n-1}\cdots\sigma_2\sigma_1 R_{n-2}(n-1\dots 2)] \\ &\quad = -(\sigma_{n-1} - q)R_1(n-2\dots 1)\cdots R_{n-2}(n-2\dots 1)\sigma_{n-1}\cdots\sigma_2\sigma_1 \\ &\quad = -(\sigma_{n-1} - q)P_{n-1}\sigma_{n-1}\cdots\sigma_2\sigma_1. \end{aligned}$$

Q.E.D.

There is also a representation theoretic proof of Corollary 3.4.4 that amounts to verification of the indicated equation in all relevant $\mathcal{H}\otimes_{\mathbb{Z}}\mathbb{Q}$ representations. Because P_1 and P_{n-1} are trivial in all $\mathcal{H}\otimes_{\mathbb{Z}}\mathbb{Q}$ representations other than the index and natural representations, one need only verify Corollary 3.4.4 for the index representation, where it is immediate, and for the matrices appearing in Section 3.2.

In fact Corollary 3.4.4 is a combinatorial identity that was first obtained by a representation theoretic method. I leave it to the reader's taste which argument is more elementary.

3.2. The natural $\mathcal{H}$ representation and its irreducible $\mathbb{Z}$-forms

The representations of $\mathcal{H}\otimes\mathbb{C}$ are deformations of the irreducible representations of the associated Weyl group by a celebrated theorem of Tits. Hoefsmit [7] gives explicit formulas in case the building arises from a classical group. Because we are concerned with incidence maps involving the points $\mathcal{P}_1$ of $PG_{n-1}(q)$, we must deal explicitly with the $\mathcal{H}(A_{n-1})$-representations arising from $\mathbb{Z}\,\mathrm{typ}^{-1}(\{1\})$. These correspond to the S_n representations labelled by partitions $\{n\}$ and $\{n-1, 1\}$ [3, Theorem 2.1].

The *index* representation *Ind* corresponds to the trivial representation $\{n\}$ of S_n. *Ind* is the homomorphism $\mathcal{H} \to \mathbb{Z}$ that maps each generator σ_i to its index paramater q_i. Thus the Poincaré polynomial is just $Ind(P_\phi)$. The *Steinberg* representation *St* is the homomorphism $\mathcal{H} \to \mathbb{Z}$ that maps each generator σ_i to -1.

The *natural* representation corresponds to the natural representation $\{n-1, 1\}$ of S_n and is a homomorphism $\mathcal{H} \to \mathrm{Mat}_{n-1}(\mathbb{C})$. Hoefsmit's version of this representation is a natural generalization of Young's "semi-normal" form for the symmetric group representation and is:

$$(\sigma_1) = \left(\begin{array}{c|c} qI_{n-2} & \\ \hline & -1 \end{array}\right); \quad (\sigma_i) = \left(\begin{array}{c|cc|c} qI_{n-i-1} & & & \\ \hline & \frac{q-1}{1-q^i} & \frac{1-q^{i+1}}{1-q^i} & \\ & \frac{q-q^i}{1-q^i} & \frac{q^i-q^{i+1}}{1-q^i} & \\ \hline & & & qI_{i-2} \end{array}\right)$$

for $i = 2, \ldots, n-1$. We record a few easy consequences of this explicit representation.

Lemma 3.5. *Under the natural representation of* $\mathcal{H}$ *we have*:

1. $(\sigma_k - q)(\sigma_i - q) = 0$ *whenever* $|i-k| > 1$,
2. *For all* $j \notin J \subseteq I$, $\mathrm{Im}(P_J) \cap \ker(\sigma_j + 1) = 0$.

Proof. In each case this follows from a simple matrix computation. The result is a little more transparent if a general $\mathbb{Z}$-form presented below is used. Q.E.D.

The restriction of the centralizer in $\mathcal{H}$ of P_J to the image of P_J is called the J-th *parabolic* subalgebra of $\mathcal{H}$. The adjacency algebras of many distance transitive graphs arise as parabolic subalgebras for $|J| = 1$. One can in principle compute the eigenvalues of these graphs [11] from the explicit $\mathbb{C}$-representations of $\mathcal{H}$.

Say a $\mathbb{Z}$-form for an irreducible representation of $\mathcal{H} \otimes \mathbb{C}$ is of *type* $\mathcal{N}$ (respectively of *type* $\mathcal{N}'$) if the associated $\mathcal{H}$-module is generated by the -1-eigenspaces (respectively q_i-eigenspaces) of $\{\sigma_i\}$.

Of particular interest to us are the irreducible $\mathbb{Z}$-forms for the natural representation of $\mathcal{H}(A_n)$. One can show that the irreducible $\mathbb{Z}$-forms of the natural representation of type $\mathcal{N}$ are parameterized by integers $a_1, \ldots, a_{n-1}$, all divisors of q and the block diagonal matrices:

$$(\sigma_1) = \left(\begin{array}{cc|c} -1 & q/a_1 & \\ 0 & q & \\ \hline & & qI_{n-2} \end{array}\right); (\sigma_n) = \left(\begin{array}{c|cc} qI_{n-2} & & \\ \hline & q & 0 \\ & a_{n-1} & -1 \end{array}\right)$$

$$(\sigma_i) = \left(\begin{array}{c|ccc|c} qI_{i-1} & & & & \\ \hline & q & 0 & 0 & \\ & a_{i-1} & -1 & q/a_i & \\ & 0 & 0 & q & \\ \hline & & & & qI_{n-i-2} \end{array}\right); \text{ for } i = 2, \dots, n-1.$$

In particular, distinct $a_1, \dots, a_{n-1}$ give $\mathbb{Z}$-inequivalent forms. Denote the $\mathbb{Z}$-form with parameters $a_1, \dots, a_{n-1}$ by $\mathcal{N}(a_1, \dots, a_{n-1})$. The same computation shows that there is but one $\mathbb{Z}[1/p]$-form of type $\mathcal{N}$ for the natural representation when q is a power of the prime p.

Example 3.6. The reader must be warned that there do exist indecomposable $\mathbb{Z}$-forms having all natural $\mathcal{H}$-composition factors that are reducible. Indeed, the matrices

$$\sigma_1 = \begin{pmatrix} -1 & q/a & 0 & 0 \\ 0 & q & 0 & 0 \\ 0 & 0 & -1 & q/b \\ 0 & 0 & 0 & q \end{pmatrix}, \ \sigma_2 = \begin{pmatrix} q & 0 & bq & -(q+1)q \\ a & -1 & ab(q+1) & -aq \\ 0 & 0 & q & 0 \\ 0 & 0 & b & -1 \end{pmatrix}$$

satisfy the relations defining $\mathcal{H}(A_2)$. Moreover any transition matrix $P \in SL(4, \mathbb{Z})$, taking these matrices to a sum of the above forms, must leave invariant the -1-eigenspace of σ_1 and take the -1-eigenspace of σ_2 to $\langle e_2, e_4 \rangle$. Computation of $P^{-1}\sigma_i P$ shows that P is forced to have the form

$$P = \begin{pmatrix} g(vb - uab(q+1)) & uq & h(fb - wab(q+1)) & wq \\ 0 & v & 0 & x \\ uag & 0 & wah & 0 \\ 0 & u & 0 & w \end{pmatrix},$$

for integers g, h, u, v, w, x. But $\det(P) = abgh(xu - vw)^2$, so this representation does not split over $\mathbb{Z}$ whenever $|ab| \neq 1$.

The reader must also be warned that not every irreducible natural $\mathcal{H}$-module is of type $\mathcal{N}$. For example, the transposes of the above displayed $\mathbb{Z}$-form of $\mathcal{H}(A_n)$ is again a $\mathbb{Z}$-form of $\mathcal{H}(A_n)$, but of type $\mathcal{N}'$.

When one attempts to transform such a represenation to one of type $\mathcal{N}$, exploitation of eigenspaces as above leads to an essentially unique transition matrix and to the form $\mathcal{N}(q/a_1, \dots, q/a_{n-1})$. When $n = 4$ the transition matrix is

$$\begin{pmatrix} (1+q)(1+q^2) & (1+q+q^2)a_1 & a_1a_2(1+q) & a_1a_2a_3 \\ (1+q+q^2)q/a_1 & (1+q)(1+q+q^2) & (1+q)^2a_2 & (1+q)a_2a_3 \\ (1+q)q^2/(a_1a_2) & (1+q)^2q/a_2 & (1+q)(1+q+q^2) & (1+q+q^2)a_3 \\ q^3/(a_1a_2a_3) & (1+q)q^2/(a_2a_3) & (1+q+q^2)q/a_3 & (1+q)(1+q^2) \end{pmatrix}$$

and has determinant $(1+q+q^2+q^3+q^4)^3$. Therefore these forms are not $\mathbb{Z}$-equivalent. Thanks to L. Solomon for sharing this observation.

Let $\mathcal{L}$ be the pure closure in $\mathbb{Z}\mathcal{C}h$ of $\mathcal{H}(\sigma_1 - q)P_{\{1\}}\mathcal{C}h$. The module $\mathcal{L}$ is fundamental to our study of the incidence maps $\mathcal{M}_k(q)$. It will turn out that $\mathcal{L}$ is the largest $\mathcal{H}$-submodule of the standard module affording only the natural representation of $\mathcal{H}$. It seems unlikely, however, that $\mathcal{L} \otimes_{\mathbb{Z}} R$ is a summand of $R\mathcal{C}h$ unless the Poincaré polynomial $p_{\{1\}}$ is a unit in the domain R.

Theorem 3.7. *The sum of the components of $\mathbb{Q}\mathcal{C}h$ affording the natural representations of $\mathcal{H} \otimes \mathbb{Q}$ is $\mathcal{L} \otimes \mathbb{Q}$.*

Let $\mathcal{L}_i \leq \mathcal{L}$ be the -1-eigenspace of σ_i. Then

1. *$\sum \mathcal{L}_i$ is a direct sum of irreducible $\mathcal{H}$-modules of type $\mathcal{N}$,*
2. *$\mathcal{L}/\sum \mathcal{L}_i$ is torsion of exponent dividing $1+q+\cdots+q^n$ and affords (a modular version of) the index representation of $\mathcal{H}$.*

Proof. As already mentioned, $\mathcal{H}P_{\{1\}}\mathbb{Z}\mathcal{C}h \otimes \mathbb{Q}$ is the sum of the components of $\mathbb{Q}\mathcal{C}h$ associated with the index and the natural representations of $\mathcal{H} \otimes \mathbb{Q}$. By Equation (1) $\sigma_1 v = -v$, for all $v \in (\sigma_1 - q)\mathbb{Z}\mathcal{C}h$ so $\mathcal{L} \otimes \mathbb{Q}$ affords only the natural representation of $\mathcal{H} \otimes \mathbb{Q}$ and to full multiplicity.

Define $m_i : \mathcal{L}_1 \to \mathcal{L}_i$ by $m_i(v) = (\sigma_i - q)m_{i-1}(v)$. When m_{n-1} is regarded as an homomorphism of rational vector spaces, it is invertible. Therefore m_{n-1} is of maximal rank. Pick bases $\{\ell_{11}, \ldots, \ell_{1m}\}$, of $\mathcal{L}_1$ and $\{\ell_{n-1\,1}, \ldots, \ell_{n-1\,m}\}$ of $\mathcal{L}_{n-1}$ so that m_{n-1} is in Smith normal form. Let ℓ_{ij} span the pure $\mathbb{Z}$-submodule generated by $m_i(\ell_{1j})$ and define $a_{ij} \in \mathbb{Z}$ by $(\sigma_i - q)(\ell_{i-1\,j}) = a_{ij}\ell_{ij}$. Finally set $N_j := \mathbb{Z}\{\ell_{ij} | i = 1, \ldots, n-1\}$.

Notice that m_{n-1} appears in Corollary 3.4.4. This result implies that the pure closure of $\operatorname{Im}(\sigma_{n-1} - q)$ within $\mathcal{H}\operatorname{Im}(P_{\{n\}})$ is exactly $\mathcal{L}_{n-1}$, the pure closure of $\operatorname{Im}(m_{n-1})$. Thus, we could as well have started with hyperplanes rather than points, the same module $\mathcal{L}$ would arise. Moreover, m_{n-1} is in Smith normal form with respect to the bases $\{\ell_{11}, \ldots, \ell_{1m}\}$, of $\mathcal{L}_1$ and $\{\ell_{n-1\,1}, \ldots, \ell_{n-1\,m}\}$ of $\mathcal{L}_{n-1}$ if and only if its transpose is in Smith normal form with respect to the bases $\{\ell_{n-1\ 1}, \ldots, \ell_{n-1\ m}\}$ of $\mathcal{L}_{n-1}$ and $\{\ell_{11}, \ldots, \ell_{1m}\}$, of $\mathcal{L}_1$. Therefore, this point - hyperplane symmetry leaves invariant the set $\{\ell_{ij}\}$ as well.

In order to show that N_j is $\mathcal{H}$-invariant it is sufficient to show $(\sigma_k - q)\ell_{ij} \in N_j$ for $|k-i| = 1$, by Lemma 3.5.1. But this follows from the above noted symmetry and the definition $\ell_{kj} = (\sigma_k - q)(\ell_{k-1j})$. This completes the proof that N_j is $\mathcal{H}$-invariant.

To see that N_j affords the $\mathbb{Z}$-form with parameters a_{ij} relative to $\{\ell_{ij}\}$, note that $(\sigma_i - q)(\sigma_k - q) = 0$ in $\mathcal{L} \otimes \mathbb{Q}$, whenever $|i-k| > 1$, by Lemma 3.5.1. Since $\ell_{jk} \in \mathrm{Im}(\sigma_k - q)$, this and the definition of a_{ij} forces each of the columns of the matrix of σ_i to be as specified except perhaps the $i+1$-st column. This column is forced by the others and the Iwahori relations. Since $\mathcal{N} := \sum N_j = \sum \mathcal{L}_i$, this establishes Part 1.

Suppose $\ell \in \mathcal{L} \setminus \mathcal{N}$ and set $n_i = (\sigma_i(\ell) - q)(\ell)$. Then $\sigma_i(n_i) = \sigma_i(\sigma_i - q)(\ell) = -(\sigma_i - q)(\ell) = -n_i$, so $n_i \in \mathcal{N}$. Now $\mathcal{L}/\mathcal{N}$ affords the index representation because $\sigma_i(\ell) \equiv q\ell \pmod{\mathcal{N}}$.

Clearly $\mathcal{L} \leq \mathcal{N} \otimes \mathbb{Q}$ and it suffices to show that $\mathcal{L}/(\sum_{j \neq k} N_j)$ has the indicated exponent. But this finite abelian group has order dividing the volume of $N_k \leq N_k \otimes \mathbb{Q}$, which in turn equals the absolute value of the determinant, say k_n, of

$$\sum(\sigma_i - q) = \begin{pmatrix} -q-1 & q/a_{1k} & & & & \\ a_{1k} & -q-1 & q/a_{2k} & & & \\ & a_{2k} & \ddots & \ddots & & \\ & & \ddots & -q-1 & q/a_{n-1k} \\ & & & a_{n-1k} & -q-1 \end{pmatrix}.$$

The result follows from the recurrence $k_n = -(1+q)k_{n-1} - qk_{n-2}$.
Q.E.D.

If we use the common q-eigenspaces of all but one of the σ_i's (as in the last part of Example 3.6) instead of -1-eigenspaces, the ideas of this proof lead to $\mathbb{Z}$-forms of the natural $\mathcal{H}$-representation of type $\mathcal{N}'$ and a submodule $\mathcal{N}'$ of $\mathcal{L}$ of finite volume such that $\mathcal{L}/\mathcal{N}'$ affords a finite version of the Steinberg representation.

Proposition 3.8. *The Smith normal form of an integral incidence map π coincides with that computed using any submodule of finite volume v in* domain(π).

Proof. Because any free $\mathbb{Z}$-module M is $\mathbb{Z}$-isomorphic to its submodule vM, the incidence map has the same algebraic invariants as their restriction to $v \cdot \mathrm{domain}(\pi)$. Q.E.D.

Example 3.9. For $\mathcal{H}$-modules of type $\mathcal{N}'$, $\mathrm{Im}(P_{\{i\}})$ is contained in the span of the i-th natural basis vector. In case $n = 4$, the i,j entry of the matrix $(P_1 + P_2 + P_3 + P_4)$ gives the diagonal entry of the i-space versus j-space incidence matrix of $PG(4,q)$ arising from the $\mathbb{Z}$-form of transposes of $\mathcal{N}(a_1, a_2, a_3)$ and is, amazingly enough, the matrix appearing in Example 3.6 left multiplied by $\mathrm{diag}\{(q^2+q+1)(q+1), q+1, q+1, (q^2+q+1)(q+1)\}$.

In case $q = p$ is prime one can use the known p-rank formulae for incidence maps to compute the multiplicites of each of the $\mathbb{Z}$-forms. All multiplicities are zero except: $m_{1,1,1} = m_{p,p,p} = (p+3)(p+2)(p+1)p/24$; $m_{1,1,p} = m_{1,p,p} = p(p+1)(11p^2 - 5p + 6)/24$.

The fact that Propositon 3.8 can be applied to either $\mathcal{N} \leq \mathcal{L}$ or to $\mathcal{N}' \leq \mathcal{L}$ seems to imply a well known symmetry in the multiplicities of the p-power elementary divisors of $\mathcal{M}_{n-1}(q)$, $p|q$.

Also, as noted just before Example 3.6, any two irreducible $\mathbb{Z}$-forms for the natural $\mathcal{H}$-module are equivalent if coefficients are extended to $\mathbb{Z}[1/p]$ so that p is a unit. Some results of Frumkin and Yakir [5] about certain incidence maps' modular rank therefore extend to statements about elementary divisors.

Theorem 3.10. *Let $\mathcal{V}$ be the sublattice of $\mathbb{Z}\mathcal{P}_1$ whose elements have coordinate sum equal to zero. Then the p'-part of the finite abelian group $\mathcal{V}/(\mathcal{V} \cap$ column space$\mathcal{M}_k(q))$ is homocyclic of exponent e where e is the $(1,k)$ entry of the $\mathcal{H}(A_{n-1})$-form $\mathcal{N}(q,\ldots,q_{n-1})$.*

It is natural to extend the above arguments to more general incidence maps from, say from i-flats to j-flats, or to other buildings. For example, one may consider the algebra generated by $\{P_{\{i\}}|i \in I\}$ acting on the $\mathbb{Z}$-span $\mathcal{T}$ of $\{\operatorname{Im} P_{\{i\}}|i \in I\}$. This is almost the incidence algebra of the associated uniform poset in the sense of Terwilliger [13], so the spectral nature of these incidence algebras may also be studied using the $\mathbb{C}$-representations of $\mathcal{H}$. If the coefficient ring is only localized at a prime dividing an index parameter, then $\mathcal{T}$ a summand of standard module. In general, $\mathcal{T}$ is a pure submodule of $\mathbb{Z}\mathcal{C}h$ and therefore its direct sum decomposition as $\mathbb{Z}$-forms exhibits each of the geometrically significant eigenpotents $P_{\{i\}}$ in a block diagonal form over $\mathbb{Z}$. Because $P_{\{i\}}$ is rank one in each relevant representation, each of the level schemes is what Terwilliger calls "thin" and no blocks are bigger than 1 by 1. This means that all incidence maps are simultaneously in a diagonal form over $\mathbb{Z}$ and all important information conveyed by their Smith normal form is visible from the direct sum decomposition of $\mathbb{Z}\mathcal{T}$ as $\mathbb{Z}$-forms. One serious difficulty in this enterprise seems to be finding the relevant $\mathbb{Z}$-forms.

From the nature of the $\mathbb{Z}$-forms on other pure submodules of $\mathbb{Z}\mathcal{C}h$ and their multiplicities (whatever that means), one can compute a variety of additional arithmetic invariants of the building beyond incidence map invariant factors. Moreover much of the above machinery also applies to finite Tits geometries of "type M". So there is some hope that these additional arithmetic invariants might help resolve the celebrated question of existence of non-building Tits geometries of type B_3.

§4. Incidence matrices and finite local rings

This section is concerned with the multiplicities of powers of p as elementary divisors of the incidence matrices $\mathcal{M}_{n-k}(p^d)$. The main tools are finite local rings and the discrete Fourier transform.

Here is one (not particularly standard) way of constructing the finite field $GF(p^m)$. Let ζ be primitive a complex p^m-1-th root of unity. The integral domain $\mathbb{Z}[\zeta] \leq \mathbb{C}$ of $\mathbb{Z}$-linear combinations of powers of ζ is called a ring of *Cyclotomic Integers* and is a basic construct of algebraic number theory. Although $\mathbb{Z}[\zeta]$ is not a principal ideal domain, its ideal structure is well understood. Any prime ideal π of $\mathbb{Z}[\zeta]$ that contains (p) has the form $(p, f(\zeta))$, where $f(x) \in \mathbb{Z}[x]$ is congruent modulo p to an irreducible polynomial of degree m in $GF(p)[x]$. Since the quotient ring $\mathbb{Z}[\zeta]/\pi$ is an integral domain of order p^m, it must be $GF(p^m)$.

The finite local rings are most naturally constructed by considering $\mathbb{Z}[\zeta]/(\pi)^s$ instead. We record the basic facts about these rings.

Theorem 4.1. *Let $p \in \mathbb{N}$ be a prime, ζ a primitive complex (p^m-1)-th root of unity. Let $f(x) \in \mathbb{Z}[x]$ be an integral polynomial whose reduction modulo (p) is irreducible of degree m. (For $m=0$, set $\zeta=1$ and $f=0$.)*

1. *Then $\pi := (p, f(\zeta))$ is a prime ideal of $\mathbb{Z}[\zeta]$ over (p).*
2. *Then the ring $R_s := \mathbb{Z}[\zeta]/\pi^s$ has order p^{ms} and characteristic p^s.*
3. *The multiplicative group $R_s^\star$ of R_s contains $\langle\zeta\rangle$ and $1+r\pi$ for all $r \in R$.*
4. *$R_s^\star$ consists of all elements of R_s not in π and is of exponent $p^s(p^m-1)$.*
5. *The ring R_s is a local ring.*

The basic idea of the method of Brouwer is to present the incidence matrix to be studied as the table of values of a function in two variables. Then one uses a finite Fourier transform to substitute the coefficient matrix for the incidence matrix.

Lemma 4.2 (cf. [2]). *Let U be a cyclic multiplicative group of order u in the commutative ring $\mathcal{R}$. For $X \subseteq U$, let V_X denote a Vandermonde matrix with rows indexed by $x \in X$ with x-th row consists of the successive powers $1, x, x^2, \ldots, x^{u-1}$ of x. Suppose $p(x,y) = \sum\sum c_{ij}x^iy^j$, is a polynomial function with coefficients in $\mathcal{R}$ of x-degree* *and y-degree less than u and C the u by u matrix of coefficients of* $p(x, y)$. *Then for* $A, B \subset U$, *the matrix of values of*

$$M := (p(a,b))_{a\in A\ b\in B} \text{ is given by } M = V_A C V_B^t.$$

In particular M and C are equivalent whenever V_A *and* V_B *possess right-inverses having coefficients in* $\mathcal{R}$,

Proof. Just multiply out the right hand side. Q.E.D.

Adopt the notation of Theorem 4.1, so ζ is a primitive complex $(p^m - 1)$-th root of unity, π is a prime ideal of $\mathbb{Z}[\zeta]$ over (p) and $R_s := \mathbb{Z}[\zeta]/\pi^s$ is a finite local ring of order p^{ms} and characteristic p^s.

We intend to apply Lemma 4.2 with $\mathcal{R} = R_s$, and M one of the incidence matrices $\mathcal{M}_k(p^e)$. In view of Theorem 4.1, the most natural choice for U is perhaps a maximal cyclic subgroup of $R_s^\star$, but then V_U does not have a right inverse defined over R_s. If one chooses $U \leq R_s^\star$ small enough for V_A and V_B to have right inverses over R_s, then one is forced to deal with polynomial function $p(x, y)$ of uncomfortably small degree. Either way, one is led to the same results. We choose to present the later line of argument because it avoids explicit computation of the Smith normal form of the character table of a cyclic p-group as appears in [1].

Suppose $m = ef$, and $r = p^e$, define the function:

$$\operatorname{tr}_r(z) := z + z^r + \cdots + z^{r^{f-1}}; \quad z \in U := \langle\zeta\rangle.$$

Observe $\operatorname{tr}_r(z) \in \pi$ if and only if z is in the kernel of the trace map from the field R_s/π to $E := GF(r)$. This trace map gives rise to the nondegenerate symmetric bilinear form on the E-vector space R_s/π given by $(x, y) = \operatorname{tr}_r(xy)$.

Organize the elements of U into cosets of the subgroup $E^\star$ of order $(r - 1)$ and consider the matrix M whose rows and columns are labelled by the elements of U in this order and whose x, y entry is $\operatorname{tr}_r(xy)$. Then the blocks of this matrix can be labelled with the points and hyperplanes of $PG_{f-1}(r)$ in such a way that the point labelled by $\bar{x}$ is incident with the hyperplane $\bar{y}$ if and only if some entry of the $(\bar{x}, \bar{y})$ block of M is in π. This occurs if and only if all entries of the $(\bar{x}, \bar{y})$ block of M are in π. Since $\operatorname{Im}(\operatorname{tr}_r)/\pi \leq E$, $M^{\circ(r-1)} \pmod{\pi}$ is a zero one matrix (here $\circ$ denotes the Schur, or termwise, product). And, in fact,

$$M^{\circ(r-1)} \equiv J_{p^m-1} - \mathcal{M}_{f-1}(r) \otimes J_{r-1} \pmod{\pi},$$

where J denotes the appropriate matrix of all ones, $\otimes$ denotes Kronecker product of matrices, and $\mathcal{M}_{f-1}(r)$ a point-hyperplane incidence matrix of $PG_{f-1}(r)$ as above.

By Theorem 4.1.5 R_s is a local ring and all of its elements not in π are units, so this also implies:

$$M^{\circ(r-1)p^s} \equiv J_{p^n-1} - \mathcal{M}_{f-1}(r) \otimes J_{r-1} \pmod{\pi^s}.$$

This equation is sufficient to study the point-hyperplane matrix $\mathcal{M}_{f-1}(r)$ because left hand side is the table of values of the polynomial function $(\mathrm{tr}_r(xy))^{(r-1)p^s}$.

But in order to deal with the matrices $\mathcal{M}_{k-1}(q)$, we need to take R_s/π to be much larger than the underlying vector space and replace the field E with two fields.

Let $d = \gcd(e, f)$, and let $e = dk, f = nd$ define k and n. Set $q = p^d$, so $r = p^e = q^k$. Let $F^\star \leq L^\star$ be the subgroups of U having order $q^n - 1$, $r^n - 1$, respectively. Then, modulo π, the groups $L^\star$, $F^\star$, $E^\star$ and $F^\star \cap E^\star$ are the multiplicative groups of fields having like names (without the *) and isomorphic to $GF(r^n) \geq GF(q^n)$, $GF(r) \geq GF(q)$, respectively. We collect some required geometric facts in a lemma.

Lemma 4.3.

1. *Any $E \cap F$-subspace $W \leq F$ generates an E-subspace EW and $EW \cap F = W$. Moreover $\dim_{E\cap F} W = \dim_E EW$.*
2. *Suppose V is an E-hyperplane in R_s/π. Then $L \cap V$ is an $E \cap F$-subspace containing E and of codimension less than or equal to k in F.*
3. *Conversely, if $W \leq F$, has $E \cap F$-codimension less than or equal to k then $W = F \cap V$ for some E-hyperplane V in R_s/π.*

Proof. Part 1 follows from the fact that k and n are relatively prime. Of course an E-hyperplane has $E \cap F$-codimension $k = [E : E \cap F]$ in R_s/π, so Part 2 is also clear. Finally, Part 1 implies that EW is and E-subspace of L having like codimension, from which Part 3 follows. Q.E.D.

Recall that the matrix M has blocks of size $r - 1$ and the blocks of this matrix can be labelled with the points and hyperplanes of $PG_{f-1}(r)$ in such a way that the point labelled by $\bar{x}$ is incident with the hyperplane $\bar{y}$ if and only if all entries of the $(\bar{x}, \bar{y})$ block of M are in π. Let M_F be the submatrix of M with rows labelled by elements of $F^\star$. Then

$$M_F^{\circ(r-1)} \equiv J - A_F \otimes J_r \pmod{\pi},$$

where A_F is a zero-one matrix whose columns are the characteristic functions of $F \cap V$ where V is an E-hyperplane in R_s/π. By Lemma 4.3, the column space of A_F is the R_s-span of (the characteristic vectors of) the $E \cap F$-subspaces of codimension less than or equal to k in F. By Lemma 3.4.2, the R_s-incidence map from $\mathcal{P}_1$ to $\mathcal{P}_{n-j}$ factors through the incidence map from $\mathcal{P}_1$ to $\mathcal{P}_{n-k}$, whenever $k \geq j$. Therefore the R_s-column space of A_F is the R_s-module spanned by the columns of $\mathcal{M}_{n-k}(q) \otimes J_r$.

Again, since R_s is a local ring, this also implies:

$$M_F^{\circ(r-1)p^s} \equiv J - A_L \otimes J_r \pmod{\pi^s}. \tag{4}$$

Now M_F is just the table of values of the function $(\mathrm{tr}_{q^k}(xy))^{(q^k-1)p^s}$ for $x \in F^\star$, $y \in U$. Unfortunately the degree of this polynomial is too high to apply Lemma 4.2 directly.

Proposition 4.4. *Suppose n and k are relatively prime and $q = p^d$. Let C be the coefficient matrix of $1 - f(x,y)$ where*

$$f(x,y) \equiv (\mathrm{tr}_{q^k}(xy))^{(q^k-1)p^s} \quad \mathrm{mod}\ (x^{q^n-1} - 1, y^{q^{nkd}-1} - 1)$$

has x-degree $\leq q^n - 2$ and y degree $\leq q^{nkd} - 2$. Then $\mathcal{M}_{n-k}(q)$ has the same non-zero R_s-elementary divisors as C.

Proof. By definition of $F^\star$ and U, the two polynomials take exactly the same values for $x \in F^\star$, $y \in U$. Lemma 4.2, Equation 4 and the fact that the Vandermonde matrix V_U is the character table of the group U having order not divisible by p, imply that $\mathcal{M}_{n-k}(q) \otimes J_{q^k-1}$ and C are R_s-equivalent. The result follows from the fact that J_{q^k-1} has rank 1 and that $q^k - 1$ is a unit in R_s. Q.E.D.

Corollary 4.5. *Suppose n and k are relatively prime and $q = p^d$. Set $p_0 = 1$ and for $0 \leq i \leq q^n - 2$, define $d_i = \gcd\{p_j | j \equiv i \pmod{q^n - 1}\}$, where p_j is the multinomial coefficient sum:*

$$p_j := \sum_{\sum r_\alpha q^{k\alpha} \equiv j \pmod{q^{nkd}-1}} \binom{(q^k-1)p^s}{r_0, r_1, \ldots}.$$

Then $\mathcal{M}_{n-k}(q)$ has an R_s-diagonal form $D := \mathrm{diag}\{1, d_i\}$. In particular, p^d , $d < s$ occurs as an elementary divisor over $\mathbb{Z}$ to like multiplicity in $\mathcal{M}_{n-k}(q)$ and in D.

Proof. The coefficient matrix C of $(\mathrm{tr}_{q^k}(xy))^{(q^k-1)p^s}$ is diagonal since each term in the expansion of this expression has like degree in x

and y. Moreover $c_{jj} \equiv p_j \pmod{(xy)^{q^{nkd}-1} - 1}$. Since $q^n - 1$ divides $q^{nkd}-1$, the coefficient matrix of $f(x,y)-1$ appearing in Proposition 4.4 is

$$\left(\begin{array}{ccc|ccc|ccc} -1 & & & p_{q^n-1} & & & p_{q^{nkd}-q^n-1} & & \\ & p_1 & & & \ddots & & & \ddots & \\ & & \ddots & & & \ddots & & & \ddots \\ & & p_{q^n-2} & & & & & & p_{q^{nkd}-2} \end{array}\right).$$

Since each column has only one nonzero entry, the result follows from permuting the columns so those having nonzero in a given row are adjacent. Q.E.D.

The actual multiplicities of p^k as an elementary divisor of $\mathcal{M}_{n-1}$ arising from this formula seem to be difficult to compute. Thanks to Maple, here are some numerical results with new values in **boldface**:

	1	p	p^2	p^3	p^4	p^5	p^6
$PG_2(4)$	9	2	9				
$PG_3(4)$	16	**8**	**36**	**8**	16		
$PG_4(4)$	25	**20**	**100**	**50**	**100**	**20**	25
$PG_5(4)$	36	**40**	**225**	**180**	**402**	**180**	**225**
$PG_6(4)$	49	**70**	**441**	**490**	**1239**	**882**	**1239**
$PG_2(8)$	27	9	9	27			
$PG_3(8)$	64	**48**	**72**	**216**	**72**	**48**	64
$PG_4(8)$	125	**150**	**400**	**1015**	**750**	**750**	**1015**
$PG_2(9)$	36	8	36				
$PG_3(9)$	100	**128**	**363**	**128**	100		
$PG_4(9)$	225	**510**	**2055**	**1800**	**2055**	**510**	225

References

[1] S. C. Black and R. J. List, On certain Abelian groups associated with finite projective geometries, Geom. Dedicata, **33** (1990), 13–19.

[2] A. E. Brouwer and H. Wilbrink, Block designs, in "The Handbook of Incidence Geometries", (F. Buekenhout, ed.), North-Holland, 1994.

[3] C. W. Curtis, N. Iwahori and R. Kilmoyer, Hecke algebras and characters of parabolic type of finite groups with (B,N)-pairs, Inst. Hautes Études Sci. Publ. Math., **40** (1972), 81–116.

[4] C. W. Curtis and I. Reiner, Methods of Representation Theory, J. Wiley, New York, 1987.

[5] A. Frumkin and A. Yakir, Rank of inclusion matrices and modular representaion theory, Israel J. Math., **71** (1990), 309–320.

[6] N. Hamada, The rank of the incidence matrix of points and d-flats in finite geometries, J. Sci. Hiroshima Univ. Ser A-1, **32** (1968), 381–396.

[7] P. N. Hoefsmit, Representations of Hecke algebras of finite groups with BN-pairs of classical type, Ph. D. Thesis, Univ. British Columbia, 1974.

[8] R. Howlet and G. Lehrer, Representations of generic algebras of finite groups of Lie type, Trans. Amer. Math. Soc., **280** (1983), 753–777.

[9] N. Iwahori, On the structure of the Hecke ring of a Chevalley group over a finite field, J. Fac. Sci. Univ. Tokyo, **10** (1964), 215–236.

[10] E. S. Lander, "Symmetric Designs: an Algebraic Approach", London Math. Soc. Lecture Notes **74**, Cambridge Univ Press, 1983.

[11] R. A. Liebler, A representation theoretic appoach to finite geometries of spherical type, (1985), unpublished.

[12] K. J. C. Smith, On the p-rank of the incidence matrix of points and hyperplanes in a finite geometry, J. Combin. Theory, **7** (1969), 122–129.

[13] P. Terwilliger, The incidence algebra of a uniform poset, in "Coding Theory and Design Theory part I: Coding Theory", IMA volumes in Mathematics and its applications, Vol. 20, Springer, New York, 1990, pp. 193–212.

Department of Mathematics
Colorado State University
Fort Collins, Colorado 80523
U.S.A.

Advanced Studies in Pure Mathematics 24, 1996
Progress in Algebraic Combinatorics
pp. 285–308

Spin Models and Almost Bipartite 2-Homogeneous Graphs

Kazumasa Nomura

Abstract.

A connected graph of diameter d is said to be almost bipartite if it contains no cycle of length $2\ell+1$ for all $\ell < d$. An almost bipartite distance-regular graph $\Gamma = (X, E)$ is 2-homogeneous if and only if there are constants $\gamma_1, \ldots, \gamma_d$ such that $|\Gamma_{i-1}(u) \cap \Gamma_1(x) \cap \Gamma_1(y)| = \gamma_i$ holds for all $u \in X$ and for all x, $y \in \Gamma_i(u)$ with $\partial(x, y) = 2$ $(i = 1, \ldots, d)$.

In this paper, almost bipartite 2-homogeneous distance-regular graphs are classified. This determines triangle-free connected graphs affording spin models (for link invariants) with certain weights.

§1. Introduction

A spin model is one of the statistical mechanical models which were introduced by Vaughan Jones to construct invariants of knots and links [12]. A spin model is defined as a complex-valued symmetric function w on $X \times X$, where X is a finite set of "spins", satisfying several axioms. Each spin model S gives a corresponding link invariant through its partition function. Three examples of spin models are mentioned in Jones' paper [12]; Potts models, cyclic models and square models. It must be remarked that the Jones polynomial can be obtained from the partition function of the Potts models.

A connection between spin models and distance-regular graphs was found by François Jaeger [9] by constructing a new spin model on the Higman-Sims graph, a distance-regular graph of diameter $d = 2$ with $n = 100$ vertices, which was discovered by D. Higman and C. Sims [8], where we say that a spin model $S = (X, w)$ is constructed on a connected graph $\Gamma = (X, E)$ if $w(x, y)$ depends only on the distance $\partial(x, y)$ in the graph Γ. Jaeger [9] proved that the corresponding link invariant

Received April 10, 1995.
Revised June 7, 1995.

of the Higman-Sims model becomes a specialization of the Kauffman polynomial [14]. After Jaeger's discovery, a new infinite family of spin models were constructed on Hadamard graphs by the author [14]. The corresponding link invariants of the Hadamard models were determined by Jaeger [10,11], and then Jones [13] gave a pair of two links which can be detected by this invariant but not by Jones polynomial.

These examples of spin models can be constructed on almost bipartite distance-regular graphs. Moreover these graphs have extra regularity which we call 2-homogeneity; an almost bipartite distance-regular graph $\Gamma = (X, E)$ is 2-homogeneous if and only if $|\Gamma_{i-1}(u) \cap \Gamma_1(x) \cap \Gamma_1(y)|$ is a constant for all u, x, $y \in X$ with $\partial(u,x) = \partial(u,y) = i$, $\partial(x,y) = 2$ ($i = 1, \ldots, d$, where d denotes the diameter of Γ). In fact it was shown [21] that if a triangle-free connected graph affords a spin model with certain weights then the graph must be distance-regular and almost bipartite.

This paper contains two main results. At first we show that if a spin model is constructed on an almost bipartite distance-regular graph then the graph must be 2-homogeneous (under some conditions, see Theorem 4.3). Next we classify almost bipartite 2-homogeneous distance-regular graphs (Theorem 5.1). The proofs of these results are given in Section 4 and Section 5. In Section 2, some preliminaries on spin models and distance-regular graphs are given. In Section 3, two necessary and sufficient conditions (H1), (H2) for 2-homogeneity of almost bipartite distance-regular graphs are given. Then we slightly generalize of Yamazaki's sufficient condition for 2-homogeneity [22].

There are two generalizations of Jones' spin models by Kawagoe-Munemasa-Watatani [15] and Bannai-Bannai [1] (see also [2, 3, 16]). In this paper we restrict our interest to the original spin models defined in [12].

§2. Spin models and distance-regular graphs

2.1. Definition and examples of spin models

A *spin model* is a pair $S = (X, w)$ of a finite set of size $|X| = n > 0$ and a complex-valued function w on $X \times X$ such that (for all a, b, c in X)

$$\text{(S1)} \qquad w(a,b) = w(b,a) \neq 0,$$

$$\text{(S2)} \qquad \sum_{x \in X} w(a,x)w(b,x)^{-1} = n\delta_{a,b},$$

$$\text{(S3)} \quad \sum_{x \in X} w(a,x)w(b,x)w(c,x)^{-1} = \sqrt{n}\, w(a,b)w(a,c)^{-1}w(b,c)^{-1}.$$

The equation (S3) is called the "star-triangle" relation. The elements of X is called the *spins*, and the function w is called the (*Boltzmann*) *weight*. Putting $a = c$ in (S3), we have

$$\sum_{x \in X} w(b,x) = \sqrt{n}\, w(a,a)^{-1},$$

so that $w(a,a) = \alpha$ is a constant, called the *modulus* of S, which is independent of the choice of a in X.

The *weight matrix* of a spin model $S = (X, w)$, $|X| = n$, is a $n \times n$ matrix W, indexed by $X \times X$, whose (x,y)-entry is $W_{x,y} = w(x,y)$. For b, c in X, we consider a vector $\mathbf{u}_{bc}$ in the n-space $V = \mathbf{C}^n$, where the entries of the vectors are indexed by X, whose x-entry is given by

$$(\mathbf{u}_{bc})_x = \frac{w(b,x)}{w(c,x)}, \qquad (x \in X).$$

Then the condition (S3) can be written as

$$W\mathbf{u}_{bc} = \sqrt{n}\, w(b,c)^{-1}\mathbf{u}_{bc}.$$

This means the vector $\mathbf{u}_{bc}$ is an eigenvector of W for the eigenvalue $\sqrt{n}\, w(b,c)^{-1}$. It can be easily shown from (S2) that, for a fixed $b \in X$, the vectors $\mathbf{u}_{bc}$, $c \in X$ are linearly independent and hence form a basis of V. Therefore the values $\sqrt{n}\, w(b,c)^{-1}$, $c \in X$ give all the eigenvalues of W, where multiplicities are counted. This means that the multiplicity of an eigenvalue $\sqrt{n}\,\lambda^{-1}$ agrees with the number of $x \in X$ such that $w(b,x) = \lambda$ (thus this number does not depend on the choice of b). The vector $\mathbf{u}_{bb}$ becomes the all one vector $\mathbf{j}$, and it is an eigenvector of W corresponding the eigenvalue $\sqrt{n}\,\alpha^{-1}$ (α is the modulus). From condition (S2), the other vectors $\mathbf{u}_{bc}$, $b \neq c$ are orthogonal to $\mathbf{j}$.

Now we give three basic examples of spin models.

Potts model. Let X be a finite set with $n > 1$ elements. Let β be a solution of $\beta^2 + \beta^{-2} + \sqrt{n} = 0$ and put $\alpha = -\beta^{-3}$. Define a function w on $X \times X$ by

$$w(x,y) = \begin{cases} \alpha & x = y, \\ \beta & \text{otherwise.} \end{cases}$$

Then (X, w) is a spin model called the *Potts model* [12]. Potts model with $n = 2$ is also called the *Ising model.*

Cyclic model. Let $X = \{0, 1, \ldots, n-1\}$, and let θ be a primitive n-root of unity when n is odd, or a primitive $2n$-root of unity when n is even. Define a function w on $X \times X$ by

$$w(x,y) = \alpha \theta^{(x-y)^2},$$

where

$$\alpha^2 = \frac{\sqrt{n}}{\sum_{i=0}^{n-1} \theta^{i^2}}.$$

Then (X, w) becomes a spin model, called the *cyclic model* [2,6,12].

Square model. Let $X = \{1, 2, 3, 4\}$ and let α be an arbitrary non-zero complex number. Let us consider the following matrix:

$$W = \begin{pmatrix} \alpha & \alpha^{-1} & -\alpha & \alpha^{-1} \\ \alpha^{-1} & \alpha & \alpha^{-1} & -\alpha \\ -\alpha & \alpha^{-1} & \alpha & \alpha^{-1} \\ \alpha^{-1} & -\alpha & \alpha^{-1} & \alpha \end{pmatrix},$$

and define a function w on $X \times X$ by $w(x, y) = W_{x,y}$. Then (X, w) becomes a spin model, called the *square model* [7,12].

2.2. Preliminaries for distance-regular graphs

Let $\Gamma = (X, E)$ be a connected (undirected simple) graph of diameter d with the vertex set X and the edge set E with the usual metric ∂ on X. For vertices u, v and for integers r, s, define

$$\Gamma_r(u) = \{x \in X \mid \partial(u,x) = r\},$$

$$\mathrm{D}^r_s(u,v) = \Gamma_r(u) \cap \Gamma_s(v).$$

Γ is said to be *distance-regular* if there are integers b_r, c_r such that for any two vertices u, x at distance $r = \partial(u,x)$, there are precisely c_r neighbours of x in $\Gamma_{r-1}(u)$ and b_r neighbours of x in $\Gamma_{r+1}(u)$. In particular Γ is regular of valency $k = b_0$, and there are $a_r = k - c_r - b_r$ neighbours of x in $\Gamma_r(u)$. The parameters c_r, b_r, a_r $(r = 0, \ldots, d)$ satisfy (see [5], Proposition 4.1.6)

$$1 = c_1 \leq c_2 \leq \cdots \leq c_{d-1} \leq c_d,$$

$$k = b_0 \geq b_1 \geq \cdots \geq b_{d-1} \geq b_d = 0.$$

The array

$$\left\{\begin{matrix} 0 & c_1 & c_2 & \cdots & c_{d-1} & c_d \\ 0 & a_1 & a_2 & \cdots & a_{d-1} & a_d \\ k & b_1 & b_2 & \cdots & b_{d-1} & 0 \end{matrix}\right\}$$

is called the *intersection array* of Γ.

It is known (see [5], Section 4.1) that the parameters

$$p^t_{r,s} = |\mathrm{D}^r_s(u,v)|, \qquad (t = \partial(u,v))$$

are well-defined, i.e., these parameters depends only on r, s and $t = \partial(u,v)$, rather than on the individual vertices u, v with $t = \partial(u,v)$. The parameters $p^t_{r,s}$ are called the *intersection numbers* of Γ. Clearly $c_r = p^r_{r-1,1}$, $a_r = p^r_{r,1}$ and $b_r = p^r_{r+1,1}$ hold.

Let A_i $(i = 0, 1, \ldots, d)$ denote the i-th adjacency matrix of Γ, i.e., A_i is the $n \times n$ matrix, indexed by $X \times X$, whose (x,y)-entry is

$$(A_i)_{x,y} = \begin{cases} 1 & \partial(x,y) = i, \\ 0 & \text{otherwise.} \end{cases}$$

In particular, $A_0 = I$ the identity matrix of size n and $A_1 = A$ the usual adjacency matrix of Γ. The matrices $A_0, A_1, \ldots, A_d$ satisfy

$$A_iA_j = A_jA_i = \sum_{\ell=0}^{d} p^{\ell}_{ij} A_{\ell}.$$

In particular,

$$AA_i = b_{i-1}A_{i-1} + a_iA_i + c_{i+1}A_{i+1}$$

holds. Using this relation recursively, A_i can be written as a polynomial in A, i.e., there are polynomials $v_i(x)$ of degree i such that $A_i = v_i(A)$ holds for $i = 0, 1, \ldots, d$.

It is known that the adjacency matrix A has distinct eigenvalues $\theta_0 = k, \theta_1, \ldots, \theta_d$, and the corresponding eigenspaces $V_0, V_1, \ldots, V_d$ in $V = \mathbf{C}^n$ $(n = |X|)$ are mutually orthogonal (see [5], Section 4.1):

$$V = V_0 \oplus V_1 \oplus \cdots \oplus V_d \quad \text{(orthogonal sum).}$$

Remark that V_0 is the 1-dimensional subspace spanned by $\mathbf{j}$.

More precise descriptions of distance-regular graphs can be found in [4,5].

2.3. Spin models on distance-regular graphs

Let $\Gamma = (X, E)$ be a connected graph of diameter d with the usual metric ∂ on X. Let R_i $(i = 0, 1, \ldots, d)$ be the set of pairs (x, y) in $X \times X$ such that $\partial(x, y) = i$. Then $X \times X$ is partitioned into $d + 1$ relations:

$$X \times X = R_0 \cup R_1 \cup \cdots \cup R_d.$$

We consider spin models $S = (X, w)$ such that w takes a constant value t_i on R_i $(i = 0, 1, \ldots, d)$, i.e., $w(x, y) = t_i$ holds for all x, y in X at distance $\partial(x, y) = i$. In this case we say that the spin model $S = (X, w)$ is constructed on the graph $\Gamma = (X, E)$. We are particularly interested in spin models which are constructed on distance-regular graphs.

For three vertices x, y, z and for integers i, j, ℓ, define

$$P_{i,j,\ell}(x, y, z) = |\Gamma_i(x) \cap \Gamma_j(y) \cap \Gamma_\ell(z)|.$$

Lemma 2.1. *Let $\Gamma = (X, E)$ be a distance-regular graph of diameter d with the intersection numbers $p^\ell_{i,j}$, and let $t_0, \ldots, t_d$ be non-zero complex numbers. Define a function w on $X \times X$ by $w(x, y) = t_{\partial(x,y)}$. Then $S = (X, w)$ is a spin model if and only if the following conditions hold:*

(S2') *For $\ell = 1, \ldots, d$,*

$$\sum_{i=0}^{d} \sum_{j=0}^{d} p^\ell_{i,j} t_i t_j^{-1} = 0,$$

(S3') *For all x, y, z in X,*

$$\sum_{i=0}^{d} \sum_{j=0}^{d} \sum_{\ell=0}^{d} P_{i,j,\ell}(x, y, z) t_i t_j t_\ell^{-1} = \sqrt{n}\, t_{\partial(x,y)} t^{-1}_{\partial(x,z)} t^{-1}_{\partial(y,z)}.$$

Proof. It is not difficult to show that (S2), (S3) are equivalent to (S2'), (S3') respectively. Remark that (S1) holds for a spin model constructed on a connected graph. Q.E.D.

Now we give two examples which are constructed on distance-regular graphs.

Jaeger's Higman-Sims model. The Higman-Sims graph, which was discovered by D. Higman and C. Sims [8], is the unique distance-regular graph $\Gamma = (X, E)$ of diameter $d = 2$ with the following intersection array:

$$\left\{\begin{matrix} 0 & 1 & 6 \\ 0 & 0 & 16 \\ 22 & 21 & 0 \end{matrix}\right\}.$$

Γ has $|X| = 100$ vertices.

A spin model was constructed on the Higman-Sims graph by F. Jaeger [9] (see also [7]). Let $\tau = (1+\sqrt{5})/2$ and put

$$t_0 = (5\tau + 3)\sqrt{-1}, \quad t_1 = \tau\sqrt{-1}, \quad t_2 = (-\tau + 1)\sqrt{-1}.$$

Define a function w on $X \times X$ by $w(x, y) = t_{\partial(x,y)}$ for x, $y \in X$. Then $S = (X, w)$ becomes a spin model. The corresponding link invariant becomes a specialization of the Kauffman polynomial [7].

Hadamard model. Hadamard graphs are distance-regular graphs of diameter $d = 4$ with the following intersection array:

$$\left\{\begin{matrix} 0 & 1 & 2m & 4m-1 & 4m \\ 0 & 0 & 0 & 0 & 0 \\ 4m & 4m-1 & 2m & 1 & 0 \end{matrix}\right\},$$

where m is a positive integer. There is a natural correspondence between Hadamard graphs of valency $4m$ and Hadamard matrices of size $4m$ (see [5], Theorem 1.8.1). Let s, t_0, t_1 be complex numbers such that

$$s^2 + 2(2m-1)s + 1 = 0, \qquad t_0^2 = \frac{2\sqrt{m}}{(4m-1)s+1}, \qquad t_1^4 = 1,$$

and put

$$t_2 = st_0, \quad t_3 = -t_1, \quad t_4 = t_1.$$

Define a function w on $X \times X$ by $w(x, y) = t_{\partial(x,y)}$ for x, $y \in X$. Then $S = (X, w)$ is a spin model [17]. The corresponding link invariants of these models were determined by Jaeger [10,11].

§3. 2-Homogeneous distance-regular graphs

3.1. Definition of 2-homogeneity

Let $\Gamma = (X, E)$ be a distance-regular graph of diameter d. For a vertex x in X and for a subset A of X, let $e(x, A)$ denote the number of edges from x into A; $e(x, A) = |\Gamma_1(x) \cap A|$. Γ is said to be *t-homogeneous* (where t is an non-negative integer) if the following condition holds for all integers r, s, i, j and for all vertices u, v, u', v' with $\partial(u,v) = \partial(u',v') = t$:

$$x \in \mathrm{D}^r_s(u,v),\ x' \in \mathrm{D}^r_s(u',v') \Longrightarrow e(x,\, \mathrm{D}^i_j(u,v)) = e(x', \mathrm{D}^i_j(u',v')).$$

This means that, for two vertices u, v at distance t and for x in $\mathrm{D}^r_s(u,v)$, the number of edges from x into $\mathrm{D}^i_j(u,v)$ depends only on r, s, i, j rather than on the individual vertices u, v, x with $\partial(u,v) = t$ and $x \in \mathrm{D}^r_s(u,v)$.

It was shown [18] that, for a distance-regular graph Γ of diameter d in which $\mathrm{D}^1_1(u,v)$ is a (non-empty) clique for every edge uv, Γ is 1-homogeneous if and only if Γ is isomorphic to a regular near $2d$-gon (see [5], Section 6.4 for the definition).

Now we restrict our interest to the case $t = 2$. Let us consider the following conditions for a distance-regular graph Γ of diameter d:

(H1) There are integers $\delta_2, \ldots, \delta_d$ such that, for every pair of vertices u, v at distance $\partial(u,v) = 2$, and for every x in $\Gamma_r(u) \cap \Gamma_r(v)$, there are precisely δ_r neighbours of x in $\Gamma_{r-1}(u) \cap \Gamma_{r-1}(v)$ $(r = 2, \ldots, d)$.

(H2) There are integers $\gamma_1, \ldots, \gamma_d$ such that, for every vertex x and for every u, v in $\Gamma_r(x)$ with $\partial(u,v) = 2$, there are precisely γ_r common neighbours of u and v in $\Gamma_{r-1}(x)$ $(r = 1, \ldots, d)$.

Lemma 3.1. *Let $\Gamma = (X, E)$ be a distance-regular graph of diameter d. Then* (H1) *is equivalent to* (H2).

Proof. First assume Γ satisfies (H1). We must show that the size

$$|\Gamma_{r-1}(x) \cap \Gamma_1(u) \cap \Gamma_1(v)|$$

does not depend on the choice of x in X and u, v in $\Gamma_r(x)$ with $\partial(u,v) = 2$. Clearly this holds for $r = 1$. Assume $r > 1$. Fix a vertex x and fix two vertices u, v in $\Gamma_r(x)$ with $\partial(u,v) = 2$, and put

$$\mathrm{D}^i_j = \mathrm{D}^i_j(u,v) = \Gamma_i(u) \cap \Gamma_j(v)$$

for all integers i, j. We count the number N of paths of length $r-1$ from x to D^1_1. Let $x = x_r, x_{r-1}, \ldots, x_2, x_1$ be a path of length $r-1$ such that $x_1 \in \mathrm{D}^1_1$. Then we have $x_i \in \mathrm{D}^i_i$ for $i = 1, \ldots, r$. By (H1), there are precisely δ_i edges from x_i to D^{i-1}_{i-1} $(i = 2, \ldots, r)$. Hence we have

$$N = \delta_r \delta_{r-1} \cdots \delta_2.$$

On the other hand, for a fixed vertex y in $\Gamma_{r-1}(x) \cap \mathrm{D}^1_1$, there are precisely $c_{r-1}c_{r-2} \cdots c_2 c_1$ paths of length $r-1$ connecting x and y, since we have $\partial(x,y) = r-1$. Hence we have

$$N = |\Gamma_{r-1}(x) \cap \mathrm{D}^1_1| c_{r-1}c_{r-2} \cdots c_2 c_1.$$

So we obtain

$$|\Gamma_{r-1}(x) \cap \mathrm{D}_1^1| = \frac{\delta_r \delta_{r-1} \cdots \delta_2}{c_{r-1} c_{r-2} \cdots c_2 c_1}.$$

This means the number of common neighbours of u and v in $\Gamma_{r-1}(x)$ does not depend on the choice of x in X and u, v in $\Gamma_r(x)$ with $\partial(u, v) = 2$. Thus Γ satisfies (H2).

Next assume Γ satisfies (H2). We show by induction on r that the number of edges $e(x, \mathrm{D}_{r-1}^{r-1}(u, v))$ does not depend on the choice of u, v with $\partial(u, v) = 2$ and x in $\mathrm{D}_r^r(u, v)$ $(r = 2, \ldots, d)$. This holds when $r = 2$, since for $x \in \mathrm{D}_2^2(u, v)$ we have u, $v \in \Gamma_2(x)$ and so

$$e(x, \mathrm{D}_1^1(u, v)) = |\Gamma_{r-1}(x) \cap \Gamma_1(u) \cap \Gamma_1(v)| = \gamma_2.$$

Assume $r > 2$ and assume that there are constants $\delta_2, \ldots, \delta_{r-1}$ such that $e(x, \mathrm{D}_{i-1}^{i-1}(u, v)) = \delta_r$ holds for every $x \in \mathrm{D}_i^i(u, v)$ $(i = 2, \ldots, r-1)$. Fix two vertices u, $v \in X$ at distance $\partial(u, v) = 2$ and put $\mathrm{D}_j^i = \mathrm{D}_j^i(u, v)$. Pick a vertex $x \in \mathrm{D}_r^r$ and put

$$\delta(x) = e(x, \mathrm{D}_{r-1}^{r-1}).$$

We count the number N of paths $x = x_r, x_{r-1}, \ldots, x_1$ of length $r-1$ with $x_1 \in \mathrm{D}_1^1$. Since $x_i \in \mathrm{D}_i^i$ $(i = 1, \ldots, r)$ holds for every path $x = x_r, \ldots, x_1$ with $x_1 \in \mathrm{D}_1^1$,

$$N = \delta(x) \delta_{r-1} \delta_{r-2} \cdots \delta_2.$$

On the other hand, since there are precisely γ_r common neighbours y of u, v in $\Gamma_{r-1}(x)$ by (H2),

$$|\mathrm{D}_1^1 \cap \Gamma_{r-1}(x)| = \gamma_r.$$

Since for each vertex y in $\mathrm{D}_1^1 \cap \Gamma_{r-1}(x)$ there are precisely $c_{r-1} c_{r-2} \cdots c_1$ paths of length $r - 1$ connecting y and x, the number of paths is given by

$$N = |\mathrm{D}_1^1 \cap \Gamma_{r-1}(x)|\, c_{r-1} c_{r-2} \cdots c_2 c_1 = \gamma_r c_{r-1} c_{r-2} \cdots c_2 c_1.$$

Therefore we obtain

$$\delta(x) = \frac{\gamma_r c_{r-1} c_{r-2} \cdots c_2 c_1}{\delta_{r-1} \delta_{r-2} \cdots \delta_2}.$$

Thus Γ satisfies (H1). Q.E.D.

A connected graph Γ is said to be *bipartite* if there is no cycle of odd length, and *almost bipartite* if there is no cycle of odd length ℓ with $\ell < 2d+1$ (where d is the diameter of Γ). Let Γ be a distance-regular graph of diameter d with intersection numbers c_r, a_r, b_r ($r = 0, \dots, d$). Clearly Γ is bipartite if and only if $a_r = 0$ for $r = 0, \dots, d$, and Γ is almost bipartite if and only if $a_r = 0$ for $r = 0, \dots, d-1$.

Lemma 3.2. *Let Γ be an almost bipartite distance-regular graph of diameter d. Then Γ is 2-homogeneous if and only if Γ satisfies* (H1).

Proof. The condition (H1) says that $e(x, \mathrm{D}^{r-1}_{r-1}(u,v)) = \delta_r$ holds for every u, v, x with $\partial(u,v) = 2$ and $x \in \mathrm{D}^r_r(u,v)$. Hence (H1) holds if Γ is 2-homogeneous.

Fix two vertices u, v at distance $\partial(u,v) = 2$ and let us denote $\mathrm{D}^i_j = \mathrm{D}^i_j(u,v)$ for all i, j. Remark that D^i_j is empty for all i, j with $|i-j| > 2$ since $\partial(u,v) = 2$. Also remark that D^i_j is empty for all i, j with $i+j \equiv 1 \pmod 2$ and $i+j < 2d-1$ since there is no cycle of odd length $\ell < 2d+1$. Therefore the vertex set of Γ is partitioned into the following subsets:

$$\begin{array}{ccccccccc}
\mathrm{D}^0_2 & \mathrm{D}^1_3 & \mathrm{D}^2_4 & \cdots & \mathrm{D}^{d-3}_{d-1} & \mathrm{D}^{d-2}_d & & & \\
 & & & & & & \mathrm{D}^{d-1}_d & & \\
\mathrm{D}^1_1 & \mathrm{D}^2_2 & \mathrm{D}^3_3 & \cdots & \mathrm{D}^{d-2}_{d-2} & \mathrm{D}^{d-1}_{d-1} & & \mathrm{D}^d_d & \\
 & & & & & & \mathrm{D}^d_{d-1} & & \\
\mathrm{D}^2_0 & \mathrm{D}^3_1 & \mathrm{D}^4_2 & \cdots & \mathrm{D}^{d-1}_{d-3} & \mathrm{D}^d_{d-2} & & &
\end{array}$$

Remark that there is no edge connecting D^i_j and $\mathrm{D}^{i'}_{j'}$ if $|i-i'| > 1$ or $|j-j'| > 1$. Remark also that there is no edge inside D^i_j for all i, j with $i < d$ or $j < d$ since $a_1 = \cdots = a_{d-1} = 0$.

First we show that the number of edges $e(x, \mathrm{D}^i_j)$ ($x \in \mathrm{D}^r_s$) is determined by the intersection numbers for all r, s with $r \neq s$. For x in D^{r-2}_r we have

$$e(x, \mathrm{D}^{r-3}_{r-1}) = e(x, \Gamma_{r-3}(u)) = c_{r-2},$$

$$e(x, \mathrm{D}^{r-1}_{r-1}) = e(x, \Gamma_{r-1}(v)) - e(x, \mathrm{D}^{r-3}_{r-1}) = c_r - c_{r-2}.$$

Moreover when $r < d$ we have

$$e(x, \mathrm{D}^{r-1}_{r+1}) = e(x, \Gamma_{r+1}(v)) = b_r,$$

and when $r = d$ we have

$$e(x, \mathrm{D}^{d-1}_d) = e(x, \Gamma_{d-1}(u)) - e(x, \mathrm{D}^{d-1}_{d-1}) = b_{d-2} - (c_d - c_{d-2}).$$

For x in D_d^{d-1} we have

$$e(x, \mathrm{D}_d^{d-2}) = e(x, \Gamma_{d-1}(u)) = c_{d-1},$$

$$e(x, \mathrm{D}_{d-1}^{d-1}) \leq e(x, \Gamma_{d-1}(u)) = a_{d-1} = 0,$$

$$e(x, \mathrm{D}_{d-1}^{d}) = e(x, \Gamma_{d-1}(v)) - e(x, \mathrm{D}_{d-1}^{d-1}) = c_d,$$

$$e(x, \mathrm{D}_d^d) = e(x, \Gamma_d(u)) - e(x, \mathrm{D}_{d-1}^d) = b_{d-1} - c_d.$$

Thus $e(x, \mathrm{D}_j^i)$ is determined by the intersection numbers for $x \in \mathrm{D}_s^r$ with $r \neq s$. Moreover for x in D_1^1 we have

$$e(x, \mathrm{D}_2^0) = e(x, \mathrm{D}_0^2) = c_1, \qquad e(x, \mathrm{D}_2^2) = b_1 - c_1.$$

Now we assume Γ satisfies (H1) and let $x \in \mathrm{D}_r^r$ $(2 \leq r \leq d)$. Then by (H1) we have

$$e(x, \mathrm{D}_{r-1}^{r-1}) = \delta_r.$$

When $r < d$ we have

$$e(x, \mathrm{D}_{r+1}^{r-1}) = e(x, \Gamma_{r-1}(u)) - e(x, \mathrm{D}_{r-1}^{r-1}) = c_r - \delta_r,$$

$$e(x, \mathrm{D}_{r-1}^{r+1}) = e(x, \Gamma_{r-1}(v)) - e(x, \mathrm{D}_{r-1}^{r-1}) = c_r - \delta_r,$$

$$e(x, \mathrm{D}_{r+1}^{r+1}) = e(x, \Gamma_{r+1}(u)) - e(x, \mathrm{D}_{r-1}^{r+1}) = b_r - (c_r - \delta_r),$$

here remark that there is no edge between D_{d-1}^{d-1} and D_{d-1}^d. For $x \in \mathrm{D}_d^d$ we have

$$e(x, \mathrm{D}_d^{d-1}) = e(x, \Gamma_{d-1}(u)) - e(x, \mathrm{D}_{d-1}^{d-1}) = c_d - \delta_d,$$

$$e(x, \mathrm{D}_{d-1}^d) = e(x, \Gamma_{d-1}(v)) - e(x, \mathrm{D}_{d-1}^{d-1}) = c_d - \delta_d,$$

$$e(x, \mathrm{D}_d^d) = e(x, \Gamma_d(u)) - e(x, \mathrm{D}_{d-1}^d) = a_d - (c_d - \delta_d).$$

Therefore Γ is 2-homogeneous. Q.E.D.

3.2. A sufficient condition for 2-homogeneity

Yamazaki [22] proved that every bipartite distance-regular graph with an eigenvalue of multiplicity k (k is the valency) satisfies condition (H1). Here we give a slight generalization.

Proposition 3.3. *Let Γ be an almost bipartite distance-regular graph of valency k. If the adjacency matrix A of Γ has an eigenvalue θ of multiplicity f with $1 < f \leq k$, then Γ is 2-homogeneous.*

In the following we prove the above proposition in a similar way as Yamazaki's proof [22].

Let $\Gamma = (X, E)$ be an almost bipartite distance-regular graph of diameter d and valency k. We may assume $d > 1$ and $k > 2$ since the graph is clearly 2-homogeneous if $d = 1$ or $k \leq 2$. Let c_i, b_i and a_i $(i = 0, 1, \ldots, d)$ be the usual intersection numbers of Γ. We have $a_1 = \cdots = a_{d-1} = 0$ since Γ is almost bipartite. In particular Γ has no triangle. Assume that the adjacency matrix A of Γ has an eigenvalue θ of multiplicity f with $1 < f \leq k$. By [5] Proposition 4.4.1, we have a mapping $^- : X \longrightarrow \mathbf{R}^f$ such that $\langle \overline{x}, \overline{y} \rangle = u_i$ holds for all x, y at distance $\partial(x, y) = i$, where $\langle \overline{x}, \overline{y} \rangle$ denote the ordinary inner product of the Euclidean space $\mathbf{R}^f$, and $(u_0, u_1, \ldots, u_d)$ is the standard sequence corresponding to θ, i.e., it is the sequence defined by the recurrence: $u_0 = 1$, $u_1 = \theta/k$, $c_i u_{i-1} + b_i u_{i+1} = \theta u_i$ $(i = 1, \ldots, d-1)$. It is known that an eigenvalue η of A has multiplicity 1 if and only if $\eta = \pm k$ [5] Proposition 4.4.8. So $\theta \neq \pm k$ by our assumption $f > 1$. Then we obtain $u_2 \neq u_0 = 1$ from the above recurrence. Hence $\overline{x} \neq \overline{y}$ holds for all vertices x, y with $\partial(x, y) = 2$.

Lemma 3.4. *Let $\sigma : Y \longrightarrow X$ be a mapping from a subset Y of X which preserves distances. Then for real numbers λ_y $(y \in Y)$, $\sum_{y \in Y} \lambda_y \overline{y} = 0$ if and only if $\sum_{y \in Y} \lambda_y \overline{\sigma(y)} = 0$.*

Proof. Use $\langle \overline{x}, \overline{y} \rangle = u_{\partial(x,y)}$ to show

$$\| \textstyle\sum_{y \in Y} \lambda_y \overline{\sigma(y)} \| = \| \sum_{y \in Y} \lambda_y \overline{y} \| = 0.$$

Q.E.D.

For a subset Y of X, we denote $\overline{Y} = \{ \overline{y} \mid y \in Y \}$, $\widetilde{Y} = \sum_{y \in Y} \overline{y}$.

Lemma 3.5. *For every $x \in X$, $\overline{\Gamma_1(x) \cup \{x\}}$ spans a k-dimensional subspace of $\mathbf{R}^f$. In particular $f = k$.*

Proof. Assume that the subspace U spanned by $\overline{\Gamma_1(x) \cup \{x\}}$ has dimension $m + 1 < k$. Choose m vertices $y_1, \ldots, y_m$ in $\Gamma_1(x)$ such that $\overline{x}, \overline{y_1}, \ldots, \overline{y_m}$ form a basis of U, and choose two distinct vertices y, $y' \in \Gamma_1(u)$ which are different from $y_1, \ldots, y_m$ (here remark that $m \leq k - 2$). Write $\overline{y} = \lambda \overline{x} + \sum_{i=1}^{m} \lambda_i \overline{y_i}$ $(\lambda, \lambda_i \in \mathbf{R})$. Applying Lemma 3.4 for $Y = \{x, y, y_1, \ldots, y_m\}$ and $\sigma : Y \longrightarrow X$ such that $\sigma(y) = y'$,

$\sigma(x) = x$, $\sigma(y_i) = y_i$ $(i = 1, \ldots, m)$, we obtain $\overline{y'} = \lambda\overline{x} + \sum_{i=1}^{m} \lambda_i \overline{y_i}$. Hence $\overline{y} = \overline{y'}$, contradicting $\partial(y, y') = 2$. Q.E.D.

Lemma 3.6. *There are constants λ_i, μ_i, ν_i $(i = 2, \ldots, d)$ such that $\overline{v} = \lambda_i \overline{x} + \nu_i \widetilde{C} + \mu_i \widetilde{B}$ holds for all v, x with $i = \partial(v, x)$, where $C = \Gamma_1(x) \cap \Gamma_{i-1}(v)$ and $B = \Gamma_1(x) \setminus C$.*

Proof. Remark that $B = \Gamma_1(x) \cap \Gamma_{i+1}(v)$ when $i < d$, and $B = \Gamma_1(x) \cap \Gamma_i(v)$ when $i = d$. From Lemma 3.5, $\overline{v}$ can be written as

$$\overline{v} = \lambda\overline{x} + \sum_{y \in C} \nu_y \overline{y} + \sum_{z \in B} \mu_z \overline{z}$$

for some λ, ν_y, $\mu_z \in \mathbf{R}$ $(y \in C,\ z \in B)$. We would like to show that $\nu_{y_1} = \nu_{y_2}$ holds for all y_1, $y_2 \in C$. Let y_1, $y_2 \in C$ with $y_1 \neq y_2$. We use Lemma 3.4 for $Y = \{v, x\} \cup B \cup C$ and $\sigma : Y \longrightarrow X$ which fixes all vertices in Y except $\sigma(y_1) = y_2$, $\sigma(y_2) = y_1$. Clearly σ preserves distances. Then the above equation implies

$$\overline{v} = \lambda\overline{x} + \nu_{y_1}\overline{y_2} + \nu_{y_2}\overline{y_1} + \sum_{y \in C \setminus \{y_1, y_2\}} \nu_y \overline{y} + \sum_{z \in B} \mu_z \overline{z}.$$

These two equations imply $\nu_{y_1}\overline{y_1} + \nu_{y_2}\overline{y_2} = \nu_{y_1}\overline{y_2} + \nu_{y_2}\overline{y_1}$, and this becomes $(\nu_{y_1} - \nu_{y_2})(\overline{y_1} - \overline{y_2}) = 0$. Here we have $\overline{y_1} \neq \overline{y_2}$ by $\partial(y_1, y_2) = 2$, so $\nu_{y_1} = \nu_{y_2}$. This means $\nu_y = \nu$ is a constant for $y \in C$. In the same way, $\mu_z = \mu$ is a constant for $z \in B$. Thus $\overline{v} = \lambda\overline{x} + \nu\widetilde{C} + \mu\widetilde{B}$. Use Lemma 3.4 again to show that λ, μ, ν do not depend on v and x with $\partial(v, x) = i$. Q.E.D.

Fix two vertices v, w with $\partial(v, w) = 2$ and put $\mathrm{D}_s^r = \mathrm{D}_s^r(v, w)$. We have

$$\|\overline{v} - \overline{w}\|^2 = \langle \overline{v}, \overline{v} \rangle + \langle \overline{w}, \overline{w} \rangle - 2\langle \overline{v}, \overline{w} \rangle = u_0 + u_0 - 2u_2 = 2(u_0 - u_2).$$

First take $x \in \mathrm{D}_i^i$ $(1 < i < d)$ and put $A = \Gamma_1(x) \cap \mathrm{D}_{i-1}^{i-1}$, $B = \Gamma_1(x) \cap \mathrm{D}_{i+1}^{i-1}$, $C = \Gamma_1(x) \cap \mathrm{D}_{i-1}^{i+1}$, $D = \Gamma_1(x) \cap \mathrm{D}_{i+1}^{i+1}$. Then we have a partition $\Gamma_1(x) = A \cup B \cup C \cup D$. Clearly we have $|A| + |B| = |A| + |C| = c_i$, so that $|B| = |C|$. By Lemma 3.6, we have

$$\overline{v} = \lambda_i \overline{x} + \nu_i(\widetilde{A} + \widetilde{B}) + \mu_i(\widetilde{C} + \widetilde{D}),$$

$$\overline{w} = \lambda_i \overline{x} + \nu_i(\widetilde{A} + \widetilde{C}) + \mu_i(\widetilde{B} + \widetilde{D}).$$

Hence

$$\|\overline{v}-\overline{w}\|^2 = \|(\nu_i-\mu_i)(\widetilde{B}-\widetilde{C})\|^2 = (\nu_i-\mu_i)^2\left(\|\widetilde{B}\|^2+\|\widetilde{C}\|^2-2\langle\widetilde{B},\,\widetilde{C}\rangle\right).$$

Here we have $\|\widetilde{B}\|^2 = \|\widetilde{C}\|^2 = |B|u_0 + |B|(|B|-1)u_2$ and $\langle\widetilde{B},\,\widetilde{C}\rangle = |B|^2u_2$. Hence $\|\overline{v}-\overline{w}\|^2 = 2(\nu_i-\mu_i)^2|B|(u_0-u_2)$. Therefore we obtain $(\nu_i-\mu_i)^2|B| = 1$ and hence $|A| = c_i - |B| = c_i - (\nu_i-\mu_i)^{-2}$. This means the size of $\Gamma_1(x)\cap \mathrm{D}_{i-1}^{i-1}$ depends only on i.

Next take $x \in \mathrm{D}_d^d$ and put $A = \Gamma_1(x)\cap\mathrm{D}_{d-1}^{d-1}$, $B = \Gamma_1(x)\cap\mathrm{D}_d^{d-1}$, $C = \Gamma_1(x)\cap\mathrm{D}_{d-1}^{d}$, $D = \Gamma_1(x)\cap\mathrm{D}_d^d$. Then we can show that $|A| = c_i - (\nu_i-\mu_i)^{-2}$ in the same way.

Thus Γ satisfies (H1) and hence Γ is 2-homogeneous by Lemma 3.2.

§4. Graphs with spin model structure

4.1. An observation

Here we observe that the examples of spin models given in Section 2 can be constructed on distance-regular graphs. Jaeger's Higman-Sims model and the Hadamard models are constructed on distance-regular graphs with the intersection arrays:

$$\left\{\begin{matrix} 0 & 1 & 6 \\ 0 & 0 & 16 \\ 22 & 21 & 0 \end{matrix}\right\},$$

and

$$\left\{\begin{matrix} 0 & 1 & 2m & 4m-1 & 4m \\ 0 & 0 & 0 & 0 & 0 \\ 4m & 4m-1 & 2m & 1 & 0 \end{matrix}\right\}.$$

The Potts models with n spins is constructed on a complete graph K_n, which is a distance-regular graph of diameter $d=1$ with the intersection array

$$\left\{\begin{matrix} 0 & 1 \\ 0 & k-1 \\ k & 0 \end{matrix}\right\}, \qquad k = n-1.$$

The weights are given by $t_0 = \alpha$, $t_1 = \beta$, where $\beta^2+\beta^{-2}+\sqrt{n} = 0$ and $\alpha = -\beta^{-3}$.

The cyclic model with n spins is constructed on the n-cycle C_n which is a distance-regular graph of diameter d with the intersection array:

$$\left\{\begin{matrix} 0 & 1 & 1 & \cdots & 1 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 2 & 1 & 1 & \cdots & 1 & 0 \end{matrix}\right\} \quad \text{when } n = 2d+1,$$

or

$$\begin{Bmatrix} 0 & 1 & 1 & \cdots & 1 & 2 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ 2 & 1 & 1 & \cdots & 1 & 0 \end{Bmatrix} \quad \text{when } n = 2d.$$

The weights are given by $t_i = \alpha\theta^{i^2}$ $(i = 0, \ldots, d)$, where θ is a primitive n-root of unity if $n = 2d + 1$, a primitive $2n$-root of unity if $n = 2d$, and $\alpha = \sqrt{n}/(\sum_{i=0}^{n-1} \theta^{i^2})$.

The square model is constructed on the 4-cycle C_4 with $t_0 = \alpha$, $t_1 = \alpha^{-1}$, $t_2 = -\alpha$, where α is a non-zero complex number.

Observe that all the above distance-regular graphs are almost bipartite. Moreover, as easily observed, each successive three terms t_{i-1}, t_i, t_{i+1} are distinct $(0 < i < d)$ in each of the above spin models except the square model with $\alpha = \pm 1$.

Motivated by the above observation, the author obtained the following result [21].

Theorem 4.1. *Let $\Gamma = (X, E)$ be a connected graph of diameter d which has no 3-cycle. Let $t_0, \ldots, t_d$ be non-zero complex numbers such that $t_1 \neq t_i$ and $t_{i-2} \neq t_i \neq t_{i-1}$ for $i = 2, \ldots, d$. Define a function w on $X \times X$ by $w(x, y) = t_{\partial(x,y)}$ for x, $y \in X$. If $S = (X, w)$ is a spin model, then Γ is an almost bipartite distance-regular graph.*

This was obtained by "localizing" the star-triangle relation (S3). This technique of localization was introduced in [19].

4.2. 2-homogeneity

Lemma 4.2. *Let $\Gamma = (X, E)$ be a distance-regular graph of diameter $d > 1$ and valency k, and let $t_0, \ldots, t_d$ be non-zero complex numbers such that $t_i \neq t_1$ for $i = 2, \ldots, d$. Assume $S = (X, w)$ is a spin model, where w is a function on $X \times X$ defined by $w(x, y) = t_{\partial(x,y)}$ for x, $y \in X$. Then the adjacency matrix A of Γ has an eigenvalue θ of multiplicity f with $1 < f \leq k$.*

Proof. Let $\theta_0 = k, \theta_1, \ldots, \theta_d$ be the eigenvalues of the adjacency matrix A of Γ and let V_i be the eigenspace corresponding to $\theta_i, i = 0, \ldots, d$, where V_0 is the 1-dimensional subspace of $V = \mathbf{C}^n$ spanned by the all 1 vector $\mathbf{j}$. V splits into an orthogonal direct sum:

$$V = V_0 \oplus V_1 \oplus \cdots \oplus V_d \qquad \text{(orthogonal)}.$$

On the other hand, let $\mathbf{u}_{bc}$, b, $c \in X$ be the vector defined in Section 2.1, which is an eigenvector of the weight matrix W of S for the eigenvalue $\sqrt{n}\, w(b, c)^{-1}$.

Now fix a vertex $b \in X$. Then the vectors $\mathbf{u}_{bc}$, $c \in X$, form a basis of V. Let V_i' be the subspace of V spanned by the vectors $\mathbf{u}_{bc}$, $c \in \Gamma_i(b)$ $(i = 0, \ldots, d)$. Remark that $V_0' = \langle \mathbf{j} \rangle = V_0$. So V splits into a direct sum:

$$V = V_0 \oplus V_1' \oplus \cdots \oplus V_d',$$

where we have $V_i' \subset V_0^\perp$ for $i = 1, \ldots, d$. Since $\mathbf{u}_{bc}$ is an eigenvector of W for the eigenvalue $\sqrt{n}\, w(b,c)^{-1}$, V_i' is included in the eigenspace of W for the eigenvalue $\sqrt{n}\, t_i^{-1}$, $i = 0, \ldots, d$. Since $t_1 \neq t_i$ for $i = 2, \ldots, d$, the eigenspace of W for the eigenvalue $\sqrt{n}\, t_1^{-1}$ is included in $V_0 \oplus V_1'$. Now consider the action of W on

$$V_0^\perp = V_1' \oplus \cdots \oplus V_d'.$$

Then V_1' is the eigenspace of W in $V_0^\perp$ for the eigenvalue $\sqrt{n}\, t_1^{-1}$.

On the other hand, W is written as

$$W = \sum_{i=0}^{d} t_i A_i,$$

where A_i denotes the i-th adjacency matrix of the distance-regular graph Γ $(i = 0, \ldots, d)$. Since A_i is a polynomial in A, $A_i = v_i(A)$, W is written as a polynomial in A:

$$W = \sum_{i=0}^{d} t_i v_i(A).$$

Hence for each eigenvector $\mathbf{x}$ of A for the eigenvalue θ_j of A, $j > 0$, we have

$$W\mathbf{x} = \sum_{i=0}^{d} t_i v_i(A)\mathbf{x} = \sum_{i=0}^{d} t_i v_i(\theta_j)\mathbf{x},$$

so $\mathbf{x}$ is an eigenvector of W for the eigenvalue $\sum_{i=0}^{d} t_i v_i(\theta_j)$. Since $\mathbf{x} \in V_0^\perp$, $\mathbf{x}$ must belong to some eigenspace (in $V_0^\perp$) of W.

Therefore we can conclude that V_1' is a sum of some eigenspaces of A, say:

$$V_1' = V_1 \oplus \cdots \oplus V_\ell,$$

so that

$$k = \dim V_1' = f_1 + \cdots + f_\ell,$$

where $f_i = \dim V_i$. This implies $f_i \leq k$ $(i = 1, \ldots, \ell)$. We must show that $1 < f_i \leq k$ holds for some i $(1 \leq i \leq \ell)$. If $\ell = 1$ then we have $f_1 = k$ and $f_1 > 1$ since $k > 1$ by our assumption $d > 1$. So we may assume $\ell > 1$. If $f_i > 1$ holds for some i, then we have the conclusion. So we may assume $f_1 = \cdots = f_\ell = 1$. Now it is known that an eigenvalue θ of a distance-regular graph has multiplicity 1 if and only if $\theta = \pm k$ [5] Proposition 4.4.8. Hence $f_i = 1$ occurs at most one i, that is when $\theta_i = -k$ (remark that $\theta_i \neq k$ since $\theta_0 = k$). This implies $\ell = 1$, a contradiction. Q.E.D.

Theorem 4.3. *Let $\Gamma = (X, E)$ be an almost bipartite distance-regular graph of diameter d, and let $t_0, t_1, \ldots, t_d$ be non-zero complex numbers such that $t_1 \neq t_i$ for $i = 2, \ldots, d$. If $S = (X, w)$ is a spin model with the weight w defined by $w(x, y) = t_{\partial(x,y)}$, x, $y \in X$, then Γ is 2-homogeneous.*

Proof. It is obtained from Lemma 4.2 and Proposition 3.3. Q.E.D.

Corollary 4.4. *Let $\Gamma = (X, E)$ be a triangle-free connected graph of diameter d, and let $t_0, \ldots, t_d$ be non-zero complex numbers such that $t_1 \neq t_i$ and $t_{i-2} \neq t_i \neq t_{i-1}$ for $i = 2, \ldots, d$. If $S = (X, w)$ is a spin model with the weight w defined by $w(x, y) = t_{\partial(x,y)}$, x, $y \in X$, then Γ is an almost bipartite 2-homogeneous distance-regular graph.*

Proof. It is obtained from Theorem 4.1 and Theorem 4.3. Q.E.D.

Remark. The assumption 'triangle-free' in Corollary 4.4 is essential. Actually there exists a distance-regular graph Γ (with triangles) such that Γ affords a spin model structure with weights $t_0, \ldots, t_d$ satisfying the same conditions but Γ is not 2-homogeneous. Also remark that every connected graph can have a spin model structure with the weights $t_1 = \cdots = t_d$ (Potts model), and so we need some conditions on the weights $t_0, \ldots, t_d$ in Corollary 4.4.

§5. Classification of almost bipartite 2-homogeneous graphs

In this section we determine the intersection arrays of almost bipartite 2-homogeneous distance-regular graphs.

Theorem 5.1. *Let Γ be an almost bipartite 2-homogeneous distance-regular graph of diameter $d > 0$ and valency k. Then Γ has one of the following intersection arrays:*

(1) $\left\{ \begin{matrix} 0 & 1 \\ 0 & k-1 \\ k & 0 \end{matrix} \right\}, \quad k > 0,$

(2) $\left\{\begin{matrix} 0 & 1 & k \\ 0 & 0 & 0 \\ k & k-1 & 0 \end{matrix}\right\}$, $k > 1$,

(3) $\left\{\begin{matrix} 0 & 1 & c \\ 0 & 0 & k-c \\ k & k-1 & 0 \end{matrix}\right\}$, $\begin{matrix} k = \gamma(\gamma^2+3\gamma+1), \\ c = \gamma(\gamma+1),\ \gamma > 0, \end{matrix}$

(4) $\left\{\begin{matrix} 0 & 1 & k-1 & k \\ 0 & 0 & 0 & 0 \\ k & k-1 & 1 & 0 \end{matrix}\right\}$, $k > 1$,

(5) $\left\{\begin{matrix} 0 & 1 & 2\gamma & 4\gamma-1 & 4\gamma \\ 0 & 0 & 0 & 0 & 0 \\ 4\gamma & 4\gamma-1 & 2\gamma & 1 & 0 \end{matrix}\right\}$, $\gamma > 0$,

(6) $\left\{\begin{matrix} 0 & 1 & c & k-c & k-1 & k \\ 0 & 0 & 0 & 0 & 0 & 0 \\ k & k-1 & k-c & c & 1 & 0 \end{matrix}\right\}$, $\begin{matrix} k = \gamma(\gamma^2+3\gamma+1), \\ c = \gamma(\gamma+1),\ \gamma > 0, \end{matrix}$

(7) $\left\{\begin{matrix} 0 & 1 & \cdots & 1 & 2 \\ 0 & 0 & \cdots & 0 & 0 \\ 2 & 1 & \cdots & 1 & 0 \end{matrix}\right\}$, $d > 1$,

(8) $\left\{\begin{matrix} 0 & 1 & \cdots & 1 & 1 \\ 0 & 0 & \cdots & 0 & 1 \\ 2 & 1 & \cdots & 1 & 0 \end{matrix}\right\}$, $d > 1$,

(9) $\left\{\begin{matrix} 0 & 1 & 2 & 3 & \cdots & k-1 & k \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ k & k-1 & k-2 & k-3 & \cdots & 1 & 0 \end{matrix}\right\}$, $k = d$,

(10) $\left\{\begin{matrix} 0 & 1 & 2 & 3 & \cdots & d-1 & d \\ 0 & 0 & 0 & 0 & \cdots & 0 & d+1 \\ 2d+1 & 2d & 2d-1 & 2d-2 & \cdots & d+2 & 0 \end{matrix}\right\}$ $d > 1$.

Remark. The intersection arrays in the above list are realized by the following graphs:

(1) complete graph K_{k+1},
(2) complete bipartite graph $K_{k,k}$,
(3) antipodal quotient of 5-dimensional hypercube when $\gamma = 1$, Higman-Sims graph when $\gamma = 2$, the existence of graphs is unknown when $\gamma > 2$,
(4) complement of $2 \times (k+1)$-grid,
(5) Hadamard graph of valency $k = 4\gamma$,
(6) antipodal double cover of (3),
(7) cycle C_{2d+1} of length $2d+1$,
(8) cycle C_{2d} of length $2d$,
(9) d-dimensional hypercube,

(10) antipodal quotient of $(2d+1)$-dimensional hypercube.

Now we prove Theorem 5.1. Let $\Gamma = (X, E)$ be an almost bipartite 2-homogeneous distance-regular graph of diameter d and valency k with the intersection array:

$$\left\{ \begin{array}{cccccc} 0 & c_1 & c_2 & \cdots & c_{d-1} & c_d \\ 0 & a_1 & a_2 & \cdots & a_{d-1} & a_d \\ b_0 & b_1 & b_2 & \cdots & b_{d-1} & 0 \end{array} \right\}.$$

We have $a_i = 0$ $(i = 1, \ldots, d-1)$, $c_1 = 1$, $b_0 = k$, $b_1 = k-1$ and $a_d = k - c_d$. If $k \le 2$ or $d \le 1$, then Γ is isomorphic to a cycle or a complete graph and the intersection array of Γ becomes (1), (7) or (8). So in the following we assume $k > 2$ and $d > 1$. In particular we have $a_1 = 0$ and hence Γ has no 3-cycle.

By Lemma 3.1 and Lemma 3.2, Γ satisfies condition (H2), so that there are constants $\gamma_1, \ldots, \gamma_d$ such that

$$\gamma_i = |\Gamma_{i-1}(u) \cap \Gamma_1(x) \cap \Gamma_1(y)|$$

holds for all vertices u, x, $y \in X$ with $\partial(u,x) = \partial(u,y) = i$ and $\partial(x,y) = 2$ $(i = 1, \ldots, d)$.

Lemma 5.2. (i) $c_2 > 1$,
(ii) $(k-2)(\gamma_2 - 1) = (c_2 - 1)(c_2 - 2)$,
(iii) $\gamma_i(c_{i+1} - 1) = c_i(c_2 - 1)$, $(0 < i < d)$,
(iv) $(c_2 - 1)(\gamma_i - 1) = (c_i - 1)(\gamma_2 - 1)$, $(0 < i < d)$.

Proof. Fix a vertex u in X.

(i) We claim that $\gamma_i > 0$ if $c_i = 1$. Pick a vertex w in $\Gamma_{i-1}(u)$. Then w has at least two neighbours x, y in $\Gamma_i(u)$, since we have $b_{i-1} = k - c_{i-1} \ge k - c_i = k - 1 > 1$. So we have $\partial(x,y) = 2$ and $w \in \Gamma_{i-1}(u) \cap \Gamma_1(x) \cap \Gamma_1(y)$, and hence $\gamma_i > 0$.

First assume $c_d = 1$. We have $\gamma_d > 0$ as shown above. Each vertex v in $\Gamma_d(u)$ has at least two distinct neighbours x, y in $\Gamma_d(u)$ since $a_d = k - c_d = k - 1 \ge 2$. Then $\partial(x,y) = 2$ since Γ has no 3-cycle, and hence x and y has at least one common neighbour z in $\Gamma_{d-1}(u)$ by $\gamma_d > 0$. We have $\partial(v,z) = 2$ and x, y are common neighbours of v and z, so that $c_2 > 1$.

Next assume $c_d > 1$. Since $1 = c_1 \le c_2 \le \cdots \le c_d$ and $c_d > 1$, there is an integer r such that $1 = c_1 = c_2 = \cdots = c_r < c_{r+1}$. Pick a vertex z in $\Gamma_{r+1}(u)$. Since $c_{r+1} > 1$, z has at least two distinct neighbours x', $y' \in \Gamma_r(u)$. Since $\partial(x', y') = 2$ and $\gamma_r > 0$ by the above claim, x' and y'

have a common neighbour v in $\Gamma_{r-1}(u)$. Then $z \in \Gamma_2(v)$, and z has two distinct neighbours x', y' in $\Gamma_1(v)$. This implies $c_2 \geq 2$.

(ii) Fix an edge vw with $v \in \Gamma_1(u)$ and $w \in \Gamma_2(u)$. We count the number N of edges xy with $x \in \Gamma_1(u) \cap \Gamma_1(w) \cap \Gamma_2(v)$ and $y \in \Gamma_1(v) \cap \Gamma_2(u) \cap \Gamma_2(w)$ in two different ways. Since $w \in \Gamma_2(u)$, there are precisely $c_2 - 1$ vertices $x \in \Gamma_1(u) \cap \Gamma_1(w)$ with $x \neq v$. Fix such a vertex x. Since $x \in \Gamma_2(v)$, there are precisely $c_2 - 2$ vertices $y \in \Gamma_1(v) \cap \Gamma_1(x)$ with $y \neq u$, $y \neq w$. So we have $N = (c_2 - 1)(c_2 - 2)$. On the other hand, there are precisely $k - 2$ vertices $y \in \Gamma_2(u) \cap \Gamma_1(v)$ with $y \neq w$. Fix such a vertex y. Since w, $y \in \Gamma_2(u)$ and $\partial(w, y) = 2$, w and y have precisely $\gamma_2 - 1$ common neighbours x in $\Gamma_1(u)$ with $x \neq v$. So we obtain $N = (k - 2)(\gamma_2 - 1)$.

(iii) Fix an edge vw with $v \in \Gamma_i(u)$ and $w \in \Gamma_{i+1}(u)$. We count the number N of edges xy with $x \in \Gamma_{i-1}(u) \cap \Gamma_1(v)$ and $y \in \Gamma_i(u) \cap \Gamma_1(w) \cap \Gamma_2(v)$ in two different ways. Since $v \in \Gamma_i(u)$, v has precisely c_i neighbours x in $\Gamma_{i-1}(u)$. Fix such a vertex x. Since $w \in \Gamma_2(x)$, w has precisely $c_2 - 1$ neighbours y in $\Gamma_1(x)$ with $y \neq v$. Hence we have $N = c_i(c_2 - 1)$. On the other hand, since $w \in \Gamma_{i+1}(u)$, w has precisely $c_{i+1} - 1$ neighbours y in $\Gamma_i(u)$ with $y \neq v$. Fix such a vertex y. Since v, $y \in \Gamma_i(u)$ and $\partial(v, y) = 2$, v and y have precisely γ_i common neighbours x in $\Gamma_{i-1}(u)$. So we obtain $N = (c_{i+1} - 1)\gamma_i$.

(iv) Fix a path zvw with $z \in \Gamma_{i-1}(u)$, $v \in \Gamma_i(u)$, $w \in \Gamma_{i+1}(u)$, and count the number of edges xy with $x \in \Gamma_{i-1}(u) \cap \Gamma_1(v) \cap \Gamma_2(z)$ and $y \in \Gamma_i(u) \cap \Gamma_1(z) \cap \Gamma_1(w) \cap \Gamma_2(v)$ in two different ways. Since $v \in \Gamma_i(u)$, v has precisely $c_i - 1$ neighbours x in $\Gamma_{i-1}(u)$ with $x \neq z$. Fix such a vertex x. Since x, $z \in \Gamma_2(w)$ and $\partial(x, z) = 2$, x and z have precisely $\gamma_2 - 1$ common neighbours y in $\Gamma_1(w)$ with $y \neq v$. So we have $N = (c_i - 1)(\gamma_2 - 2)$. On the other hand, since $w \in \Gamma_2(z)$, w has precisely $c_2 - 1$ neighbours y in $\Gamma_1(z)$ with $y \neq v$. Fix such a vertex y. Since v, $y \in \Gamma_i(u)$ and $\partial(v, y) = 2$, v and y have precisely $\gamma_i - 1$ common neighbours x in $\Gamma_{i-1}(u)$ with $x \neq z$. So we obtain $N = (c_2 - 1)(\gamma_i - 1)$.
Q.E.D.

Lemma 5.3. *If* $a_d > 0$,
(v) $c_d(c_2 - 1) = (k - c_d - 1)\gamma_d$,
(vi) $k \geq 2c_d$.

Proof. (v) Since $a_d > 0$, there is an edge vw in $\Gamma_d(u)$. We count the number N of edges xy with $x \in \Gamma_{d-1}(u) \cap \Gamma_1(v)$ and $y \in \Gamma_d(u) \cap \Gamma_1(w) \cap \Gamma_2(v)$ in two different ways. Since $v \in \Gamma_d(u)$, v has precisely c_d neighbours x in $\Gamma_{d-1}(u)$. Fix such a vertex x. Since $x \in \Gamma_2(w)$, x has precisely $c_2 - 1$ neighbours y in $\Gamma_1(w)$ with $y \neq v$, where we have

$y \in \Gamma_d(u)$ since there is no edge in $\Gamma_{d-1}(u)$. So we have $N = c_d(c_2 - 1)$. On the other hand, since $w \in \Gamma_d(u)$, w has precisely $a_d - 1$ neighbours y in $\Gamma_d(u)$ with $y \neq v$. Fix such a vertex y. Since $v, y \in \Gamma_d(u)$ and $\partial(v, y) = 2$, v and y have precisely γ_d common neighbours x in $\Gamma_{d-1}(u)$. So we obtain $N = (a_d - 1)\gamma_d = (k - c_d - 1)\gamma_d$.

(vi) Let vw be an edge in $\Gamma_d(u)$. If there is a vertex x in $\Gamma_1(u) \cap \Gamma_{d-1}(v) \cap \Gamma_{d-1}(w)$, then uv is an edge in $\Gamma_{d-1}(x)$, contradicting $a_{d-1} = 0$. Hence $\Gamma_1(u) \cap \Gamma_{d-1}(v)$ and $\Gamma_1(u) \cap \Gamma_{d-1}(w)$ are mutually disjoint, each of which has size c_d since $u \in \Gamma_d(v)$ and $u \in \Gamma_d(w)$. Hence $k = |\Gamma_1(u)| \geq 2c_d$. Q.E.D.

To simplify notations, we put

$$c = c_2, \quad \gamma = \gamma_2.$$

When $\gamma = 1$, we have $c > 1$ by Lemma 5.2 (i), and hence $c = 2$ by Lemma 5.2 (ii). Then $\gamma_i = 1$ $(i = 1, \ldots, d-1)$ by Lemma 5.2 (iv) and this implies $c_i = i$ $(i = 1, \ldots, d)$ by Lemma 5.2 (iii). If $a_d = 0$ then we have $k = c_d = d$, so that the intersection array becomes of type (9). If $a_d > 0$ then Lemma 5.3 (v) implies $d = (k - d - 1)\gamma_d$, here we have $k \geq 2c_d = 2d$ by Lemma 5.3 (vi). Hence we must have $\gamma_d = 1$ and $k = 2d + 1$ so that the intersection array becomes of type (10).

Now we assume $\gamma > 1$. By Lemma 5.2 (i), (ii), we have $c > 1$ and

$$k = \frac{(c-1)(c-2)}{\gamma - 1} + 2.$$

First we consider the case $a_d > 0$.

When $d = 2$, Lemma 5.3 (v) becomes

$$k = \frac{c(c-1)}{\gamma} + c + 1,$$

and hence we have

$$\frac{(c-1)(c-2)}{\gamma - 1} + 2 = \frac{c(c-1)}{\gamma} + c + 1.$$

This becomes

$$c = \gamma(\gamma + 1),$$

and hence

$$k = \frac{(c-1)(c-2)}{\gamma - 1} + 2 = \gamma(\gamma^2 + 3\gamma + 1),$$

so that the intersection array becomes of type (3) in the case $d = 2$.

Assume $d > 2$. We have $2c_3 \le k$ by Lemma 5.3 (vi) and by $c_3 \le c_d$. By Lemma 5.2 (iii), we have

$$c_3 = \frac{c(c-1)}{\gamma} + 1.$$

So $2c_3 \le k$ implies

$$2\left(\frac{c(c-1)}{\gamma} + 1\right) \le \frac{(c-1)(c-2)}{\gamma - 1} + 2,$$

and this becomes

$$2(\gamma-1)c(c-1) \le \gamma(c-1)(c-2).$$

By Lemma 5.2 (i), we have $c - 1 > 0$, so the above inequality implies

$$2(\gamma - 1)c \le \gamma(c-2)$$

and hence

$$(\gamma - 2)c + 2\gamma \le 0.$$

This is impossible by our assumption $\gamma \ge 2$. Thus the case $d > 2$ does not occur.

Next we consider the case $a_d = 0$. Since $b_0 = k$, $b_1 = k - 1$ and $b_2 = k - c$, we have

$$b_0 = \frac{(c-1)(c-2)}{\gamma - 1} + 2, \qquad b_1 = \frac{(c-1)(c-2)}{\gamma-1} + 1,$$

$$b_2 = \frac{(c-1)(c-2)}{\gamma-1} + 2 - c = \frac{(c-\gamma)(c-2)}{\gamma - 1}.$$

From Lemma 5.2 (iii) with $i = 2$, we obtain

$$c_3 = \frac{c(c-1)}{\gamma} + 1 = \frac{c^2 - c + \gamma}{\gamma},$$

and $b_3 = k - c_3$ implies

$$b_3 = \frac{(c-1)(c-2)}{\gamma - 1} + 2 - \frac{c^2 - c + \gamma}{\gamma} = \frac{(c-\gamma)(c-\gamma-1)}{\gamma(\gamma-1)}.$$

When $d > 3$, Lemma 5.2 (iii), (iv) and $c_3 = (c^2 - c + \gamma)/\gamma$ imply

$$c_4 = \frac{c(c^2 - 2c + 2\gamma)}{\gamma + \gamma c - c},$$

and $b_4 = k - c_4$ implies

$$b_4 = \frac{(c-\gamma)(c-2\gamma)}{(\gamma-1)(c\gamma+\gamma-c)}.$$

When $d > 4$, Lemma 5.2 (iii), (iv) imply

$$c_5 = \frac{c^4 - 3c^3 + c^2 + 3\gamma c^2 - 2\gamma c + \gamma^2}{\gamma c^2 + \gamma^2 - c^2},$$

and $b_5 = k - c_5$ implies

$$b_5 = \frac{(c-\gamma)(c-\gamma-\gamma^2)}{(\gamma-1)(c^2\gamma+\gamma^2-c^2)}.$$

If $d > 5$, we have $b_5 \geq 1$, so the above equation implies (noting that the denominator is positive since $\gamma > 1$)

$$(c-\gamma)(c-\gamma-\gamma^2) \geq (\gamma-1)(c^2\gamma+\gamma^2-c^2),$$

and this becomes

$$\gamma(c-1)(2c-2\gamma-c\gamma) \geq 0.$$

This implies a contradiction since $c \geq 2$ and $\gamma \geq 2$. Hence we have $d \leq 5$.

When $d = 5$, we have $b_5 = 0$, and this implies $c = \gamma$ or $c = \gamma^2 + \gamma$. But $c = \gamma$ does not occur by Lemma 5.2 (ii) since $c_2 = k - b_2 < k$. So we have $c = \gamma^2 + \gamma$. Substituting this value of c in the above equations, we obtain $k = \gamma(\gamma^2 + 3\gamma + 1)$, $c_3 = k - c$, $c_4 = k - 1$. So the intersection array becomes of type (6).

When $d = 4$, we have $b_4 = 0$, and this implies $c = 2\gamma$ ($c = \gamma$ is impossible as above). So we obtain $k = 4\gamma$, $c_3 = k - 1$, so the intersection array becomes of type (5).

When $d = 3$, we have $b_3 = 0$, and this implies $c = \gamma + 1$. So we obtain $c = k - 1$, and the intersection array becomes of type (4).

This completes the proof of Theorem 5.1.

References

[1] E. Bannai and Et. Bannai, Generalized generalized spin models (four-weight spin models), Pacific J. Math., to appear.

[2] E. Bannai and Et. Bannai, Spin models on finite cyclic groups, J. Algebraic Combin., **3** (1994), 243–259.

[3] E. Bannai, Et. Bannai and F. Jaeger, On spin models, modular invariance, and duality, submitted.

[4] E. Bannai and T. Ito, "Algebraic Combinatorics I", Benjamin/Cummings, 1984.

[5] A. E. Brouwer, A. M. Cohen and A. Neumaier, "Distance-Reguar Graphs", Springer-Verlag, Berlin, Heidelberg, 1989.

[6] D. M. Goldschmidt and V. Jones, Metaplectic link invariants, Geom. Dedicata, **31** (1989), 165–191.
[7] P. de la Harpe, Spin models for link polynomials, strongly regular graphs and Jaeger's Higman-Sims model, Pacific J. Math, **162** (57–96).
[8] D. Higman and C. Sims, A simple group of order 44,352,000, Math. Z., **105** (1968), 110–113.
[9] F. Jaeger, Strongly regular graphs and spin models for the Kauffman polynomial, Geom. Dedicata, **44** (1992), 23–52.
[10] F. Jaeger, On spin models, triply regular association schemes, and duality, J. Algebraic Combin., **4** (1995), 103–144.
[11] F. Jaeger, New constructions of models for link invariants, Pacific J. Math., to appear.
[12] V. Jones, On knot invariants related to some statistical mechanical models, Pacific J. Math., **137** (1989), 311–336.
[13] V. Jones, private communication.
[14] L. H. Kauffman, An invariant of regular isotopy, Trans. Amer. Math. Soc., **318** (1990), 417–471.
[15] K. Kawagoe, A. Munemasa and Y. Watatani, Generalized spin models, J. Knot Theory Ramifications, **3** (1994), 465–475.
[16] V. G. Kac and M. Wakimoto, A construction of generalized spin models, preprint.
[17] K. Nomura, Spin models constructed from Hadamard matrices, J. Combin. Theory Ser. A, **68** (1994), 251–261.
[18] K. Nomura, Homogeneous graphs and regular near polygons, J. Combin. Theory Ser. B, **60** (1994), 63–71.
[19] K. Nomura, Spin models with an eigenvalue of small multiplicity, J. Combin. Theory Ser. A, **71** (1995), 293–315.
[20] K. Nomura, Spin models on bipartite distance-regular graphs, J. Combin. Theory Ser. B, **64** (1995), 300–313.
[21] K. Nomura, Spin models on triangle-free connected graphs, submitted.
[22] N. Yamazaki, Bipartite distance-regular graphs with an eigenvalue of multiplicity k, J. Combin. Theory Ser. B, to appear.

Tokyo Ikashika University
Ichikawa, Chiba 272, Japan

Advanced Studies in Pure Mathematics 24, 1996
Progress in Algebraic Combinatorics
pp. 309–321

Spherical Designs and Tensors

J.J. Seidel

§1. Introduction

The set X of the 12 vertices of a regular icosahedron on the unit sphere Ω in $\mathbf{R}^3$ provides a first example of a spherical 5-design (of strength 5). It satisfies

$$\frac{1}{n}\sum_{x\in X} h(x) = \int_{\Omega} h(u)\, d\sigma(u),$$

short, $\operatorname*{Ave}_X h = \operatorname*{Ave}_\Omega h$ for all polynomials h in 3 variables of degree ≤ 5 and $n = 12$. If the defining relation only refers to the homogeneous polynomials of degree q, then we use the term spherical design of index q. Thus strength q means index 1, 2, ..., and q.

The second part of the title refers to symmetric tensors, and to the desire to express symmetric polynomials as the inner products of tensors, for instance

$$\sum_{i,j,k=1}^{d} h_{ijk}a_i a_j a_k = \langle h, a\otimes a\otimes a\rangle,$$

where $a = (a_1, a_2, \dots, a_d) \in \mathbf{R}^d =: V$. The linear space $S^q(V)$ of the symmetric q-tensors on V is spanned by the q-fold tensor powers $\otimes^q a := a\otimes a\otimes\cdots\otimes a$. This space is isomorphic to the space $\operatorname{Hom}_q(V)$ of homogeneous polynomials of degree q in d variables.

Section 2 deals with tensors in $\mathbf{R}^d$, in particular with the distribution q-tensor D, and the Sidelnikov inequality. In Section 3 this leads to the tensor-definition of spherical designs of index q and of strength q. This notion was introduced by Delsarte, Goethals and Seidel, and was further developed by Bannai. We recall that the combinatorial t-(v,k,λ) design can be phrased in analogous terms. Generalization to t-(v,K,λ)

Received October 27, 1994.
Revised February 1, 1995.

designs with unequal block sizes from a set K suggests generalization to multispherical designs as defined in joint papers with Neumaier and with Delsarte. In the present Section 4 this is put into the natural framework of isometric linear maps from real spaces with the 2-norm into real Banach spaces with the q-norm, following Reznick and Lyubich-Vaserstein. This leads to the main Theorem 4.4 on the existence of cubature formulae, isometric embeddings, and (multi-)spherical designs of index q. These all are corollaries of Hilbert's solution of the Waring problem. Their equivalence provides a link between combinatorics and local Banach theory [LV93]

In Section 5 the notions of Eutactic star and Euclidean t-design are phrased in the present terminology. The final Section 6 comes back to spherical designs of strength t, and reviews the main results before and after the omni-existence theorem by Seymour and Zaslavsky. The paper ends with a new proof, by A. Blokhuis and the author, for the construction by Hardin and Sloane of spherical 4-designs in 3-space, by use of tensors.

§2. Symmetric q-tensors

We consider real d-dimensional Euclidean space $V = \mathbf{R}^d$ with standard basis $B = \{e_1, \dots, e_d\}$. Vectors $a \in V$ are indicated by their coordinates $(a_1, \dots, a_d)$ with respect to B, their standard inner product by $\langle a, b\rangle$, and their q-norm by $\|a\|_q$. The coordinates of the q-fold tensor power $\otimes^q a = a \otimes \cdots \otimes a$, are the monomials of degree q in $a_1, \dots, a_d$. The space $S^q(V)$ of the *symmetric q-tensors* over V is spanned by the tensor powers $\otimes^q a$, $a \in V$, and has dimension $\binom{d+q-1}{d-1}$. Each basis B defines an isomorphism of the symmetric algebra $S(V)$ with the polynomial algebra over $\mathbf{R}$ in d variables. [For these elementary facts of tensor algebra see [Sha82], Chapter 10, in particular 10.5.2 and 10.5.1]. For the inner products in $S^q(V)$ it follows that

$$\langle \otimes^q a, \otimes^q b\rangle = \langle a, b\rangle^q, \quad \langle h, \otimes^q x\rangle = h(x) \in \mathrm{Hom}_q(V),$$

for $a, b \in V$, $h \in S^q(V)$, $h(x)$ the corresponding homogeneous polynomial of degree q in d variables. From now on we fix the dimension d and the degree q. As a special symmetric q-tensor we define the *distribution q-tensor D* as follows.

Definition 2.1.

$$D := \int_\Omega \otimes^q u \, d\sigma(u).$$

Here Ω is the unit sphere in V, $d\sigma$ is the normalized standard measure on Ω, and the integral is explained in terms of coordinates. Clearly, $D = 0$ if q is odd.

Lemma 2.2.

$$\langle D, D\rangle = \delta, \quad \langle D, \otimes^q a\rangle = \delta \, \|a\|_2^q,$$

where $\delta = \dfrac{1 \cdot 3 \cdots (q-1)}{d(d+2)\cdots(d+q-2)}$ *if q is even, $\delta = 0$ if q is odd.*

This follows from the properties of the inner product in $S^q(V)$ referred to above, and from the well-known formula

$$\int_\Omega \langle u, v\rangle^q \, d\sigma(u) = \int_\Omega u_1^q \, d\sigma(u) = \delta.$$

Another useful formula is the inequality of *Sidelnikov* [Sid74] for a finite set $U \subset \Omega \subset V$, $|U| = n$.

Lemma 2.3.

$$0 \leq \|D - \frac{1}{n}\sum_{u\in U} \otimes^q u\|_2^2 = -\delta + \frac{1}{n^2}\sum_{u,v\in U} \langle u, v\rangle^q.$$

Proof. The square of the 2-norm of $D - \frac{1}{n}\sum_{u\in U} \otimes^q u$ is the inner product of that tensor with itself. Now evaluate and use Lemma 2.2. Q.E.D.

§3. Designs in Euclidean space

Definition 3.1. A *spherical design* of index q is a finite subset U of size n of the unit sphere Ω in $V = \mathbf{R}^d$, such that

$$D = \frac{1}{n}\sum_{u\in U} \otimes^q u.$$

Thus, a spherical design represents an extremal case of Sidelnikov's inequality, cf. [GS81], and by Lemma 2.3

$$\frac{1}{n^2}\sum_{u,v\in U} \langle u, v\rangle^q = \delta;$$

is an equivalent definition. Also equivalent is:

$$\frac{1}{n}\sum_{u\in U} h(u) = \int_\Omega h(u) \, d\sigma(u),$$

for any homogeneous polynomial h of degree q; just take tensor inner products of both sides of 3.1 with any q-tensor h. Hence, the present definition 3.1 is equivalent to the original definition of the notion in [DGS77]. For the case of odd q the same definition is accepted with $D = 0$.

Remark. The name design is justified for the following reason. A combinatorial t-(v, k, λ) design consists of a v-set V and a collection X of k-subsets (blocks) of V such that any t-subset of V is in exactly λ blocks. Describing the blocks by v-vectors having coordinates $x_i \in \{0, 1\}$ with $\sum_{i=1}^{v} x_i = k$, we can prove [Sei90] that the definition above is equivalent to requiring that

$$\frac{1}{n} \sum_{x \in X} f(x) = \binom{v}{k}^{-1} \sum_{x \in \binom{V}{k}} f(x),$$

for all square-free monomials

$$f(x) = x_{\sigma(1)} x_{\sigma(2)} \cdots x_{\sigma(t)}, \quad \sigma \in \mathrm{Sym}(v).$$

As a consequence, both spherical designs of strength t and combinatorial t-(v, k, λ) designs can be defined by similar formulae of the type

$$\operatorname*{Ave}_{X} h = \operatorname*{Ave}_{\Omega} h$$

the average over the vectors of the finite set X equals the average over all vectors, for given sets of test functions h. The analogy still goes further. The combinatorial t-(v, K, λ) designs with unequal block sizes from the set K serve as a model for the multispherical designs to be introduced in the next section.

§4. Linear maps subject to condition (D)

Spherical designs are often produced by linear maps of Euclidean spaces. Let

$$F \,:\, \mathbf{R}^d \to \mathbf{R}^N \qquad x \mapsto y = Fx$$

be a linear map with standard $N \times d$ matrix F. We denote again by F the set of the vectors in $\mathbf{R}^d$ which correspond to the rows of this standard matrix. We take q even and consider the following condition (D).

Definition 4.1. A linear map F from $\mathbf{R}^d$ into $\mathbf{R}^N$ satisfies *condition* (D) whenever

$$D = \delta \sum_{f \in F} \otimes^q f,$$

where D is the distribution q-tensor for $\mathbf{R}^d$.

If all row vectors f of the standard matrix F have equal length, then condition (D) says that F is proportional to a spherical design of index q in $\mathbf{R}^d$. However, if the row vectors of F are not requested to have equal length, then they are arranged in $\mathbf{R}^d$ on various concentric spheres; we then say that these vectors form a *multispherical design* of index q in $\mathbf{R}^d$, cf. [NS88, Sei90]. The tensor formula in Definition 4.1 can be phrased in different terms, yielding equivalent notions.

Theorem 4.2. *The existence of a linear map F satisfying condition* (D) *is equivalent to the existence of any of the following formulae*:

$$\text{Cubature:} \qquad \int_\Omega h(u)\, d\sigma(u) = \delta \sum_{f \in F} h(f), \quad h \in \operatorname{Hom}_q(\mathbf{R}^d).$$

$$\text{Waring:} \qquad \langle x, x \rangle^{q/2} = \|x\|_2^q = \sum_{f \in F} \langle f, x \rangle^q, \quad x \in \mathbf{R}^d.$$

$$\text{Isometry:} \qquad \|x\|_2 = \|y\|_q := \Big(\sum_{\nu=1}^{N} y_\nu^q\Big)^{1/q}, \quad x \in \mathbf{R}^d.$$

Proof. Take tensor inner products of both sides of (4.1), first with any symmetric q-tensor h, and then with any tensor power $\otimes^q x$. This gives the cubature and the Waring formulas. The last formula is a rewriting of its predecessor since the coordinates of $y = Fx$ are the inner products of x with the row vectors f of the matrix F.

Cubature formulae for homogeneous polynomials, as above, can also be expressed by use of weights for a finite number of points on the unit sphere $\Omega \subset \mathbf{R}^d$.

Waring's formulae refers to the classical problem of expressing integers as sums of q-th powers of integers. This problem was solved by Hilbert, cf. [Ell71, Rie53, Rez92].

Isometry refers to the linear map F from $\mathbf{R}^d$ with 2-norms, into $\mathbf{R}^N$ with equal q-norms, so to isometric embedding of $\mathbf{R}^d$ into $\mathbf{R}^N$. The existence of such an isometry is the $\varepsilon = 0$ case of Dvoretzky's theorem for finite dimensional Banach space, cf. [LV93, p.329], [Sei94]. Q.E.D.

The solution of Waring's problem by Hilbert (1909) was later simplified by Stridsberg (1916) and others. In the formulation of G. J. Rieger the essential *Lemma of Hilbert* reads as follows; its proof in [Rie53] covers less than 2 pagesc.

Lemma 4.3. *Given $d \in \mathbf{N}$, $q \in 2\mathbf{N}$, there exist $N \in \mathbf{N}$ and an identity*

$$\langle x, x\rangle^{q/2} = \sum_{\nu=1}^{N} r_\nu \langle a_\nu, x\rangle^q, \qquad \textit{for } x \in \mathbf{R}^d,$$

with positive rational r_ν and nonzero integral $a_{\nu,i}$.

In addition, for $N = N(d,q)$ we have, [LV93],

$$\binom{d+\frac{1}{2}q-1}{d-1} \leq N(d,q) \leq \binom{d+q-1}{d-1}.$$

As a consequence of Hilbert's Lemma 4.3, and of Theorem 4.2, we now have the following existence theorem.

Theorem 4.4. *Given $d \in \mathbf{N}$ and $q \in 2\mathbf{N}$, there exist $N \in \mathbf{N}$ and a linear $F : \mathbf{R}^d \to \mathbf{R}^N$ satisfying* (D), *hence satisfying Cubature, Waring, Isometry, and forming a multispherical design of index q.*

§5. Eutactic stars

Let G, of size n, denote a symmetric idempotent matrix, that is,

$$G^t = G = G^2.$$

Then G has the eigenvalues 1 and 0, of multiplicities d and $n-d$, say. We can write $G = HH^t$ with $H^tH = I$ of size d. Hence G is the Gram matrix of n vectors $y \in Y \subset \mathbf{R}^d$, called a *Eutactic star* Y, since its vectors are the projections into $\mathbf{R}^d$ of an orthonormal frame of a space $\mathbf{R}^n$ which contains $\mathbf{R}^d$ as a subspace, cf. [Sei76]. We show that the set Y is a multispherical design of index 2 (and of strength 2 if in addition its vectors add up to zero). Indeed,

$$\begin{aligned}\sum_{x,y\in Y} \langle x, y\rangle^2 &= \operatorname{tr} G^2 = \operatorname{tr} G = \sum_{y\in Y} \langle y, y\rangle = d = \\ &= d^2 \cdot \frac{1}{d} = (\sum_{y\in Y} \|y\|^2)^2 \int_\Omega \langle u, v\rangle^2 \, d\sigma(u),\end{aligned}$$

equivalently,

$$\sum_{y\in Y} y \otimes y = (\sum_{y\in Y} \|y\|^2) \int_\Omega u \otimes u \, d\sigma(u).$$

If G has a constant diagonal, then Y is a spherical design of index 2. If the diagonal of G is not constant, but consists of p distinct numbers, say, then Y is a multispherical design of index 2, also called a Euclidean design of index 2. Then Y has the same regularity condition as a spherical design, but its vectors are distributed over p concentric spheres. The existence of Eutactic stars is implied by the following theorem by Sofman [Sof69].

Theorem 5.1. *Symmetric idempotent matrices G exist iff*

$$\operatorname{trace} G = \operatorname{rank} G, \quad \operatorname{diag} G \geq 0.$$

In other words, stars consisting of n vectors in $\mathbf{R}^d$ at lengths $\xi_1, \ldots, \xi_n$ are eutactic iff

$$\xi_1^2 + \cdots + \xi_n^2 = d, \quad 0 \leq \xi_k \leq 1.$$

Examples of spherical 2-designs are provided by *strongly regular graphs*, since the defining equations for their adjacency matrix A define a symmetric idempotent Gram matrix G with zero row sums as follows:

$$(A - rI)(A - sI) = \mu J, \ AJ = kJ, \ G := \frac{1}{r-s}(A - sI - \frac{k-s}{n}J).$$

Likewise, the minimal idempotents of any association scheme yield spherical 2-designs, cf. [God93], Chapter 13. Some strongly regular graphs yield spherical t-designs with $t > 2$. For example, Smith graphs on 16, 112 and 162 vertices yield spherical 3-designs, and those on 27 and 275 vertices yield spherical 4-designs. In fact, any distance regular graph is represented by a spherical 3-design X in an eigenspace if and only if $\sum_{a \in X} a \otimes a \otimes a = 0$, see [CGS78], [God93, p. 275].

A straightforward generalization of a balanced eutactic star is provided by the following notion [NS88].

Definition 5.2. A *Euclidean t-design* of strength t is a finite subset Y of $\mathbf{R}^d$ subject to the conditions, for $k = 1, 2, \ldots, t$,

$$\sum_{y \in Y} \otimes^k y = (\sum_{y \in Y} \|y\|^k) \int_\Omega \otimes^k u \, d\sigma(u).$$

Similar definitions can be given for finite weighted sets (Y, w) in $\mathbf{R}^d$ and, more general, for *measures ξ of strength t* in $\mathbf{R}^d$. This last notion is defined by

$$\int_{R\Omega} p(y) \, d\xi(y) = \sum_{k=0}^{t} \int_{R\Omega} \|y\|^k \, d\xi(y) \int_\Omega p_k(x) \, d\sigma(x),$$

equivalently, see [NS88, Sei90], by

$$\int_{R\Omega} p\, d\xi = \int_{R\Omega} p\, d\xi \circ \gamma\, ; \quad \gamma \in O(d).$$

Here $O(d)$ is the orthogonal group in $\mathbf{R}^d$, and

$$R\Omega := \cup_{r\in R} rS\, ,\ rS = \{y \in \mathbf{R}\, :\, \langle y, y\rangle = r^2\}\, ,\ r \in R \subset \mathbf{R}^+,$$

denotes *any* union of concentric spheres in $\mathbf{R}^d$, and the condition should hold for all polynomials p of degree $\leq t$, restricted to $R\Omega$:

$$p = \sum_{k=0}^{t} p_k \in \mathrm{Pol}_t(R\Omega)\, ,\ p_k \in \mathrm{Hom}_k(\mathbf{R}^d).$$

Thus, a Euclidean t-design is a measure of strength t having finite support. A spherical t-design has finite support on the unit sphere, with equal weights.

It is interesting to interpret the strength t conditions

$$\int_{r\Omega} f(y)g(y)\, d\xi(y) = \int_{R\Omega} f(x)g(x)\, d\mu(x), \quad \mu = \mu(R\Omega),$$

as the equality of inner products:

$$\langle f, g\rangle_\xi = \langle f, g\rangle_{R\Omega},$$

for all polynomials f and g of degree $\leq e \leq t/2$. This is related to rotatable designs, and to Kiefer's theorem on optimal experimental designs [NS92]. This is also related to Fisher-type inequalities for Euclidean t-designs Y, [DS89]. Indeed, $\langle f, g\rangle_Y = \langle f, g\rangle_{R\Omega}$ implies the isometry

$$\mathrm{Pol}_e(Y) \cong \mathrm{Pol}_e(R\Omega),$$

and for the dimensions of these linear spaces it follows that

$$\begin{aligned} |Y| \geq \dim \mathrm{Pol}_e(R\Omega) &= \sum_{i=0}^{2p-1} \dim \mathrm{Hom}_{e-i}(\mathbf{R}^d) \\ &= \sum_{i=0}^{2p-1} \binom{d+e-i-1}{d-1}. \end{aligned}$$

For $2p \geq e+1$, and for $p = 1$, this formula reads $|Y| \geq \binom{d+e}{e}$, and $|Y| \geq \binom{d+e-1}{d-1} + \binom{d+e-2}{d-1}$ respectively. For the second case, a *tight spherical t-design* Y in $\mathbf{R}^d$ is defined by $|Y| = \binom{d+e-1}{d-1} + \binom{d+e-2}{d-1}$, for $t = 2e$, and $|Y| = 2\binom{d+e-1}{d-1}$, for $t = 2e+1$, antipodal case. We shall come back to tight spherical t-designs in the next Section 6.

§6. Spherical designs

The early constructions of spherical designs by Delsarte, Goethals, Seidel [DGS77, GS79, GS81], by Conway and Sloane [CS77], and by Bannai [Ban84b, Ban84a, Ban86] have their basis in combinatorics and in finite group theory, and use the harmonic analysis of the unit sphere. Highlights are the sphere version of Delsarte's linear programming bound (Theorem 5.10 of [DGS77]), and the proof, by analytic and number-theoretic methods by Bannai and Damerell of

Theorem 6.1. *There exist no tight spherical* $2e$*-designs in* $\mathbf{R}^d$ *for* $d \geq 3$, $e \geq 3$. *There exist no tight antipodal* $(2e+1)$*-design in* $\mathbf{R}^d$, *except for* $d = 24$, $t = 11$.

Tight spherical t-designs do exist for $t = 2, 3, 4, 5, 7, 11$. For details we refer to [DGS77, BD80] and [God93, Chapter 16].

In the *group case*, originated by Sobolev [Sob62], and rediscovered in the present context by Conway and Sloane [CS77], the designs are the orbits of a finite subgroup of the orthogonal group $O(d)$ in d-space. Bannai [Ban84b] introduced the notion of a *t-homogeneous* subgroup Γ of $O(d)$ by the property that for each point x of the unit sphere Ω the orbit x^Γ is a spherical t-design. This works if the coefficients h_1, h_2 $\ldots, h_t$ vanish in the Molien series of Γ:

$$\sum_{\gamma \in \Gamma} \frac{1-\lambda^2}{\det(I - \lambda\gamma)} = h_0 + h_1\lambda + \cdots + h_t\lambda^t + h_{t+1}\lambda^{t+1} + \cdots .$$

Here h_k is the dimension of the linear space of the Γ-invariant harmonic polynomials of degree k. This applies to finite reflection groups and beyond, [GS81].

Example. The icosahedral group A_5 has the Molien series

$$1 + \lambda^6 + \lambda^{10} + \cdots ,$$

hence there is a spherical 5-design for every orbit of the icosahedral group: the icosahedron itself, the dodecahedron, the icosidodecahedron, the football, etc. It is interesting to observe that a spherical 9-design is obtained as the orbit of any zero of the polynomial that spans the A_5-invariant harmonic polynomial of degree 6.

Bannai [Ban86] also introduced the notion of *rigidity.* A spherical

t-design in d-space is rigid whenever all sufficiently close spherical t-designs are equivalent under the orthogonal group $O(d)$. Bannai conjectures that for given d and t there are finitely many rigid t-designs mod $O(d)$. For rigidity and reflection groups cf. [Sal94]

Finally we mention the relations between representations of a finite group Γ and the spherical designs generated by Γ. A *representation* of Γ on a real vector space V is a homomorphism $\rho : \Gamma \to GL(V)$. The representation is irreducible if V contains no proper Γ-invariant subspace. So irreducibility means real irreducibility. The following theorem [GS79] deals with representations into the space Harm_k of the harmonic polynomials of degree k in d variables.

Theorem 6.2. *If the representations ρ_k of Γ on Harm_k are irreducible for $k \leq s$, then Γ is $2s$-homogeneous.*

The converse of this theorem was also claimed and proved in Theorem 6.7 of [GS79]. However, Bannai [Ban84b, Ban84a] convincingly demonstrated the falsity of both the statement and its proof, by counterexamples involving the unitary subgroup $U(d)$ of $O(2d)$.

In 1984 Seymour and Zaslavsky [SZ84] proved

Theorem 6.3. $\forall_d \forall_t \exists_{n_0} \forall_{n \geq n_0}$ *there exists a spherical design in $\mathbf{R}^d$ of strength t and size n.*

The proof of this existence theorem is not constructive, and the function $n_0(d,t)$ is important. It would be interesting to have a proof of this theorem that relates to Hilbert's lemma.

Meanwhile, several results have been obtained in this area about new constructions, as well as about the value of n_0, by Mimura, Wagner, Bajnok, Rabau-Bajnok, Grabner-Tichy, Korevaar-Meyers, Hardin-Sloane, and others. The first explicit construction of spherical t-designs for arbitrary t was given by Wagner [Wag91], cf. [GT91]. The explicit construction by Bajnok [Baj92] uses

$$n_0 = C(d)t^{O(d^3)}.$$

Korevaar and Meyers [KM93] show that in 3-space there exist spherical t-designs consisting of $O(t^3)$ points and conjecture $O(t^2)$ points. They think that for $\mathbf{R}^d$ the number $O(t^{d-1})$ should be possible, cf. also [Kor94], Remark 3.3.4.

Hardin and Sloane [HS92] used their computer program Gosset to find many new spherical 4-designs. This was achieved by minimizing the socalled average prediction variance $I(D)$ for designs D fitting a quadratic model in the unit ball B in d-space. Here $I(D) := \mathrm{trace}\, M_B M_D^{-1}$

for the moment matrices

$$M_D = \frac{1}{n} \sum_{x \in D} p_\mu(x) p_\nu(x)\,, \; M_B = \frac{1}{|B|} \int_B p_\mu(x) p_\nu(x)\, d\omega(x)$$

of the $\frac{1}{2}(d+1)(d+2)$ basic polynomials p_μ of degree ≤ 2 in d variables. Restricting to designs D with b points on the unit sphere Ω and c coinciding points in the origin $(b + c = n)$, they found that in extremal situations for $I(D)$ the b points on Ω form a spherical 4-design. They conjecture that, for $d \leq 8$, the list of their findings is complete. Thus, even in ordinary 3-space many new spherical t-designs have been discovered, cf. also [Rez95] and [HS95].

Hardin and Sloane found infinitely many distinct spherical 4-designs on 12 points in 3-space by rotating the northern hemisphere of a regular icosahedron about a diameter NS over an arbitrary angle, cf. also [Sal94]. A. Blokhuis and the author have the following independent proof by tensors: Let $X = N \cup P \cup Q \cup S$, with poles N, S and planar regular pentagons P and Q, denote the vertex set of a regular icosahedron. Let $P(\phi)$ denote the pentagon obtained from P by a rotation ϕ about NS. Decompose its vectors in and orthogonal to the equator plane following $x_i(\phi) = p_i(\phi) + u$, and observe that

$$\otimes^s (p_i(\phi) + u) = \sum_{a+b=s} \binom{s}{a} (\otimes^a p_i(\phi)) \otimes (\otimes^b u).$$

Since $p_i(\phi)$ form a spherical 4-design we know that $\sum_{i=1}^5 \otimes^a p_i(\phi)$ is independent of $\phi \in O(2)$, for $a = 1, 2, 3, 4$. Therefore, the half-rotated $N \cup P(\phi) \cup Q \cup S$ has sums of the a-th tensor products independent of ϕ, hence equal to the a-th tensor products of the icosahedron. Since the regular icosahedron is a spherical 5-design, the half-rotated icosahedron is a spherical 4-design, for any $\phi \in O(2)$.

References

[Baj92] B. Bajnok, Construction of spherical t-designs, Geom. Dedicata, **43** (1992), 167–179.

[Ban84a] E. Bannai, Spherical designs and group representations, Contemp. Math., **34** (1984), 95–107.

[Ban84b] E. Bannai, Spherical t-designs which are orbits of finite groups, J. Math. Soc. Japan, **36** (1984), 341–354.

[Ban86] E. Bannai, On extremal finite sets in the sphere and other metric spaces, London Math. Soc. Lecture Note Ser., **131** (1986), 13–38.

[BD80] E. Bannai and R. M. Damerell, Tight spherical designs II, J. London Math. Soc., **21** (1980), 13–30.

[CGS78] P. J. Cameron, J.-M. Goethals, and J. J. Seidel, The Krein condition, spherical designs, Norton algebras and permutation groups, Proc. Konink. Nederl. Akad. Wetensch. A, **81** (= Indag. Math., **40**) (1978), 196–206.

[CS77] J. H. Conway and N. J. A. Sloane, personal communication, 1977.

[DGS77] P. Delsarte, J.-M. Goethals, and J. J. Seidel, Spherical codes and designs, Geom. Dedicata, **6** (1977), 363–388.

[DS89] P. Delsarte and J. J. Seidel, Fisher-type inequalities for Euclidean t-designs, Linear Algebra Appl., **114/115** (1989), 213–230.

[Ell71] W. J. Ellison, Waring's problem, Amer. Math. Monthly, **78** (1971), 10–36.

[God93] C. D. Godsil, "Algebraic Combinatorics", Chapman-Hall, 1993.

[GS79] J.-M. Goethals and J. J. Seidel, Spherical designs, Proc. Sympos. Pure Math. (AMS), **34** (1979), 255–272.

[GS81] J.-M. Goethals and J. J. Seidel, Cubature formulae, polytopes and spherical designs, in "The Geometric Vein, Coxeter Festschrift", (B. Grünbaum C. Davis and F. A. Sherk, ed.), Springer, 1981, pp. 203–218.

[GT91] P. J. Grabner and R. F. Tichy, Spherical designs, discrepancy and numerical integration, Math. Comp., **60** (1991), 327–336.

[HS92] R. H. Hardin and N. J. A. Sloane, New spherical 4-designs, Discrete Math., **106/107** (1992), 255–264.

[HS95] R. H. Hardin and N. J. A. Sloane, An improved snub cube and other new spherical designs in three dimensions, Discrete Comput. Geom., 1995, to appear.

[KM93] J. Korevaar and J. L. H. Meyers, Spherical Faraday cage for the case of equal point charges and Chebyshev-type quadrature, J. of Integral Transforms and Special Functions, **1** (1993), 105–117.

[Kor94] J. Korevaar, Chebyshev-type quadratures: use of complex analysis and potential theory, notes by A. B. J. Kuijlaars, in "Complex Potential Theory" (P. M. Gauthier, ed.), Kluwer, 1994, pp. 325–364.

[LV93] Y. I. Lyubich and L. N. Vaserstein, Isometric embeddings between classical Banach spaces, cubature formulas and spherical designs, Geom. Dedicata, **47** (1993), 327–362.

[NS88] A. Neumaier and J. J. Seidel, Discrete measures for spherical designs, eutactic stars and lattices, Proc. Konink. Nederl. Akad. Wetensch. A, **91** (= Indag. Math. **50**) (1988), 321–334.

[NS92] A. Neumaier and J. J. Seidel, Measures of strength $2e$ and optimal designs of degree e, Sankyā, **54** (1991), 299–309.

[Rez92] B. Reznick, Sums of even powers of real linear forms, Mem. Amer. Math. Soc., **96** (1992), no. 463.

[Rez95] B. Reznick, Some constructions of spherical 5-designs, Linear Algebra Appl., **226-228** (1995), 163–196.

[Rie53] G. J. Rieger, Zur Hilbertschen Lösung des Waringschen Problems: Abschätzung von $g(n)$. Arch. Math., **4** (1953), 275–281.

[Sal94] A. Sali, On the rigidity of spherical t-designs that are orbits of finite reflection groups, Des. Codes Cryptogr., **4** (1994), 157–170.

[Sei76] J. J. Seidel, Eutactic stars, in "Combinatorics", (Hajnal A. and V. T. Sós, ed.), North-Holland, 1976, pp. 983–999.

[Sei90] J. J. Seidel, Designs and approximation, Contemp. Math. (AMS), **111** (1990), 179–186.

[Sei94] J. J. Seidel, Isometric embeddings and geometric designs, Discrete Math., **136** (1994), 281–293.

[Sha82] R. Shaw, "Linear Algebra and Group Representations, volume II", Academic Press, 1982.

[Sid74] V. M. Sidelnikov, New bounds for the density of sphere packings in an n-dimensional euclidean space, Mat. Sb., **95** (1974); English transl. Math. USSR Sbornik, **24** (1974), 147–157.

[Sob62] S. L. Sobolev, The formulas of mechanical cubature on a sphere, Sibirsk Mat. Zh., **3** (1962), 769–796.

[Sof69] L. B. Sofman, Diameters of octahedra, Math. Notes, **5** (1969), 258–262.

[SZ84] P. D. Seymour and T. Zaslavsky, Averaging sets, Adv. Math., **52** (1984), 213–240.

[Wag91] G. Wagner, On averaging sets, Monatsh. Math., **111** (1991), 69–78.

Techn. University Eindhoven
P.O. Box 513, 5600 MB Eindhoven
The Netherlands

Advanced Studies in Pure Mathematics 24, 1996
Progress in Algebraic Combinatorics
pp. 323–441

Quantum Matroids

Paul Terwilliger

Abstract.

We define a *quantum matroid* to be any finite nonempty poset P satisfying the conditions R, SL, M, AU below.

R: P is ranked.
SL: P is a (meet) semilattice.
M: For all $x \in P$, the interval $[0, x]$ is a modular atomic lattice.
AU: For all $x, y \in P$ satisfying $\text{rank}(x) < \text{rank}(y)$, there exists an atom $a \in P$ such that $a \leq y$, $a \not\leq x$, and such that $x \vee a$ exists in P.

Condition AU is the *augmentation axiom.*

We develop a theory of quantum matroids. Although we deal at length with the general case, our emphasis is on quantum matroids P with the following extra structure: We say P is *nontrivial* if P has rank $D \geq 2$, and P is not a modular atomic lattice. In what follows suppose P is nontrivial. We say P is *q-line regular* whenever each rank 2 element in P covers exactly $q+1$ elements of P. We say P is *β-dual-line regular* whenever each element in P with rank $D-1$ is covered by exactly $\beta+1$ elements of P. We say P is *α-zig-zag regular* whenever for all pairs $x, y \in P$ such that $\text{rank}(x) = D-1$, $\text{rank}(y) = D$, and such that x covers $x \wedge y$, there exists exactly $\alpha+1$ pairs $x', y' \in P$ such that y' covers x, y' covers x', and such that y covers x'. We say P is *regular* whenever P is line regular, dual-line regular, and zig-zag regular. We prove the following theorem.

Theorem. *Let D denote an integer at least 4. Then a poset P is a nontrivial regular quantum matroid of rank D if and only if P is isomorphic to one of the following:*

(i) *A truncated Boolean algebra $B(D, N)$, $(D < N)$.*
(ii) *A Hamming matroid $H(D, N)$, $(2 \leq N)$.*

Received April 6, 1995.
Revised October 15, 1995.
Research partially supported by NSF grant DMS-9200792. AMS 1991 Subject Classification: Primary 06A12, Secondary 05E25, 51E20.

(iii) *A truncated projective geometry* $L_q(D,N)$, $(D < N)$.
(iv) *An attenuated space* $A_q(D,N)$, $(D < N)$.
(v) *A classical polar space of rank* D.

Contents

§1. The definition of a $\mathcal{P}$-matroid

In this paper, we only consider finite structures.

We begin by recalling the classical notion of a (finite) matroid.

Definition 1.1. Let A denote a finite set. By an *A-matroid*, we mean a collection P of distinct subsets of A, that satisfies the following axioms NT, LI, AU:

NT: $P \neq \emptyset$.

LI: For all subsets $x, y \subseteq A$, $x \in P$ and $y \subseteq x$ implies $y \in P$.

AU: For all $x, y \in P$ such that $|x| < |y|$, there exists an element $a \in A$ such that $a \in y$, $a \notin x$, and such that $x \cup a \in P$.

AU is referred to as the *augmentation axiom.*

Here is the standard example of an A-matroid: let A denote a finite set of vectors taken from a fixed vector space, and let P denote the collection of all linearly independent subsets of A. Then P is an A-matroid.

Two more examples follow, which we refer to later in the paper.

Example 1.2. (i) The truncated Boolean algebra $B(D,N)$ $(0 \leq D \leq N)$: Let A denote a set of cardinality N, and set

$$P := \{ x \subseteq A \mid \ |x| \leq D \}.$$

Then P is an A-matroid.

(ii) $C(N_1, N_2, \ldots, N_D)$ $(0 < D,\ N_1,\ N_2,\ \ldots,\ N_D)$: Set

$$A := A_1 \cup A_2 \cup \cdots \cup A_D \qquad \text{(disjoint union)},$$

where

$$|A_i| = N_i \qquad\qquad (1 \le i \le D).$$

Set

$$P := \{\, x \subseteq A \mid \quad |x \cap A_i| \le 1 \quad \text{for all} \quad i \ \ (1 \le i \le D)\}.$$

Then P is an A-matroid.

Proof. Routine.

Note 1.3. In Example 1.2(ii), the case $N_i = N \ \ (1 \le i \le D)$ will turn out to have special significance. We refer to this matroid as the *Hamming matroid* $H(D, N)$.

We now generalize Definition 1.1, replacing subsets of a set A with subspaces of a vector space V.

Definition 1.4. Let V denote a finite vector space. By a *V-matroid*, we mean a collection P of distinct subspaces of V, that satisfies the following axioms NT, LI, AU:

NT: $P \ne \emptyset$.
LI: For all subspaces $x, y \subseteq V$, $x \in P$ and $y \subseteq x$ implies $y \in P$.
AU: For all $x, y \in P$ such that $\dim(x) < \dim(y)$, there exists a one dimensional subspace $a \subseteq V$ such that $a \subseteq y$, $a \not\subseteq x$, and such that $x + a \in P$.

We now consider some examples of V-matroids.

Example 1.5. The truncated projective geometry $L_q(D, N)$ $(0 \le D \le N)$: Let V denote a vector space of dimension N over the finite field $GF(q)$, and set

$$P := \{\, x \mid x \text{ is a subspace of } V, \quad \dim(x) \le D\}.$$

Then P is a V-matroid.

Proof. Routine.

Example 1.6. The attenuated space $A_q(D, N)$ $(0 \le D \le N)$: Let V denote an N-dimensional vector space over the finite field $GF(q)$. Fix a subspace $w \subseteq V$ such that $\dim(w) = N - D$, and set

$$P \ := \ \{\, x \mid x \text{ is a subspace of } V, \quad x \cap w = 0\}.$$

Then P is a V-matroid.

Proof. We check NT, LI, AU in Definition 1.4. NT holds since $0 \in P$, and LI holds trivially, so consider AU. Recall that for all subspaces $u,\, v \subseteq V$,

$$\text{(1.1)} \qquad \dim(u) \,+\, \dim(v) \;=\; \dim(u \cap v) \,+\, \dim(u+v).$$

Pick any $x,\, y \in P$ such that $\dim(x) < \dim(y)$. We find a one dimensional subspace $a \subseteq V$ such that $a \subseteq y$, $a \not\subseteq x$, and such that $x + a \in P$. Observe

$$\begin{aligned} \dim(x+w) &= \dim(x) \,+\, \dim(w) \\ &< \dim(y) \,+\, \dim(w) \\ &= \dim(y+w), \end{aligned}$$

so $y + w \not\subseteq x + w$. In particular,

$$\text{(1.2)} \qquad y \not\subseteq x + w.$$

By (1.2), there exists a one dimensional subspace $a \subseteq y$ such that $a \not\subseteq x + w$. Observe $a \not\subseteq x$, so it remains to show $x + a \in P$. Observe by (1.1) and the construction that

$$\begin{aligned} \dim\big((x+a) \cap w\big) &= \dim(x+a) \,+\, \dim(w) \,-\, \dim(x+a+w) \\ &= \dim(x) \,+\, 1 \,+\, \dim(w) \,-\, \dim(x+w) \,-\, 1 \\ &= 0, \end{aligned}$$

and we conclude $x + a \in P$, as desired.

Example 1.7. Polar spaces over $GF(q)$: Let V denote a finite dimensional vector space over the finite field $GF(q)$. Endow V with a form $\langle\, ,\, \rangle : V \times V \to GF(q)$ such that

$$\begin{aligned} \langle u+v,\, w\rangle &= \langle u,\, w\rangle + \langle v,\, w\rangle && (\forall u, v, w \in V), \\ \langle u,\, v+w\rangle &= \langle u,\, v\rangle + \langle u,\, w\rangle && (\forall u, v, w \in V), \\ \langle \alpha u,\, v\rangle &= \alpha\langle u,\, v\rangle && (\forall u, v \in V, \quad \forall \alpha \in GF(q)). \end{aligned}$$

Further assume

$$\langle u,\, v\rangle = \langle v,\, u\rangle \qquad (\forall u, v \in V) \quad \text{(the symmetric bilinear case)},$$

or

$$\langle u,\, u\rangle = 0 \qquad (\forall u \in V) \quad \text{(the alternating bilinear case)},$$

or

$$\langle u,\, v\rangle = \overline{\langle v,\, u\rangle} \qquad (\forall u, v \in V) \quad \text{(the Hermitean case)}.$$

(In the last case $^-$ denotes a field automorphism of $GF(q)$ of order 2.) A subspace $x \subseteq V$ is said to be *totally isotropic* (abbreviated *t.i.*) whenever $\langle u, v\rangle = 0$ for all $u, v \in x$. The set

$$P := \{x \mid x \text{ is a t.i. subspace of } V\} \tag{1.3}$$

is a V-matroid.

Proof. We verify NT, LI, AU in Definition 1.4. NT holds since $0 \in P$. LI is trivial, so consider AU. For all subspaces $x \subseteq V$, let $x^\perp$ denote the orthogonal complement

$$x^\perp := \{u \in V \mid \langle u, v\rangle = 0 \text{ for all } v \in x\}. \tag{1.4}$$

Observe x is t.i. if and only if $x \subseteq x^\perp$. By elementary linear algebra,

$$\dim(x^\perp) + \dim(x) = \dim(x \cap V^\perp) + \dim(V). \tag{1.5}$$

Pick any t.i. subspaces $x, y \subseteq V$ such that $\dim(x) < \dim(y)$. We find a one dimensional subspace $a \subseteq V$ such that $a \subseteq y$, $a \not\subseteq x$, and such that $x + a$ is t.i. To obtain a, we claim

$$x^\perp \cap y \not\subseteq x. \tag{1.6}$$

Suppose (1.6) fails. Then by the construction,

$$x^\perp \cap y = x \cap y. \tag{1.7}$$

Observe $x^\perp$, y are both contained in $(x \cap y)^\perp$, so

$$x^\perp + y \subseteq (x \cap y)^\perp. \tag{1.8}$$

Now by (1.5), (1.7), (1.8),

$$\begin{aligned}
\dim(x \cap V^\perp) + \dim(V) &= \dim(x^\perp) + \dim(x)\\
&< \dim(x^\perp) + \dim(y)\\
&= \dim(x^\perp + y) + \dim(x^\perp \cap y)\\
&\le \dim\big((x \cap y)^\perp\big) + \dim(x \cap y)\\
&= \dim(x \cap y \cap V^\perp) + \dim(V),
\end{aligned}$$

an impossibility. Hence (1.6) holds. By (1.6), there exists a one dimensional subspace $a \subseteq x^\perp \cap y$ such that $a \not\subseteq x$. Observe $a \subseteq y$, and y is t.i., so a is t.i. Also x is t.i. and $a \subseteq x^\perp$, so $x + a$ is t.i. Now a

has the desired properties, so AU holds. Now P is a V-matroid, and we are done.

Example 1.8. More polar spaces over $GF(q)$: Let V denote a finite dimensional vector space over the finite field $GF(q)$. Endow V with a quadratic form, i.e., a function $f : V \to GF(q)$ satisfying

$$f(\alpha v) = \alpha^2 f(v) \qquad (\forall \alpha \in GF(q),\ \forall v \in V),$$

$$f(u+v) = f(u) + f(v) + \langle u, v\rangle \qquad (\forall u, v \in V),$$

where $\langle\, ,\, \rangle = \langle\, ,\, \rangle_f$ is a symmetric bilinear form from Example 1.7. A subspace $x \subseteq V$ is said to be *totally singular* (abbreviated *t.s.*) whenever $f(v) = 0$ for all $v \in x$. The set

$$P := \{x \mid x \text{ is a t.s. subspace of } V\}$$

is a V-matroid.

Proof. We verify NT, LI, AU in Definition 1.4. NT, LI hold as in Example 1.7, so consider AU. First observe that any t.s. subspace $z \subseteq V$ is t.i. (with respect to $\langle\, ,\, \rangle$, in the sense of Example 1.7). Now pick any t.s. subspaces $x, y \subseteq V$ such that $\dim(x) < \dim(y)$. We find a one dimensional subspace $a \subseteq V$ such that $a \subseteq y$, $a \not\subseteq x$, and such that $x + a$ is t.s. By Example 1.7 and our preliminary comment, there exists a one dimensional subspace $a \subseteq V$ such that $a \subseteq y$, $a \not\subseteq x$, and such that $x + a$ is t.i. In fact $x + a$ is t.s. To see this, we pick any $v \in x + a$ and show $f(v) = 0$. Observe $v = v_1 + v_2$ for some $v_1 \in x$ and some $v_2 \in a$. Observe $f(v_1) = 0$, since x is t.s. Observe $f(v_2) = 0$, since $a \subseteq y$, and since y is t.s. Observe $\langle v_1, v_2\rangle = 0$, since $x + a$ is t.i. Now

$$\begin{aligned} f(v) &= f(v_1 + v_2) \\ &= f(v_1) + f(v_2) + \langle v_1, v_2\rangle \\ &= 0, \end{aligned}$$

as desired. Now a has the desired properties, so AU holds. Now P is a V-matroid, and we are done.

Note 1.9. In the nondegenerate case (Definitions 26.6, 26.8), the examples in 1.7, 1.8 are often referred to as *classical polar spaces*. This distinguishes them from the closely related *Tits polar spaces*, which we will encounter in Section 30. See [Ar], [Ca2], [Mu], [Ti] for information on the classical polar spaces.

We now seek a common generalization of Definitions 1.1, 1.4. We will use the language of partially ordered sets, so first we review some basic concepts from this area.

Let P denote a finite set. By a *partial order* on P, we mean a binary relation $\leq$ on P such that

$$
\begin{array}{llll}
\text{(i)} & x \leq x & (\forall x \in P), & \\
\text{(ii)} & x \leq y \text{ and } y \leq z & \rightarrow \quad x \leq z & (\forall\, x,\, y,\, z \in P), \\
\text{(iii)} & x \leq y \text{ and } y \leq x & \rightarrow \quad x = y & (\forall\, x,\, y \in P).
\end{array}
$$

By a *partially ordered set* (or *poset*, for short), we mean a pair $(P, \leq)$, where P is a finite set, and where $\leq$ is a partial order on P. Abusing notation, we will suppress reference to $\leq$, and just write P instead of $(P, \leq)$.

Let P, Q denote any posets. A map $\phi : P \rightarrow Q$ is said to be an *isomorphism of posets* whenever ϕ is a bijection, and for all $x, y \in P$,

$$x \leq y \quad (\text{in } P) \qquad \leftrightarrow \qquad \phi(x) \leq \phi(y) \qquad (\text{in } Q).$$

P and Q are said to be *isomorphic* whenever there exists an isomorphism of posets $\phi : P \rightarrow Q$. We do not distinguish between isomorphic posets.

Let P denote a poset, with partial order $\leq$, and let x and y denote any elements in P. As usual, we write $x < y$ whenever $x \leq y$ and $x \neq y$. We say y *covers* x whenever $x < y$, and there is no $z \in P$ such that $x < z < y$. An element $x \in P$ is said to be *maximal* (resp. *minimal*) whenever there is no $y \in P$ such that $x < y$ (resp. $y < x$). Let $\max(P)$ (resp. $\min(P)$) denote the set of all maximal (resp. minimal) elements in P. Whenever $\max(P)$ (resp. $\min(P)$) consists of a single element, we denote that element by 1 (resp. 0), and we say P *has a* 1 (resp. P *has a* 0).

Suppose P has a 0. By an *atom* (or *point*) in P, we mean an element in P that covers 0. We let A_P denote the set of atoms in P.

Suppose P has a 0. By a *rank function* on P, we mean a function $\text{rank} : P \rightarrow \mathbb{Z}$ such that $\text{rank}(0) = 0$, and such that for all $x, y \in P$,

$$\text{(1.9)} \qquad y \text{ covers } x \qquad \rightarrow \qquad \text{rank}(y) - \text{rank}(x) = 1.$$

Observe the rank function is unique if it exists. P is said to be *ranked* whenever P has a rank function. In this case, we set

$$\text{(1.10)} \qquad \text{rank}(P) := \max\{\, \text{rank}(x) \mid x \in P\},$$

$$\text{(1.11)} \qquad P_i := \{x \in P \mid \text{rank}(x) = i\} \qquad (i \in \mathbb{Z}),$$

and observe $P_0 = \{0\}$, $P_1 = A_P$. We refer to the elements of P_2 as the *lines* of P. For notational convenience set

$$\text{top}(P) := P_D, \tag{1.12}$$

where $D = \text{rank}(P)$. Observe

$$\text{top}(P) \subseteq \max(P), \tag{1.13}$$

but we might not have equality in (1.13).

Let P denote any poset, and let S denote any subset of P. Then there is a unique partial order on S such that for all $x, y \in S$,

$$x \leq y \quad (\text{in } S) \qquad \leftrightarrow \qquad x \leq y \quad (\text{in } P). \tag{1.14}$$

This partial order is said to be *induced* from P. By a *subposet* of P, we mean a subset of P, together with the partial order induced from P. Pick any $x, y \in P$ such that $x \leq y$. By the *interval* $[x, y]$, we mean the subposet

$$[x, y] := \{z \in P \mid x \leq z \leq y\}$$

of P.

Let P denote any poset, and pick any $x, y \in P$. By a *lower bound* for x, y, we mean an element $z \in P$ such that $z \leq x$ and $z \leq y$. Suppose the subposet of lower bounds for x, y has a unique maximal element. In this case we denote this maximal element by $x \wedge y$, and say $x \wedge y$ *exists.* The element $x \wedge y$ is known as the *meet* of x and y. P is said to be a (meet) *semilattice* whenever P is nonempty, and $x \wedge y$ exists for all $x, y \in P$. A semilattice has a 0. Suppose P is a semilattice, and pick any $x, y \in P$. By an *upper bound* for x, y, we mean an element $z \in P$ such that $z \geq x$ and $z \geq y$. Observe the subposet of upper bounds for x, y is closed under $\wedge$; in particular, it has a unique minimal element if and only if it is nonempty. In this case we denote this minimal element by $x \vee y$, and say $x \vee y$ *exists.* The element $x \vee y$ is known as the *join* of x and y. By a *lattice*, we mean a semilattice P such that $x \vee y$ exists for all $x, y \in P$. A lattice has a 1.

Suppose P is a semilattice. Then every interval in P is a lattice.

Suppose P is a semilattice. Then P is said to be *atomic* whenever each element of P is a join of atoms. A semilattice P is atomic if and only if each element of P that is not 0 and not an atom covers at least 2 elements of P.

Suppose P is a lattice. Then P is said to be *modular* whenever for all $x, y \in P$,

$$x, y \text{ cover } x \wedge y \qquad \leftrightarrow \qquad x \vee y \text{ covers } x, y. \tag{1.15}$$

P is modular if and only if P is ranked, and for all $x, y \in P$,

$$\text{rank}(x) + \text{rank}(y) = \text{rank}(x \wedge y) + \text{rank}(x \vee y) \tag{1.16}$$

[St, p104]. Suppose P is a modular atomic lattice. Then any interval in P is a modular atomic lattice.

We mention two examples of modular atomic lattices. (A full classification is given in Theorems 1.12, 1.13.)

Example 1.10. Let A denote a finite set. The *Boolean algebra* B_A is the poset of all subsets of A, ordered by inclusion. B_A is a modular atomic lattice. Moreover, for all x, $y \in B_A$,

$$x \wedge y = x \cap y, \tag{1.17}$$

$$x \vee y = x \cup y, \tag{1.18}$$

$$\text{rank}(x) = |x|. \tag{1.19}$$

We often write $B(D)$ to denote B_A, where $D = |A|$.

Proof. Routine.

Example 1.11. Let V denote a finite vector space. The *projective geometry* L_V is the poset of all subspaces of V, ordered by inclusion. L_V is a modular atomic lattice. Moreover, for all x, $y \in L_V$,

$$x \wedge y = x \cap y, \tag{1.20}$$

$$x \vee y = x + y, \tag{1.21}$$

$$\text{rank}(x) = \dim(x). \tag{1.22}$$

We often write $L_q(D)$ to denote L_V, where V is over the field $GF(q)$ and where $D = \dim(V)$.

Proof. Routine.

There is a classification of all modular atomic lattices essentially due to Veblen and Young, which we present below without proof.

Let q denote an integer at least 2. By a *projective plane* of order q, we mean a ranked lattice P of rank 3 such that each line in P covers exactly $q+1$ points in P, each point in P is covered by exactly $q+1$ lines in P, $a \vee b$ is a line for any distinct points $a, b \in P$, and $x \wedge y$ is a point for any distinct lines $x, y \in P$.

Let P denote a ranked poset with 0. A line $x \in P$ is said to be *thick* whenever x covers at least three points in P.

In the following theorem, we consider the modular atomic lattices that have all lines thick. In Theorem 1.13, we consider the general case.

Theorem 1.12 ([Ca2, Theorems 3.3.1, 3.4.1], [V-Y]). *For each nonnegative integer D, let Ω_D denote the class of all modular atomic lattices that have rank D and have all lines thick. Then for any poset P,*

$P \in \Omega_0$ if and only if $P = \{0\}$.
$P \in \Omega_1$ if and only if $P = \{0, 1\}$.
$P \in \Omega_2$ if and only if P is a ranked lattice with rank 2 and P has at least 3 points.
$P \in \Omega_3$ if and only if P is a projective plane of order q for some integer $q \geq 2$.
For $D \geq 4$, $P \in \Omega_D$ if and only if P is isomorphic to $L_q(D)$ for some (prime power) integer $q \geq 2$.

Let P, Q denote any posets. By the *Cartesian product* $P \times Q$, we mean the poset on the set

$$P \times Q := \{xy \mid x \in P, \ \ y \in Q\}, \tag{1.23}$$

such that for all x, $x' \in P$ and all y, $y' \in Q$,

(1.24) $xy \leq x'y'$ (in $P \times Q$) $\leftrightarrow$ $x \leq x'$ (in P) and $y \leq y'$ (in Q).

Theorem 1.13 ([Ca2, Theorems 3.3.3, 3.4.1], [V-Y]). *Let Ω denote the class of modular atomic lattices that have all lines thick. Then for any poset P, the following are equivalent.*

(i) *P is a modular atomic lattice.*
(ii) *There exists an integer $r \geq 1$ and there exists $P_1, P_2, \ldots, P_r \in \Omega$ such that $P = P_1 \times P_2 \times \cdots \times P_r$.*

A modular atomic lattice is sometimes referred to as a *generalized projective geometry.*

Let P denote any lattice. Elements $x, y \in P$ are said to be *complements* whenever $x \wedge y = 0$ and $x \vee y = 1$. P is said to be *complemented* whenever each element in P has a complement. Let P denote any semilattice. Then P is said to be *relatively complemented* whenever each interval in P is complemented. A modular atomic lattice is relatively complemented. Let P denote any semilattice, and let $I = [x, y]$ denote any interval. Then elements $u, v \in I$ are said to be *relative complements* in I whenever

$$u \wedge v = x \qquad \text{and} \qquad u \vee v = y. \tag{1.25}$$

Let P denote any poset. By a *lower ideal* in P, we mean a subposet $S \subseteq P$ such that for all $x,\ y \in P$,

$$x \in S \quad \text{and} \quad y \le x \quad \rightarrow \quad y \in S. \tag{1.26}$$

An *upper ideal* of P is defined similarly.

Definition 1.14. Let $\mathcal{P}$ denote a modular atomic lattice. By a *$\mathcal{P}$-matroid*, we mean any subposet $P \subseteq \mathcal{P}$ satisfying conditions NT, LI, AU below.

NT: $P \neq \emptyset$.
LI: P is a lower ideal in $\mathcal{P}$.
AU: For all $x, y \in P$ such that $\operatorname{rank}(x) < \operatorname{rank}(y)$, there exists an atom $a \in \mathcal{P}$ such that $a \le y$, $a \not\le x$, and such that $x \vee a \in P$.

In view of Example 1.10, for any finite set A and any subset $P \subseteq B_A$, P is an A-matroid if and only if P (together with the partial order induced from B_A) is a B_A-matroid. Similarly, in view of Example 1.11, for any finite vector space V and any subset $P \subseteq L_V$, P is a V-matroid if and only if P (together with the partial order induced from L_V) is a L_V-matroid.

We end this section with a fundamental fact about $\mathcal{P}$-matroids.

Lemma 1.15. *Let $\mathcal{P}$ denote a modular atomic lattice, and let P denote a $\mathcal{P}$-matroid. Then*

$$\max(P) = \operatorname{top}(P).$$

Proof. The inclusion $\supseteq$ is clear, so consider the inclusion $\subseteq$. Pick $x \in \max(P)$, and suppose $x \notin \operatorname{top}(P)$. Pick $y \in \operatorname{top}(P)$. Then $\operatorname{rank}(x) < \operatorname{rank}(y)$, so by AU, there exists an atom $a \in \mathcal{P}$ such that $a \le y$, $a \not\le x$, and such that $x \vee a \in P$. Now $x < x \vee a$, so $x \notin \max(P)$, a contradiction.

§2. $\mathcal{P}$-basis systems

Let P denote any poset. Elements $x, y \in P$ are said to be *comparable* whenever $x \le y$ or $y \le x$, and *incomparable* otherwise. By an *antichain* in P, we mean a subset $S \subseteq P$, where any two distinct elements of S are incomparable. There is a natural bijection from the set of all antichains in P to the set of all lower ideals in P. Indeed, for any subset $S \subseteq P$, let S^- denote the subposet

$$S^- := \{x \in P \mid x \le s \text{ for some } s \in S\}. \tag{2.1}$$

It is clear S^- is a lower ideal in P.

Lemma 2.1. *For any poset P, the map*

$$A \to A^- \tag{2.2}$$

induces a bijection from the set of all antichains of P to the set of all lower ideals of P. The inverse map is

$$L \to \max(L). \tag{2.3}$$

Proof. Routine.

Let $\mathcal{P}$ denote a modular atomic lattice. We have already considered one set of lower ideals in $\mathcal{P}$, namely the $\mathcal{P}$-matroids. The $\mathcal{P}$-matroids correspond to what set of antichains under (2.2), (2.3)? As we show in Theorem 2.5, this set consists of the *$\mathcal{P}$-basis systems*, defined as follows.

Definition 2.2. Let $\mathcal{P}$ denote a modular atomic lattice. By a *$\mathcal{P}$-basis system*, we mean any subset $B \subseteq \mathcal{P}$ that satisfies the conditions NT, AC, BA below.

NT: $B \neq \emptyset$.
AC: B is an antichain.
BA: For all $x, y \in \mathcal{P}$ such that $x \leq y$, if there exists b_1, $b_2 \in B$ such that $x \leq b_1$ and $b_2 \leq y$, then there exists $b_3 \in B$ such that $x \leq b_3 \leq y$.

The following lemma gives a second equivalent definition of a $\mathcal{P}$-basis system.

Lemma 2.3. *Let $\mathcal{P}$ denote a modular atomic lattice, and let B denote a nonempty antichain in $\mathcal{P}$. Then the following are equivalent.*

BA: *For all $x, y \in \mathcal{P}$ such that $x \leq y$, if there exists b_1, $b_2 \in B$ such that $x \leq b_1$ and $b_2 \leq y$, then there exists $b_3 \in B$ such that $x \leq b_3 \leq y$.*
BA': *For all $b_1, b_2 \in B$ and all $x \in \mathcal{P}$ such that b_1 covers x, there exists $b_3 \in B$ such that b_3 covers x and such that $b_3 \wedge b_2 > x \wedge b_2$.*

Proof. BA $\rightarrow$ BA'. We first claim that for all $b_1, b_2 \in B$ such that b_1 covers $b_1 \wedge b_2$, then b_2 covers $b_1 \wedge b_2$. To see this, observe $b_2 > b_1 \wedge b_2$ since B is an antichain, so there exists $z \in [b_1 \wedge b_2, b_2]$ such that z covers $b_1 \wedge b_2$. To prove the claim, it suffices to show

$b_2 = z$. Since b_1, z cover $b_1 \wedge b_2 = b_1 \wedge z$, we find by modularity that $b_1 \vee z$ covers b_1, z. In particular

$$b_1 \ \leq \ b_1 \vee z \ \geq \ z \ \leq \ b_2,$$

so by BA there exists $b_3 \in B$ such that

$$z \ \leq \ b_3 \ \leq \ b_1 \vee z.$$

But $b_3 \neq b_1 \vee z$ since B is an antichain, so $b_3 = z$. Now b_2 equals $b_3 = z$, since $b_3 \leq b_2$ and B is an antichain. This proves the claim.

Now pick any b_1, $b_2 \in B$, and pick any $x \in \mathcal{P}$ such that b_1 covers x. We must find $b_3 \in B$ such that b_3 covers x and $b_3 \wedge b_2 > x \wedge b_2$. We may assume $x \wedge b_2 = b_1 \wedge b_2$; otherwise we are done with $b_3 := b_1$. Set $y := x \vee b_2$. Then

$$b_1 \ \geq x \ \leq y \ \geq \ b_2,$$

so by BA, there exists $b_3 \in B$ such that $x \leq b_3 \leq y$. Observe $b_1 \not\leq y$; otherwise x and b_1 are both relative complements of b_2 in $[x \wedge b_2, y]$, contradicting (1.16). In particular $b_1 \not\leq b_3$. Now $x = b_1 \wedge b_3$, so b_3 covers x by our preliminary claim. Also $b_3 \wedge b_2 > x \wedge b_2$; otherwise x and b_3 are both relative complements of b_2 in $[x \wedge b_2, y]$, contradicting (1.16).

BA' $\rightarrow$ BA. Suppose we are given $x, y \in \mathcal{P}$ and $b_1, b_2 \in B$ such that

$$b_1 \ \geq \ x \ \leq \ y \ \geq \ b_2.$$

Of all the elements $b_3 \in B$ such that $b_3 \geq x$, pick one where $\operatorname{rank}(b_2 \wedge b_3)$ is maximal. BA will follow if we can show $b_3 \leq y$. Suppose $b_3 \not\leq y$. Then $b_3 > b_3 \wedge y$, so there exists $z \in [b_3 \wedge y, b_3]$ such that b_3 covers z. By BA', there exists $b'_3 \in B$ such that b'_3 covers z and $b'_3 \wedge b_2 > z \wedge b_2$. Now

$$\begin{aligned} b'_3 \wedge b_2 &> z \wedge b_2 \\ &\geq (b_3 \wedge y) \wedge b_2 \\ &= b_3 \wedge b_2, \end{aligned}$$

and

$$\begin{aligned} b'_3 &\geq z \\ &\geq b_3 \wedge y \\ &\geq x, \end{aligned}$$

contradicting the construction of b_3. Hence $b_3 \leq y$, as desired. This proves Lemma 2.3.

Lemma 2.4. *Let $\mathcal{P}$ denote a modular atomic lattice, and let B denote a $\mathcal{P}$-basis system. Then for all $b \in B$, $D = \operatorname{rank}(b)$ is independent of b. We refer to D as the rank of B.*

Proof. Suppose there exists $b_1, b_2 \in B$ such that $\operatorname{rank}(b_1) \neq \operatorname{rank}(b_2)$. Of all such b_1, b_2, pick a pair such that $\operatorname{rank}(b_1 \wedge b_2)$ is maximal. Observe $b_1 \wedge b_2 < b_1$, since B is an antichain, so there exists $x \in [b_1 \wedge b_2, b_1]$ such that b_1 covers x. By BA', there exists $b_3 \in B$ such that b_3 covers x and such that $b_3 \wedge b_2 > x \wedge b_2$. Now

$$\begin{aligned} b_3 \wedge b_2 &> x \wedge b_2 \\ &= b_1 \wedge b_2, \end{aligned}$$

so $\operatorname{rank}(b_2) = \operatorname{rank}(b_3)$ by the construction. Also $\operatorname{rank}(b_1) = \operatorname{rank}(b_3)$, since both b_1, b_3 cover x, so $\operatorname{rank}(b_1) = \operatorname{rank}(b_2)$, contradicting our assumptions. This proves Lemma 2.4.

Theorem 2.5. *Let $\mathcal{P}$ denote a modular atomic lattice.*

(i) *Let B denote a $\mathcal{P}$-basis system. Then B^- is a $\mathcal{P}$-matroid.*
(ii) *Let P denote a $\mathcal{P}$-matroid. Then $\max(P)$ is a $\mathcal{P}$-basis system.*

In particular, the map $B \to B^-$ is a bijection from the set of all $\mathcal{P}$-basis systems to the set of all $\mathcal{P}$-matroids.

Proof. (i) Observe B^- certainly satisfies NT, LI in Definition 1.14. To verify condition AU in that definition, pick any $x, y \in B^-$ such that

$$\operatorname{rank}(x) < \operatorname{rank}(y). \tag{2.4}$$

By the construction, there exists $b_1, b_2 \in B$ such that $x \leq b_1$, $y \leq b_2$. Observe

$$b_1 \;\geq\; x \;\leq b_2 \vee x \;\geq\; b_2,$$

so by BA, there exists $b_3 \in B$ such that

$$x \;\leq\; b_3 \;\leq b_2 \vee x.$$

It is immediate from the left inequality above that $x \wedge y \leq b_3 \wedge y$. We claim that in fact

$$x \wedge y < b_3 \wedge y. \tag{2.5}$$

To see (2.5), observe $b_3 \vee y \leq b_2 \vee x$ by the construction, so

$$\operatorname{rank}(b_3 \vee y) \leq \operatorname{rank}(b_2 \vee x). \tag{2.6}$$

Observe $x \wedge y \leq x \wedge b_2$ since $y \leq b_2$, so

$$\text{rank}(x \wedge y) \leq \text{rank}(b_2 \wedge x). \tag{2.7}$$

From (1.16) and Lemma 2.4, we also have

$$\text{rank}(b_2) = \text{rank}(b_3), \tag{2.8}$$

$$\text{rank}(b_3) + \text{rank}(y) = \text{rank}(b_3 \wedge y) + \text{rank}(b_3 \vee y), \tag{2.9}$$

$$\text{rank}(b_2 \wedge x) + \text{rank}(b_2 \vee x) = \text{rank}(x) + \text{rank}(b_2). \tag{2.10}$$

Summing (2.4), (2.6)–(2.10), we obtain

$$\text{rank}(x \wedge y) < \text{rank}(b_3 \wedge y),$$

and (2.5) follows. Now by (2.5), there exists an atom $a \in \mathcal{P}$ such that $a \leq b_3 \wedge y$ but $a \not\leq x \wedge y$. Now $a \leq y$ and $a \not\leq x$ by the construction. Also $a \vee x \in B^-$, since $a \vee x \leq b_3$ by the construction. We have now verified AU in Definition 1.14, so B^- is a $\mathcal{P}$-matroid.

(ii) Certainly $B := \max(P)$ satisfies conditions NT, AC in Definition 2.2. To show B is $\mathcal{P}$-basis system, it suffices to show BA'. First, we remark by Lemma 1.15 that $\text{rank}(b)$ is independent of $b \in B$. Now pick any $b_1, b_2 \in B$, and any $x \in \mathcal{P}$ such that b_1 covers x. We must find $b_3 \in B$ such that b_3 covers x, and such that $b_3 \wedge b_2 > x \wedge b_2$. Observe $\text{rank}(x) < \text{rank}(b_2)$ by our remark, so by AU in Definition 1.14, there exists an atom $a \in \mathcal{P}$ such that $a \leq b_2$, $a \not\leq x$, and such that $x \vee a \in P$. Observe a covers $0 = x \wedge a$, so $x \vee a$ covers x by modularity. In particular $\text{rank}(x \vee a) = \text{rank}(b_1)$, forcing $x \vee a \in B$ by our remark. Set $b_3 := x \vee a$. Since $b_3 \geq x$ we have

$$b_3 \wedge b_2 \geq x \wedge b_2.$$

In fact

$$b_3 \wedge b_2 > x \wedge b_2,$$

since $a \leq b_3 \wedge b_2$ but $a \not\leq x \wedge b_2$. Now BA' holds, so B is a $\mathcal{P}$-basis system by Lemma 2.3. This proves Theorem 2.5.

We can use Theorem 2.5 to get examples of $\mathcal{P}$-matroids.

Example 2.6. Let D denote a nonnegative integer. Let $\mathcal{P}$ denote a modular atomic lattice with rank $D+1$, and let B denote any non empty subset of $\mathcal{P}_D$. Then B is a $\mathcal{P}$-basis system of rank D. Moreover, B^- is a $\mathcal{P}$-matroid of rank D.

Proof. Routine application of Definition 2.2, Theorem 2.5.

We mention one other fact about basis systems in a modular atomic lattice.

Lemma 2.7. *Let $\mathcal{P}$ denote a modular atomic lattice, and let B denote a $\mathcal{P}$-basis system. Pick any $x, y \in \mathcal{P}$ such that $x \leq y$ and such that $B \cap [x, y] \neq \emptyset$. Then $B \cap [x, y]$ is a $[x, y]$-basis system.*

Proof. $B \cap [x, y]$ certainly satisfies conditions NT, AC in Definition 2.2. To verify BA, pick $u, v \in [x, y]$ and pick $b_1, b_2 \in B \cap [x, y]$ such that

$$b_1 \geq u \leq v \geq b_2.$$

We must find $b_3 \in B \cap [x, y]$ such that $u \leq b_3 \leq v$. Applying BA to B, we find there exists $b_3 \in B$ such that $u \leq b_3 \leq v$. But now $x \leq b_3 \leq y$ by the construction, so in fact $b_3 \in B \cap [x, y]$, as desired.

§3. The dual of a $\mathcal{P}$-matroid

Let P denote any poset. By the *poset-dual* of P, we mean the poset P^* on the same set as P, such that for all elements x, y,

$$x \leq y \ (\text{in } P^*) \quad \leftrightarrow \quad x \geq y \ (\text{in } P). \tag{3.1}$$

More generally, let S denote any subset of P. Then S^* will denote the subposet of P^* induced on S.

We mention that P is a modular atomic lattice if and only if P^* is a modular atomic lattice [St, Theorem 3.3.3].

Lemma 3.1. *Let $\mathcal{P}$ denote a modular atomic lattice. Then for all subsets $B \subseteq \mathcal{P}$, the following are equivalent.*

(i) *B is a $\mathcal{P}$-basis system.*
(ii) *B is a $\mathcal{P}^*$-basis system.*

Proof. This is immediate from the symmetry in the axioms NT, AC, BA from Definition 2.2.

Let P denote any poset. For any subset $S \subseteq P$, let S^+ denote the subposet

$$S^+ = \{ x \in P \mid x \geq s \ \text{for some} \ \ s \in S\}. \tag{3.2}$$

Observe S^+ is an upper ideal in P.

Definition 3.2. Let $\mathcal{P}$ denote a modular atomic lattice, and let P denote a $\mathcal{P}$-matroid. By the *matroid-dual* of P (with respect to $\mathcal{P}$), we mean the $\mathcal{P}^*$-matroid $(B^+)^*$, where $B = \max(P)$.

§4. The definition of a quantum matroid

In this section, we introduce the notion of a quantum matroid, and consider the examples with rank at most 2.

Definition 4.1. By a *quantum matroid*, we mean any nonempty poset P satisfying the conditions R, SL, M, AU below.

R: P is ranked.
SL: P is a (meet) semilattice.
M: For all $x \in P$, the interval $[0, x]$ is a modular atomic lattice.
AU: For all $x, y \in P$ satisfying $\operatorname{rank}(x) < \operatorname{rank}(y)$, there exists an atom $a \in P$ such that $a \leq y$, $a \not\leq x$, and such that $x \vee a$ exists in P.

Let $\mathcal{P}$ denote a modular atomic lattice, and let P denote a $\mathcal{P}$-matroid. Then the subposet P is a quantum matroid. In particular, any modular atomic lattice is a quantum matroid. We now consider the quantum matroids of rank at most 2.

A poset P is a quantum matroid of rank 0 if and only if P consists of a single element. A poset P is a quantum matroid of rank 1 if and only if P has a 0 and at least one other element, and all nonzero elements of P cover 0. The example below characterizes the quantum matroids of rank 2.

Example 4.2. A poset P is a quantum matroid of rank 2 if and only if P has a 0, and satisfies the following four conditions:

R: P is ranked and $\operatorname{rank}(P) = 2$.
SL: For any distinct points $x, y \in P$, there exists at most one line $z \in P$ such that $x \leq z$, $y \leq z$.
M: Each line in P covers at least 2 points in P.
AU: For each point $x \in P$ and each line $y \in P$ such that $x \not\leq y$, there exists a point $x' \in P$ and a line $y' \in P$ such that $x \leq y' \geq x' \leq y$.

We have already proved some facts about $\mathcal{P}$-matroids. Are there corresponding results about the more general quantum matroids? Lemma 1.15 can certainly be extended to this level.

Lemma 4.3. *Let P denote a quantum matroid. Then*

$$\max(P) = \operatorname{top}(P). \tag{4.1}$$

Proof. Similar to the proof of Lemma 1.15.

Problem 4.4. Extend the notion of the dual of a $\mathcal{P}$-matroid (Def. 3.2) to the level of quantum matroids.

§5. Prematroids and their subposets

Definition 5.1. By a *pre-quantum matroid* (or simply, a *prematroid*), we mean a nonempty poset P that satisfies conditions R, SL, M in Definition 4.1.

We will have occasion to consider subposets of prematroids that possess the following properties.

Definition 5.2. Let P denote any poset, and let S denote any subposet of P.

(i) S is said to be $\wedge$-*closed* in P whenever for all $x, y \in S$,

$$x \wedge_P y \quad \text{exists} \quad \to \quad x \wedge_P y \in S. \tag{5.1}$$

The notion of $\vee$-closure is defined similarly.

(ii) S is said to be *convex* in P whenever for all $x, y, z \in P$,

$$x, y \in S \quad \text{and} \quad x \le z \le y \quad \to \quad z \in S. \tag{5.2}$$

Lemma 5.3. *Let P denote a poset with 0, and let S denote any nonempty subposet of P. Then the following are equivalent.*

(i) *S is a lower ideal in P.*

(ii) *S is convex in P, and $0_P \in S$.*

If (i)–(ii) *hold, then S is $\wedge$-closed in P, and $0_S = 0_P$.*

Proof. (i) $\to$ (ii). Routine.

(ii) $\to$ (i). Pick any $x \in S$ and any $y \in P$ such that $y \le x$. Then $0_P \le y \le x$, so $y \in S$ by convexity.

Now assume (i)–(ii). To see S is $\wedge$-closed in P, pick any $x, y \in S$ such that $x \wedge_P y$ exists. Certainly $x \wedge_P y \le x$, so $x \wedge_P y \in S$ by the definition of a lower ideal. Hence S is $\wedge$-closed in P. It is clear that S, P have the same 0. This proves Lemma 5.3.

Lemma 5.4. *Let P denote a semilattice, and let S denote a nonempty $\wedge$-closed subposet of P. Then S is a semilattice. Moreover, for all $x, y \in S$,*

$$x \wedge_S y = x \wedge_P y. \tag{5.3}$$

Proof. Pick any $x,\, y \in S$. Then it suffices to show

$$\max(L) = \{\, x \wedge_P y \}, \tag{5.4}$$

where

$$L := \{\, z \in S \mid z \leq x, \quad z \leq y \}.$$

Certainly $x \wedge_P y \in L$, since $x \wedge_P y \in S$ by $\wedge$-closure. Also $x \wedge_P y \geq z$ for all $z \in L$, so $x \wedge_P y$ is the unique maximal element of L. this proves Lemma 5.4.

Lemma 5.5. *Let P denote a semilattice, and let S denote a convex subposet of P. Then for all $x,\, y \in S$, the following are equivalent.*

(i) $x \vee_S y$ *exists.*
(ii) $x \vee_P y$ *exists, and* $x \vee_P y \in S$.

If (i)–(ii) *hold, then*

$$x \vee_S y = x \vee_P y. \tag{5.5}$$

Proof. (i) $\to$ (ii). $x \vee_P y$ exists, since $x \vee_S y$ is an upper bound for $x,\, y$ in P. Also $x \vee_P y \in S$ by convexity, since $x \leq x \vee_P y \leq x \vee_S y$.

(ii) $\to$ (i). Clear.

Now suppose (i), (ii). We have observed $x \vee_P y \leq x \vee_S y$. Also $x \vee_P y$ is an upper bound for $x,\, y$ in S, so $x \vee_P y \geq x \vee_S y$. Hence $x \vee_P y$ and $x \vee_S y$ are identical. This proves Lemma 5.5.

Lemma 5.6. *Let P denote a ranked semilattice, and let S denote a nonempty $\wedge$-closed, convex subposet of P. Then S is ranked. Moreover, for all $x \in S$,*

$$\operatorname{rank}_S(x) = \operatorname{rank}_P(x) \; - \; \operatorname{rank}_P(0_S). \tag{5.6}$$

Proof. S is a semilattice by Lemma 5.4; in particular S has a 0. We show the function $R : S \to \mathbb{Z}$ defined by

$$R(x) = \operatorname{rank}_P(x) \; - \; \operatorname{rank}_P(0_S) \qquad\qquad (x \in S) \tag{5.7}$$

is a rank function for S. Certainly $R(0_S) = 0$. Also, for any $x,\, y \in S$ such that y covers x (in S), then y covers x (in P) by the convexity of S, forcing

$$\operatorname{rank}_P(y) \; - \; \operatorname{rank}_P(x) = 1.$$

Now

$$R(y) \; - \; R(x) = 1$$

by (5.7), so R is a rank function for S by (1.9). This proves Lemma 5.6.

Corollary 5.7. *Let P denote a prematroid. Then any nonempty $\wedge$-closed, convex subposet of P is a prematroid. In particular, any nonempty lower ideal of P is a prematroid.*

Proof. Let S denote a nonempty $\wedge$-closed, convex subposet of P. Then S satisfies axiom SL by Lemma 5.4, and axiom R by Lemma 5.6. To see that S satisfies axiom M, pick any $x \in S$. Then the interval $[0_S, x]$ of S may be viewed as an interval in the modular atomic lattice $[0_P, x]$, and is therefore a modular atomic lattice. We have now shown S is a prematroid. The last line of the present corollary follows from Lemma 5.3.

§6. Embeddable Posets

In this section, we define the notion of an *embedding* of a poset, and consider posets that are embeddable into a modular atomic lattice.

Definition 6.1. Let P and $\mathcal{P}$ denote posets. By a *$\mathcal{P}$-embedding* of P, we mean an injection $\sigma : P \to \mathcal{P}$ that satisfies (i), (ii) below.

(i) $x \le y \quad \leftrightarrow \quad \sigma(x) \le \sigma(y) \qquad\qquad (\forall\, x,\, y \in P)$.
(ii) $\sigma(P)$ is a lower ideal in $\mathcal{P}$.

Lemma 6.2. *Let $\mathcal{P}$ denote a modular atomic lattice.*

(i) *Let P denote a quantum matroid, and let $\sigma : P \to \mathcal{P}$ denote a $\mathcal{P}$-embedding of P. Then $\sigma(P)$ is a $\mathcal{P}$-matroid.*
(ii) *Let Q denote a $\mathcal{P}$-matroid, and let $\sigma : Q \to \mathcal{P}$ denote the identity map on Q. Then σ is an embedding of Q.*

Proof. Immediate from Definitions 1.14, 4.1.

Definition 6.3. A poset P is said to be *embeddable* whenever $P \ne \emptyset$, and there exists a pair $\mathcal{P}$, σ, where $\mathcal{P}$ is a modular atomic lattice, and σ is a $\mathcal{P}$-embedding of P.

Lemma 6.4. *Let P denote an embeddable poset. Then P is isomorphic to a lower ideal in a modular atomic lattice. In particular, P is a prematroid.*

Proof. Immediate from Definition 6.1, Corollary 5.7, and the observation that any modular atomic lattice is a prematroid.

We end this section with a conjecture.

Conjecture 6.5. *Let* P *denote a quantum matroid with rank at least* 4. *Then* P *is embeddable.*

We will see in Corollary 39.8 that the above conjecture is true for the *regular* quantum matroids. See also [C-J-P], [Sp1], [Ti, Theorem 8.21].

§7. The distance function ∂

For the next several sections, we will develop a theory of prematroids. We will use the following notation.

Let P denote any poset, and pick any $x, y \in P$. Let us say x, y are *adjacent* whenever x covers y or y covers x. Pick any nonnegative integer d. By a *path of length* d connecting x, y, we mean any sequence $x = x_0, x_1, \ldots, x_d = y$ $(x_0, x_1, \ldots, x_d \in P)$, such that x_i, x_{i+1} are adjacent for all i $(0 \leq i \leq d-1)$. P is said to be *connected* whenever for all $x, y \in P$, there exists a path in P connecting x and y. Suppose P has a 0. Then P is connected, since for all $x \in P$, there exists a path in P connecting x, 0.

Let P denote an arbitrary connected poset. For any $x, y \in P$, define the *distance*

$$\partial(x, y) := \min\{\, d \mid \text{there exists a path of length } d \text{ connecting } x, y\}.$$

Then for all x, y, $z \in P$,

$$\partial(x,y) + \partial(y,z) \geq \partial(x,z). \tag{7.1}$$

If P is ranked, we can say a bit more.

Lemma 7.1. *Let* P *denote a ranked poset with* 0, *and pick any* $x, y \in P$. *Then*

(i) *We have*

$$\partial(x,y) \geq \operatorname{rank}(y) - \operatorname{rank}(x). \tag{7.2}$$

(ii) *Equality holds in* (7.2) *if and only if* $y \geq x$.

(iii) *For all* $z \in P$ *such that* z *is adjacent to* x,

$$\partial(x,y) - \partial(z,y) \in \{1, -1\}. \tag{7.3}$$

Proof. Lines (i), (ii) are immediate from (1.9). Line (iii) follows from the observation that the graph structure on P is bipartite.

Let P denote an arbitrary connected poset. A path $x_0, x_1, \ldots, x_d$ in P is said to be *geodesic* whenever $\partial(x_0, x_d) = d$. More generally, any sequence $x_0, x_1, \ldots, x_d$ of elements from P is said to be *geodesic* whenever

$$\sum_{i=0}^{d-1} \partial(x_i, x_{i+1}) \quad = \quad \partial(x_0, x_d). \tag{7.4}$$

Let P denote a ranked poset with 0. Then for all $x_0, x_1, \ldots, x_d \in P$,

$$x_0 \le x_1 \le \cdots \le x_d \qquad \rightarrow \qquad x_0,\, x_1,\, \ldots,\, x_d \ \text{ is geodesic.} \tag{7.5}$$

Let P denote a ranked poset with 0, and let $p := (x_0, x_1, \ldots, x_d)$ denote a path in P. By the *shape* of P, we mean the sequence

$$\text{shape}(p) \ := \ (\text{rank}(x_0),\ \text{rank}(x_1), \ldots, \text{rank}(x_d)). \tag{7.6}$$

By the *weight* of p, we mean the scalar

$$\text{weight}(p) \ := \ \sum_{i=0}^{d} \text{rank}(x_i). \tag{7.7}$$

Lemma 7.2. *Let P denote a prematroid. Pick any nonnegative integer d, and pick any path $p = (x_0, x_1, \ldots, x_d)$ $(x_0, x_1, \ldots, x_d \in P)$. Then the following are equivalent.*

(i) *There does not exist an integer i $(1 \le i \le d-1)$ such that*

$$x_{i-1} < x_i > x_{i+1}. \tag{7.8}$$

(ii) *There exists an integer e $(0 \le e \le d)$ such that*

$$x_0 > x_1 > x_2 > \cdots > x_e,$$
$$x_e < x_{e+1} < x_{e+2} < \cdots < x_d.$$

Suppose (i)–(ii) *hold. Then we say p is down-up. We call x_e the base of p.*

Proof. Routine.

Lemma 7.3. *Let P denote a prematroid, and pick any $x, y \in P$. Then there exists a geodesic down-up path connecting x, y.*

Proof. Set $d := \partial(x, y)$, and pick a geodesic path

$$p \;:=\; (x = x_0,\, x_1,\, \ldots,\, x_d = y) \qquad\qquad (x_0,\, x_1,\, \ldots,\, x_d \in P)$$

with minimal weight in the sense of (7.7). We claim p is down-up. Suppose not. Then by Lemma 7.2, there exists an integer i $(1 \le i \le d-1)$ such that

$$x_{i-1} < x_i > x_{i+1}. \tag{7.9}$$

Observe $x_{i-1} \neq x_{i+1}$ since p is geodesic, and of course x_i covers both x_{i-1} and x_{i+1}, so $x_i = x_{i-1} \vee x_{i+1}$. It follows $x_{i-1} \vee x_{i+1}$ covers x_{i-1}, x_{i+1}, so by modularity, x_{i-1}, x_{i+1} both cover $x_{i-1} \wedge x_{i+1}$. Now the sequence

$$p' \;:= (x = x_0,\, x_1,\, \ldots,\, x_{i-1},\, x_{i-1} \wedge x_{i+1},\, x_{i+1},\, \ldots,\, x_d = y)$$

is a path. Observe p' is geodesic, since p, p' have the same length, and

$$\text{weight}(p') = \text{weight}(p) - 2.$$

This contradicts our construction, and we conclude p is down-up. This proves Lemma 7.3.

Lemma 7.4. *Let P denote a prematroid, and pick any $x, y, z \in P$. Then the following are equivalent:*

(i) *The sequence xzy is geodesic, and $z \le x$, $z \le y$.*
(ii) $z = x \wedge y$.

Proof. We set

$$N := \{\, u \in P \mid xuy \text{ is geodesic}, \quad u \le x, \;\; u \le y\},$$

and show $N = \{x \wedge y\}$. To do this, it suffices to show

$$N \subseteq \{x \wedge y\}, \tag{7.10}$$

and

$$N \neq \emptyset. \tag{7.11}$$

To obtain (7.10), pick any $u \in N$. Observe u is a lower bound for x and y, so $u \le x \wedge y$. Now $u \le x \wedge y \le x$, so u, $x \wedge y$, x is geodesic by (7.5). Similarly $u \le x \wedge y \le y$, so u, $x \wedge y$, y is geodesic. Recall xuy is geodesic, so x, $x \wedge y$, u, $x \wedge y$, y is geodesic. In particular $x \wedge y$, u, $x \wedge y$ is geodesic, so $u = x \wedge y$. We now have (7.10). To obtain (7.11), recall by Lemma 7.3, there exists a geodesic down-up path p connecting x, y.

Let u denote the base of p, in the sense of Lemma 7.2. Then $u \in N$ by construction, and (7.11) follows. This proves Lemma 7.4.

Replacing $\wedge$ by $\vee$ in Lemma 7.4, we get the following result.

Lemma 7.5. *Let P denote a prematroid, and pick any $x, y, z \in P$. Then the following are equivalent:*

(i) *The sequence xzy is geodesic, and $z \geq x$, $z \geq y$.*
(ii) *The join $x \vee y$ exists, and $z = x \vee y$.*

Proof. Similar to Lemma 7.4.

Corollary 7.6. *Let P denote a prematroid, and pick any $x, y \in P$.*

(i)

$$\partial(x,y) = \operatorname{rank}(x) + \operatorname{rank}(y) - 2\operatorname{rank}(x \wedge y). \tag{7.12}$$

(ii) *Suppose $x \vee y$ exists. Then*

$$\partial(x,y) = 2\operatorname{rank}(x \vee y) - \operatorname{rank}(x) - \operatorname{rank}(y). \tag{7.13}$$

(iii) *Suppose $x \vee y$ exists. Then*

$$\operatorname{rank}(x) + \operatorname{rank}(y) = \operatorname{rank}(x \wedge y) + \operatorname{rank}(x \vee y). \tag{7.14}$$

Proof. To see (i), observe by Lemma 7.1(i),(ii) and Lemma 7.4 that

$$\begin{aligned} \partial(x,y) &= \partial(x, x \wedge y) + \partial(x \wedge y, y) \\ &= \operatorname{rank}(x) - \operatorname{rank}(x \wedge y) + \operatorname{rank}(y) - \operatorname{rank}(x \wedge y). \end{aligned}$$

The proof of (ii) is similar. To get (iii), equate (7.12), (7.13).

Corollary 7.7. *Let P denote a prematroid, and let S denote a nonempty $\wedge$-closed, convex subposet of P. Then for all $x, y \in S$,*

$$\partial_S(x,y) = \partial_P(x,y). \tag{7.15}$$

Proof. By Corollary 7.6(i),

$$\partial_P(x,y) = \operatorname{rank}_P(x) + \operatorname{rank}_P(y) - 2\operatorname{rank}_P(x \wedge_P y). \tag{7.16}$$

Observe S is a prematroid by Corollary 5.7. Applying Corollary 7.6(i) to S, we obtain

$$\partial_S(x,y) = \operatorname{rank}_S(x) + \operatorname{rank}_S(y) - 2\operatorname{rank}_S(x \wedge_S y). \tag{7.17}$$

Evaluating the right hand side of (7.17) using (5.3), (5.6), we find the right hand sides of (7.16), (7.17) are equal. Line (7.15) follows, and Corollary 7.7 is proved.

Lemma 7.8. *Let P denote a prematroid, and pick any $x, y, z \in P$. Then the following are equivalent:*

(i) *The sequence xzy is geodesic.*

(ii) *$z \geq x \wedge y$, and $x \wedge z$, $y \wedge z$ are relative complements in the interval $[x \wedge y, z]$.*

(iii) *There exists $u \in [x \wedge y, x]$ and there exists $v \in [x \wedge y, y]$ such that $z = u \vee v$.*

Moreover, if (i)–(iii) *hold, then*

$$u = x \wedge z, \tag{7.18}$$

$$v = y \wedge z. \tag{7.19}$$

Proof. (i) $\rightarrow$ (ii). Set $u := x \wedge z$, $v := y \wedge z$. Then it suffices to show $z = u \vee v$ and $x \wedge y = u \wedge v$. Observe xuz, zvy are each geodesic by Lemma 7.4, so $xuzvy$ is geodesic. In particular uzv is geodesic. Since $z \geq u$ and $z \geq v$ by the construction, we find $z = u \vee v$ by Lemma 7.5. Observe by our remarks above that $xuvy$ is geodesic. Observe by Lemma 7.4 that u, $u \wedge v$, v is geodesic, so x, u, $u \wedge v$, v, y is geodesic. In particular x, $u \wedge v$, y is geodesic. But $u \wedge v \leq u \leq x$ and $u \wedge v \leq v \leq y$, so $u \wedge v = x \wedge y$ by Lemma 7.4.

(ii) $\rightarrow$ (iii). Set $u := x \wedge z$, $v := y \wedge z$. Observe $u \geq x \wedge y$ since $z \geq x \wedge y$, and $u \leq x$ by construction, so $u \in [x \wedge y, x]$. Similarly $v \in [x \wedge y, y]$. Also $z = u \vee v$ by (ii) and (1.25).

(iii) $\rightarrow$ (i). Observe x, $x \wedge y$, y is geodesic by Lemma 7.4. Observe both x, u, $x \wedge y$ and $x \wedge y$, v, y are geodesic by the construction, so x, u, $x \wedge y$, v, y is geodesic. In particular, $xuvy$ is geodesic. Also u, $u \vee v$, v is geodesic by Lemma 7.5, so x, u, $u \vee v$, v, y is geodesic. In particular, x, $u \vee v$, y is geodesic, so (i) holds.

Now suppose (i)–(iii) hold. We have observed in the proof of (iii) $\rightarrow$ (i) that $xuzvy$ is geodesic. Now (7.18) holds by Lemma 7.4, since xuz is geodesic, and since $u \leq x$, $u \leq z$. Line (7.19) is similar. This proves Lemma 7.8.

Interchanging the roles of $\vee$, $\wedge$ in the above lemma, we obtain the following result.

Lemma 7.9. *Let P denote a prematroid, and pick any $x, y, z \in P$ such that $x \vee y$ exists. Then the following are equivalent:*

(i) *The sequence xzy is geodesic.*

(ii) *$z \le x \vee y$, and $x \vee z$, $y \vee z$ are relative complements in the interval $[z, x \vee y]$.*

(iii) *There exists $u \in [x, x \vee y]$ and there exists $v \in [y, x \vee y]$ such that $z = u \wedge v$.*

Moreover, if (i)–(iii) *hold, then*

$$u = x \vee z, \tag{7.20}$$

$$v = y \vee z. \tag{7.21}$$

Proof. Similar to Lemma 7.8.

Lemma 7.10. *Let P denote a prematroid, and pick any x, y, z, $z' \in P$ such that both xzy and $xz'y$ are geodesic, and such that $z \vee z'$ exists. Then x, $z \vee z'$, y is geodesic.*

Proof. By Lemma 7.8, there exists $u, u' \in [x \wedge y, x]$ and there exists $v, v' \in [x \wedge y, y]$ such that $z = u \vee v$ and $z' = u' \vee v'$. Now

$$\begin{aligned} &z \vee z' = (u \vee u') \vee (v \vee v'), \\ &u \vee u' \in [x \wedge y, x], \\ &v \vee v' \in [x \wedge y, y], \end{aligned}$$

so x, $z \vee z'$, y is geodesic by Lemma 7.8.

Lemma 7.11. *Let P denote a prematroid, and pick any x, y, z, $z' \in P$ such that both xzy and $xz'y$ are geodesic, and such that $x \vee y$ exists. Then x, $z \wedge z'$, y is geodesic.*

Proof. Similar to Lemma 7.10.

Lemma 7.12. *Let P denote an embeddable poset, and pick any $x, y, z, z' \in P$ such that both xzy and $xz'y$ are geodesic. Then x, $z \wedge z'$, y is geodesic.*

Proof. By Lemma 6.4 , we may identify P with a lower ideal in some modular atomic lattice $\mathcal{P}$. By Lemmas 5.3, 5.4, $z \wedge z'$ is the same as computed in P or $\mathcal{P}$. By Lemma 5.3 and Corollary 7.7, the distance function for P equals the restriction to P of the distance function for $\mathcal{P}$. In particular, both xzy and $xz'y$ are geodesic in $\mathcal{P}$. Applying Lemma 7.11 to $\mathcal{P}$, we find $x, z \wedge z', y$ is geodesic in $\mathcal{P}$. By our above remark, $x, z \wedge z', y$ is geodesic in P. This proves Lemma 7.12.

Conjecture 7.13. *Let P denote a prematroid such that*

(i) *the rank of P is at least* 3,

(ii) *for all* $x, y, z, z' \in P$ *such that both* xzy *and* $xz'y$ *are geodesic, then* $x,\ z \wedge z',\ y$ *is geodesic.*

Then P *is embeddable.*

§8. Geodesically closed subposets in a prematroid

In this section, we introduce the notion of a geodesically closed subposet in a prematroid, and characterize these subposets in terms of the meet and join operation.

Definition 8.1. Let P denote a prematroid. A subposet $G \subseteq P$ is said to be *geodesically closed* in P, whenever G is nonempty, and for all $x,\ y,\ z \in P$,

$$x,\ y \in G \quad \text{and} \quad xzy \text{ geodesic in } P \quad \rightarrow \quad z \in G.$$

Lemma 8.2. *Let* P *denote a prematroid. Then for any subposet* $G \subseteq P$, *the following are equivalent.*

(i) G *is geodesically closed in* P.
(ii) G *is nonempty,* $\wedge$*-closed,* $\vee$*-closed, and convex in* P.

Proof. (i) $\rightarrow$ (ii). G is nonempty by Definition 8.1. Given $x, y \in G$, observe $x,\ x \wedge y,\ y$ is geodesic in P by Lemma 7.4, so $x \wedge y \in G$. Suppose $x \vee y$ exists in P. Then $x,\ x \vee y,\ y$ is geodesic in P by Lemma 7.5, so $x \vee y \in G$. Suppose $x \le y$, and pick any $z \in [x, y]$. Then xzy is geodesic in P by (7.5), so $z \in G$. We now have (ii).

(ii) $\rightarrow$ (i). Suppose we are given $x, y \in G$ and $z \in P$ such that xzy is geodesic in P. Then by Lemma 7.8(i),(iii), $z = u \vee v$ for some $u \in [x \wedge y, x]$ and some $v \in [x \wedge y, y]$. Observe $x \wedge y \in G$ by $\wedge$-closure, so now $u, v \in G$ by convexity, and now $z \in G$ by $\vee$-closure. This proves Lemma 8.2.

Corollary 8.3. *Let* P *denote a prematroid, and let* G *denote a geodesically closed subposet of* P. *Then* G *is a prematroid.*

Proof. G is $\wedge$-closed and convex by Lemma 8.2, and is therefore a prematroid by Corollary 5.7.

Lemma 8.4. *Let* P *denote a prematroid, and pick any* $x \in P$.

(i) *The subposet* $x^+ := \{y \in P \mid y \ge x\}$ *is geodesically closed in* P.
(ii) *For all* $y \in P$ *such that* $y \ge x$, *the interval* $[x, y]$ *is geodesically closed in* P.

Proof. The subposets x^+, $[x,y]$ satisfy the condition Lemma 8.2(ii), and are therefore geodesically closed in P by that lemma.

Lemma 8.5. *Let P denote a prematroid, and let G denote a geodesically closed subposet of P that is contained in an interval of P. Then G is an interval. In particular, G is a modular atomic lattice.*

Proof. Observe by Lemma 8.2 that $G = [x,y]$, where $x = \bigwedge_{w\in G} w$ and $y = \bigvee_{w\in G} w$.

§9. Submatroids and subspaces in a prematroid

In this section we introduce the notions of a *submatroid* and a *subspace* in a prematroid, and show these objects are in 1–1 correspondence.

Lemma 9.1. *Let P denote a prematroid. Then for any subposet $G \subseteq P$, the following are equivalent.*

(i) *G is a nonempty $\vee$-closed lower ideal in P.*
(ii) *G is geodesically closed in P, and $0_G = 0_P$.*

If (i)–(ii) *hold, we say G is a subprematroid of P,* (*or simply, a submatroid*).

Proof. (i) $\to$ (ii). G is convex and $\wedge$-closed in P by Lemma 5.3, so G is geodesically closed in P by Lemma 8.2. G, P share the same 0 by Lemma 5.3.

(ii) $\to$ (i). G is nonempty by Definition 8.1. G is $\vee$-closed in P by Lemma 8.2. G is convex in P by the same lemma, so G is a lower ideal in P by Lemma 5.3. This proves Lemma 9.1.

Lemma 9.2. *Let P denote a prematroid. Then for all subposets $G \subseteq P$, and for all $x \in P$, the following are equivalent.*

(i) *G is geodesically closed in P, and $0_G = x$.*
(ii) *G is a submatroid of x^+.*

Proof. (i) $\to$ (ii). G is geodesically closed in P and contained in x^+, so G is geodesically closed in x^+. The result now follows from Lemma 9.1.

(ii) $\to$ (i). By Lemma 9.1, $0_G = x$, and G is geodesically closed in x^+. But x^+ is geodesically closed in P by Lemma 8.4(i), so G is geodesically closed in P. This proves Lemma 9.2.

Let P denote a poset with 0, and recall A_P denotes the set of all atoms of P. For all $x \in P$, define

$$\text{Shadow}(x) := \{\, a \in A_P \mid a \leq x \}. \tag{9.1}$$

Observe by Definition 5.1 that any prematroid is atomic.

Lemma 9.3. *Let P denote an atomic semilattice. Then*

(i) *for all $x \in P$,*

$$x = \bigvee_{a \in \mathrm{Shadow}(x)} a. \tag{9.2}$$

(ii) *for all $x, y \in P$,*

$$x \le y \quad \leftrightarrow \quad \mathrm{Shadow}(x) \subseteq \mathrm{Shadow}(y). \tag{9.3}$$

(iii) *for all $x, y \in P$,*

$$\mathrm{Shadow}(x \wedge y) = \mathrm{Shadow}(x) \cap \mathrm{Shadow}(y). \tag{9.4}$$

Proof. (i) Immediate from the definition of an atomic semilattice. (ii), (iii) Immediate from (i).

Lemma 9.4. *Let P denote a prematroid, and let x, y denote incomparable elements in P such that $x \vee y$ exists. Then*

$$\mathrm{Shadow}(x \vee y) \backslash \mathrm{Shadow}(x \wedge y) = \bigcup \mathrm{Shadow}(a \vee b), \tag{9.5}$$

where the union is over all $a \in \mathrm{Shadow}(x) \backslash \mathrm{Shadow}(y)$ and all $b \in \mathrm{Shadow}(y) \backslash \mathrm{Shadow}(x)$.

Proof. $\supseteq$: Pick any $a \in \mathrm{Shadow}(x) \backslash \mathrm{Shadow}(y)$, any $b \in \mathrm{Shadow}(y) \backslash \mathrm{Shadow}(x)$, and any $c \in \mathrm{Shadow}(a \vee b)$. Observe $a \le x \le x \vee y$ and $b \le y \le x \vee y$, so $a \vee b \le x \vee y$. Now $c \le a \vee b \le x \vee y$, so $c \in \mathrm{Shadow}(x \vee y)$. Observe $c \notin \mathrm{Shadow}(x \wedge y)$; otherwise

$$b \le a \vee b = a \vee c \le x,$$

a contradiction.

$\subseteq$: Pick $c \in \mathrm{Shadow}(x \vee y) \backslash \mathrm{Shadow}(x \wedge y)$. We find $a \in \mathrm{Shadow}(x) \backslash \mathrm{Shadow}(y)$ and $b \in \mathrm{Shadow}(y) \backslash \mathrm{Shadow}(x)$ such that $c \in \mathrm{Shadow}(a \vee b)$. We may assume $c \not\le x$; otherwise we are done with $a := c$, and with b an arbitrary element in $\mathrm{Shadow}(y) \backslash \mathrm{Shadow}(x)$. Similarly, we may assume $c \not\le y$.

Observe $x \vee y$ is an upper bound for c, y, so $c \vee y$ exists. Set $y' := c \vee y$. Then y' covers y by modularity, and $y' \in [y, x \vee y]$ by the construction. Observe $x \wedge y' \ge x \wedge y$. In fact $x \wedge y' > x \wedge y$; otherwise y, y' are both relative complements of x in $[x \wedge y, x \vee y]$, contradicting

(7.14). Now $\text{Shadow}(x \wedge y')$ properly contains $\text{Shadow}(x \wedge y)$ by (9.3), so there exists an element $a \in \text{Shadow}(x \wedge y') \backslash \text{Shadow}(x \wedge y)$. Observe $a \in \text{Shadow}(x) \backslash \text{Shadow}(y)$. Observe y' is an upper bound for a, c, so $a \vee c$ exists. set $z := a \vee c$. Observe $a \neq c$ since $a \leq x$ and $c \not\leq x$, so z is a line by modularity. We mentioned y' is an upper bound for a, c, so $z = a \vee c \leq y'$. Now $y \vee z = y'$ covers y, so z covers $y \wedge z$ by modularity. In particular $y \wedge z$ is an atom. Set $b := y \wedge z$. Observe $b \in \text{Shadow}(y)$, so $a \neq b$. Observe z covers a, b, so $z = a \vee b$. Observe

$$\begin{aligned} c &\in \text{Shadow}(z) \\ &= \text{Shadow}(a \vee b). \end{aligned}$$

Observe $b \notin \text{Shadow}(x)$; otherwise $c \leq a \vee b \leq x$, a contradiction. this proves Lemma 9.4.

Definition 9.5. Let P denote a prematroid. By a *subspace* of P, we mean a subset $S \subseteq A_P$ such that for all lines $x \in P$,

$$|\text{Shadow}(x) \cap S| \geq 2 \quad \rightarrow \quad \text{Shadow}(x) \subseteq S. \tag{9.6}$$

Let P denote a prematroid. Our purpose for the rest of this section is to establish a 1–1 correspondence between the set $\mathcal{G}$ of all submatroids of P, and the set $\mathcal{S}$ of all subspaces of P. We proceed as follows. In Lemma 9.6, we find a map $G \rightarrow A_G$ from $\mathcal{G}$ to $\mathcal{S}$, and a map $S \rightarrow G_S$ from $\mathcal{S}$ to $\mathcal{G}$. In Theorem 9.7, we show these maps are inverses, establishing our 1–1 correspondence.

Lemma 9.6. *Let P denote a prematroid.*

(i) *Let G denote a submatroid of P. Then the set of atoms $A_G = A_P \cap G$ of G is a subspace of P.*
(ii) *Let S denote a subspace of P, and set*

$$G_S := \{\, x \in P \mid \text{Shadow}(x) \subseteq S\}. \tag{9.7}$$

Then G_S is a submatroid of P.

Proof. (i) Pick any line $x \in P$, and suppose there exists distinct points

$$y, z \in \text{Shadow}(x) \cap A_G.$$

We show

$$\text{Shadow}(x) \subseteq A_G. \tag{9.8}$$

To see (9.8), observe

$$x = y \vee z \in G$$

since G is $\vee$-closed, so now

$$\mathrm{Shadow}(x) \subseteq [0, x] \subseteq G$$

since G is convex, and now

$$\mathrm{Shadow}(x) \subseteq A_P \cap G = A_G,$$

as desired.

(ii) By Lemma 9.1, it suffices to show G_S is a nonempty, $\vee$-closed lower ideal in P. Observe $\mathrm{Shadow}(0) = \emptyset \subseteq S$, so $0 \in G_S$ by (9.7). In particular $G_S \neq \emptyset$. G_S is a lower ideal in P by the construction. To see that G_S is $\vee$-closed in P, we pick any $x, y \in G_S$ such that $x \vee y$ exists in P, and show $x \vee y \in G_S$. This will occur if $\mathrm{Shadow}(x \vee y) \subseteq S$, so we pick any $c \in \mathrm{Shadow}(x \vee y)$ and show $c \in S$. We may assume x, y are incomparable, and that $c \notin \mathrm{Shadow}(x \wedge y)$; otherwise the result is trivial. Now by Lemma 9.4, there exists $a \in \mathrm{Shadow}(x) \backslash \mathrm{Shadow}(y)$ and there exists $b \in \mathrm{Shadow}(y) \backslash \mathrm{Shadow}(x)$ such that $c \in \mathrm{Shadow}(a \vee b)$. Observe $a \vee b$ is a line, and

$$|\mathrm{Shadow}(a \vee b) \cap S| \geq |\{a, b\}| = 2,$$

so

$$\mathrm{Shadow}(a \vee b) \subseteq S$$

by (9.6). Now

$$c \in \mathrm{Shadow}(a \vee b) \subseteq S,$$

as desired. Now G_S is $\vee$-closed. We have now shown G_S is a nonempty, $\vee$-closed lower ideal in P, so G_S is a submatroid of P by Lemma 9.1.

Theorem 9.7. *Let P denote a prematroid. let $\mathcal{G}$ denote the set of all submatroids of P, and let $\mathcal{S}$ denote the set of all subspaces of P. Then the maps*

$$\begin{aligned} \mathcal{G} &\rightarrow \mathcal{S} \\ G &\rightarrow A_G \end{aligned}$$

and

$$\begin{aligned} \mathcal{S} &\rightarrow \mathcal{G} \\ S &\rightarrow G_S \end{aligned}$$

are inverses. In particular, They are both bijections.

Proof. First, let S denote a subspace of P, and write $G = G_S$. Then it is immediate from the construction that $A_G = S$. Secondly, let G denote any submatroid in P, and write $S = A_G$. We show $G = G_S$. To see $G \subseteq G_S$, pick any $x \in G$. Observe $\text{Shadow}(x) \subseteq S$ since G is a lower ideal in P, so $x \in G_S$. Hence $G \subseteq G_S$. To see $G \supseteq G_S$, pick any $x \in G_S$. Then $\text{Shadow}(x) \subseteq S$. Now

$$x = \bigvee_{a \in \text{Shadow}(x)} a \in G,$$

since G is $\vee$-closed by Lemma 9.1, and since $S \subseteq G$ by the construction. Hence $G \supseteq G_S$, so $G = G_S$. We have now established the given maps are inverses. This proves Theorem 9.7.

§10. Singular subspaces

Definition 10.1. Let P denote a prematroid. A subspace S of P is said to be *singular* whenever

$$(10.1) \qquad x \vee_P y \text{ exists for all } x, y \in S.$$

Lemma 10.2. *Let P denote a prematroid, and pick any $x \in P$. Then* $\text{Shadow}(x)$ *is a singular subspace of P.*

Proof. To show $\text{Shadow}(x)$ is a subspace of P, consider the interval $G = [0, x]$. Observe G is a submatroid of P by Lemma 8.4(ii), Lemma 9.1(ii), so A_G is a subspace of P by Lemma 9.6(i). Observe $\text{Shadow}(x) = A_G$ by construction, so $\text{Shadow}(x)$ is a subspace of P. It is clear that $\text{Shadow}(x)$ is singular. This proves Lemma 10.2.

Let P denote a prematroid, and let S denote a singular subspace of P. Must there exist an element $x \in P$ such that $\text{Shadow}(x) = S$? The answer is "no" in general, but "yes" in the following special case.

Theorem 10.3. *Let P denote a prematroid such that: for all $a \in A_P$, and all $u \in P$,*

$$(10.2) \qquad \text{if } a \vee_P b \text{ exists for all } b \in \text{Shadow}(u), \text{ then } a \vee_P u \text{ exists.}$$

Then for all singular subspaces S of P, there exists an element $x \in P$ such that

$$(10.3) \qquad \text{Shadow}(x) = S.$$

Proof. Let the singular subspace S be fixed, and let $G = G_S$ be the corresponding submatroid from (9.7). Pick any $x \in \max(G)$, and recall

$$\text{Shadow}(x) \subseteq S \tag{10.4}$$

by (9.7). We show equality holds in (10.4). Suppose not. Then there exists a point $a \in S \backslash \text{Shadow}(x)$. Since a, $\text{Shadow}(x)$ are contained in a common singular subspace, $a \vee_P b$ exists for all $b \in \text{Shadow}(x)$. It follows by (10.2) that $a \vee_P x$ exists. Observe $a \vee_P x \in G$, (since G is $\vee$-closed by Lemma 9.1(i)), and $a \vee_P x > x$ (since $a \not\leq x$), and we have contradicted the maximality of x in G. We conclude equality holds in (10.4), and the theorem is proved.

§11. More on the distance function

Lemma 11.1. *Let P denote a prematroid, and pick any x, x', $y, y' \in P$ such that both $xx'y$ and $xy'y$ are geodesic. Then the following* (i)–(iv) *are equivalent.*

(i) $xx'y'y$ *is geodesic.*
(ii) $x \wedge x' \geq x \wedge y'$ *and* $y \wedge y' \geq y \wedge x'$.
(iii) $x' \wedge y' \geq (x \wedge y') \vee (x' \wedge y)$.
(iv) $x' \wedge y' = (x \wedge y') \vee (x' \wedge y)$.

Proof. (i) $\to$ (ii). Applying Lemma 7.8(ii) to the geodesic sequence $xx'y'$, we find $x' \geq x \wedge y'$. Of course $x \geq x \wedge y'$, so $x \wedge x' \geq x \wedge y'$. Interchanging the roles of x, y, we find $y \wedge y' \geq y \wedge x'$.

(ii) $\to$ (iii). Observe $x' \geq x \wedge x' \geq x \wedge y'$ by (ii), and certainly $y' \geq x \wedge y'$, so $x' \wedge y' \geq x \wedge y'$. Interchanging the roles of x, y, we find $x' \wedge y' \geq x' \wedge y$, and line (iii) follows.

(iii) $\to$ (iv). The sequence x, $x \wedge y'$, y' is geodesic by Lemma 7.4, and we assume $xy'y$ is geodesic, so x, $x \wedge y'$, y', y is geodesic. Since

$$x \wedge y' \leq (x \wedge y') \vee (x' \wedge y) \leq x' \wedge y' \leq y',$$

the four elements in the above line form a geodesic sequence. Now

$$x,\ x \wedge y',\ (x \wedge y') \vee (x' \wedge y),\ x' \wedge y',\ y',\ y \qquad \text{is geodesic,}$$

so in particular,

$$x,\ (x \wedge y') \vee (x' \wedge y),\ x' \wedge y',\ y \qquad \text{is geodesic.} \tag{11.1}$$

Interchanging the roles of x, y, we find

$$(11.2) \qquad x,\ x'\wedge y',\ (x\wedge y')\vee(x'\wedge y),\ y \qquad \text{is geodesic,}$$

and (11.1), (11.2) imply (iv).

(iv) $\to$ (i). By Lemma 7.4 and our assumptions, we observe x, $x\wedge y'$, y', y is geodesic. Observe $x\wedge y' \leq x'\wedge y' \leq y'$, so x, $x\wedge y', x'\wedge y'$, y', y is geodesic. In particular x, $x'\wedge y'$, y', y is geodesic. Interchanging the roles of x, y, we find x, x', $x'\wedge y'$, y is geodesic. Combining the above information, we find

$$(11.3) \qquad x,\ x',\ x'\wedge y',\ y',\ y \qquad \text{is geodesic.}$$

In particular, $xx'y'y$ is geodesic, so (i) holds. We have now proved Lemma 11.1.

Corollary 11.2. *Let P denote a prematroid, and pick $x, x', y, y' \in P$ such that $x \leq x'$ and $y \leq y'$. Then the following are equivalent.*

(i) *$xx'y'y$ is geodesic.*

(ii) *$xx'y$ and $xy'y$ are both geodesic.*

Proof. (i) $\to$ (ii). Clear.

(ii) $\to$ (i). We assume $x \leq x'$, so

$$\begin{aligned} x\wedge x' &= x \\ &\geq x\wedge y'. \end{aligned} \qquad (11.4)$$

Interchanging the roles of x, y, we obtain

$$(11.5) \qquad y\wedge y' \geq y\wedge x'.$$

The condition in Lemma 11.1(ii) is now satisfied by (11.4), (11.5), so $xx'y'y$ is geodesic by that lemma.

§12. The function δ and the sets $x \star y$

Definition 12.1. Let P denote a prematroid. For all $x, y \in P$, define

$$(12.1) \qquad \delta(x,y) := \min\{\,\partial(x,z) \mid z \in P, \quad z \vee y \ \text{exists}\},$$

$$(12.2) \qquad x \star y := \{z \in P \mid \partial(x,z) = \delta(x,y), \quad z \vee y \ \text{exists}\}.$$

To get a feel for the above definition, we consider a very special case.

Lemma 12.2. *Let P denote a prematroid, and pick any $x, y \in P$. Then the following are equivalent.*

(i) $x \vee y$ *exists.*
(ii) $\delta(x, y) = 0$.
(iii) $x \star y = \{x\}$.
(iv) $\delta(y, x) = 0$.
(v) $y \star x = \{y\}$.

Proof. Immediate from Definition 12.1.

Let P denote a prematroid, and pick any $x, y \in P$. In general

$$\delta(x, y) \neq \delta(y, x). \tag{12.3}$$

It turns out (as we will show in Sections 18, 19) that δ is symmetric in its arguments precisely when P satisfies the augmentation axiom, and in this case $x \star y$ is a $[x \wedge y, x]$-basis system. For now, we will lay some groundwork with a general fact about $x \star y$, and an interpretation of δ.

Theorem 12.3. *Let P denote a prematroid, and pick any x, $y \in P$. Then*

$$x \star y \subseteq [x \wedge y,\, x]. \tag{12.4}$$

Proof. Pick any $z \in x \star y$. We first show $z \leq x$. Suppose $z \not\leq x$. We will obtain a contradiction by showing

$$(x \wedge z) \vee y \qquad \text{exists} \tag{12.5}$$

and

$$\partial(x, x \wedge z) \;<\; \partial(x, z). \tag{12.6}$$

To see (12.5), observe $x \wedge z \leq z \leq y \vee z$ and $y \leq y \vee z$, so $x \wedge z$, y have an upper bound. To see (12.6), recall $x, x \wedge z, z$ is geodesic by Lemma 7.4, and $x \wedge z \neq x$ by the construction. Line (12.6) follows. Now (12.5), (12.6) contradict Definition 12.1, so $z \leq x$. It remains to show $x \wedge y \leq z$. Suppose $x \wedge y \not\leq z$. We will obtain a contradiction by showing

$$\big((x \wedge y) \vee z\big) \vee y \qquad \text{exists} \tag{12.7}$$

and

$$\partial\big(x,\, (x \wedge y) \vee z\big) \;<\; \partial(x, z). \tag{12.8}$$

To see (12.7), observe $x \wedge y \leq y \leq y \vee z$ and $z \leq y \vee z$, so $(x \wedge y) \vee z \leq y \vee z$. Of course $y \leq y \vee z$, so $y \vee z$ is an upper bound for $(x \wedge y) \vee z$, y. We now have (12.7). To see (12.8), observe x is an upper bound for $x \wedge y$, z, so $(x \wedge y) \vee z \leq x$. Now $z \leq (x \wedge y) \vee z \leq x$, so $z, (x \wedge y) \vee z, x$ is geodesic. $z \neq (x \wedge y) \vee z$ by construction, and (12.8) follows. Now (12.7), (12.8) contradict Definition 12.1, so $x \wedge y \leq z$. We conclude $z \in [x \wedge y,\, x]$, and the theorem is proved.

We now establish an interpretation of δ. We begin with a technical lemma, and proceed to our main result Theorem 12.5.

Lemma 12.4. *Let P denote a prematroid, and pick any $x, y \in P$. Then for all $x' \in x^+$ and all $y' \in y^+$,*

(i) $\partial(x', x' \wedge y') \geq \delta(x, y)$,
(ii) $\partial(y', x' \wedge y') \geq \delta(y, x)$,
(iii) $\partial(x', y') \geq \delta(x, y) + \delta(y, x)$.

Proof. (i) Observe y' is an upper bound for $x \wedge y'$, y, so $(x \wedge y') \vee y$ exists. Now

$$\partial(x, x \wedge y') \geq \delta(x, y) \tag{12.9}$$

by Definition 12.1. Observe x' is an upper bound for x, $x' \wedge y'$, so $x \vee (x' \wedge y')$ exists. Clearly

$$x' \wedge y' \quad \leq \quad x \vee (x' \wedge y') \quad \leq \quad x',$$

so

$$\begin{aligned}
\partial(x', x' \wedge y') &\geq \partial(x \vee (x' \wedge y'),\, x' \wedge y') \\
&= \partial(x,\, x \wedge (x' \wedge y')) && \text{(modularity)} \\
&= \partial(x,\, x \wedge y') && \text{(since } x \leq x'\text{)} \\
&\geq \delta(x, y)
\end{aligned}$$

by (12.9).

(ii) Similar.

(iii) Recall x', $x' \wedge y'$, y' is geodesic by Lemma 7.4, so

$$\begin{aligned}
\partial(x', y') &= \partial(x', x' \wedge y') + \partial(x' \wedge y', y') \\
&\geq \delta(x, y) + \delta(y, x)
\end{aligned}$$

by (i), (ii). This proves Lemma 12.4.

Theorem 12.5. *Let P denote a prematroid. Then for all $x, y \in P$,*

$$\delta(x,y) + \delta(y,x) \quad = \quad \min\{\, \partial(x',y') \mid x' \in x^+, \;\; y' \in y^+\}. \tag{12.10}$$

Proof. In view of Lemma 12.4, it suffices to find some $x' \in x^+$ and some $y' \in y^+$ such that

$$\partial(x',y') \;=\; \delta(x,y) \;+\; \delta(y,x). \tag{12.11}$$

To this end, pick $u \in x \star y$ and $v \in y \star x$, and set

$$x' := x \vee v,$$
$$y' := y \vee u.$$

Clearly $x' \in x^+$, $y' \in y^+$, so it remains to check (12.11). To obtain it, we decompose $\partial(x,y)$ in two ways. First, observe both $xx'y$ and $xy'y$ are geodesic by Lemma 7.8(i), (iii). Now $xx'y'y$ is geodesic by Corollary 11.2, so

$$\partial(x,y) = \partial(x,x') \;+\; \partial(x',y') \;+\; \partial(y',y). \tag{12.12}$$

Second, observe $x, u, x \wedge y, v, y$ is geodesic by Lemma 7.4 and Theorem 12.3, so

$$\partial(x,y) = \partial(x,u) \;+\; \partial(u, x \wedge y) \;+\; \partial(x \wedge y, v) \;+\; \partial(v,y). \tag{12.13}$$

We now evaluate the terms on the right in (12.13). By Definition 12.1,

$$\partial(x,u) = \delta(x,y), \tag{12.14}$$

$$\partial(v,y) = \delta(y,x). \tag{12.15}$$

Observe u, y are relative complements in $[x \wedge y, y']$ by Lemma 7.8(ii),(iii), so by modularity

$$\partial(u, x \wedge y) = \partial(y', y). \tag{12.16}$$

Similarly,

$$\partial(x \wedge y, v) = \partial(x, x'). \tag{12.17}$$

Subtracting (12.12) from the sum of (12.13)–(12.17), we obtain (12.11), as desired. This proves the theorem.

§13. The functions ρ and γ

Definition 13.1. Let P denote a prematroid. For all $x, y \in P$, define

(i) $\rho(x,y) := \mathrm{rank}(x \wedge y)$,
(ii) $\gamma(x,y) := \partial(x \wedge y, z)$,

where z is any element of $x \star y$. (Observe $\gamma(x,y)$ is independent of the choice of z by Definition 12.1, Theorem 12.3.)

Let P denote a prematroid, and pick any $x, y \in P$. In this section, we consider the triple $\rho(x,y), \gamma(x,y), \delta(x,y)$. In our main result Theorem 13.5, we determine how this triple changes as we replace x by an element in P adjacent x.

Lemma 13.2. *Let P denote a prematroid. Then for all $x, y \in P$,*

$$\mathrm{rank}(x) \;=\; \rho(x,y) \;+\; \gamma(x,y) \;+\; \delta(x,y). \tag{13.1}$$

Proof. Pick any $z \in x \star y$. Observe $0,\ x \wedge y,\ z,\ x$ is geodesic by Theorem 12.3, so

$$\begin{aligned} \mathrm{rank}(x) &= \partial(0, x \wedge y) \;+\; \partial(x \wedge y, z) \;+\; \partial(z, x) \\ &= \rho(x,y) \;+\; \gamma(x,y) \;+\; \delta(x,y) \end{aligned}$$

by Lemma 7.1(ii) and Definitions 12.1, 13.1.

Lemma 13.3. *Let P denote a prematroid of rank D. Then for all $x, y \in P$,*

(i) *$\rho(x,y),\ \gamma(x,y),\ \delta(x,y)$ are nonnegative integers,*
(ii) *$\rho(x,y) \;+\; \gamma(x,y) \;+\; \gamma(y,x) \;+\; \delta(x,y) \;\le\; D$.*

Proof. (i) Immediate.

(ii) Pick any $v \in y \star x$. Then $x \vee v$ exists by Definition 12.1. Observe x, v are relative complements in the interval $[x \wedge y, x \vee v]$ by Lemma 7.8(ii),(iii) (with $u := x$), so

$$\begin{aligned} \mathrm{rank}(x \vee v) \;-\; \mathrm{rank}(x) &= \mathrm{rank}(v) \;-\; \mathrm{rank}(x \wedge y) \\ &= \gamma(y,x) \end{aligned} \tag{13.2}$$

by Lemma 7.1(ii). Now by (13.2), Lemma 13.2, and the construction,

$$\begin{aligned} D &\ge \mathrm{rank}(x \vee v) \\ &= \mathrm{rank}(x) \;+\; \gamma(y,x) \\ &= \rho(x,y) \;+\; \gamma(x,y) \;+\; \delta(x,y) \;+\; \gamma(y,x), \end{aligned}$$

as desired. This proves Lemma 13.3.

Before proceeding to the main theorem of this section, we mention a result about ρ.

Lemma 13.4. *Let P denote a prematroid. Then for all x, y, $z \in P$, the following are equivalent.*

(i) $z \in [x \wedge y, x]$.
(ii) $z \leq x$, *and* xzy *is geodesic.*
(iii) $z \leq x$, *and* $z \wedge y = x \wedge y$.
(iv) $z \leq x$, *and* $\rho(z, y) = \rho(x, y)$.

Proof. (i) $\to$ (ii). Clearly $z \leq x$. We may view $z = u \vee v$, where $u := z$, $v := x \wedge y$, so xzy is geodesic by Lemma 7.8 (i),(iii).

(ii) $\to$ (iii). By Lemma 7.8(i),(ii), $x \wedge z$, $z \wedge y$ are relative complements in the interval $[x \wedge y, z]$. But $x \wedge z = z$, so $z \wedge y = x \wedge y$.

(iii) $\to$ (i). Observe $x \wedge y = z \wedge y \leq z$.

(iii) $\to$ (iv). Immediate from Definition 13.1(i).

(iv) $\to$ (iii). Observe $z \wedge y \leq z \leq x$ and $z \wedge y \leq y$, so $z \wedge y \leq x \wedge y$. But $z \wedge y, x \wedge y$ have the same rank, so they are equal. This proves Lemma 13.4.

Theorem 13.5. *Let P denote a prematroid, and pick any x, y, $z \in P$ such that x, z are adjacent. Then*

(i)

$$\operatorname{rank}(x) - \operatorname{rank}(z) = \Delta\rho + \Delta\gamma + \Delta\delta, \tag{13.3}$$

(ii)

$$\partial(x, y) - \partial(z, y) = -\Delta\rho + \Delta\gamma + \Delta\delta, \tag{13.4}$$

(iii)

$$|\Delta\rho| + |\Delta\gamma| + |\Delta\delta| = 1, \tag{13.5}$$

where

$$\Delta\rho = \rho(x, y) - \rho(z, y), \tag{13.6}$$

$$\Delta\gamma = \gamma(x, y) - \gamma(z, y), \tag{13.7}$$

$$\Delta\delta = \delta(x, y) - \delta(z, y). \tag{13.8}$$

Proof. (i) Immediate from Lemma 13.2.

(ii) By Corollary 7.6(i),

$$\partial(x,y) = \operatorname{rank}(x) + \operatorname{rank}(y) - 2\operatorname{rank}(x \wedge y), \tag{13.9}$$

$$\partial(z,y) = \operatorname{rank}(z) + \operatorname{rank}(y) - 2\operatorname{rank}(z \wedge y). \tag{13.10}$$

Subtracting (13.10) from (13.9), and evaluating $\operatorname{rank}(x)-\operatorname{rank}(z)$ using (13.3), we obtain (13.4).

(iii) Since $\delta(x,y)$, $\delta(z,y)$ measure the distance from x, z to the same set, and since x, z are adjacent, we have

$$\Delta\delta \in \{-1,\ 0,\ 1\} \tag{13.11}$$

by the triangle inequality. First assume $\Delta\delta \neq 0$. Interchanging x, z if necessary, we may assume that

$$\Delta\delta = 1. \tag{13.12}$$

We show

$$\Delta\rho = 0, \qquad \Delta\gamma = 0. \tag{13.13}$$

This will follow by Definition 13.1 if we can show

$$z \star y \subseteq x \star y, \tag{13.14}$$

$$z \wedge y = x \wedge y. \tag{13.15}$$

To see (13.14), pick any $w \in z \star y$. Then $w \vee y$ exists, so

$$\partial(x,w) \geq \delta(x,y). \tag{13.16}$$

Also $\partial(z,w) = \delta(z,y)$, so

$$\begin{aligned} \partial(x,w) &\leq \partial(x,z) + \partial(z,w) && (13.17)\\ &= 1 + \delta(z,y) \\ &= \delta(x,y) && (13.18) \end{aligned}$$

by (13.12). We must now have equality in (13.16)–(13.18). In particular $w \in x \star y$ by the construction. We now have (13.14). To see (13.15), pick any $w \in z \star y$. Then $w \in x \star y$ by (13.14), so $w \in [x \wedge y, x]$ by Theorem 12.3. Also xzw is geodesic by (13.17), so $z \in [x \wedge y, x]$ by Lemma 8.4. Now $z \wedge x = x \wedge y$ by Lemma 13.4(i),(iii), so (13.15) holds. Line (13.13) follows from (13.14), (13.15), and (13.5) follows from (13.12), (13.13). Next assume $\Delta\delta = 0$. By (7.3), the pair

$$(\operatorname{rank}(x) - \operatorname{rank}(z),\ \partial(x,y) - \partial(z,y))$$

is one of $(1,1)$, $(1,-1)$, $(-1,1)$, $(-1,-1)$. In each of these four cases, one readily solves (13.3), (13.4) for $\Delta\rho$, $\Delta\gamma$, and finds (13.5) holds in each case. We have now proved (iii), and the theorem.

Corollary 13.6. *Let P denote a prematroid, and pick $x, y, z \in P$ such that $z \geq x$. Then*

(i) $\rho(x,y) \leq \rho(z,y)$,
(ii) $\gamma(x,y) \leq \gamma(z,y)$,
(iii) $\delta(x,y) \leq \delta(z,y)$,
(iv) $\delta(y,x) \leq \delta(y,z)$,
(v) $\gamma(y,x) \geq \gamma(y,z)$.

Proof. (i)–(iii) It suffices to assume z covers x. By Theorem 13.5(i),(iii), exactly one of $\rho(x,y)-\rho(z,y)$, $\gamma(x,y)-\gamma(z,y)$, $\delta(x,y)-\delta(z,y)$ equals -1, and the other two equal 0. The inequalities follow.

(iv) Pick any $w \in y \star z$. Then by Definition 12.1,

$$\partial(y,w) = \delta(y,z) \tag{13.19}$$

and $z \vee w$ exists. Observe $x \vee w$ exists, since $x \leq z \leq z \vee w$ and $w \leq z \vee w$, so

$$\partial(y,w) \geq \delta(y,x) \tag{13.20}$$

by Definition 12.1. Our result follows from (13.19), (13.20).

(v) By Lemma 13.2, parts (i), (iv) above, and since ρ is symmetric in its arguments,

$$\begin{aligned}\gamma(y,x) &= \operatorname{rank}(y) - \rho(y,x) - \delta(y,x)\\ &\geq \operatorname{rank}(y) - \rho(y,z) - \delta(y,z)\\ &= \gamma(y,z),\end{aligned}$$

as desired. This proves Corollary 13.6.

Lemma 13.7. *Let P denote a prematroid, and let G denote a geodesically closed subposet of P. Then for all x, $y \in G$,*

(i) $x \star_G y = x \star_P y$,
(ii) $\delta_G(x,y) = \delta_P(x,y)$,
(iii) $\gamma_G(x,y) = \gamma_P(x,y)$,
(iv) $\rho_G(x,y) = \rho_P(x,y) - \operatorname{rank}(0_G)$.

Proof. (i), (ii) Immediate from Lemma 8.2 and Definition 12.1.
(iii) Immediate from Lemma 5.4, Definition 13.1(ii).
(iv) Immediate from Lemma 5.6, Definition 13.1(i).

§14. The posets $[x \wedge y, x \star y]$, $[x \star y, x]^*$

For notational convenience, we expand our notion of an interval in a prematroid.

Definition 14.1. Let P denote a prematroid. For any nonempty subsets $H \subseteq P$, $K \subseteq P$, define the subposet

$$[H, K] := \{ z \in P \mid \exists x \in H, \ \exists y \in K \quad \text{such that} \quad xzy \quad \text{is geodesic}\}.$$

Let P denote a prematroid, and pick any $x, y \in P$. In this section, we consider the posets $[x \wedge y, x \star y]$, $[x \star y, x]^*$. (Recall from (3.1), the $*$ means we *reverse* the usual partial order. As we mentioned in Section 12, it will turn out that if P satisfies the augmentation axiom, then $x \star y$ is a $[x \wedge y, x]$-basis system. In this case $[x \wedge y, x \star y]$ becomes a $[x \wedge y, x]$-matroid, $[x \star y, x]^*$ becomes a $[x \wedge y, x]^*$-matroid, and these matroids are duals in the sense of Definition 3.2. In this section, we establish a few facts about these posets.

Lemma 14.2. *Let P denote a prematroid, and pick any $x, y \in P$.*

(i) *The zero of $[x \wedge y, x \star y]$ is $x \wedge y$.*
(ii) $\text{top}([x \wedge y, x \star y]) = x \star y$.
(iii) $\text{rank}([x \wedge y, x \star y]) = \gamma(x, y)$.
(iv) *The zero of $[x \star y, x]^*$ is x.*
(v) $\text{top}([x \star y, x]^*) = x \star y$.
(vi) $\text{rank}([x \star y, x]^*) = \delta(x, y)$.

Proof. Immediate from Definitions 12.1, 13.1, 14.1, and Theorem 12.3.

Lemma 14.3. *Let P denote a prematroid, and fix any $x, y \in P$. Then for all $z \in P$, the following are equivalent.*

(i) $z \in [x \star y, x]^*$.
(ii) $z \leq x$, *and*

$$\delta(x, y) - \delta(z, y) = \text{rank}(x) - \text{rank}(z). \tag{14.1}$$

(iii) $z \leq x$, *and* $z \star y \subseteq x \star y$.
(iv) $z \leq x$, *and*

$$\rho(x, y) = \rho(z, y), \qquad \gamma(x, y) = \gamma(z, y). \tag{14.2}$$

Proof. (i) $\to$ (ii). Observe $z \le x$ by Theorem 12.3. By assumption, there exists $u \in x \star y$ such that $u \le z \le x$. Now

$$\begin{aligned}\operatorname{rank}(x) - \operatorname{rank}(z) &= \partial(x,z)\\ &= \partial(x,u) - \partial(z,u)\\ &= \delta(x,y) - \delta(z,y)\end{aligned}$$

since $u \in z \star y$.

(ii) $\to$ (iii). Pick any $w \in z \star y$. We show $w \in x \star y$. Observe $w \vee y$ exists by assumption, so it suffices to show $\partial(x,w) = \delta(x,y)$. Observe $w \le z \le x$ by Theorem 12.3 and the construction, so

$$\begin{aligned}\partial(x,w) &= \partial(x,z) + \partial(z,w)\\ &= \operatorname{rank}(x) - \operatorname{rank}(z) + \delta(z,y)\\ &= \delta(x,y),\end{aligned}$$

as desired.

(iii) $\to$ (i). Pick any $w \in z \star y$. Then $w \le z \le x$ by Theorem 12.3, and $w \in x \star y$ by assumption, so $z \in [x \star y, x]^*$.

(ii) $\leftrightarrow$ (iv). The three scalars $\rho(x,y) - \rho(z,y)$, $\gamma(x,y) - \gamma(z,y)$, $\delta(x,y) - \delta(z,y)$ are non-negative by Corollary 13.6, and sum to $\operatorname{rank}(x)$ $-\operatorname{rank}(z)$ by Theorem 13.5(i). The result is now immediate. This proves Lemma 14.3.

§15. The posets x_y^+, x_y^-

Let P denote a prematroid. For all x, $y \in P$, define the subposets $x_y^+, x_y^- \subseteq P$ by

(15.1) $$x_y^+ := \{ z \in P \mid x \le z, \quad xzy \text{ is geodesic}\},$$

(15.2) $$x_y^- := \{ z \in [x \wedge y, x] \mid z \vee y \text{ exists in } P\}.$$

Observe x is the zero of x_y^+, and $x \wedge y$ is the zero of x_y^-. In particular x_y^+, x_y^- are not empty.

Example 15.1. Let P denote a prematroid, and pick any $x, y \in P$ such that $x \vee y$ exists. Then

(i) $x_y^+ = [x, x \vee y]$,

(ii) $x_y^- = [x \wedge y, x]$.

Proof. (i) First consider the inclusion $\subseteq$. Pick any $z \in x_y^+$. Then xzy is geodesic, so $z \leq x \vee y$ by Lemma 7.9(i),(ii). Also $x \leq z$, so $z \in [x, x \vee y]$. Now consider the inclusion $\supseteq$. Pick any $z \in [x, x \vee y]$. Then $x, z, x \vee y$ is geodesic by (7.5). Also $x, x \vee y, y$ is geodesic by Lemma 7.5, so $x, z, x \vee y, y$ is geodesic. In particular, xzy is geodesic, so $z \in x_y^+$.

(ii) Immediate from (15.2).

Lemma 15.2. *Let P denote a prematroid, and pick any $x, y \in P$.*

(i) *x_y^+ is a submatroid of x^+.*
(ii) *x_y^+ is geodesically closed in P.*

Proof. (i) First, we claim x_y^+ is a lower ideal in x^+. Pick any $z \in x_y^+$ and any $w \in x^+$ such that $w \leq z$. We show $w \in x_y^+$. Of course $x \leq w$, so it remains to show xwy is geodesic. to this end, observe xwz is geodesic, since $x \leq w \leq z$, and xzy is geodesic by assumption, so $xwzy$ is geodesic. In particular xwy is geodesic, as desired. Hence x_y^+ is a lower ideal in x^+.

Next, we claim x_y^+ is $\vee$-closed in P. To see this, pick any $z, z' \in x_y^+$ such that $z \vee z'$ exists in P. Recall xzy, $xz'y$ are both geodesic, so $x, z \vee z', y$ is geodesic by Lemma 7.10. Recall $x \leq z$, $x \leq z'$, so $x \leq z \vee z'$. Now $z \vee z' \in x_y^+$, so x_y^+ is $\vee$-closed in P. Now x_y^+ is a nonempty $\vee$-closed lower ideal in x^+, so x_y^+ is a submatroid of x^+ by Lemma 9.1.

(ii) Immediate from (i) above and Lemma 9.2.

Lemma 15.3. *Let P denote a prematroid, and fix any x, $y \in P$. Then for all $z \in P$, the following are equivalent.*

(i) *$z \in x_y^+$.*
(ii) *$x \leq z$, and x, $y \wedge z$ are relative complements in the interval $[x \wedge y, z]$.*
(iii) *There exists an element $v \in [x \wedge y, y]$ such that $z = x \vee v$.*
(iv) *$x \leq z$, and*

$$\rho(z, y) - \rho(x, y) = \operatorname{rank}(z) - \operatorname{rank}(x). \tag{15.3}$$

(v) *$x \leq z$, and*

$$\gamma(x, y) = \gamma(z, y), \qquad \delta(x, y) = \delta(z, y). \tag{15.4}$$

Moreover, if (i)–(v) *hold, then*

$$v = y \wedge z. \tag{15.5}$$

Proof. (i) $\leftrightarrow$ (ii) $\leftrightarrow$ (iii). This is a special case of Lemma 7.8.

(ii) $\rightarrow$ (iv). Immediate from (7.14) and Definition 13.1(i).

(iv) $\rightarrow$ (i). By (7.12) and the observation $x \wedge z = x$, we find $\partial(x,z) + \partial(z,y) - \partial(x,y)$ equals twice

$$\mathrm{rank}(z) \ - \ \mathrm{rank}(x) \ + \ \rho(x,y) - \rho(z,y),$$

and is therefore 0 by (15.3). Now xzy is geodesic by (7.4), and (i) follows.

(iv) $\leftrightarrow$ (v). The scalars $\rho(z,y) - \rho(x,y)$, $\gamma(z,y) - \gamma(x,y)$, $\delta(z,y) - \delta(x,y)$ are nonnegative by Corollary 13.6, and sum to $\mathrm{rank}(z) - \mathrm{rank}(x)$ by Theorem 13.5(i). The result is now immediate.

Now suppose (i)–(v) hold. Then (15.5) holds by (7.19). This proves Lemma 15.3.

We now turn to x_y^-.

Lemma 15.4. *Let* P *denote a prematroid, and pick any* $x, y \in P$. *Then*

(i) x_y^- *is a lower ideal in the interval* $[x \wedge y, x]$,
(ii) $\mathrm{top}(x_y^-) = x \star y$,
(iii) $[x \wedge y, x \star y] \subseteq x_y^-$.

Proof. (i) Suppose we are given some $z \in x_y^-$ and some $z' \in [x \wedge y, x]$ such that $z' \leq z$. We show $z' \in x_y^-$. To do this, we must show $y \vee z'$ exists in P. But this is the case, since $y \vee z \geq y$ and $y \vee z \geq z \geq z'$.

(ii) x_y^- is a lower ideal in $[x \wedge y, x]$ by (i), so we may regard $\mathrm{top}(x_y^-)$ as the set of elements $z \in x_y^-$ with $\partial(x,z)$ minimal. Recall $x \star y$ consists of the elements $z \in P$ such that $z \vee y$ exists, and such that $\partial(x,z)$ is minimal subject to this existence. Observe $x \star y \subseteq x_y^-$ by Theorem 12.3 and the construction, and $w \vee y$ exists for all $w \in x_y^-$, so we may regard $x \star y$ as the set of elements $z \in x_y^-$ with $\partial(x,z)$ minimal. Our result follows.

(iii) Immediate from (i), (ii).

Our next goal is to show the posets x_y^+, y_x^- are isomorphic.

Theorem 15.5. *Let P denote a prematroid, and pick any $x, y \in P$. Then there exists poset isomorphisms $\sigma : x_y^+ \to y_x^-$, $\varepsilon : y_x^- \to x_y^+$ such that*

$$(15.6) \qquad \sigma(z) = y \wedge z \qquad (\forall z \in x_y^+),$$

$$(15.7) \qquad \varepsilon(v) = x \vee v \qquad (\forall v \in y_x^-).$$

Moreover, σ, ε are inverses.

Proof. Pick any $z \in x_y^+$. By Lemma 15.3(i),(iii), and (15.5),

$$(15.8) \qquad y \wedge z \in y_x^-$$

and

$$(15.9) \qquad x \vee (y \wedge z) = z.$$

Pick any $v \in y_x^-$. Then by Lemma 15.3(i),(iii) and (15.5),

$$(15.10) \qquad x \vee v \in x_y^+$$

and

$$(15.11) \qquad y \wedge (x \vee v) = v.$$

By (15.8), there exists a map $\sigma : x_y^+ \to y_x^-$ satisfying (15.6). By (15.10), there exists a map $\varepsilon : y_x^- \to x_y^+$ satisfying (15.7). Observe σ, ε are inverses by (15.9), (15.11); in particular, these maps are bijections. It remains to check σ, ε respect the partial order. But this follows, since for all $z, z' \in x_y^+$,

$$z \leq z' \quad \to \quad y \wedge z \leq y \wedge z',$$

and for all $v, v' \in y_x^-$,

$$v \leq v' \quad \to \quad x \vee v \leq x \vee v'.$$

This proves Theorem 15.5.

Corollary 15.6. *Let P denote a prematroid. Then for all $x, y \in P$, the subposet x_y^+ is embeddable.*

Proof. It is clear y_x^- is embeddable. Indeed the identity map is an embedding of y_x^- into the modular atomic lattice $[x \wedge y, y]$. Now x_y^+ is embeddable, since x_y^+, y_x^- are isomorphic by Theorem 15.5.

Corollary 15.7. *Let P denote a prematroid. Then for all $x, y \in P$,*

(i) $\mathrm{rank}(x_y^-) = \gamma(x, y)$,
(ii) $\mathrm{rank}(x_y^+) = \gamma(y, x)$,
(iii) $|\mathrm{top}(x_y^+)| = |y \star x|$.

Proof. (i) Immediate from Lemma 14.2(iii) and Lemma 15.4(i),(ii).

(ii) Immediate from (i) above and the fact that x_y^+, y_x^- are isomorphic.

(iii) Immediate from Lemma 15.4(ii) and the fact that x_y^+, y_x^- are isomorphic.

This proves Corollary 15.7.

We finish this section with two technical results.

Lemma 15.8. *Let P denote a prematroid. Pick any $x, y \in P$ and pick any $z \in x_y^+$. Then*

(i) $z_y^+ = z^+ \cap x_y^+$,
(ii) $y_z^- = [y \wedge z, y] \cap y_x^-$,
(iii) $y_z^+ = y_x^+$.

Proof. (i) Recall $x \leq z$ and xzy is geodesic. To see the inclusion $\subseteq$, pick any $w \in z_y^+$. Then $z \leq w$, so $w \in z^+$. We show $w \in x_y^+$. Certainly $x \leq z \leq w$. Observe xzy and zwy are geodesic, so $xzwy$ is geodesic. In particular xwy is geodesic, so $w \in x_y^+$. To see the inclusion $\supseteq$, pick any $w \in z^+ \cap x_y^+$. Observe $x \leq z \leq w$ so xzw is geodesic. Also xwy is geodesic, so $xzwy$ is geodesic. In particular zwy is geodesic, so $w \in z_y^+$.

(ii) By Theorem 15.5, the map $\sigma : w \to y \wedge w$ induces an isomorphism of posets $x_y^+ \to y_x^-$ such that $\sigma(z_y^+) = y_z^-$. Now by (i) above,

$$\begin{aligned} y_z^- &= \sigma(z_y^+) \\ &= \sigma(z^+ \cap x_y^+) \\ &= (y \wedge z)^+ \cap y_x^- \\ &= [y \wedge z, y] \cap y_x^-, \end{aligned}$$

as desired.

(iii) This is immediate from Corollary 11.2. This proves Lemma 15.8.

Lemma 15.9. *Let P denote a premotroid, and pick any $x, y \in P$. Then for all $u \in y^+$,*

$$x_y^+ \text{ is a submatroid of } x_u^+. \tag{15.12}$$

Proof. By Lemma 15.2(i), it suffices to show $x_y^+ \subseteq x_u^+$, and without loss, we may assume u covers y. By (7.3), there are two possibilities;

$$\partial(x,u) = \partial(x,y) - 1 \tag{15.13}$$

or

$$\partial(x,u) = \partial(x,y) + 1. \tag{15.14}$$

First suppose (15.13). Then xuy is geodesic, so $u \in y_x^+$. Now $x_y^+ = x_u^+$ by Lemma 15.8(iii). Next assume (15.14). To see $x_y^+ \subseteq x_u^+$ in this case, we pick any $w \in x_y^+$ and show $w \in x_u^+$. Observe xwy is geodesic by assumption, and xyu is geodesic by (15.14), so $xwyu$ is geodesic. In particular xwu is geodesic, so $w \in x_u^+$, as desired. This proves Lemma 15.9.

§16. Projection into a submatroid

Lemma 16.1. *Let P denote a prematroid, and let G, H denote submatroids of P. Then $G \cap H$ is a submatroid of P.*

Proof. $G \cap H$ is geodesically closed in P, and contains 0_P, so we are done by Lemma 9.1.

Lemma 16.2. *Let P denote a prematroid. Let G denote a submatroid of P, and pick any $x \in P$. Then there exists a unique element $p = p(x,G)$ in G such that*

$$[0,x] \cap G = [0,p]. \tag{16.1}$$

We call p the projection of x into G, and write

$$p = \mathrm{proj}_G x. \tag{16.2}$$

Proof. $[0,x]$ is a submatroid of P by Lemma 8.4(ii), Lemma 9.1, so $[0,x] \cap G$ is a submatroid of P by Lemma 16.1. $[0,x] \cap G$ is contained in the interval $[0, x]$, and is therefore an interval by Lemma 8.5.

Lemma 16.3. *Let P denote a prematroid. Let G denote any submatroid of P, and pick any $x, y \in P$.*

(i) $\mathrm{proj}_G x \leq x$.
(ii) *Equality holds in* (i) *if and only if* $x \in G$.
(iii) $\mathrm{proj}_G(x \wedge y) = x \wedge \mathrm{proj}_G y$.
(iv) $\mathrm{proj}_G(x \wedge y) = \mathrm{proj}_G x \wedge \mathrm{proj}_G y$.

Proof. (i), (ii) Immediate from Lemma 16.2.
(iii) By (16.1), (16.2),

$$\begin{aligned}[0, \mathrm{proj}_G(x \wedge y)] &= [0, x \wedge y] \cap G \\ &= [0, x] \cap [0, y] \cap G \\ &= [0, x] \cap [0, \mathrm{proj}_G y],\end{aligned}$$

and our result follows.
(iv) Interchanging the roles of x, y in (iii),

$$\mathrm{proj}_G(x \wedge y) = y \wedge \mathrm{proj}_G x.$$

By this and (iii), we may view

$$\begin{aligned}\mathrm{proj}_G(x \wedge y) &= \mathrm{proj}_G(x \wedge y) \wedge \mathrm{proj}_G(x \wedge y) \\ &= x \wedge \mathrm{proj}_G x \wedge y \wedge \mathrm{proj}_G y \\ &= \mathrm{proj}_G x \wedge \mathrm{proj}_G y,\end{aligned}$$

since $\mathrm{proj}_G x \leq x$, $\mathrm{proj}_G y \leq y$. This proves Lemma 16.3.

Theorem 16.4. *Let P denote a prematroid, and let G denote a submatroid of P. Then for all $x \in P$, and for all $y \in G$, the sequence*

$$x, \mathrm{proj}_G x, y \quad \textit{is geodesic.} \tag{16.3}$$

Proof. G is a lower ideal in P that contains y, and $x \wedge y \leq y$, so $x \wedge y \in G$. Certainly $x \wedge y \leq x$, so

$$x \wedge y \in [0, x] \cap G = [0, \mathrm{proj}_G x],$$

forcing $x \wedge y \leq \mathrm{proj}_G x$. Now $x \wedge y \leq \mathrm{proj}_G x \leq x$ by the construction, so $x, \mathrm{proj}_G x, y$ is geodesic by Lemma 13.4(i),(ii). This proves Theorem 16.4.

§17. The projection $x^+ \to x_y^+$

Let P denote a prematroid. Pick any $x, y \in P$, and write $G = x_y^+$. Observe x^+ is a prematroid by Corollary 8.3, Lemma 8.4, and G is a submatroid of x^+ by Lemma 15.2(i), so there exists a projection map $\mathrm{proj}_G : x^+ \to G$ by Lemma 16.2.

Lemma 17.1. *Let P denote a prematroid. Pick any $x, y \in P$, and write $G = x_y^+$. Then for all $z \in x^+$,*

$$\mathrm{proj}_G z = x \vee (y \wedge z). \tag{17.1}$$

Proof. We first show

$$\mathrm{proj}_G z \geq x \vee (y \wedge z). \tag{17.2}$$

Observe $x \leq z$, so $x \wedge y \leq y \wedge z$. Also $y \wedge z \leq y$, so $y \wedge z \in [x \wedge y, y]$. Observe z is an upper bound for x, $y \wedge z$, so $x \vee (y \wedge z)$ exists. Now $y \wedge z \in y_x^-$ by (15.2), so

$$x \vee (y \wedge z) \in G \tag{17.3}$$

by Theorem 15.5. Observe

$$x \vee (y \wedge z) \leq z \tag{17.4}$$

by our remarks above, and (17.2) follows from (17.3), (17.4), and Lemma 16.2. Next, we show

$$\mathrm{proj}_G z \leq x \vee (y \wedge z). \tag{17.5}$$

Write $p := \mathrm{proj}_G z$. Then $p \leq z$ by Lemma 16.3(i), so $y \wedge p \leq y \wedge z$. Also $p \in G = x_y^+$ by Lemma 16.2, so by Theorem 15.5,

$$p = x \vee (y \wedge p) \leq x \vee (y \wedge z).$$

Line (17.5) follows, and we are done by (17.2), (17.5).

Theorem 17.2. *Let P denote a prematroid. Pick any $x, y \in P$, and write $G = x_y^+$. Then for all $z \in x^+$, and for all $p \in P$, the following are equivalent.*

(i) $p = \mathrm{proj}_G z$.
(ii) $p \in [x, z]$, *and* wpy *is geodesic for all* $w \in [x, z]$.
(iii) $p \in [x, z]$, *and both* xpy *and* zpy *are geodesic.*
(iv) $p \in [x, z]$, *and*

$$\rho(p, y) = \rho(z, y), \qquad \gamma(p, y) = \gamma(x, y), \qquad \delta(p, y) = \delta(x, y). \tag{17.6}$$

Proof. (i) $\to$ (ii). Observe $p \in [x, z]$ by Lemma 16.3(i). Pick any $w \in [x, z]$. We show wpy is geodesic. By Lemma 7.8(i),(iii), it suffices to find

$$u \in [y \wedge w, w], \tag{17.7}$$

$$v \in [y \wedge w, y], \tag{17.8}$$

such that

$$p = u \vee v. \tag{17.9}$$

Set

$$u := \mathrm{proj}_G w.$$

Then (17.7) holds, since $u \leq w$ by Lemma 16.3(i), and since

$$\begin{aligned} u &= x \vee (y \wedge w) \\ &\geq y \wedge w \end{aligned} \tag{17.10}$$

by Lemma 17.1. Set

$$v := y \wedge z. \tag{17.11}$$

Then $v \leq y$ by construction. Recall $w \leq z$, so

$$y \wedge w \leq v \tag{17.12}$$

by (17.11). Now (17.8) holds. Now

$$\begin{aligned} p &= x \vee (y \wedge z) && \text{(Lemma 17.1)} \\ &= x \vee v && (17.11) \\ &= x \vee (y \wedge w) \vee v && (17.12) \\ &= u \vee v, && (17.10) \end{aligned}$$

so (17.9) holds. Now wpy is geodesic by (17.7)–(17.9), and we are done.

(ii) $\to$ (iii). Clear.

(iii) $\to$ (i). We assume $x \leq p$ and xpy is geodesic, so $p \in x_y^+$ by (15.1). Now

$$p = x \vee (y \wedge p) \tag{17.13}$$

by Theorem 15.5. We assume $p \leq z$ and zpy is geodesic, so

$$y \wedge p = y \wedge z \tag{17.14}$$

by Lemma 13.4(ii),(iii). Now

$$p = x \vee (y \wedge z) \tag{17.15}$$

by (17.13), (17.14), so $p = \mathrm{proj}_G z$ by Lemma 17.1.

(iii) $\leftrightarrow$ (iv). We assume $p \leq z$, so by Lemma 13.4(ii),(iv), zpy is geodesic if and only if

$$\rho(p, y) = \rho(z, y).$$

We assume $x \leq p$, so by Lemma 15.3(i),(v), xpy is geodesic if and only if

$$\gamma(p, y) = \gamma(x, y), \qquad \delta(p, y) = \delta(x, y).$$

This proves Theorem 17.2.

Theorem 17.3. *With the notation of Theorem* 17.2, *suppose* (i)–(iv) *hold. Then for all* $w \in [x, z]$, *the following are equivalent.*

(i) wxy *and* wzy *are both geodesic.*
(ii) wxp *and* wzp *are both geodesic.*
(iii) w, p *are relative complements in the interval* $[x, z]$.
(iv) $\rho(w, y) = \rho(x, y), \qquad \gamma(w, y) = \gamma(z, y), \qquad \delta(w, y) = \delta(z, y).$

Proof. (i) $\to$ (ii). Observe xpy is geodesic by Theorem 17.2(iii), and wxy is geodesic, so $wxpy$ is geodesic. In particular, wxp is geodesic. Similarly, zpy is geodesic by Theorem 17.2(iii), and wzy is geodesic, so $wzpy$ is geodesic. In particular, wzp is geodesic.

(ii) $\to$ (i). Recall wpy is geodesic by Theorem 17.2(ii), and wxp is geodesic, so $wxpy$ is geodesic. In particular wxy is geodesic. Similarly wzp is geodesic, so $wzpy$ is geodesic. In particular wzy is geodesic.

(ii) $\leftrightarrow$ (iii). We assume $x \leq w$, $x \leq p$, so by Lemma 7.4, $x = w \wedge p$ if and only if wxp is geodesic. We assume $w \leq z$, $p \leq z$, so by Lemma 7.5, $z = w \vee p$ if and only if wzp is geodesic.

(i) $\leftrightarrow$ (iv). We assume $x \leq w$, so by Lemma 13.4(ii),(iv), wxy is geodesic if and only if

$$\rho(w, y) = \rho(x, y).$$

We assume $w \leq z$, so by Lemma 15.3(i),(v), wzy is geodesic if and only if both

$$\gamma(w, y) = \gamma(z, y), \qquad \delta(w, y) = \delta(z, y).$$

This proves Theorem 17.3.

§18. The augmentation axiom

Let P denote a prematroid, and recall by Definition 4.1 that P is a quantum matroid if and only if P satisfies the augmentation axiom AU from that definition. In this section, we show this occurs if and only if the function δ from Definition 12.1 is symmetric in its arguments.

We begin with some notation.

Lemma 18.1. *Let P denote a prematroid. Pick any nonnegative integer d, and pick any path $p = (x_0, x_1, \ldots, x_d)$ $(x_0, x_1, \ldots, x_d \in P)$. Then the following* (i), (ii) *are equivalent.*

(i) *There does not exist an integer i $(0 \leq i \leq d-3)$ such that*

$$x_i > x_{i+1} < x_{i+2} < x_{i+3}, \tag{18.1}$$

or such that

$$x_i > x_{i+1} > x_{i+2} < x_{i+3}. \tag{18.2}$$

(ii) *There exists integers e, f $(0 \leq e \leq f \leq d,\ \ f-e$ is even), such that*

$$x_0 < x_1 < \cdots < x_{e-1} < x_e,$$
$$x_e > x_{e+1} < x_{e+2} > x_{e+3} < \cdots > x_{f-3} < x_{f-2} > x_{f-1} < x_f,$$
$$x_f > x_{f+1} > \cdots > x_{d-1} > x_d.$$

Suppose (i), (ii) *hold. Then we say p is up-flat-down. If $e = 0$ and $f = d$, we say p is flat.*

Proof. Routine.

Theorem 18.2. *Let P denote a prematroid of rank D. Then the following* (i)–(vi) *are equivalent.*

(i) *P satisfies the augmentation axiom AU in Definition* 4.1.
(ii) *For all integers i $(2 \leq i \leq D)$, and for all $x, y \in P$ such that* $\mathrm{rank}(x) = i-1$, $\mathrm{rank}(y) = i$, *and* $\partial(x, y) = 3$, *there exists a path in P with endpoints x, y and shape $(i-1, i, i-1, i)$.*
(iii) *For all integers i $(2 \leq i \leq D)$, and all geodesic paths p in P of shape $(i-1, i-2, i-1, i)$, there exists a path p' in P that has the same endpoints as p, and has shape $(i-1, i, i-1, i)$.*
(iv) *For all $x, y \in P$, there exists a geodesic up-flat-down path connecting x, y.*
(v) *For all $x, y \in P$, for all $x' \in \mathrm{top}(x_y^+)$, and for all $y' \in \mathrm{top}(y_x^+)$, there exists a geodesic flat path connecting x', y'.*
(vi) $\delta(x, y) = \delta(y, x)$ *for all $x, y \in P$.*

Proof. (i) $\to$ (ii). Observe $\operatorname{rank}(x) < \operatorname{rank}(y)$, so by assumption, there exists an atom $a \in P$ such that $a \leq y$, $a \not\leq x$, and such that $x \vee a$ exists in P. Set $u := x \vee a$. Observe a covers $0 = x \wedge a$, so u covers x by modularity. In particular $\operatorname{rank}(u) = i$. We show $\partial(u, y) = 2$. Suppose not. Then $\partial(u, y) = 4$ by (7.3), so uxy is geodesic. In this case $u \wedge y = x \wedge y$ by Lemma 13.4(ii), (iii), contradicting the fact that $a \leq u \wedge y$, $a \not\leq x \wedge y$. We have now shown $\partial(u, y) = 2$, so u, y cover $u \wedge y$ by Lemma 7.4. Now $x, u, u\wedge y, y$ is a path with shape $(i-1, i, i-1, i)$, as desired.

(ii) $\to$ (iii). Immediate.

(iii) $\to$ (iv). Set $d := \partial(x, y)$, and pick a geodesic path

$$p := (x = x_0,\ x_1,\ \ldots,\ x_d = y) \qquad\qquad (x_0,\ x_1,\ \ldots,\ x_d \in P)$$

with maximal weight in the sense of (7.7). We claim p is up-flat-down. Suppose not. Then by Lemma 18.1, there exists an integer i $(0 \leq i \leq d-3)$ such that either

$$x_i > x_{i+1} < x_{i+2} < x_{i+3} \tag{18.3}$$

or

$$x_i > x_{i+1} > x_{i+2} < x_{i+3}. \tag{18.4}$$

Interchanging the roles of x, y if necessary, we may assume (18.3). The path x_i, x_{i+1}, x_{i+2}, x_{i+3} is geodesic, with shape $(j-1, j-2, j-1, j)$ for an appropriate integer j $(2 \leq j \leq D)$, so by (iii), there exists a path x_i, x'_{i+1}, x'_{i+2}, x_{i+3} of shape $(j-1, j, j-1, j)$. Observe the sequence

$$p' := (x_0,\ x_1,\ \ldots,\ x_i,\ x'_{i+1},\ x'_{i+2},\ x_{i+3},\ \ldots,\ x_{d-1},\ x_d = y)$$

is a path. p' is geodesic, since p, p' have the same length, and

$$\operatorname{weight}(p') = \operatorname{weight}(p) + 2.$$

This contradicts our construction, and we conclude p is up-flat-down.

(iv) $\to$ (v). Let x, y, x', y' be given. By (iv), there exists a geodesic up-flat-down path p connecting x', y'. We claim p is flat. Set $d := \partial(x', y')$, and write $p = (x' = x_0,\ x_1,\ \ldots,\ x_d = y')$. Suppose p is not flat. Then $d \geq 1$, and either $x' < x_1$ or $x_{d-1} > y'$. Interchanging the roles of x, y if necessary, we may assume $x' < x_1$. Recall $xx'y'y$ is geodesic by Corollary 11.2. Now $xx'x_1y'y$ is geodesic by the construction, and in particular xx_1y is geodesic. Also $x \leq x' < x_1$, so $x_1 \in x_y^+$ by (15.1). But this is inconsistent with $x' < x_1$ and the assumption $x' \in \operatorname{top}(x_y^+)$. We conclude p is flat.

(v) $\to$ (vi). Let x, y be given, and pick any $x' \in \mathrm{top}(x_y^+)$, $y' \in \mathrm{top}(y_x^+)$. By (v) there exists a flat path connecting x', y', so

$$\mathrm{rank}(x') = \mathrm{rank}(y'). \tag{18.5}$$

By Lemma 13.2, Lemma 15.2(i), and Corollary 15.7(ii),

$$\begin{aligned} \mathrm{rank}(x') &= \mathrm{rank}(x) + \mathrm{rank}(x_y^+) \\ &= \rho(x,y) + \gamma(x,y) + \delta(x,y) + \gamma(y,x), \end{aligned} \tag{18.6}$$

and similarly

$$\mathrm{rank}(y') = \rho(y,x) + \gamma(y,x) + \delta(y,x) + \gamma(x,y). \tag{18.7}$$

The result now follows from (18.5)–(18.7), since ρ is symmetric in its arguments.

(vi) $\to$ (i). Pick any $x, y \in P$ such that $\mathrm{rank}(x) < \mathrm{rank}(y)$. We find an atom $a \in P$ such that $a \leq y$, $a \not\leq x$, and such that $x \vee a$ exists in P. To this end, observe by Lemma 13.2, Corollary 15.7(i), and (vi) above that

$$\begin{aligned} \mathrm{rank}(y_x^-) &= \gamma(y,x) \\ &= \mathrm{rank}(y) - \rho(y,x) - \delta(y,x) \\ &> \mathrm{rank}(x) - \rho(x,y) - \delta(x,y) \\ &= \gamma(x,y) \\ &\geq 0, \end{aligned}$$

so there exists $v \in y_x^-$ such that $v > x \wedge y$. By Lemma 9.3(ii), there exists an atom $a \in P$ such that $a \leq v$ but $a \not\leq x \wedge y$. Observe $a \leq v \leq y$. Observe $a \not\leq x$; otherwise $a \leq x \wedge y$, a contradiction. Observe $x \vee a$ exists in P, since $x \vee v$ is an upper bound for a, x. The element a now has the desired properties, so we are done. This proves Theorem 18.2.

§19. $x \star y$ is a $[x \wedge y, x]$-basis system

Let P denote a quantum matroid, and pick any $x, y \in P$. Our goal in this section is to establish the related facts that $x \star y$ is a $[x \wedge y, x]$-basis system, $[x \wedge y, x \star y]$ is a $[x \wedge y, x]$-matroid, and that $[x \star y, x]^*$ is a $[x \wedge y, x]^*$-matroid.

Lemma 19.1. *Let P denote a quantum matroid, and let G denote a geodesically closed subposet of P. Then G is a quantum matroid.*

Proof. G is a prematroid by Corollary 8.3, so it remains to show G satisfies the augmentation axiom. By Theorem 18.2, it suffices to show the function δ_G is symmetric in its arguments. But this is the case, since δ_G is a restriction of δ_P by Lemma 13.7(ii), and since δ_P is symmetric in its arguments by Theorem 18.2(vi). This proves Lemma 19.1.

Corollary 19.2. *Let P denote a quantum matroid, and pick any $x, y \in P$. Then x_y^+, x^+ are both quantum matroids.*

Proof. These subposets are geodesically closed in P by Lemma 8.4, Lemma 15.2(ii), so they are quantum matroids by Lemma 19.1. This proves Corollary 19.2.

Theorem 19.3. *Let P denote a quantum matroid, and pick any $x, y \in P$.*

(i) $x_y^- = [x \wedge y, x \star y]$.
(ii) $x \star y$ *is a* $[x \wedge y, x]$*-basis system.*
(iii) $[x \wedge y, x \star y]$ *is a* $[x \wedge y, x]$*-matroid.*
(iv) $[x \star y, x]^*$ *is a* $[x \wedge y, x]^*$*-matroid.*

Proof. (i) Observe the posets x_y^-, y_x^+ are isomorphic by Theorem 15.5, and y_x^+ is a quantum matroid by Lemma 19.2, so x_y^- is a quantum matroid. Now by Lemma 4.3 and Lemma 15.4(ii),

$$\max(x_y^-) = \mathrm{top}(x_y^-) = x \star y,$$

and our result follows since x_y^- is a lower ideal in $[x \wedge y, x]$.

(ii) We have seen x_y^- is both a quantum matroid and a lower ideal in $[x \wedge y, x]$, so x_y^- is a $[x \wedge y, x]$-matroid. Now $x \star y = \max(x_y^-)$ is a $[x \wedge y, x]$-basis system by Theorem 2.5(ii).

(iii) Immediate from (ii) and Theorem 2.5(i).

(iv) Immediate from (ii), Theorem 2.5(i), and Lemma 3.1.

§20. The notion of relative closeness

Lemma 20.1. *Let P denote a quantum matroid. Then for all $x, y \in P$,*

$$\rho(x, y) + \delta(x, y) \leq \mathrm{rank}(y). \tag{20.1}$$

Proof. Observe by Definition 13.1(i), Lemma 13.2, and Theorem 18.2(vi) that

$$\begin{aligned} \rho(x,y)+\delta(x,y) &= \rho(y,x)+\delta(y,x) \\ &= \operatorname{rank}(y)-\gamma(y,x) \\ &\le \operatorname{rank}(y). \end{aligned} \tag{20.2}$$

We now consider the case of equality.

Lemma 20.2. *Let* P *denote a quantum matroid, and pick any* $x, y \in P$. *Then the following* (i)–(v) *are equivalent.*

(i) *Equality holds in (20.1).*
(ii) $\gamma(y,x) = 0$.
(iii) x *is the unique element in* x_y^+.
(iv) $x \wedge y$ *is the unique element in* y_x^-.
(v) $y \star x = x \wedge y$.
(vi) zxy *is geodesic for all* $z \in x^+$.

If (i)–(vi) *hold, we say* x *is relatively close to* y.

Proof. (i) $\leftrightarrow$ (ii). Immediate from (20.2).

(ii) $\leftrightarrow$ (iii) $\leftrightarrow$ (iv). The posets x_y^+, y_x^- both have rank $\gamma(y,x)$ by Corollary 15.7.

(iv) $\leftrightarrow$ (v). Recall $y_x^- = [x \wedge y, y \star x]$ by Theorem 19.3(i).

(iii) $\rightarrow$ (vi). Pick any $z \in x^+$, and write $p := x \vee (y \wedge z)$. Then zpy is geodesic by Lemma 17.1, Theorem 17.2(iii). But $p \in x_y^+$ by Lemma 17.1, so $p = x$ by (iii). We conclude zxy is geodesic, as desired.

(vi) $\rightarrow$ (iii). Pick any $z \in x_y^+$. Then certainly $z \in x^+$, so zxy is geodesic by (vi). Also xzy is geodesic by (15.1), so $z = x$. This proves Lemma 20.2.

Lemma 20.3. *Let* P *denote a quantum matroid, and pick any* $x, y, z \in P$.

(i) *If* x *is relatively close to* y, *and* $x \le z$, *then* z *is relatively close to* y.
(ii) *Suppose* $z \in \operatorname{top}(x_y^+)$. *Then* z *is relatively close to* y.

Proof. (i) Observe $0 \le \gamma(y,z) \le \gamma(y,x)$ by Lemma 13.3(i), Corollary 13.6(v), and $\gamma(y,x) = 0$ by Lemma 20.2(ii), so $\gamma(y,z) = 0$. Now z is relatively close to y by Lemma 20.2.

(ii) z is the unique element in z_y^+ by Lemma 15.8(i), so z is relatively close to y by Lemma 20.2(iii).

Theorem 20.4. *Let P denote a quantum matroid, and pick any $x, y \in P$ such that x is relatively close to y. Then for all $z \in x^+$, the following* (i)–(iii) *hold.*

(i) $\rho(z, y) = \rho(x, y)$,
(ii) $\delta(z, y) = \delta(x, y)$,
(iii) $\gamma(z, y) - \gamma(x, y) = \text{rank}(z) - \text{rank}(x)$.

Proof. (i) $x \le z$ by assumption, and zxy is geodesic by Lemma 20.2(vi), so $\rho(z, y) = \rho(x, y)$ by Lemma 13.4(ii),(iv).

(ii) z is relatively close to y by Lemma 20.3(i), so by Lemma 20.2(i) and (i) above,

$$\begin{aligned}\delta(z, y) &= \text{rank}(y) - \rho(z, y)\\ &= \text{rank}(y) - \rho(x, y)\\ &= \delta(x, y).\end{aligned}$$

(iii) By Lemma 13.2,

$$\text{rank}(x) = \rho(x, y) + \gamma(x, y) + \delta(x, y), \tag{20.3}$$

$$\text{rank}(z) = \rho(z, y) + \gamma(z, y) + \delta(z, y). \tag{20.4}$$

Our result is immediate upon subtracting (20.3) from (20.4), and evaluating the result using (i), (ii) above.

§21. The staircase theorem

In this section, we describe a quantum matroid in a way that may help the reader visualize its structure. Theorem 21.3 is our main result. First, we need a few definitions.

Recall a *directed graph* (or *di-graph*) is a pair $\mathcal{D} := (V\mathcal{D}, E\mathcal{D})$, where $V\mathcal{D}$ is a nonempty finite set (of *vertices*) and $E\mathcal{D} \subseteq V\mathcal{D} \times V\mathcal{D}$ (the *edges*). For all $u, v \in V\mathcal{D}$, we write $u \to v$ whenever $uv \in E\mathcal{D}$. Observe possibly both $u \to v$ and $v \to u$, possibly only one of these occurs, or possibly neither. We may also have $u \to u$.

Definition 21.1. Let $\mathcal{D}$ denote any di-graph, and let P denote any poset. By a *$\mathcal{D}$-partition* of P, we mean a map $\sigma : P \to V\mathcal{D}$, such that (i), (ii) hold below:

(i) For all $x, y \in P$, if x, y are adjacent then $\sigma(x) \to \sigma(y)$ or $\sigma(y) \to \sigma(x)$ (or both).

(ii) For all $u, v \in V\mathcal{D}$ such that $u \to v$, and for all $x \in P$ such that $\sigma(x) = u$, then there exists $y \in P$ such that x, y are adjacent and such that $\sigma(y) = v$.

(Caution: We do not require σ be onto $V\mathcal{D}$).

The quantum matroids have $\mathcal{D}$-partitions for certain di-graphs $\mathcal{D}$, described below.

Definition 21.2. For any nonnegative integers a, b, define the di-graph $\mathcal{D} = \mathcal{D}(a, b)$ as follows: The vertex set $V\mathcal{D}$ is the set of three tuples

(21.2) $V\mathcal{D} := \{(\rho, \gamma, \delta) | \rho, \gamma, \delta \in \mathbb{Z}, 0 \le \rho, 0 \le \delta, \rho + \delta \le a, 0 \le \gamma \le b\}.$

For all pairs of vertices (ρ, γ, δ), $(\rho', \gamma', \delta') \in V\mathcal{D}$, there is an edge $(\rho, \gamma, \delta) \to (\rho', \gamma', \delta')$ in $\mathcal{D}$ whenever one of the following rows holds:

(21.2)	$\rho' = \rho + 1$	$\gamma' = \gamma$	$\delta' = \delta$
(21.3)	$\rho' = \rho - 1$	$\gamma' = \gamma$	$\delta' = \delta$
(21.4)	$\rho' = \rho$	$\gamma' = \gamma + 1$	$\delta' = \delta$
(21.5)	$\rho' = \rho$	$\gamma' = \gamma - 1$	$\delta' = \delta = 0$
(21.6)	$\rho' = \rho$	$\gamma' = \gamma$	$\delta' = \delta - 1$

Observe the "shape" of $\mathcal{D}(a, b)$ resembles that of a staircase of height a and width b. For example, $\mathcal{D}(2, 3)$ looks as follows (we abbreviate $u \leftrightarrow v$ whenever $u \to v$ and $v \to u$):

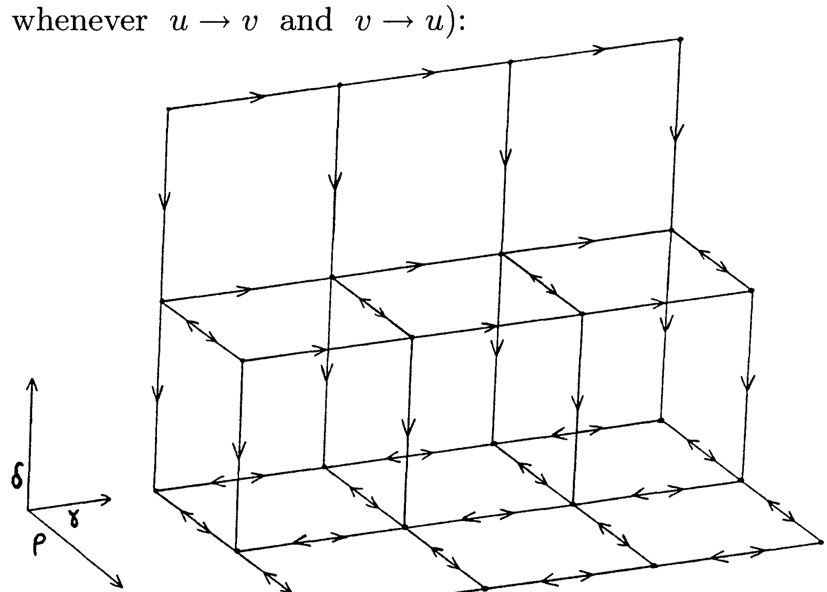


Fig. 1.

Theorem 21.3. *Let P denote a quantum matroid with rank D. Pick any integer a $(0 \leq a \leq D)$, fix any $y \in P$ such that* $\operatorname{rank}(y) = a$, *and set*

$$\sigma(x) := \big(\rho(x,y),\ \gamma(x,y),\ \delta(x,y)\big) \qquad (\forall x \in P). \tag{21.7}$$

Then σ is a $\mathcal{D}(a,\ D-a)$-partition of P.

Proof. Abbreviate $\mathcal{D} := \mathcal{D}(a, D-a)$. Pick any $x \in P$, and abbreviate

$$\begin{aligned} \rho &:= \rho(x,y), \\ \gamma &:= \gamma(x,y), \\ \delta &:= \delta(x,y). \end{aligned}$$

Let us first check $\sigma(x) \in V\mathcal{D}$. To do this, we verify ρ, γ, δ satisfy the inequalities in (21.1) (with $b = D - a$). Observe $0 \leq \rho$, $0 \leq \gamma$, $0 \leq \delta$ by Lemma 13.3(i), and $\rho + \delta \leq a$ by Lemma 20.1. To see $\gamma \leq D - a$, observe

$$\begin{aligned} \gamma(x,y) &\leq D - \rho(x,y) - \gamma(y,x) - \delta(x,y) \\ &= D - \operatorname{rank}(y) \end{aligned}$$

by Lemmas 13.2, 13.3 and Theorem 18.2(vi). We have now shown $\sigma(x) \in V\mathcal{D}$.

Next, let us verify that σ satisfies (i) of Definition 21.1. To this end, pick any $z \in P$ such that x, z are adjacent, and set

$$\begin{aligned} \rho' &:= \rho(z,y), \\ \gamma' &:= \gamma(z,y), \\ \delta' &:= \delta(z,y). \end{aligned}$$

We must show $(\rho,\gamma,\delta) \to (\rho',\gamma',\delta')$ or $(\rho',\gamma',\delta') \to (\rho,\gamma,\delta)$. Interchanging x, z if necessary, we may assume x covers z. But then by Theorem 13.5(i),(iii), the three tuple $(\rho-\rho', \gamma-\gamma', \delta-\delta')$ equals either

$$(1,0,0) \quad \text{(in which case } (\rho,\gamma,\delta) \to (\rho',\gamma',\delta') \text{ by (21.3))},$$

or

$$(0,1,0) \quad \text{(in which case } (\rho',\gamma',\delta') \to (\rho,\gamma,\delta) \text{ by (21.4))},$$

or

$$(0,0,1) \quad \text{(in which case } (\rho,\gamma,\delta) \to (\rho',\gamma',\delta') \text{ by (21.6))}.$$

We have now shown σ satisfies (i) of Definition 21.1.

It remains to show σ satisfies part (ii) of Definition 21.1. To this end, let x, ρ, γ, δ be as above, and pick any $(\rho',\gamma',\delta') \in V\mathcal{D}$ such that $(\rho,\gamma,\delta) \to (\rho',\gamma',\delta')$. We must find $z \in P$ such that x, z are adjacent, and such that $\sigma(z) = (\rho',\gamma',\delta')$. We consider the 5 cases (21.2)–(21.6) in turn.

First assume $(\rho',\gamma',\delta') = (\rho+1,\gamma,\delta)$. Observe by Lemma 13.2 and Corollary 15.7(ii) that

$$\begin{aligned}\operatorname{rank}(x_y^+) &= \gamma(y,x)\\ &= a-\rho-\delta\\ &= a-\rho'-\delta'+1\\ &\geq 1,\end{aligned}$$

so there exists $z \in x_y^+$ such that z covers x. Observe $\sigma(z) = (\rho+1,\gamma,\delta)$ by Lemma 15.3(iv),(v).

Next assume $(\rho',\gamma',\delta') = (\rho-1,\gamma,\delta)$. Then

$$\rho = \rho'+1 \geq 1,$$

so $x \wedge y \neq 0$. Let z' denote a relative complement of $x \wedge y$ in the interval $[0,x]$. Then $z' < x$ by modularity, so there exists $z \in [z',x]$ such that x covers z. Observe $x \in z'^+_y$ by Lemma 15.3(i),(ii), so $z'xy$ is geodesic. Observe $z'zx$ is geodesic by the construction, so $z'zxy$ is geodesic. In particular zxy is geodesic, so $x \in z_y^+$ by (15.1). Now $\sigma(z) = (\rho-1,\gamma,\delta)$ by Lemma 15.3(iv),(v).

Next assume $(\rho',\gamma',\delta') = (\rho,\gamma+1,\delta)$. Pick any $u \in \operatorname{top}(x_y^+)$, and observe

$$\begin{aligned}\operatorname{rank}(u) &= \operatorname{rank}(x) + \operatorname{rank}(x_y^+)\\ &= \rho(x,y)+\gamma(x,y)+\delta(x,y)+\gamma(y,x)\\ &= a+\gamma\\ &= a+\gamma'-1\\ &< D,\end{aligned}$$

so by Lemma 4.3, there exists $v \in P$ such that v covers u. Let z denote a relative complement of u is $[x,v]$. Then z covers x by modularity. We now compute $\sigma(z)$. Observe by Lemma 15.3(iv),(v) and our choice of u that

$$\sigma(u) = (\rho_1,\gamma,\delta), \tag{21.8}$$

where $\rho_1 = \rho + \mathrm{rank}(u) - \mathrm{rank}(x)$. Observe u is relatively close to y by Lemma 20.3(ii), so

$$\sigma(v) = (\rho_1, \gamma + 1, \delta) \tag{21.9}$$

by (21.8) and Theorem 20.4. Observe $u = x \vee (y \wedge v)$ by (21.8), (21.9), Lemma 17.1, and Theorem 17.2(i),(iv), so

$$\sigma(z) = (\rho, \gamma + 1, \delta)$$

by Theorem 17.3(iii),(iv).

Next assume $\delta = 0$ and $(\rho', \gamma', \delta') = (\rho, \gamma - 1, 0)$. Observe by Lemma 13.2 that

$$\begin{aligned} \mathrm{rank}(x) - \rho &= \gamma \\ &= \gamma' + 1 \\ &\geq 1, \end{aligned}$$

so $x > x \wedge y$. Hence there exists $z \in [x \wedge y, x]$ such that x covers z. Now $\sigma(z) = (\rho, \gamma - 1, 0)$ by Definition 13.1, since $z \wedge y = x \wedge y$ by Lemma 13.4(i),(iii) and $z \star y = z$ by Lemma 12.2.

Finally assume $(\rho', \gamma', \delta') = (\rho, \gamma, \delta - 1)$. Observe by Lemma 14.2(vi) that

$$\begin{aligned} \mathrm{rank}\big([x \star y, x]^*\big) &= \delta \\ &= \delta' + 1 \\ &\geq 1, \end{aligned}$$

so there exists $z \in [x \star y, x]^*$ that is covered by x (in the poset P). Observe $\sigma(z) = (\rho, \gamma, \delta - 1)$ by Lemma 14.3(ii),(iv). This proves Theorem 21.3.

§22. The graph on top(P)

In this section we consider a graph defined on the top of a quantum matroid.

Theorem 22.1. *Let P denote a quantum matroid with rank D.*

(i) *For all $x, y \in top(P)$, there exists a geodesic flat path in P connecting x, y.*

For (ii)–(iv) *below, we view* $\mathrm{top}(P)$ *as the vertex set of an undirected graph, where vertices* $x, y \in \mathrm{top}(P)$ *are declared adjacent whenever* x, y *cover* $x \wedge y$.

(ii) *The graph* $\mathrm{top}(P)$ *is connected.*

(iii) *For all* $x, y \in \mathrm{top}(P)$,

$$
\begin{aligned}
\partial_{\mathrm{top}}(x,y) &= \partial(x,y)/2 && (22.1)\\
&= D - \rho(x,y) && (22.2)\\
&= \delta(x,y), && (22.3)
\end{aligned}
$$

where ∂_{top} *denotes the path length distance function for the graph* $\mathrm{top}(P)$.

(iv)

$$
\begin{aligned}
\mathrm{diam}_{\mathrm{top}}(P) &= \max\{\delta(x,y) \mid x, y \in P\} && (22.4)\\
&\le D, && (22.5)
\end{aligned}
$$

where $\mathrm{diam}_{\mathrm{top}}(P)$ *denotes the diameter of the graph* $\mathrm{top}(P)$.

Proof. (i) By Theorem 18.2(iv), there exists a geodesic up-flat-down path p connecting x, y. But x, $y \in \mathrm{top}(P)$, so p is flat.

(ii) Immediate from (i).

(iii) Line (22.1) is immediate from (i) and the definition of a flat path in Lemma 18.1. To see (22.2), observe by Corollary 7.6(i) that

$$
\begin{aligned}
\partial(x,y) &= \mathrm{rank}(x) + \mathrm{rank}(y) - 2\mathrm{rank}(x \wedge y)\\
&= 2\big(D - \rho(x,y)\big).
\end{aligned}
$$

To see (22.3), observe x is the unique element in x_y^+, so x is relatively close to y by Lemma 20.2(iii). Now

$$\rho(x,y) + \delta(x,y) = \mathrm{rank}(y) = D$$

by Lemma 20.2(i).

(iv) To see (22.4), observe by (22.3) and Corollary 13.6(iii),(iv) that

$$
\begin{aligned}
\mathrm{diam}_{\mathrm{top}}(P) &= \max\{\, \partial_{\mathrm{top}}(x,y) \mid x, y \in \mathrm{top}(P)\}\\
&= \max\{\, \delta(x,y) \mid x, y \in \mathrm{top}(P)\}\\
&= \max\{\, \delta(x,y) \mid x, y \in P\}.
\end{aligned}
$$

We now have (22.4). Line (22.5) is immediate from (22.2). This proves Theorem 22.1.

For the remainder of this section, we investigate the quantum matroids P such that $\mathrm{diam}_{\mathrm{top}}(P) \le 1$.

Lemma 22.2. *For any poset* P, *the following are equivalent.*

(i) P *is a quantum matroid, and*

$$\operatorname{diam}_{\mathrm{top}}(P) = 0. \tag{22.6}$$

(ii) P *is a quantum matroid, and*

$$|\operatorname{top}(P)| = 1. \tag{22.7}$$

(iii) P *is a modular atomic lattice.*

Proof. (i) $\leftrightarrow$ (ii). Clear

(ii) $\rightarrow$ (iii). Observe $\max(P) = \operatorname{top}(P)$ consists of a single element, so P has a 1. Now $P = [0,1]$ is a modular atomic lattice by condition M in Definition 4.1.

(iii) $\rightarrow$ (ii). Clear.

Lemma 22.3. *For any poset* P, *the following are equivalent.*

(i) P *is a quantum matroid, and*

$$\operatorname{diam}_{\mathrm{top}}(P) \leq 1. \tag{22.8}$$

(ii) P *is a quantum matroid, and*

$$x, y \text{ cover } x \wedge y \text{ for all distinct } x, y \in \operatorname{top}(P). \tag{22.9}$$

(iii) P *is a prematroid, and*

$$x, y \text{ cover } x \wedge y \text{ for all distinct } x, y \in \max(P). \tag{22.10}$$

(iv) P *is a prematroid, and*

$$\delta(x, y) \leq 1 \qquad \text{for all } x, y \in P. \tag{22.11}$$

If (i)–(iv) *hold, we call* P *a design matroid.*

Proof. (i) $\rightarrow$ (ii). Let x, y denote distinct elements in $\operatorname{top}(P)$. Then $\partial_{\mathrm{top}}(x,y) = 1$ by (22.8), so $\partial(x,y) = 2$ by (22.1). Now x, y cover $x \wedge y$ by Lemma 7.4.

(ii) $\rightarrow$ (iii). It is clear P is a prematroid. Also $\max(P) = \operatorname{top}(P)$ by Lemma 4.3, so (22.10) follows from (22.9).

(iii) $\rightarrow$ (iv). Let x, $y \in P$ be given. We show $\delta(x,y) \leq 1$. There exists x', $y' \in \max(P)$ such that $x \leq x'$, $y \leq y'$. Observe $\delta(x,y) \leq \delta(x',y')$ by Corollary 13.6(iii),(iv), so it suffices to show $\delta(x',y') \leq 1$. Assume $\delta(x',y') \neq 0$; otherwise we are done. Then x', y' are distinct, so x', y' covers $x' \wedge y'$ by (22.10). Now $\delta(x', y') = 1$ by Definition 12.1, and we are done.

(iv) $\to$ (i). To show P is a quantum matroid, we show P satisfies the augmentation axiom. Pick any $x, y \in P$. By Theorem 18.2(vi), it suffices to show

$$\delta(x, y) = \delta(y, x). \tag{22.12}$$

First suppose $x \vee y$ exists. Then $\delta(x, y) = 0$, $\delta(y, x) = 0$ by Lemma 12.2, so (22.12) holds. Now suppose $x \vee y$ does not exist. Then $\delta(x, y) \neq 0$, $\delta(y, x) \neq 0$ by Lemma 12.2, so $\delta(x, y)$, $\delta(y, x)$ are both 1 by (22.11). Again (22.12) holds, so (22.12) holds in general. Now P satisfies the augmentation axiom by Theorem 18.2, so P is a quantum matroid. To see (22.8), observe by (22.4), (22.11) that

$$\begin{aligned} \mathrm{diam}_{\mathrm{top}}(P) &= \max\{\delta(x, y) \mid x, y \in P\} \\ &\leq 1. \end{aligned}$$

This proves Lemma 22.3.

Lemma 22.4. *Let P denote a quantum matroid. Then x_y^+, x_y^- are design matroids for all $x, y \in P$ such that $\delta(x, y) = 1$.*

Proof. We show x_y^- is a design matroid by showing it satisfies condition (ii) in Lemma 22.3. Observe x_y^- is a quantum matroid by Theorem 19.3(i),(iii), so it remains to show u, v cover $u \wedge v$ for all distinct $u, v \in \mathrm{top}(x_y^-)$. By Theorem 12.3, Lemma 15.4(ii), and our assumption $\delta(x, y) = 1$, we find x covers u, v. Now $u \vee v = x$ covers u, v, so u, v covers $u \wedge v$ by modularity. We have now shown x_y^- satisfies condition (ii) in Lemma 22.3, so x_y^- is a design-matroid.

To see that x_y^+ is a design-matroid, recall $\delta(y, x) = 1$ by Theorem 18.2(vi), so y_x^- is a design-matroid by our above remarks. Recall x_y^+ is isomorphic to y_x^- by Theorem 15.5, so x_y^+ is a design-matroid. We have now proved Lemma 22.4.

§23. Quantum matroids and diagram geometries

In this section we obtain a characterization of a quantum matroid that might be useful to people doing research on diagram geometries. We do not explicitly introduce the language of diagram geometries in order to avoid cumbersome terminology, but a reader familiar with these geometries should have no trouble translating our result into that language.

Theorem 23.1. *Let D denote a nonnegative integer. Then a poset P is a quantum matroid of rank D if and only if* (i)–(iv) *hold below.*

(i) *P is a prematroid of rank D.*
(ii) *For all $x \in P$, there exists $x' \in \mathrm{top}(P)$ such that $x \le x'$.*
(iii) *For all $x \in P$ with $\mathrm{rank}(x) \le D-2$, the subposet of P induced on $x^+ \backslash \{x\}$ is connected.*
(iv) *For all $x \in P$ with $\mathrm{rank}(x) = D-2$, the poset x^+ is a quantum matroid of rank* 2.

(The quantum matroids of rank 2 are described in Example 4.2.)

Proof. We first assume P is a quantum matroid, and verify (i)–(iv). Line (i) is immediate from Definition 5.1, and (ii) is just Lemma 4.3. To see (iii), pick any $x \in P$ such that $\mathrm{rank}(x) \le D-2$. We show the subposet of P induced on $x^+ \backslash \{x\}$ is connected. By (ii) above, any element in $x^+ \backslash \{x\}$ is connected by a path in $x^+ \backslash \{x\}$ to some element in $\mathrm{top}(x^+)$. Hence it suffices to pick any $u, v \in \mathrm{top}(x^+)$, and show u, v are connected by a path in $x^+ \backslash \{x\}$. By Theorem 22.1(i), there exists a geodesic flat path p in P that connects u, v. Recall x^+ is geodesically closed in P by Lemma 8.4(i), so p is contained in x^+. The elements of p all have rank $D-1$ or D by Lemma 18.1, and $\mathrm{rank}(x) \le D-2$, so x is not included in p. It follows p is contained in $x^+ \backslash \{x\}$, as desired. We now have (iii). To see (iv), recall x^+ is a quantum matroid by Corollary 19.2, and $\mathrm{rank}(x^+) = 2$ by part (ii) above. We have now proved the theorem in one direction, so we now consider the converse.

Let P denote a poset satisfying (i)–(iv) in the present theorem. We show P is a quantum matroid of rank D by induction on D. The case $D \le 1$ is trivial, and the case $D = 2$ is immediate from assumption (iv), so from now on assume $D \ge 3$.

P is a prematroid of rank D by assumption (i), so we need only show P satisfies the augmentation axiom. To do this, we show P satisfies condition (iii) in Theorem 18.2. For the rest of this proof, we use the following terminology: For any paths p, p' in P, we say *p' replaces p* whenever p, p' share the same endpoints. For each integer i $(2 \le i \le D)$, let E_i denote the proposition that any geodesic path in P of shape $(i-1, i-2, i-1, i)$ can be replaced by a path in P that has shape $(i-1, i, i-1, i)$. The condition (iii) in Theorem 18.2 will follow if we can show $E_2, E_3, \ldots, E_D$. We do this in two steps.

Claim 1. $E_3, E_4, \ldots, E_D$ *hold.*

Proof of Claim 1. Pick an integer i $(3 \leq i \leq D)$, and pick a path $xyzw$ in P of shape $(i-1, i-2, i-1, i)$. We show $xyzw$ can be replaced by a path in P that has shape $(i-1, i, i-1, i)$. The poset y^+ satisfies the conditions (i)–(iv) of the present theorem (with D replaced by $D-i+2$), and $D-i+2 \leq D-1$, so y^+ is a quantum matroid by induction. Observe the path $xyzw$ is contained in y^+, and has shape 1012 (in y^+). Applying Theorem 18.2(iii) to y^+, we find the path $xyzw$ can be replaced by a path $xy'z'w$ in y^+ that has shape 1212 (in y^+). Observe the path $xy'z'w$ has shape $(i-1, i, i-1, i)$ (in P), so we are done. This proves Claim 1.

Claim 2. *E_2 holds.*

Proof of Claim 2. Pick any $x, y \in P$ such that $\operatorname{rank}(x) = 1$, $\operatorname{rank}(y) = 2$, and $x \not\leq y$. We show x, y are connected by a path in P that has shape 1212. To do this, we show (i), (ii) below:

(i) There exists a path in P with endpoints x, y and shape $1212 \cdots 12$.
(ii) Any path in P with shape 121212 can be replaced by a path in P with shape 1212.

To see (i), recall by (iii) of the present theorem that there exists a path in $P \backslash \{0\}$ connecting x, y. Of all such paths, pick a path

$$p = (x = x_0, x_1, \ldots, x_d = y) \qquad (x_0, x_1, \ldots, x_d \in P)$$

with minimal weight in the sense of (7.7). Set

$$r := \max\{\operatorname{rank}(x_i) \mid 0 \leq i \leq d\},$$

and observe $r \geq \operatorname{rank}(x_d) = 2$. p will have the desired shape $1212 \cdots 12$ if we can show $r = 2$. Suppose $r \geq 3$, and pick any integer i $(0 \leq i \leq d)$ such that $\operatorname{rank}(x_i) = r$. Then $1 \leq i \leq d-1$, and $x_{i-1} < x_i > x_{i+1}$. Observe $x_0, x_1, \ldots, x_d$ are distinct by the construction, so $x_{i-1} \vee x_{i+1} = x_i$ covers x_{i-1}, x_{i+1}. Now x_{i-1}, x_{i+1} cover $x_{i-1} \wedge x_{i+1}$ by modularity. Now p can be replaced by a path

$$p' = (x = x_0, x_1, \ldots, x_{i-1}, x_{i-1} \wedge x_{i+1}, x_{i+1}, \ldots, x_d = y)$$

that is contained in $P \backslash \{0\}$, and has

$$\operatorname{weight}(p') = \operatorname{weight}(p) - 2.$$

This contradicts the construction, so $r = 2$. The path p now has the desired shape $1212 \cdots 12$, so (i) holds.

To see part (ii) in the present claim, consider the following sequences:

(s1) 121212
(s2) 12323212
(s3) 12323232
(s4) 12321232
(s5) 121232
(s6) 123232
(s7) 123212
(s8) 1232
(s9) 1212

To show part (ii) in the present claim, we show that any path p in P whose shape is one of (s1)–(s8), can be replaced by a path in P whose shape is included below shape(p) in the above list. We write $p = (x_0, x_1, x_2, \ldots)$ and consider each of the shapes (s1)–(s8) in turn.

Case s1. By assumption (ii) of the present theorem, and since $D \geq 3$, there exists $y \in P$ such that y covers x_3. first suppose $x_1 \leq y$. Then $(x_0, x_1, y, x_3, x_4, x_5)$ is a path of shape 123212 (s7). Next suppose $x_1 \not\leq y$. Then (x_1, x_2, x_3, y) is a geodesic path of shape 2123, so by E_3, there exists a path (x_1, z, w, y) in P of shape 2323. Now $(x_0, x_1, z, w, y, x_3, x_4, x_5)$ is a path of shape 12323212 (s2).

Case s2. First suppose $x_4 \geq x_7$. Then $(x_0, x_1, x_2, x_3, x_4, x_7)$ is a path of shape 123232 (s6). Next suppose $x_4 \not\geq x_7$. Then (x_7, x_6, x_5, x_4) is a geodesic path of shape 2123, so by E_3, there exists a path (x_7, y, z, x_4) in P of shape 2323. Now $(x_0, x_1, x_2, x_3, x_4, z, y, x_7)$ is a path of shape 12323232 (s3).

Case s3. First suppose $x_3 = x_5$. Then $(x_0, x_1, x_2, x_3, x_6, x_7)$ is a path of shape 123232 (s6). Next suppose $x_3 \neq x_5$. Then $x_3 \vee x_5 = x_4$ covers x_3, x_5, so x_3, x_5 cover $x_3 \wedge x_5$ by modularity. Now $(x_0, x_1, x_2, x_3, x_3 \wedge x_5, x_5, x_6, x_7)$ is a path of shape 12321232 (s4).

Case s4. First suppose $x_0 = x_4$. Then (x_0, x_5, x_6, x_7) is a path of shape 1232 (s8). Next suppose $x_0 \neq x_4$. Observe x_2 is an upper bound for x_0, x_4, so $x_0 \vee x_4$ exists. $x_0 \vee x_4$ covers x_0, x_4 by modularity, so $(x_0, x_0 \vee x_4, x_4, x_5, x_6, x_7)$ is a path of shape 121232 (s5).

Case s5. First suppose $x_1 \leq x_4$. Then (x_0, x_1, x_4, x_5) is a path of shape 1232 (s8). Next suppose $x_1 \not\leq x_4$. Then (x_1, x_2, x_3, x_4) is a geodesic path of shape 2123, so by E_3, there exists a path (x_1, y, z, x_4) in P of shape 2323. Now $(x_0, x_1, y, z, x_4, x_5)$ is a path of shape 123232 (s6).

Case s6. First suppose $x_3 = x_5$. Then (x_0, x_1, x_2, x_5) is a path of shape 1232 (s8). Next suppose $x_3 \neq x_5$. Then $x_3 \vee x_5 =$

x_4 covers x_3, x_5, so x_3, x_5 cover $x_3 \wedge x_5$ by modularity. Now $(x_0, x_1, x_2, x_3, x_3 \wedge x_5, x_5)$ is a path of shape 123212 (s7).

Case s7. First suppose $x_0 = x_4$. Then (x_0, x_5, x_0, x_5) is a path of shape 1212 (s9). Next suppose $x_0 \neq x_4$. Observe x_2 is an upper bound for x_0, x_4, so $x_0 \vee x_4$ exists. $x_0 \vee x_4$ covers x_0, x_4 by modularity, so $(x_0, x_0 \vee x_4, x_4, x_5)$ is a path of shape 1212 (s9).

Case s8. First suppose $x_1 = x_3$. Then (x_0, x_3, x_0, x_3) is a path of shape 1212 (s9). Next suppose $x_1 \neq x_3$. Then $x_1 \vee x_3 = x_2$ covers x_1, x_3, so x_1, x_3 cover $x_1 \wedge x_3$ by modularity. Now $(x_0, x_1, x_1 \wedge x_3, x_3)$ is a path of shape 1212 (s9).

We have now shown part (ii) in the present claim, so E_2 holds. Now the propositions $E_2, E_3, \ldots, E_D$ hold by Claims 1, 2, so P satisfies condition (iii) in Theorem 18.2. Now P satisfies the augmentation axiom by that theorem, and we conclude P is a quantum matroid of rank D.

§24. A Characterization of quantum matroids

In this section we obtain a characterization of a quantum matroid that is related to the material on Tits polar spaces in Sections 29, 30.

Definition 24.1. Let us say a prematroid P is *transversal* whenever $\max(P) = \mathrm{top}(P)$.

Theorem 24.2. *Let P denote a prematroid of rank D. Then the following are equivalent.*

(i) *P satisfies the augmentation axiom AU in Definition* 4.1.

(ii) *For all atoms $x \in P$, and for all $y \in \mathrm{top}(P)$ such that $x \not\leq y$, x_y^+ is transversal and has rank $D-1$.*

(iii) *For all $u \in P$, and for all $v \in top(P)$ such that u covers $u \wedge v$, there exists $v' \in \mathrm{top}(P)$ such that $u \leq v'$, and such that v, v' cover $v \wedge v'$.*

Proof. (i) $\to$ (ii). Let x, y be given. Observe x_y^+ is a quantum matroid by Corollary 19.2, so it is transversal by Lemma 4.3. Observe $\mathrm{top}(x_y^+) \subseteq \mathrm{top}(P)$ by Theorem 18.2(v), so $\mathrm{rank}(x_y^+) = D-1$.

(ii) $\to$ (iii). Pick any $u \in P$ and any $v \in \mathrm{top}(P)$ such that u covers $u \wedge v$. Let x denote a relative complement of $u \wedge v$ in $[0, u]$. Then x is an atom by modularity. Observe $x \not\leq v$; otherwise

$$\begin{aligned} u &= x \vee (u \wedge v) \\ &\leq v, \end{aligned}$$

a contradiction. Now $x \wedge v = 0$. We may now view $x, u \wedge v$ as relative complements in $[x \wedge v, u]$, so $u \in x_v^+$ by Lemma 15.3(i),(ii). Pick any $v' \in \max(x_v^+)$ such that $u \leq v'$. We check v' has the required properties. We mentioned x is an atom such that $x \not\leq v$, so by assumption x_v^+ is transversal and has rank $D-1$. It follows

$$\begin{aligned} v' &\in \max(x_v^+) \\ &= \mathrm{top}(x_v^+) \\ &\subseteq \mathrm{top}(P). \end{aligned}$$

Observe $x, v \wedge v'$ are relative complements in $[0, v']$ by Lemma 15.3(i),(ii), and x covers 0, so v' covers $v \wedge v'$ by modularity. In particular $\mathrm{rank}(v \wedge v') = D - 1$, and it follows v covers $v \wedge v'$.

(iii) $\rightarrow$ (i). To show P satisfies the augmentation axiom, we first show $\max(P) = \mathrm{top}(P)$. Suppose $\max(P) \neq \mathrm{top}(P)$. Then $[0, \mathrm{top}(P)]$ is a proper subset of P. Pick any element

$$u \in P \backslash [0, \mathrm{top}(P)] \tag{24.1}$$

with $\mathrm{rank}(u)$ minimal. Certainly $u \neq 0$, so there exists $x \in P$ such that u covers x. Of course $\mathrm{rank}(x) < \mathrm{rank}(u)$, so $x \in [0, \mathrm{top}(P)]$ by construction. Pick any $v \in \mathrm{top}(P)$ such that $x \leq v$. Observe $u \not\leq v$ by (24.1), so $u \wedge v = x$. Now u covers $u \wedge v$, so by (iii), there exists $v' \in \mathrm{top}(P)$ such that $u \leq v'$ (and such that v, v' cover $v \wedge v'$). Now $u \in [0, \mathrm{top}(P)]$, contradicting (24.1). We conclude $\max(P) = \mathrm{top}(P)$.

We are now ready to show P satisfies the augmentation axiom. To do this, we show P satisfies condition (iii) in Theorem 18.2. Pick any integer i $(2 \leq i \leq D)$, and pick any geodesic path $xyzw$ in P of shape $(i-1, i-2, i-1, i)$. We find y', $z' \in P$ such that $xy'z'w$ is a path of shape $(i-1, i, i-1, i)$. We may assume $x \vee w$ does not exist; otherwise, we are done with $y' := x \vee z$, $z' := z$. Since $\max(P) = \mathrm{top}(P)$, there exists $v \in \mathrm{top}(P)$ such that $w \leq v$. Observe $x \not\leq v$ (otherwise v is an upper bound for x, w), and it follows $x \wedge v = y$. Now x covers $x \wedge v = y$, so by assumption, there exists $v' \in \mathrm{top}(P)$ such that $x \leq v'$, and such that v, v' cover $v \wedge v'$. Observe $w \not\leq v \wedge v'$; otherwise $w \leq v \wedge v' \leq v'$, making v' an upper bound for x, w. Now $v = (v \wedge v') \vee w$, so $v \wedge v'$, w are relative complements in $[w \wedge v', v]$. Recall v covers $v \wedge v'$, so

$$w \text{ covers } w \wedge v' \tag{24.2}$$

by modularity. Observe v' is an upper bound for $x, w \wedge v'$, so $x \vee$

$(w \wedge v')$ exists. Observe $x,\ w \wedge v'$ cover $x \wedge (w \wedge v') = y$, so

$$x \vee (w \wedge v') \text{ covers } x,\ w \wedge v' \tag{24.3}$$

by modularity. Set

$$\begin{aligned} y' &:= x \vee (w \wedge v'), \\ z' &:= w \wedge v'. \end{aligned}$$

Then $xy'z'w$ is a path of shape $(i-1, i, i-1, i)$ by (24.2), (24.3), and the construction. We have now shown P satisfies condition (iii) in Theorem 18.2, so P satisfies the augmentation axiom by that theorem. We have now proved Theorem 24.2.

§25. Any Cartesian product of quantum matroids is a quantum matroid

In this section, we show the property of being a quantum matroid is closed under the Cartesian product operation mentioned above line (1.23). First, a word about notation. Let P, Q denote any posets, and let S, T denote subposets of P, Q, respectively. Then the poset $S \times T$ is isomorphic to the subposet of $P \times Q$ induced on

$$\{xy \mid x \in S,\ \ y \in T\};$$

consequently, we do not distinguish between these posets.

We mention a few elementary facts about the Cartesian product. Let P, Q denote any nonempty posets. Pick any $x, y \in P$, and any $x', y' \in Q$. Then $xx' \wedge_{P\times Q} yy'$ exists if and only if both $x \wedge_P y$, $x' \wedge_Q y'$ exist. In this case,

$$xx' \wedge_{P\times Q} yy' \;=\; x \wedge_P y, x' \wedge_Q y'. \tag{25.1}$$

Similarly, $xx' \vee_{P\times Q} yy'$ exists if and only if both $x \vee_P y$, $x' \vee_Q y'$ exist, and in this case,

$$xx' \vee_{P\times Q} yy' \;=\; x \vee_P y, x' \vee_Q y'. \tag{25.2}$$

Let P, Q denote semilattices. Then $P \times Q$ is a semilattice.

Let P, Q denote posets with 0. Then $P \times Q$ has a 0. Moreover,

$$0_{P\times Q} \;=\; 0_P 0_Q. \tag{25.3}$$

Let P, Q denote ranked posets with 0. Then $P \times Q$ is ranked. Moreover, for all $x \in P$ and for all $x' \in Q$,

$$\operatorname{rank}_{P\times Q}(xx') \;=\; \operatorname{rank}_P(x) \;+\; \operatorname{rank}_Q(x'). \tag{25.4}$$

Let P, Q denote modular atomic lattices. Then $P \times Q$ is a modular atomic lattice.

Lemma 25.1. *Let P, Q denote prematroids. Then $P \times Q$ is a prematroid.*

Proof. It is immediate from our remarks above that $P \times Q$ satisfies R, SL. To see that $P \times Q$ satisfies M, pick any $x \in P$ and any $x' \in Q$. Then

$$[0_{P\times Q}, xx'] \;=\; [0_P, x] \times [0_Q, x']$$

is a Cartesian product of modular atomic lattices, and is therefore a modular atomic lattice.

Lemma 25.2. *Let P, Q denote prematroids. Then for all x, $y \in P$ and for all x', $y' \in Q$,*

(i) $\partial_{P\times Q}(xx', yy') \;=\; \partial_P(x, y) \;+\; \partial_Q(x', y')$,

(ii) $\delta_{P\times Q}(xx', yy') \;=\; \delta_P(x, y) \;+\; \delta_Q(x', y')$.

Proof. (i) Expand each side using (7.12), and evaluate the results using (25.1), (25.4).

(ii) First, we show the inequality $\leq$ holds. By Definition 12.1, there exists $z \in P$ such that $z \vee_P y$ exists, and such that

$$\partial_P(x, z) = \delta_P(x, y).$$

Similarly, there exists $z' \in Q$ such that $z' \vee_Q y'$ exists, and such that

$$\partial_Q(x', z') = \delta_Q(x', y').$$

Now $zz' \vee_{P\times Q} yy'$ exists by our preliminary remarks, so in view of Definition 12.1,

$$\begin{aligned}\delta_{P\times Q}(xx', yy') &\leq \partial_{P\times Q}(xx', zz')\\ &= \partial_P(x, z) + \partial_Q(x', z')\\ &= \delta_P(x, y) + \delta_Q(x', y'),\end{aligned}$$

as desired. Next, we show the inequality $\geq$ holds. By Definition 12.1, there exists an element $zz' \in P \times Q$ such that $zz' \vee_{P\times Q} yy'$ exists, and such that

$$\partial_{P\times Q}(xx', zz') = \delta_{P\times Q}(xx', yy').$$

Observe $z \vee_P y$ exists by our preliminary remarks, so

$$\partial_P(x, z) \geq \delta_P(x, y)$$

by Definition 12.1. Similarly $z' \vee_Q y'$ exists, so

$$\partial_Q(x', z') \geq \delta_Q(x', y').$$

Now

$$\begin{aligned}\delta_{P\times Q}(xx', yy') &= \partial_{P\times Q}(xx', zz') \\ &= \partial_P(x, z) + \partial_Q(x', z') \\ &\geq \delta_P(x, y) + \delta_Q(x', y'),\end{aligned}$$

as desired. We conclude equality holds in (ii), and we are done.

Theorem 25.3. *Let P, Q denote quantum matroids. Then the Cartesian product $P \times Q$ is a quantum matroid.*

Proof. Observe $P \times Q$ is a prematroid by Lemma 25.1, so it remains to show $P \times Q$ satisfies the augmentation axiom. To do this, it suffices by Theorem 18.2(vi) to show the function $\delta_{P\times Q}$ is symmetric in its arguments. But this is an immediate consequence of Lemma 25.2(ii), since δ_P, δ_Q are each symmetric in their arguments by Theorem 18.2(vi).

§26. The radical of a quantum matroid

Definition 26.1. Let P denote a quantum matroid. By the *radical* of P, we mean the element

$$\text{Rad}(P) := \bigwedge_{x\in \text{top}(P)} x. \tag{26.1}$$

We say P is *degenerate* whenever $\text{Rad}(P) > 0$, and *nondegenerate* whenever $\text{Rad}(P) = 0$.

Let P denote a quantum matroid, and write $R = \text{Rad}(P)$. Recall by Corollary 19.2 that the subposet R^+ is a quantum matroid.

Lemma 26.2. *Let P denote a quantum matroid, and write $R =$ $\text{Rad}(P)$. Then*

(i) *R^+ is nondegenerate,*
(ii) $\text{top}(R^+) = \text{top}(P)$.

Proof. Routine.

Lemma 26.3. *Let P denote a quantum matroid, and pick any $x \in P$. Then the following are equivalent.*

(i) $x \leq \text{Rad}(P)$.

(ii) $x \le y$ *for all* $y \in \mathrm{top}(P)$.
(iii) $x \vee y$ *exists for all* $y \in P$.

Proof. (i) $\to$ (ii). Observe $x \le \mathrm{Rad}(P) \le y$ for all $y \in \mathrm{top}(P)$.
(ii) $\to$ (iii). Pick any $y \in P$. By Lemma 4.3, there exists $u \in \mathrm{top}(P)$ such that $y \le u$. Observe $x \le u$ by assumption, so u is an upper bound for x, y.
(iii) $\to$ (i). Pick any $y \in \mathrm{top}(P)$. Observe $x \vee y$ exists by assumption, so $x \le y$. Now $x \le \mathrm{Rad}(P)$ by Definition 26.1. We have now proved Lemma 26.3.

Lemma 26.4. *Let* P *denote a quantum matroid, and suppose the condition* (10.2) *holds. Then*

$$\mathrm{Shadow}(\mathrm{Rad}(P)) = \{x \in A_P | x \vee a \text{ exists for all } a \in A_P\}. \tag{26.2}$$

Proof. $\subseteq$: Pick any $x \in \mathrm{Shadow}(\mathrm{Rad}(P))$. Then by Lemma 26.3(i), (iii), $x \vee a$ exists for all $a \in A_P$
$\supseteq$: Pick any $x \in A_P$, and assume $x \vee a$ exists for all $a \in A_P$. We show x satisfies condition (iii) of Lemma 26.3. Pick any $y \in P$. Certainly $x \vee a$ exists for all $a \in \mathrm{Shadow}(y)$, so $x \vee y$ exists by (10.2). Now x satisfies condition (iii) of Lemma 26.3, so $x \in \mathrm{Shadow}(\mathrm{Rad}(P))$ by Lemma 26.3(i),(iii).

Lemma 26.5. *Let* P *denote a quantum matroid with rank* D. *Suppose that for each* $x \in P$ *such that* $\mathrm{rank}(x) = D - 1$, x *is covered by at least two elements in* $\mathrm{top}(P)$. *Then* $\mathrm{Rad}(P) = 0$.

Proof. Suppose $R := \mathrm{Rad}(P) > 0$, and pick any $z \in \mathrm{top}(P)$. Since $[0, z]$ is relatively complemented, there exists $x \in P$ such that z covers x and $R \not\le x$. By assumption, there exists $z' \in \mathrm{top}(P)$ such that z' covers x and $z' \ne z$. Observe z' is an upper bound for x, R, forcing $z' \ge x \vee R = z$, an impossibility. Hence $\mathrm{Rad}(P) = 0$.

We finish this section with some results concerning the polar spaces from Examples 1.7, 1.8.

Definition 26.6. Let V, $\langle\,,\rangle$ be as in Example 1.7, but assume q is odd in the symmetric bilinear case.

(i) By the *radical* of $\langle\,,\rangle$, we mean

$$\mathrm{Rad}(\langle\,,\rangle) := \{\, u \in V \mid \langle u, v\rangle = 0 \quad \text{for all} \quad v \in V\}.$$

(ii) $\langle\,,\rangle$ is said to be *degenerate* if $\mathrm{Rad}(\langle\,,\rangle) = 0$, and *nondegenerate* otherwise.

Lemma 26.7. *Let* V, $\langle\, , \rangle$, P *be as in Example* 1.7, *but assume* q *is odd in the symmetric bilinear case.*

(i) $\mathrm{Rad}(\langle\, , \rangle) = \mathrm{Rad}(P)$.
(ii) $\langle\, , \rangle$ *is degenerate if and only if* P *is degenerate.*

Proof. Routine.

Definition 26.8. Let V, f, $\langle\, , \rangle_f$ be as in Example 1.8.

(i) By the *radical* of f, we mean

$$\mathrm{Rad}(f) := \{v \in \mathrm{Rad}(\langle\, , \rangle_f) \mid f(v) = 0\}.$$

(ii) f is said to be *degenerate* if $\mathrm{Rad}(f) = 0$, and *nondegenerate* otherwise.

Lemma 26.9. *Let* V, f, P *be as in Example* 1.8.

(i) $\mathrm{Rad}(f) = \mathrm{Rad}(P)$.
(ii) f *is degenerate if and only if* P *is degenerate.*

Proof. Routine.

§27. Line regularity and dual-line regularity

Definition 27.1. Let P denote a quantum matroid of rank D, and let q denote an integer.

(i) Suppose $D \geq 2$. Then P is said to be *q-line regular* whenever for all lines $x \in P$,

$$|\mathrm{Shadow}(x)| = q + 1. \tag{27.1}$$

(ii) Suppose $D \leq 1$. Then P is said to be *q-line regular* whenever q is positive.

Lemma 27.2. *Let* P *denote a* q*-line regular quantum matroid. Then*

$$q \geq 1. \tag{27.2}$$

Proof. If P has rank at least 2, then (27.2) follows from condition M in Definition 4.1. If P has rank 0 or 1, then (27.2) is immediate from Definition 27.1(ii).

Definition 27.3. Let q, j denote integers. We define $\left[{j \atop 1}\right] = \left[{j \atop 1}\right]_q$ by

$$\left[\begin{matrix} j \\ 1 \end{matrix}\right] := \frac{q^j - 1}{q - 1} \qquad \text{if} \quad q \neq 1, \tag{27.3}$$

and

$$\left[\begin{matrix} j \\ 1 \end{matrix}\right] := j \qquad \text{if} \quad q = 1. \tag{27.4}$$

Lemma 27.4. *Let P denote a modular atomic lattice of rank D, and let q denote an integer. Then the following are equivalent.*

(i) *P is q-line-regular.*
(ii) *All intervals in P are q-line-regular.*
(iii) *P^* is q-line-regular.*

Suppose (i)-(iii) *hold. Then*

$$|A_P| = \left[\begin{matrix} D \\ 1 \end{matrix}\right], \tag{27.5}$$

where [ ] *is from* (27.3), (27.4).

Proof. Routine.

Lemma 27.5. *Let P denote a q-line regular quantum matroid. Then for all intervals $I = [x, y]$ in P,*

$$\left[\begin{matrix} i \\ 1 \end{matrix}\right] = \big|\{z \in I \mid z \text{ covers } x\}\big|, \tag{27.6}$$

$$= \big|\{z \in I \mid y \text{ covers } z\}\big|, \tag{27.7}$$

where $i := \operatorname{rank}(y) - \operatorname{rank}(x)$, and where [ ] *is from* (27.3), (27.4).

Proof. The interval $[0, y]$ is q-line-regular by construction, so I, I^* are both q-line regular by Lemma 27.4. I, I^* each have $\left[{i \atop 1}\right]$ atoms by (27.5), and (27.6), (27.7) follow.

Lemma 27.6. *Let P denote a quantum matroid with rank $D \geq 3$, and suppose all the lines of P are thick. Then P is q-line regular for some integer $q \geq 2$. If $D \geq 4$ then q is a prime power.*

Proof. Fix any $x \in \operatorname{top}(P)$. Then $[0, x]$ is a modular atomic lattice, all of whose lines are thick. Applying Theorem 1.12, we find there exists an integer $q \geq 2$ such that $[0,x]$ is isomorphic to a projective plane of order q (if $D=3$) or $L_q(D)$ (if $D \geq 4$). In any case $[0,x]$ is q-line regular. We show P is q-line regular. To this end, suppose there exists $y \in \mathrm{top}(P)$ such that $\partial(x,y)=2$, i.e. x,y are adjacent in the graph $\mathrm{top}(P)$. Then by our preliminary remarks, $[0,y]$ is q'-line regular for some integer $q' \geq 2$. But the intervals $[0,x]$, $[0,y]$ share at least one line in common since $\mathrm{rank}(x \wedge y) = D-1 \geq 2$, so $q=q'$. Since the graph $\mathrm{top}(P)$ is connected by Theorem 22.1(ii), we conclude $[0,z]$ is q-line regular for all $z \in \mathrm{top}(P)$. Now pick any line $u \in P$. By Lemma 4.3, there exists $z \in \mathrm{top}(P)$ such that $u \leq z$. $[0,z]$ is q-line-regular by our above remarks, so u covers exactly $q+1$ points. We have now shown P is q-line regular. Now suppose $D \geq 4$. Then $[0,x]$ is isomorphic to $L_q(D)$, so q is a prime power. We have now proved Lemma 27.6.

Definition 27.7. Let P denote a quantum matroid with rank D.

(i) For all $x \in P$, define

$$\mathrm{Shadow}_D(x) := \{y \mid y \in \mathrm{top}(P),\ \ y \geq x\}. \tag{27.8}$$

(ii) By a *dual-line* in P, we mean any element $x \in P$ such that

$$\mathrm{rank}(x) = D-1. \tag{27.9}$$

Definition 27.8. Let P denote a quantum matroid with rank D, and let β denote an integer.

(i) Suppose $D \geq 1$. Then P is said to be *β-dual-line regular* whenever

$$|\mathrm{Shadow}_D(x)| = \beta + 1 \tag{27.10}$$

for all dual-lines $x \in P$.

(ii) Suppose $D=0$. Then P is said to be *β-dual-line regular* whenever β is nonnegative.

Lemma 27.9. *Let P denote a β-dual-line regular quantum matroid. Then*

$$\beta \geq 0. \tag{27.11}$$

Proof. Immediate from Definition 27.8(i),(ii).

In the next lemma, we consider the case of equality in (27.11).

Lemma 27.10. *Let P denote a quantum matroid. Then the following are equivalent.*

(i) *P is a 0-dual-line regular.*
(ii) *P is a modular atomic lattice.*

Proof. (i) $\rightarrow$ (ii). Suppose P is not a modular atomic lattice. Then by Lemma 22.2, there exists at least two elements in $\mathrm{top}(P)$. The graph on $\mathrm{top}(P)$ is connected by Theorem 22.1(ii), so there exists $x, y \in \mathrm{top}(P)$ that are adjacent in the graph on $\mathrm{top}(P)$. Observe x, y cover $x \wedge y$, so $x \wedge y$ is a dual-line. But x and y are both in $\mathrm{Shadow}_D(x \wedge y)$, contradicting our assumption that P is 0-dual-line regular. We conclude P is a modular atomic lattice.
(ii) $\rightarrow$ (i). Clear.

We close this section with a theorem concerning design matroids that are both line regular and dual-line regular.

Theorem 27.11. *Let P denote a q-line regular, β-dual-line regular design matroid of rank D. Then*

$$|\mathrm{top}(P)| = 1 + \beta \begin{bmatrix} D \\ 1 \end{bmatrix}, \tag{27.12}$$

where [ ] *is from* (27.3), (27.4).

Proof. Fix any element $x \in \mathrm{top}(P)$. Set

$$\Lambda := \{ y \in P \mid x \text{ covers } y \},$$

and observe by Lemma 27.5 that

$$|\Lambda| = \begin{bmatrix} D \\ 1 \end{bmatrix}. \tag{27.13}$$

We now count adjacencies between $\mathrm{top}(P)\backslash\{x\}$ and Λ. Observe by Lemma 22.3(ii), and since P is a semilattice, each element $z \in \mathrm{top}(P)\backslash\{x\}$ covers exactly one element in Λ; namely $x \wedge z$. By (27.10), each element in Λ is covered by exactly β elements in $\mathrm{top}(P)\backslash\{x\}$. It follows

$$\begin{aligned} \beta|\Lambda| &= |\mathrm{top}(P)\backslash\{x\}| \\ &= |\mathrm{top}(P)| - 1, \end{aligned} \tag{27.14}$$

and the result follows from (27.13), (27.14).

§28. Zig-zag regularity

Definition 28.1. Let P denote a quantum matroid of rank $D \geq 2$. For each integer i $(2 \leq i \leq D)$, let Δ_i denote the set of ordered pairs

$$\Delta_i := \{xy | x, y \in P, \text{rank}(x) = i-1, \text{rank}(y) = i, \partial(x,y) = 3,$$

$$x \vee y \text{ does not exist}\}$$

$$= \{xy | x, y \in P, \rho(x,y) = i-2, \gamma(x,y) = 0, \delta(x,y) = 1, \gamma(y,x) = 1\}.$$

In Lemma 28.4, we define a function *zig-zag* : $\Delta_i \to \mathbb{Z}$, but first, let us consider when $\Delta_i \neq \emptyset$.

Lemma 28.2. *Let P denote a quantum matroid with rank $D \geq 2$. For each integer i $(1 \leq i \leq D)$, let $\tilde{\Delta}_i$ denote the set of ordered pairs*

$$\tilde{\Delta}_i : = \{xy \mid x, y \in P, \quad \text{rank}(x) = i, \quad \text{rank}(y) = i, \quad \partial(x,y) = 2,$$
$$x \vee y \textit{ does not exist}\}.$$

Then the following statements (i)–(iii) *hold.*

(i) $\Delta_i \neq \emptyset \quad \leftrightarrow \quad \tilde{\Delta}_i \neq \emptyset \qquad (2 \leq i \leq D)$.

(ii) $\tilde{\Delta}_{i-1} \neq \emptyset \quad \to \quad \Delta_i \neq \emptyset \qquad (2 \leq i \leq D)$.

(iii) *Suppose there exists atoms x, $y \in P$ such that $x \vee y$ does not exist. Then*

$$\Delta_i \neq \emptyset \qquad (2 \leq i \leq D). \tag{28.1}$$

Proof. (i) $\to$: Pick any $xy \in \Delta_i$. By Theorem 18.2(ii) and Definition 28.1, there exists elements $z, w \in P$ such that $xzwy$ is a path with shape $(i-1, i, i-1, i)$. Recall $\partial(x,y) = 3$ by Definition 28.1, so $\partial(z,y) = 2$. Recall $x \vee y$ does not exist by Definition 28.1. It follows $z \vee y$ does not exist; otherwise $z \vee y$ is an upper bound for x, y. Now $zy \in \tilde{\Delta}_i$.

$\leftarrow$: Pick any $uv \in \tilde{\Delta}_i$. Observe u covers at least two elements of P since $\text{rank}(u) = i \geq 2$; in particular there exists an element $x \in P$ such that u covers x and $x \neq u \wedge v$. Observe $\partial(x,v) \in \{1,3\}$ by (7.3). In fact $\partial(x,v) = 3$; otherwise v covers x, making x a lower bound for u, v, and forcing $x = u \wedge v$. We claim $x \vee v$ does not exist. Suppose $x \vee v$ exists. Observe xuv is geodesic by our above remarks; it follows $u \leq x \vee v$ by Lemma 7.9(i),(ii). In this case $x \vee v$ is an upper bound for u, v, so $u \vee v$ exists, contradicting our assumptions. We conclude $x \vee v$ does not exist. Now $xv \in \Delta_i$ by Definition 28.1.

(ii) Pick $uv \in \tilde{\Delta}_{i-1}$. By Lemma 4.3, there exists $y \in P$ such that y covers v. Observe $u \vee y$ does not exist; otherwise $u \vee y$ is an upper bound for u, v. In particular $u \not\leq y$. Now $\partial(u, y) = 3$ by (7.3), and it follows $uy \in \Delta_i$ by Definition 28.1.

(iii) Observe $xy \in \tilde{\Delta}_1$, so $\tilde{\Delta}_1 \neq \emptyset$. The result now follows from (i), (ii).

Lemma 28.3. *Let P denote a quantum matroid with rank $D \geq 2$. Then the following are equivalent.*

(i) *P is a modular atomic lattice.*
(ii) *$x \leq y$ for all $x \in A_P$ and all $y \in \mathrm{top}(P)$.*
(iii) *$\tilde{\Delta}_D = \emptyset$.*
(iv) *$\Delta_D = \emptyset$.*

Proof. (i) $\rightarrow$ (ii). Clear since $\mathrm{top}(P) = \{1\}$.

(ii) $\rightarrow$ (iii). Suppose there exists $uy \in \tilde{\Delta}_D$, and let x denote a relative complement of $u \wedge y$ is $[0, u]$. Observe u covers $u \wedge y$ by the definition of $\tilde{\Delta}_D$, so x is an atom by modularity. Observe $x \not\leq y$; otherwise x is a lower bound for u, y, forcing $x \leq u \wedge y$, and contradicting the construction.

(iii) $\rightarrow$ (i). Suppose P is not a modular atomic lattice. Then P is not 0-dual-line regular by Lemma 27.10, so there exists a dual-line $w \in P$, and distinct elements $u, v \in \mathrm{top}(P)$ such that $w \leq u$, $w \leq v$. Observe $\partial(u, v) = 2$ by the construction, so $uv \in \tilde{\Delta}_D$.

(iii) $\leftrightarrow$ (iv). Immediate from Lemma 28.2(i).

Lemma 28.4. *Let P denote a quantum matroid with rank $D \geq 2$. For all integers i $(2 \leq i \leq D)$, and for all elements $xy \in \Delta_i$, the sets*

(i) $\mathrm{top}(x_y^+)$
(ii) $y \star x$
(iii) $\{p \mid p$ *is a path in P with endpoints x, y and shape* $(i-1, i, i-1, i)\}$

all have the same cardinality.

We denote this cardinality by *zig-zag*(x, y).

Proof. The sets (i), (ii) have the same cardinality by Corollary 15.7(iii). The sets (ii), (iii) also have the same cardinality, since the map $u \rightarrow x, x \vee u, u, y$ is a bijection from the set $y \star x$ to the set in (iii).

Definition 28.5. Let P denote a quantum matroid with rank D, and let α denote an integer.

(i) Assume $D \geq 2$, and that P is not a modular atomic lattice. Then P is said to be *α-zig-zag regular* whenever

$$\text{zig-zag}(x, y) = \alpha + 1 \tag{28.2}$$

for all $xy \in \Delta_D$.

(ii) Assume $D \leq 1$, or that P is a modular atomic lattice. Then P is said to be *α-zig-zag regular* whenever α is nonnegative.

Lemma 28.6. *Let P denote an α-zig-zag regular quantum matroid. Then*

$$\alpha \geq 0. \tag{28.3}$$

Proof. Immediate from Definition 28.5.

In the next section, we consider the case of equality in (28.3). For now, we mention a few other inequalities concerning α.

Lemma 28.7. *Let P denote an α-zig-zag regular quantum matroid with rank $D \geq 2$, and assume P is not a modular atomic lattice.*

(i) *Suppose P is q-line regular. Then*

$$\alpha \leq q. \tag{28.4}$$

(ii) *Suppose P is β-dual-line regular. Then*

$$\alpha \leq \beta. \tag{28.5}$$

Proof. By Lemma 28.3(i),(iv), there exists an element $xy \in \Delta_D$. To see (i), observe the interval $[x \wedge y, y]$ has rank 2, so by Lemma 28.4(ii), Corollary 27.5,

$$\begin{aligned} \alpha + 1 &= \text{zig-zag}(x, y) \\ &= |y \star x| \\ &\leq \left|\{v \mid v \in P, \;\; x \wedge y < v < y\}\right| \\ &= q + 1. \end{aligned}$$

To see (ii), observe by Lemma 28.4(i) that

$$\begin{aligned} \alpha + 1 &= \text{zig-zag}(x, y) \\ &= |\text{top}(x_y^+)| \\ &\leq |\text{Shadow}_D(x)| \\ &= \beta + 1. \end{aligned}$$

§29. The 0-zig-zag regular quantum matroids

The purpose of this section and the next is to establish that a nondegenerate 0-zig-zag regular quantum matroid is the same thing as a Tits polar space.

Theorem 29.1. *Let P denote a quantum matroid. Then the following are equivalent.*

(i) *P is 0-zig-zag regular.*
(ii) *x_y^+ is a modular atomic lattice for all $x, y \in P$.*
(iii) *$|x \star y| = 1$ for all $x, y \in P$.*

Proof. Let D denote the rank of P.

(i) $\rightarrow$ (ii). Suppose there exists a pair $x, y \in P$ such that x_y^+ is not a modular atomic lattice. We may assume

$$\text{rank}(y) - \partial(x, x \wedge y) \quad \text{is maximal} \tag{29.1}$$

among all such pairs.

We first claim $y \in \text{top}(P)$. Suppose not. Then by Lemma 4.3, there exists an element $u \in P$ such that $y < u$. Observe

$$x \wedge y \leq x \wedge u \leq x,$$

so

$$\partial(x, x \wedge u) \leq \partial(x, x \wedge y).$$

Now

$$\text{rank}(u) - \partial(x, x \wedge u) > \text{rank}(y) - \partial(x, x \wedge y),$$

so x_u^+ is a modular atomic lattice by (29.1). Observe x_y^+ is a submatroid of x_u^+ by Lemma 15.9, so x_y^+ is a modular atomic lattice by Lemma 8.5, Lemma 9.1. This contradicts our assumptions, so $y \in \text{top}(P)$.

Next, we claim $\partial(x, x \wedge y) \geq 2$. Certainly $\partial(x, x \wedge y) \neq 0$; otherwise $x \leq y$, implying $x_y^+ = [x, y]$ is a modular atomic lattice. Suppose $\partial(x, x \wedge y) = 1$. We obtain a contradiction to Lemma 27.10 by showing x_y^+ is 0-dual-line regular. To do this, we pick any dual-line w in x_y^+, and show

$$|\text{top}(x_y^+) \cap w^+| = 1.$$

Observe $\text{top}(x_y^+) \subseteq \text{top}(P)$ by Theorem 18.2(v), so

$$\text{rank}_P(w) = D - 1. \tag{29.2}$$

Observe $x, w \wedge y$ are relative complements in $[x \wedge y, w]$ by Lemma 15.3(ii), and x covers $x \wedge y$, so

$$w \text{ covers } w \wedge y \tag{29.3}$$

by modularity. Now $wy \in \Delta_D$ by (29.2), (29.3), so by Lemma 15.8(i),

$$\begin{aligned}|\mathrm{top}(x_y^+) \cap w^+| &= |\mathrm{top}(w_y^+)| \\ &= \text{zig-zag}(wy) \\ &= 1,\end{aligned}$$

as desired. We have now shown x_y^+ is 0-dual-line regular, so x_y^+ is a modular atomic lattice by Lemma 27.10. This contradicts the construction, so $\partial(x, x \wedge y) \neq 1$. We conclude $\partial(x, x \wedge y) \geq 2$.

Since $\partial(x, x \wedge y) \geq 2$, there exists an element $s \in P$ such that $x \wedge y < s < x$. Observe $s \wedge y = x \wedge y$ by Lemma 13.4(i),(iii), so

$$\begin{aligned}\mathrm{rank}(y) - \partial(s, s \wedge y) &= \mathrm{rank}(y) - \partial(s, x \wedge y) \\ &> \mathrm{rank}(y) - \partial(x, x \wedge y),\end{aligned}$$

implying s_y^+ is a modular atomic lattice by (29.1). Let z denote a maximal element of s_y^+. Observe $z \in \mathrm{top}(P)$ by Theorem 18.2(v), and $s = x \wedge z$ by Lemma 7.4, so

$$\begin{aligned}\mathrm{rank}(z) - \partial(x, x \wedge z) &= \mathrm{rank}(y) - \partial(x, s) \\ &> \mathrm{rank}(y) - \partial(x, x \wedge y),\end{aligned}$$

implying x_z^+ is a modular atomic lattice by (29.1).

We claim $x_y^+ \subseteq x_z^+$ (in fact equality holds, but we will not need this). To prove the claim, we pick $u \in x_y^+$ and show $u \in x_z^+$. Set

$$p := s \vee (u \wedge y).$$

Observe $p \in s_y^+$ by Lemma 17.1, so spz is geodesic. Observe szy is geodesic since $z \in s_y^+$, so $spzy$ is geodesic. In particular pzy is geodesic. Observe upy is geodesic by Lemma 17.1, Theorem 17.2(i),(iii), so $upzy$ is geodesic. In particular uzy is geodesic. Observe xuy is geodesic since $u \in x_y^+$, so $xuzy$ is geodesic. In particular xuz is geodesic, so $u \in x_z^+$, as desired. We have now shown $x_y^+ \subseteq x_z^+$. Now x_y^+ is geodesically closed in x_z^+ by Lemma 15.2(ii), and we saw x_z^+ is a modular atomic lattice, so x_y^+ is a modular atomic lattice by Lemma 8.5. This contradicts our assumption, and we are done.

(ii) $\to$ (i). For all $xy \in \Delta_D$,

$$\begin{aligned}\text{zig-zag}(xy) &= |\text{top}(x_y^+)| \\ &= 1.\end{aligned}$$

(ii) $\leftrightarrow$ (iii). Recall by Lemma 22.2(ii),(iii) that for all $x, y \in P$, x_y^+ is a modular atomic lattice if and only if $|\text{top}(x_y^+)| = 1$. But $|\text{top}(x_y^+)| = |y \star x|$ by Lemma 15.7(iii), so the result follows. This proves Theorem 29.1.

We now modify Theorem 24.2 using the above theorem, to obtain a characterization of the 0-zig-zag regular quantum matroids.

Theorem 29.2. *Let P denote a prematroid of rank D. Then the following are equivalent.*

(i) *P satisfies the augmentation axiom AU in Definition 4.1, and P is 0-zig-zag regular.*

(ii) *For all atoms $x \in P$ and for all $y \in \text{top}(P)$ such that $x \not\leq y$, x_y^+ is a modular atomic lattice with rank $D-1$.*

(iii) *For all $u \in P$ and all $v \in \text{top}(P)$ such that u covers $u \wedge v$, there exists a unique $v' \in \text{top}(P)$ such that $u \leq v'$ and such that v, v' cover $v \wedge v'$.*

Proof. (i) $\to$ (ii). Let x, y be given. P is a quantum matroid by Definition 4.1, and 0-zig-zag regular by assumption, so x_y^+ is a modular atomic lattice by Theorem 29.1(ii). x_y^+ has rank $D-1$ by Theorem 24.2.

(ii) $\to$ (iii). Very similar to the proof of Theorem 24.2 (ii) $\to$ (iii).

(iii) $\to$ (i). Observe P satisfies condition (iii) of Theorem 24.2, so P satisfies the augmentation axiom by that theorem. We show P is 0-zig-zag regular. Pick $xy \in \Delta_D$. Then x covers $x \wedge y$ by Definition 28.1, so by assumption, there exists a unique $y' \in \text{top}(P)$ such that $x \leq y'$, and such that y, y' cover $y \wedge y'$. Put another way, there exists a unique path with endpoints x, y and shape $(D-1, D, D-1, D)$, so zig-zag$(x, y) = 1$. We have now shown P is 0-zig-zag regular. This proves Theorem 29.2.

Next, we consider when a 0-zig-zag regular quantum matroid is nondegenerate.

Theorem 29.3. *Let P denote a 0-zig-zag regular quantum matroid with rank D. Then the following are equivalent.*

(i) $\text{Rad}(P) = 0$.

(ii) *For all atoms* $a \in P$*, there exists an atom* $b \in P$ *such that* $a \vee b$ *does not exist.*
(iii) *For all* $x \in P$*, there exists* $y \in \mathrm{top}(P)$ *such that* $x \wedge y = 0$.
(iv) *For all dual-lines* $x \in P$*,* x *is covered by at least two elements in* $\mathrm{top}(P)$.
(v) *There exists* $x, y \in \mathrm{top}(P)$ *such that* $x \wedge y = 0$.

Proof. (i) $\rightarrow$ (ii). Let the atom a be given. Certainly $a \not\leq \mathrm{Rad}(P)$, so by Lemma 26.3(iii), there exists an element $y \in P$ such that $a \vee y$ does not exist. Observe $\delta(a, y) = 1$ by construction, so $\delta(y, a) = 1$ by Theorem 18.2(vi). Now y covers $y \star a$ by Definition 12.1 and Theorem 12.3. Let b denote a relative complement of $y \star a$ in $[0, y]$. Then b is an atom by modularity. To show $a \vee b$ does not exist, we show $b \notin y_a^-$. But this holds, since $y_a^- = [0, y \star a]$ by Theorem 19.3(i), and $b \notin [0, y \star a]$ by construction.

(ii) $\rightarrow$ (iii). Let x be given, and pick an element $z \in \mathrm{top}(P)$ with

$$\rho(x, z) \text{ minimal.} \tag{29.4}$$

We assume $\rho(x, z) > 0$ and get a contradiction. By construction, there exists an atom $a \in P$ such that $a \leq x \wedge z$. By (ii), there exists an atom $b \in P$ such that $a \vee b$ does not exist. Observe $b \not\leq z$; otherwise z is an upper bound for a, b. Now b covers $b \wedge z = 0$, so by Theorem 29.2(iii), there exists an element

$$y \in \mathrm{top}(P) \tag{29.5}$$

such that $b \leq y$, and such that y, z cover $y \wedge z$. Set $h := y \wedge z$. By Theorem 13.5, $\rho(x, z) - \rho(x, h)$ equals 0 or 1. Suppose for the moment $\rho(x, z) = \rho(x, h)$. Then $x \wedge z \leq h$ by Lemma 13.4(i),(iv), implying y is an upper bound for a, b, a contradiction. Hence

$$\rho(x, z) = \rho(x, h) + 1, \tag{29.6}$$

and this forces

$$z \in h_x^+ \tag{29.7}$$

by Theorem 15.3(i),(iv). Observe $y \vee z$ does not exist. Now by (29.7), and since h_x^+ is a modular atomic lattice by Theorem 29.1,

$$y \notin h_x^+. \tag{29.8}$$

By Theorem 13.5 and since y covers h, $\rho(x, y) - \rho(x, h)$ equals 0 or 1. Suppose for the moment $\rho(x, y) = \rho(x, h) + 1$. Then $y \in h_x^+$ by

Lemma 15.3(iv), contradicting (29.8). Hence

$$\rho(x, y) = \rho(x, h). \tag{29.9}$$

Now (29.5), (29.6), (29.9) contradict (29.4).

(iii) $\to$ (iv). Pick any dual-line $x \in P$. We show x is covered by at least two elements in $\operatorname{top}(P)$. By Lemma 4.3, there exists an element $y \in \operatorname{top}(P)$ that covers x. By assumption, there exists an element $z \in P$ such that $y \wedge z = 0$. By Theorem 29.1, x_z^+ is a modular atomic lattice. Let y' denote the unique maximal element in x_z^+. Then $y' \in \operatorname{top}(P)$ by Theorem 18.2(v). It remains to check $y \neq y'$. Observe $x \geq y \wedge z = 0$, so yxz is geodesic by Lemma 13.4(i),(ii). But $xy'z$ is geodesic by the construction and (15.1), so $y \neq y'$, as desired.

(iv) $\to$ (i). This is just Lemma 26.5.

(iii) $\to$ (v). Clear.

(v) $\to$ (i). Pick any $x, y \in P$ such that $x \wedge y = 0$. Then by Definition 26.1,

$$\begin{aligned} \operatorname{Rad}(P) &= \bigwedge_{u \in \operatorname{top}(P)} u \\ &\leq x \wedge y \\ &= 0, \end{aligned}$$

so $\operatorname{Rad}(P) = 0$. We have now proved Theorem 29.3.

§30. Tits polar spaces

In this section, we show that a nondegenerate 0-zig-zag regular quantum matroid is the same thing as a Tits polar space.

Our first result concerns atomic semilattices. It will allow us to shift our point of view a bit, bringing it into line with how Tits polar spaces are traditionally viewed.

Lemma 30.1. *Let P denote an atomic semilattice, and define a poset*

$$\tilde{P} := \{\operatorname{Shadow}(x) \mid x \in P\}, \tag{30.1}$$

with partial order by inclusion. Then the map

$$\begin{aligned} P &\to \tilde{P} \\ x &\to \operatorname{Shadow}(x) \end{aligned}$$

is an isomorphism of posets.

Proof. The map is clearly onto $\tilde{P}$. The map is 1–1, and respects the partial order, by Lemma 9.3(ii). This proves Lemma 30.1.

Passing from P to $\tilde{P}$, Lemma 30.1 allows us to view any atomic semilattice as a collection of distinct subsets of A_P, with partial order defined by inclusion. We adopt this point of view for the remainder of the section.

Definition 30.2 [Ti, p102]. A *Tits polar space* is a collection P of distinct subsets of a set A (of *points*), partially ordered by inclusion, such that the following axioms hold.

PS1: P is closed under taking intersections, and has all the single points of A as its collection of minimal non-empty members.

PS2: All unrefinable chains in P have the same length D.

PS3: If x is a maximal member of P, then x, together with all the elements of P which it contains, is a modular atomic lattice of rank D.

PS4: Given a point $x \in P$, and a maximal member y of P that does not contain x, there exists a unique maximal member y' of P such that y, y' cover $y \wedge y'$. $y \wedge y'$ contains all the elements of y that lie together with x in some element of P.

PS5: There exists two maximal elements in P that have empty intersection.

The scalar D is the *rank* of P.

We are now ready for the main theorem of this section.

Theorem 30.3. *For any nonnegative integer D, the following are equivalent.*

(i) *P is a nondegenerate, 0-zig-zag regular quantum matroid of rank D.*

(ii) *P is a Tits polar space of rank D.*

Proof. (i) → (ii). Recall by Lemma 4.3 that

$$\max(P) = \mathrm{top}(P). \tag{30.2}$$

To show P is a Tits polar space of rank D, we check P satisfies PS1–PS5.

PS1: P is closed under taking intersections by Lemma 9.3(iii). P has $A = A_P$ as its collection of minimal nonempty members by the definition of A_P.

PS2: Immediate since P is ranked, $\mathrm{rank}(P) = D$, and since (30.2) holds.

PS3: Pick any $x \in \max(P)$. Then $\mathrm{rank}(x) = D$ by (30.2), so $[0, x]$ is a modular atomic lattice of rank D by condition M of Definition 4.1.

PS4: Pick any point $x \in A$, and any element $y \in \max(P)$ such that $x \not\leq y$. Then x covers $x \wedge y = 0$, so by Theorem 29.2(iii) and (30.2), there exists a unique element $y' \in \max(P)$ such that $x \leq y'$, and such that y, y' cover $y \wedge y'$. To see that $y \wedge y'$ contains all the elements of y that lie together with x in some element of P, observe by Theorem 19.3(i), Theorem 29.1(iii) that

$$\begin{aligned}\{a \in A \mid a \leq y, \quad a \vee x \text{ exists in } P\} &= y_x^- \cap A \\ &= [0, y \star x] \cap A \\ &= [0, y \wedge y'] \cap A.\end{aligned}$$

PS5: Immediate from Theorem 29.3(i),(v), and (30.2).

We have now shown P satisfies PS1–PS5, so P is a Tits polar space of rank D.

(ii) $\to$ (i). We first show P is a prematroid, by showing P satisfies the conditions SL, R, M in Definition 4.1.

SL: P is a semilattice by PS1.

R: P is ranked by PS2.

M: Pick any $x \in P$, and pick any $y \in \max(P)$ such that $x \leq y$. Observe $[0, y]$ is a modular atomic lattice by PS3, so the interval $[0, x]$ is a modular atomic lattice.

We have now shown P is a prematroid. In fact P is transversal and has rank D by PS2.

We now show P satisfies the augmentation axiom, and is 0-zig-zag regular. To do this, we show P satisfies condition (ii) in Theorem 29.2. Pick any atom $x \in P$ and any element $y \in \mathrm{top}(P)$ such that $x \not\leq y$. We show x_y^+ is a modular atomic lattice of rank $D - 1$. By PS4, and since P is transversal, there exists a unique $y' \in \mathrm{top}(P)$ such that $x \leq y'$, and such that y, y' cover $y \wedge y'$. Moreover, $y \wedge y'$ contains all the elements of y that lie together with x in some element of P. It follows

$$y \star x = y \wedge y' \tag{30.3}$$

and

$$\mathrm{Shadow}(y \wedge y') = y_x^- \cap A. \tag{30.4}$$

We show $y_x^- = [0, y \wedge y']$. To see the inclusion $\supseteq$, observe by Lemma 15.4(iii) and (30.3) that

$$\begin{aligned} y_x^- &\supseteq [0, y \star x] \\ &= [0, y \wedge y']. \end{aligned}$$

To see the inclusion $\subseteq$, pick any $z \in y_x^-$. Then

$$\begin{aligned} \mathrm{Shadow}(z) &\subseteq y_x^- \cap A \\ &= \mathrm{Shadow}(y \wedge y') \end{aligned}$$

by (30.4), so $z \in [0, y \wedge y']$ by Lemma 9.3(ii). We have now shown $y_x^- = [0, y \wedge y']$. By this, and since y covers $y \wedge y'$, it follows y_x^- is a modular atomic lattice of rank $D-1$. Recall x_y^+, y_x^- are isomorphic by Theorem 15.5; in particular x_y^+ is a modular atomic lattice of rank $D-1$, as desired. We have now shown P satisfies condition (ii) in Theorem 29.2; it follows P satisfies the augmentation axiom, and is 0-zig-zag regular. Now P is a 0-zig-zag regular quantum matroid by Definition 4.1. P is nondegenerate by Theorem 29.3(i),(v) and PS5. This proves Theorem 30.3.

The Tits polar spaces (and hence the nondegenerate 0-zig-zag regular quantum matroids) are essentially classified by J. Tits. In the following two theorems we present the classification in the line regular case.

Theorem 30.4. *Let D denote an integer at least 2. Then the following are equivalent.*

(i) *P is a 1-line regular Tits polar space of rank D.*
(ii) *There exists integers $N_1, N_2, \ldots, N_D$, all at least 2, such that P is isomorphic to $C(N_1, N_2, \ldots, N_D)$.*

Suppose (i), (ii) *hold. Then P is β-dual-line regular if and only if*

$$N_i = \beta + 1 \qquad (1 \leq i \leq D).$$

In this case P is isomorphic to the Hamming matroid $H(D, \beta+1)$ listed in Example 40.1(2).

Proof. Routine.

Theorem 30.5([Ti]). *Let D denote an integer at least 4, and let q denote an integer at least 2. Then the following are equivalent.*

(i) *P is a q-line regular Tits polar space of rank D.*
(ii) *q is a prime power, and P is isomorphic to a classical polar space of rank D over the field $GF(q)$.*

(*See Example* 40.1(5).)

Suppose (i), (ii) *hold. Then* P *is* $q^{1+\varepsilon}$*-dual-line regular, where* ε *is given in Example* 40.1(5).

§31. The α-zig-zag regular quantum matroids with $\alpha > 0$

In the previous two sections, we considered the 0-zig-zag regular quantum matroids. In this section, we say a bit about the α-zig-zag regular quantum matroids with $\alpha > 0$.

Theorem 31.1. *Let* P *denote a quantum matroid with rank* $D \geq 2$. *Suppose* P *is* α*-zig-zag regular for some integer* $\alpha > 0$, *but assume* P *is not a modular atomic lattice. Then*

(i) P *is dual-line regular,*
(ii) $\mathrm{Rad}(P) = 0$.

Proof. (i) Pick any $x, y \in P$ such that $\mathrm{rank}(x) = D - 1$, $\mathrm{rank}(y) = D - 1$. To show P is dual-line regular, it suffices to show

$$|\mathrm{Shadow}_D(x)| = |\mathrm{Shadow}_D(y)|. \tag{31.1}$$

First, consider the special case where $x \vee y$ exists. Here we may assume $x \neq y$; otherwise (31.1) clearly holds. Observe by Lemma 28.4(i) and Definition 28.5 that for each $u \in \mathrm{Shadow}_D(x) \backslash \{x \vee y\}$, there exists exactly α elements $v \in \mathrm{Shadow}_D(y) \backslash \{x \vee y\}$ such that $\partial(u, v) = 2$. This remains true if we interchange the roles of x and y, so

$$|\mathrm{Shadow}_D(x) \backslash \{x \vee y\}| \alpha = |\mathrm{Shadow}_D(y) \backslash \{x \vee y\}| \alpha.$$

Line (31.1) is immediate since $\alpha > 0$. We now have (31.1) in our special case. To show (31.1) holds in general, we construct a path p with endpoints x, y and shape

$$(D-1, D, D-1, D, \dots, D-1, D, D-1). \tag{31.2}$$

To obtain p, recall by Lemma 4.3 that there exists an element $y' \in \mathrm{top}(P)$ such that y' covers y. By Theorem 18.2(iv), there exists a geodesic up-flat-down path p' in P with endpoints x, y'. By the construction, p' must have shape $(D-1, D, D-1, D, \dots, D-1, D)$. Appending y to the end of p', we obtain a path p with endpoints x, y and shape (31.2), as desired. Write $p = (x = x_0, x_1, \dots, x_{d-1}, x_d = y)$ $(x_0, x_1, \dots, x_d \in P)$, and observe by (31.2) that d is even. Moreover

$$x_i \vee x_{i+2} \text{ exists} \qquad (0 \leq i \leq d-2, \ i \text{ even}),$$

so by the above special case,

$$|\mathrm{Shadow}_D(x_i)| = |\mathrm{Shadow}_D(x_{i+2})| \qquad (0 \le i \le d, \quad i \text{ even}).$$

Line (31.1) follows, so P is dual-line regular, as desired.

(ii) By (i), P is β-dual-line regular for some integer β. Observe $\beta \ge \alpha > 0$ by Lemma 28.7(ii), so $\mathrm{Rad}(P) = 0$ by Lemma 26.5.

§32. The definition of a regular quantum matroid

Definition 32.1. A quantum matroid P is said to be *regular*, with *parameters* (D, q, α, β), whenever the following four conditions hold.

(i) P has rank D.
(ii) P is q-line regular.
(iii) P is α-zig-zag regular.
(iv) P is β-dual-line regular.

Let us consider a few very special cases. Any quantum matroid with rank $D \le 1$ is regular. However, the parameters q, α are not uniquely defined in this case. Similarly, any q-line regular modular atomic lattice is regular, but the parameter α is not uniquely defined in this case. In contrast to the above two cases, consider a regular quantum matroid P with rank $D \ge 2$, that is not a modular atomic lattice. Then the parameters q, α, β are uniquely defined.

Some results concerning regular quantum matroids do not hold unless the parameters are uniquely defined, so we make the following definition.

Definition 32.2. A quantum matroid P is said to be *trivial* whenever P has rank $D \le 1$, or P is a modular atomic lattice.

In Example 4.2, we characterized the quantum matroids of rank 2. Below we present a similar result concerning the regular quantum matroids of rank 2.

Example 32.3. Let q, α, β denote integers, and let P denote a poset. Then P is a regular quantum matroid with parameters $(2, q, \alpha, \beta)$ if and only if P has a 0, and the following conditions (i)–(vi) hold.

(i) P is ranked and $\mathrm{rank}(P) = 2$.
(ii) For any distinct points $x, y \in P$, there exists at most one line $z \in P$ such that $x \le z$, $y \le z$.
(iii) Each line in P covers exactly $q + 1$ points in P.

(iv) Each point in P is covered by exactly $\beta+1$ lines in P.
(v) For each point $x \in P$ and each line $y \in P$ such that $x \not\leq y$, there exists exactly $\alpha+1$ pairs $x'y'$, such that x' is a point in P, y' is a line in P, and $x \leq y' \geq x' \leq y$.
(vi) $q \geq 1$, $\alpha \geq 0$, $\beta \geq 0$.

Note. A regular quantum matroid with parameters $(2, q, \alpha, \beta)$ is essentially the same thing as an (R, K, T)-*partial geometry*, where $R := \beta + 1$, $K := q+1$, and $T := \alpha + 1$ [Bo].

The following theorem gives a characterization of the regular quantum matroids of arbitrary rank $D \geq 2$. Compare this with Theorem 23.1.

Theorem 32.4. *Let D, q, α, β denote integers with $D \geq 2$, and let P denote a poset. Then P is a regular quantum matroid with parameters (D, q, α, β) if and only if* (i)–(iv) *hold below.*

(i) *P is a prematroid of rank D.*
(ii) *For all $x \in P$, there exists $x' \in \mathrm{top}(P)$ such that $x \leq x'$.*
(iii) *For all $x \in P$ such that $\mathrm{rank}(x) \leq D-2$, the poset induced on $x^+ \backslash \{x\}$ is connected.*
(iv) *For all $x \in P$ such that $\mathrm{rank}(x) = D-2$, the poset x^+ is a regular quantum matroid with parameters $(2, q, \alpha, \beta)$.*

Proof. First suppose P is a regular quantum matroid with parameters (D, q, α, β). Then the above conditions (i)–(iii) hold by Theorem 23.1. To see that the above condition (iv) holds, pick any $x \in P$ such that $\mathrm{rank}(x) = D-2$. Then x^+ is a quantum matroid of rank 2 by Theorem 23.1(iv). Observe x^+ is q-line regular by Lemma 27.4(i),(ii), and since P is q-line regular. Observe x^+ is α-zig-zag regular by Lemma 28.4, Definition 28.5, and since P is α-zig-zag regular. Observe x^+ is β-dual-line regular by Definition 27.7, and since P is β-dual-line regular. Now x^+ is a regular quantum matroid with parameters $(2, q, \alpha, \beta)$, as desired.

Conversely, suppose the above conditions (i)–(iv) hold. Observe P satisfies the conditions (i)–(iv) in Theorem 23.1, so P is a quantum matroid by that theorem. We check P is regular. To show P is q-line regular, it suffices to show $[0, y]$ is q-line regular for all $y \in \mathrm{top}(P)$. Observe $[0, y]^*$ is q-line regular by condition (iv) above, so $[0, y]$ is q-line regular by Lemma 27.4(i),(iii). We have now shown P is q-line regular. It is immediate from the construction that P is α-zig-zag regular and β-dual-line regular, so P is a regular quantum matroid with parameters (D, q, α, β), as desired.

§33. Formulae for $|A_P|$, $|\mathbf{top}(P)|$

In this section, we assume P is a regular quantum matroid with parameters (D, q, α, β), and compute $|A_P|$, $|\mathrm{top}(P)|$ in terms of D, q, α, β. First, we introduce some notation.

Definition 33.1. For all integers j, q, α, we define $t_j = t_j(q, \alpha)$ by

$$t_j := 1 + \alpha \begin{bmatrix} j \\ 1 \end{bmatrix}, \tag{33.1}$$

where $\left[\begin{smallmatrix} j \\ 1 \end{smallmatrix}\right]$ is from (27.3), (27.4).

Lemma 33.2. *With the notation of Definition* 33.1:

(i) $t_{j+1} - t_j = \alpha q^j \qquad (j \in \mathbb{Z})$.

(ii) *Assume* $\alpha \geq 0$, $q \geq 1$. *Then*

$$1 = t_0 \leq t_1 \leq t_2 \leq \ldots. \tag{33.2}$$

Proof. (i) Immediate from (27.3), (27.4), (33.1).
(ii) Immediate from (33.1) and (i) above.

Lemma 33.3. *Let P denote a regular quantum matroid with parameters (D, q, α, β). Then for all $x \in P$, and for all $y \in \mathrm{top}(P)$ such that x covers $x \wedge y$,*

$$|\mathrm{top}(x_y^+)| = t_{D-i}, \tag{33.3}$$

where

$$i := \mathrm{rank}(x). \tag{33.4}$$

Proof. By Definition 33.1 and Theorem 27.11, it suffices to show x_y^+ is a q-line regular, α-dual-line regular design-matroid, with rank $D - i$. Observe by Lemma 27.4(i),(ii) that x_y^+ is q-line regular. To see that x_y^+ is α-dual-line regular, we pick any dual-line z of x_y^+, and show

$$|z^+ \cap \mathrm{top}(x_y^+)| = \alpha + 1. \tag{33.5}$$

Observe

$$\mathrm{top}(x_y^+) \subseteq \mathrm{top}(P) \tag{33.6}$$

by Theorem 18.2(v), so z is a dual-line of P. Now $\partial(y,z)=3$ by the construction, so $zy \in \Delta_D$ by Definition 28.1. Now by Lemma 28.4(i) and Lemma 15.8(i),

$$\begin{aligned}\alpha+1 &= \text{zig-zag}(zy)\\ &= |\text{top}(z_y^+)|\\ &= |z^+ \cap \text{top}(x_y^+)|,\end{aligned}$$

as desired. We have now shown x_y^+ is α-dual-line regular. Observe $\delta(x,y)=1$ by (12.1) and the construction, so x_y^+ is a design-matroid by Lemma 22.4. It is clear from (33.6) and the construction that x_y^+ has rank $D-i$. We have now proved Lemma 33.3.

Theorem 33.4. *Let P denote a regular quantum matroid with parameters (D,q,α,β). Then for all integers i $(0 \le i < D)$, and for all $x \in P$ with* $\text{rank}(x)=i$,

$$\big|\{\, z \in P \mid z \text{ covers } x\}\big| = \eta_i, \tag{33.7}$$

where

$$\eta_i \;:=\; \Big(1 \;+\; \beta\frac{q^{D-i-1}}{t_{D-i-1}}\Big)\begin{bmatrix} D-i \\ 1 \end{bmatrix}, \tag{33.8}$$

and where t_j is from (33.1). *In particular,*

$$|A_P| \;=\; \Big(1+\beta\frac{q^{D-1}}{t_{D-1}}\Big)\begin{bmatrix} D \\ 1 \end{bmatrix} \tag{33.9}$$

if $D \ge 1$.

Proof. Pick any $x \in P$ such that $\text{rank}(x)=i$. By Lemma 4.3, there exists $y \in \text{top}(P)$ such that $x \le y$. To show (33.7), (33.8), it suffices to show

$$|X| = \begin{bmatrix} D-i \\ 1 \end{bmatrix}, \tag{33.10}$$

$$|Y| = \beta\frac{q^{D-i-1}}{t_{D-i-1}}\begin{bmatrix} D-i \\ 1 \end{bmatrix}, \tag{33.11}$$

where

$$X := \{z \in P \mid z \text{ covers } x, \quad z \le y\}, \tag{33.12}$$

$$Y := \{z \in P \mid z \text{ covers } x, \quad z \not\le y\}. \tag{33.13}$$

Line (33.10) is immediate from Lemma 27.5, so consider (33.11). Set

$$X' := \{z \in [x, y] \mid y \text{ covers } z\},$$

and observe

$$|X'| = \begin{bmatrix} D-i \\ 1 \end{bmatrix} \tag{33.14}$$

by Lemma 27.5. Set

$$Y' := \{z \in \operatorname{top}(P) \mid x \le z, \ \ \partial(y, z) = 2\}.$$

Observe each element of X' is covered by exactly β elements of Y'. Also, observe each element of Y' covers a unique element of X'. It follows

$$|Y'| = |X'|\beta. \tag{33.15}$$

Next, we claim each element of Y is less than or equal to exactly t_{D-i-1} elements of Y'. To see this, pick any $z \in Y$. Observe z covers $z \wedge y = x$; in particular $\operatorname{rank}(z) = i + 1$. Now by Lemma 33.3,

$$\begin{aligned} t_{D-i-1} &= |\operatorname{top}(z_y^+)| \\ &= |Y' \cap z^+|, \end{aligned}$$

as desired.

Next, we claim each element in Y' is greater than or equal to exactly q^{D-i-1} elements in Y. To see this, pick any $z \in Y'$. Then z is greater than or equal to exactly $\left[\begin{smallmatrix} D-i \\ 1 \end{smallmatrix}\right]$ elements in $X \cup Y$, but

$$\begin{aligned} \big|\{v \in X \mid v \le z\}\big| &= \big|\{v \in X \mid v \le z \wedge y\}\big| \\ &= \begin{bmatrix} D-i+1 \\ 1 \end{bmatrix} \end{aligned}$$

by Lemma 27.5. It follows z is greater than or equal to exactly

$$\begin{bmatrix} D-i \\ 1 \end{bmatrix} - \begin{bmatrix} D-i+1 \\ 1 \end{bmatrix} = q^{D-i-1}$$

elements in Y, as desired.

Combining our claims, we find

$$|Y'|q^{D-i-1} = |Y|t_{D-i-1}, \tag{33.16}$$

and (33.11) follows from (33.14), (33.15), (33.16). We have now established (33.10), (33.11), and (33.7), (33.8) follow.

To obtain (33.9), set $i = 0$ in (33.7), (33.8). This proves Theorem 33.4.

Corollary 33.5. *Let P denote a regular quantum matroid with parameters (D, q, α, β). Then*

$$|\mathrm{top}(P)| = \prod_{i=0}^{D-1} (1 + \beta q^i t_i^{-1}), \tag{33.17}$$

where t_i is from (33.1).

Proof. We compute the number of paths

$$x_0,\ x_1,\ \ldots,\ x_D \qquad (x_i \in P,\ \ \mathrm{rank}(x_i) = i,\ \ (0 \le i \le D))$$

in two ways. Constructing these paths from left to right, we find by Theorem 33.4 that the number is

$$\eta_0 \eta_1 \cdots \eta_{D-1}. \tag{33.18}$$

Constructing the above paths from right to left, we find by Lemma 27.5 that the number is

$$|\mathrm{top}(P)| \prod_{i=1}^{D} \begin{bmatrix} i \\ 1 \end{bmatrix}. \tag{33.19}$$

Now equate (33.18), (33.19), and evaluate the result using (33.8). This proves Corollary 33.5.

§34. The zig-zag function, revisited

Let P denote a regular quantum matroid, with parameters (D, q, α, β). In this section, we show the zig-zag function is constant over each Δ_i $(2 \le i \le D)$, and we compute these constants in terms of D, q, α, β. First, we prove an extension of Lemma 33.3.

Theorem 34.1. *Let P denote a regular quantum matroid with parameters (D, q, α, β), and pick any $x, y \in P$. Assume*

$$x \text{ covers } x \wedge y, \tag{34.1}$$

and

$$x \vee y \text{ does not exist.} \tag{34.2}$$

Then

$$|\mathrm{top}(x_y^+)| = \frac{t_{D-i}}{t_{D-j}}, \tag{34.3}$$

where

$$i := \mathrm{rank}(x), \tag{34.4}$$

$$j := \mathrm{rank}(y). \tag{34.5}$$

Proof. By Lemma 4.3, there exists an element $u \in \mathrm{top}(P)$ such that $y \le u$. To get our result, it suffices to show

$$|\mathrm{top}(x_u^+)| = t_{D-i} \tag{34.6}$$

and

$$|\mathrm{top}(x_y^+)| t_{D-j} = |\mathrm{top}(x_u^+)|. \tag{34.7}$$

Observe $x \not\le u$; otherwise u is an upper bound for x, y, contradicting (34.2). Now x covers $x \wedge y = x \wedge u$ by (34.1), so (34.6) follows from Lemma 33.3.

Now consider (34.7). Recall by Lemma 15.9 that x_y^+ is a submatroid of x_u^+. We claim each element in $\mathrm{top}(x_y^+)$ is less than or equal to exactly t_{D-j} elements in $\mathrm{top}(x_u^+)$. To see this, pick any $z \in \mathrm{top}(x_y^+)$. Observe z covers $z \wedge u$ by Lemma 15.3(i),(ii), and $\mathrm{rank}(z) = j$ by Theorem 18.2(v), so by Lemma 33.3 and Lemma 15.8(i),

$$\begin{aligned} t_{D-j} &= |\mathrm{top}(z_u^+)| \\ &= |z^+ \cap \mathrm{top}(x_u^+)|, \end{aligned}$$

as desired.

Next, we claim each element in $\mathrm{top}(x_u^+)$ is greater than or equal to a unique element in $\mathrm{top}(x_y^+)$. To see this, pick any $z \in \mathrm{top}(x_u^+)$. We show

$$p := x \vee (y \wedge z)$$

is the unique element in $\mathrm{top}(x_y^+)$ that is less than or equal to z. Observe $p \in x_y^+$ by Lemma 17.1. In fact $p \in \mathrm{top}(x_y^+)$, since y covers $y \wedge z$ by the construction. Observe $p \le z$ by Lemma 16.3(i), Lemma 17.1. Now suppose there exists an element $p' \in \mathrm{top}(x_y^+) \backslash \{p\}$ such that $p' \le z$. Then z is an upper bound for p, p'; consequently $p \vee p'$ exists. But this is impossible, since x_y^+ is $\vee$-closed by Lemma 8.2, Lemma

15.2, and since p, p' are both in $\mathrm{top}(x_y^+)$. Combining the above two claims, we obtain (34.7). We have now established (34.6), (34.7), and (34.3) follows. This proves Theorem 34.1.

Corollary 34.2. *Let P denote a regular quantum matroid with parameters (D, q, α, β). Then for all integers j $(2 \le j \le D)$ and for all $x, y \in P$ such that $xy \in \Delta_j$,*

$$\text{zig-zag}(x,y) = \frac{t_{D-j+1}}{t_{D-j}} \tag{34.8}$$

$$= q + \frac{\alpha + 1 - q}{t_{D-j}}. \tag{34.9}$$

(Caution: Δ_j may be empty).

Proof. Observe x, y satisfy (34.1), (34.2) by the definition of Δ_j, so by Lemma 28.4(i), Theorem 34.1,

$$\begin{aligned}\text{zig-zag}(x,y) &= |\mathrm{top}(x_y^+)| \\ &= \frac{t_{D-j+1}}{t_{D-j}}.\end{aligned}$$

Line (34.9) follows from (33.1).

§35. The ζ-uniform $\mathcal{P}$-basis systems

Let $\mathcal{P}$ denote a modular atomic lattice, and let B denote a $\mathcal{P}$-basis system from Definition 2.2. Recall by Theorem 2.5(i) that B^- is a $\mathcal{P}$-matroid, and by Lemma 3.1 that B^{+*} is a $\mathcal{P}^*$-matroid. In this section we introduce a condition on B that forces both of these quantum matroids to be regular. This condition will play a role in our subsequent work on regular quantum matroids.

Definition 35.1. Let $\mathcal{P}$ denote a modular atomic lattice of rank N, let B denote a $\mathcal{P}$-basis system of rank D, and let ζ denote an integer.

(i) Suppose $1 \le D \le N-1$. Then B is said to be *ζ-uniform* whenever

$$\big|B \cap [x,y]\big| = \zeta + 1 \tag{35.1}$$

for all $x \in B^-$ and for all $y \in B^+$ such that

$$x \leq y, \tag{35.2}$$

$$\operatorname{rank}(x) = D - 1, \tag{35.3}$$

$$\operatorname{rank}(y) = D + 1. \tag{35.4}$$

(ii) Suppose $D = 0$ or $D = N$. Then B is said to be ζ*-uniform* whenever ζ is nonnegative.

Lemma 35.2. *Let $\mathcal{P}$ denote a modular atomic lattice, and let B denote a ζ-uniform $\mathcal{P}$-basis system. Then*

$$\zeta \geq 0. \tag{35.5}$$

Proof. With the notation of Definition 35.1, suppose $1 \leq D \leq N - 1$. Then (35.5) follows from condition BA in Definition 2.2. Next suppose $D = 0$ or $D = N$. Then (35.5) is immediate from Definition 35.1(ii).

Theorem 35.3. *Let $\mathcal{P}$ denote a q-line regular modular atomic lattice of rank N, and let B denote a ζ-uniform $\mathcal{P}$-basis system of rank D. Then*

(i) *The $\mathcal{P}$-matroid B^- is regular, with parameters*

$$\left(D,\ q,\ \zeta,\ \zeta \begin{bmatrix} N-D \\ 1 \end{bmatrix}\right).$$

(ii) *The $\mathcal{P}^*$-matroid B^{+^*} is regular, with parameters*

$$\left(N-D,\ q,\ \zeta,\ \zeta \begin{bmatrix} D \\ 1 \end{bmatrix}\right).$$

Proof. (i) By the construction, B^- has rank D, and is q-line regular. To see that B^- is ζ-zig-zag regular, suppose there exists $x, y \in B^-$ such that $\operatorname{rank}(x) = D - 1$, $\operatorname{rank}(y) = D$, and $\partial(x, y) = 3$. We show

$$\text{zig-zag}(x, y) = \zeta + 1. \tag{35.6}$$

Observe $x \vee_{\mathcal{P}} y \geq y \in B$, so

$$x \vee_{\mathcal{P}} y \in B^+. \tag{35.7}$$

Observe x covers $x \wedge y$ by the construction, so $x \vee_{\mathcal{P}} y$ covers y by the modularity of $\mathcal{P}$. In particular

$$\text{rank}_{\mathcal{P}}(x \vee_{\mathcal{P}} y) = D + 1. \tag{35.8}$$

Now by (35.7), (35.8), and Definition 35.1,

$$\begin{aligned}\zeta + 1 &= \left|B \cap [x, x \vee_{\mathcal{P}} y]\right| \\ &= \text{zig-zag}(x, y),\end{aligned}$$

as desired. We have now shown B^- is ζ-zig-zag regular. It remains so show B^- is $\zeta[{N-D \atop 1}]$-dual-line regular. To this end, suppose there exists $x \in B^-$ such that $\text{rank}(x) = D - 1$. We show

$$|B \cap [x, 1_{\mathcal{P}}]| = 1 + \zeta \begin{bmatrix} N-D \\ 1 \end{bmatrix}. \tag{35.9}$$

To see (35.9), observe by Lemma 27.4(i),(ii), and Definition 35.1 that $(B^+ \cap [x, 1_{\mathcal{P}}])^*$ is a q-line regular, ζ-dual-line regular design-matroid of rank $N - D$. Applying Theorem 27.11 to this matroid, we get (35.9).

(ii) Observe $\mathcal{P}^*$ is q-line regular by Lemma 27.4. By Definition 35.1 and the construction, B is a ζ-uniform $\mathcal{P}^*$-basis system of rank $N - D$. The result now follows from part (i) above.

§36. $x \star y$ is a uniform $[x \wedge y, x]$-basis system

Let P denote a quantum matroid, and pick any x, $y \in P$. Recall by Theorem 19.3 that $x \star y$ is a $[x \wedge y, x]$-basis system. In this section, we assume P is regular, and show $x \star y$ is ζ-uniform, for some integer ζ that depends on $\rho(x, y)$, $\gamma(x, y)$, $\delta(x, y)$, $\gamma(y, x)$, and the parameters of P. We combine this information with results from Sections 33, 35 to compute the number of atoms in $[x \wedge y, x \star y]$, $[x \star y, x]^*$, and x_y^+.

Theorem 36.1. *Let P denote a regular quantum matroid with parameters (D, q, α, β). Pick $x, y \in P$, and set*

$$\rho := \rho(x, y), \tag{36.1}$$

$$\gamma := \gamma(x, y), \tag{36.2}$$

$$\delta := \delta(x, y), \tag{36.3}$$

$$\gamma^t := \gamma(y, x). \tag{36.4}$$

Then the $[x \wedge y, x]$-basis system $x \star y$ is ζ-uniform, where

$$\zeta := \alpha \frac{q^{D-\rho-\gamma-\delta-\gamma^t}}{t_{D-\rho-\gamma-\delta-\gamma^t}}, \tag{36.5}$$

and where t_j *is from* (33.1).

Proof. Recall by Lemma 13.2 and Theorem 18.2(vi) that

$$\text{rank}(x) = \rho + \gamma + \delta, \tag{36.6}$$

$$\text{rank}(y) = \rho + \gamma^t + \delta. \tag{36.7}$$

Pick $b \in x \star y$. We may assume $x \wedge y < b < x$; otherwise there is nothing to prove by Lemma 13.3(ii), Definition 35.1(ii). Now there exists $u, v \in [x \wedge y, x]$ such that b covers u and v covers b. We must show

$$\big|x \star y \cap [u, v]\big| = \zeta + 1. \tag{36.8}$$

Observe $u \in x_y^-$ by Theorem 19.3(i), so $u \vee y$ exists. Set $r := u \vee y$. Observe by Lemma 15.3(ii),(iii) that u, y are relative complements in the interval $[x \wedge y, r]$; if follows by Lemma 14.2(iii), (36.7), and the construction that

$$\begin{aligned} \text{rank}(r) &= \text{rank}(y) + \text{rank}(u) - \text{rank}(x \wedge y) \\ &= \rho + \gamma + \delta + \gamma^t - 1. \end{aligned} \tag{36.9}$$

Observe by the construction that $\delta(v, y) = 1$. Observe $r \in y_v^+$ by Lemma 15.3(i),(iii), so

$$\begin{aligned} \delta(v, r) &= \delta(v, y) \\ &= 1 \end{aligned}$$

by Lemma 15.3(v). Now there exists $w \in [u, r]$ such that r covers w and such that $v \vee w$ exists. Set $s := v \vee w$. Observe by Lemma 15.3(i),(ii) that v, w are relative complements in the interval $[u, s]$, so by (36.9),

$$\begin{aligned} \text{rank}(s) &= \text{rank}(w) + \text{rank}(v) - \text{rank}(u) \\ &= \text{rank}(w) + 2 \\ &= \text{rank}(r) + 1 \\ &= \rho + \gamma + \delta + \gamma^t. \end{aligned} \tag{36.10}$$

We compute $|\text{top}(r_s^+)|$ in two ways. On one hand, observe r covers $r \wedge s = w$ and $r \vee s$ does not exist, so by (36.9), (36.10), and Theorem 34.1,

$$\begin{aligned} |\text{top}(r_s^+)| &= \frac{t_{D-\rho-\gamma-\delta-\gamma^t+1}}{t_{D-\rho-\gamma-\delta-\gamma^t}} \\ &= \zeta + 1 \end{aligned} \tag{36.11}$$

by (33.1), (36.5). On the other hand, one finds by Theorem 15.5, Lemma 15.8(ii),(iii), and the observations $s \in v_r^+$, $r \in y_v^+$, that

$$\begin{aligned} |\mathrm{top}(r_s^+)| &= |\mathrm{top}(r_v^+)| \\ &= |\mathrm{top}(v_r^-)| \\ &= |\mathrm{top}(v_y^-) \cap [u,v]| \\ &= \big|x \star y \cap [u,v]\big|. \end{aligned} \tag{36.12}$$

Line (36.8) follows from (36.11), (36.12), so we are done.

Corollary 36.2. *With the assumptions and notation of Theorem 36.1:*

(i) *The* $[x \wedge y, x]$*-matroid* $[x \wedge y, x \star y]$ *is regular, with parameters*

$$\left(\gamma,\ q,\ \zeta,\ \zeta \begin{bmatrix} \delta \\ 1 \end{bmatrix}\right). \tag{36.13}$$

(ii) *The* $[x \wedge y, x]^*$*-matroid* $[x \star y, x]^*$ *is regular, with parameters*

$$\left(\delta,\ q,\ \zeta,\ \zeta \begin{bmatrix} \gamma \\ 1 \end{bmatrix}\right). \tag{36.14}$$

(iii) *The quantum matroid* x_y^+ *is regular, with parameters*

$$\left(\gamma^t,\ q,\ \zeta,\ \zeta \begin{bmatrix} \delta \\ 1 \end{bmatrix}\right). \tag{36.15}$$

Proof. (i), (ii) Immediate from Lemma 14.2(iii),(vi), Theorem 35.3, and Theorem 36.1.

(iii) Interchanging the roles of x, y in (i) above, we find $[x \wedge y, y \star x]$ is regular, with parameters (36.15). The result now follows, since x_y^+ is isomorphic to $y_x^- = [x \wedge y, y \star x]$ by Theorem 15.5 and Theorem 19.3(i). We have now proved Corollary 36.2.

Theorem 36.3. *With the assumptions and notation of Theorem 36.1:*

(i) *The number of atoms in the* $[x \wedge y, x]$*-matroid* $[x \wedge y, x \star y]$ *is*

$$\frac{t_{D-\rho-\gamma^t-1}}{t_{D-\rho-\delta-\gamma^t-1}} \begin{bmatrix} \gamma \\ 1 \end{bmatrix} \tag{36.16}$$

if $\gamma \geq 1$*, and* 0 *if* $\gamma = 0$*.*

(ii) *The number of atoms in the* $[x \wedge y, x]^*$*-matroid* $[x \star y, x]^*$ *is*

$$\frac{t_{D-\rho-\gamma^t-1}}{t_{D-\rho-\gamma-\gamma^t-1}} \begin{bmatrix} \delta \\ 1 \end{bmatrix} \tag{36.17}$$

if $\delta \geq 1$, *and* 0 *if* $\delta = 0$.

(iii) *The number of atoms in the quantum matroid* x_y^+ *is*

$$\frac{t_{D-\rho-\gamma-1}}{t_{D-\rho-\gamma-\delta-1}} \begin{bmatrix} \gamma^t \\ 1 \end{bmatrix} \tag{36.18}$$

if $\gamma^t \geq 1$, *and* 0 *if* $\gamma^t = 0$.

Proof. Apply the formula (33.9) in each of the three cases of Corollary 36.2, and simplify the result using (33.1), (36.5).

§37. The staircase theorem, revisited

Theorem 37.1. *Let* P *denote a regular quantum matroid with parameters* (D, q, α, β). *Pick any integer* a $(0 \leq a \leq D)$, *and fix an element* $y \in P$ *such that* $\operatorname{rank}(y) = a$. *Let* $\mathcal{D} = \mathcal{D}(a, D-a)$ *and* $\sigma : P \to V\mathcal{D}$ *be as in Theorem* 21.3. *Then for all* (ρ, γ, δ), $(\rho', \gamma', \delta') \in V\mathcal{D}$, *and for all* $x \in \sigma^{-1}(\rho, \gamma, \delta)$, *the number*

$$\left|\{z \in \sigma^{-1}(\rho', \gamma', \delta') \mid z \text{ is adjacent to } x\}\right| \tag{37.1}$$

is given as follows.

Case $(\rho', \gamma', \delta') = (\rho - 1, \gamma, \delta)$:

$$q^{\gamma+\delta} \begin{bmatrix} \rho \\ 1 \end{bmatrix} \tag{37.2}$$

Case $(\rho', \gamma', \delta') = (\rho + 1, \gamma, \delta)$:

$$\frac{t_{D-\rho-\gamma-1}}{t_{D-\rho-\gamma-\delta-1}} \begin{bmatrix} a-\rho-\delta \\ 1 \end{bmatrix} \tag{37.3}$$

Case $(\rho', \gamma', \delta') = (\rho, \gamma - 1, \delta)$:

$$q^{\delta} \frac{t_{D-a-\gamma-1}}{t_{D-a-\gamma+\delta-1}} \begin{bmatrix} \gamma \\ 1 \end{bmatrix} \tag{37.4}$$

if $\delta \geq 1$, *and*

$$\begin{bmatrix} \gamma \\ 1 \end{bmatrix} \tag{37.5}$$

if $\delta = 0$.

Case $(\rho', \gamma', \delta') = (\rho, \gamma + 1, \delta)$:

$$(37.6) \qquad q^{a-\rho-\delta} \frac{t_{D-\rho-\gamma-1}}{t_{D-\rho-\gamma-\delta-1}} \frac{t_{D-a-\gamma-1}}{t_{D-a-\gamma+\delta-1}} \eta_{a+\gamma}$$

Case $(\rho', \gamma', \delta') = (\rho, \gamma, \delta - 1)$:

$$(37.7) \qquad \frac{t_{D-a+\delta-1}}{t_{D-a-\gamma+\delta-1}} \begin{bmatrix} \delta \\ 1 \end{bmatrix}$$

Case $(\rho', \gamma', \delta') = (\rho, \gamma, \delta + 1)$:

$$(37.8) \qquad \frac{t_{\delta-1}}{q^{\delta-1}} \frac{q^{D-\rho-\gamma-\delta-1}}{t_{D-\rho-\gamma-\delta-1}} \frac{q^{D-a-\gamma+\delta-1}}{t_{D-a-\gamma+\delta-1}} (\beta - \alpha \begin{bmatrix} \delta \\ 1 \end{bmatrix}) \begin{bmatrix} a-\rho-\delta \\ 1 \end{bmatrix}$$

if $a + \gamma < D$, *and*

$$(37.9) \qquad \frac{q^{D-\rho-\gamma-\delta-1}}{t_{D-\rho-\gamma-\delta-1}} (\beta - \alpha \begin{bmatrix} \delta \\ 1 \end{bmatrix}) \begin{bmatrix} D-\rho-\gamma-\delta \\ 1 \end{bmatrix}$$

if $a + \gamma = D$.

The number (37.1) *equals* 0 *in all other cases.*

Proof. We set

$$\Lambda := \{z \in \sigma^{-1}(\rho', \gamma', \delta') \mid z \text{ is adjacent to } x\},$$

and compute $|\Lambda|$ in each of the above cases.

Case $(\rho', \gamma', \delta') = (\rho - 1, \gamma, \delta)$. By Lemma 13.4 and Lemma 27.5,

$$\begin{aligned} |\Lambda| &= \left|\{z \in P \mid x \text{ covers } z, \ z \not\geq x \wedge y\}\right| \\ &= \left|\{z \in P \mid x \text{ covers } z\}\right| \ - \ \left|\{z \in P \mid x \text{ covers } z, \ z \geq x \wedge y\}\right| \\ &= \begin{bmatrix} \rho+\gamma+\delta \\ 1 \end{bmatrix} - \begin{bmatrix} \gamma+\delta \\ 1 \end{bmatrix} \\ &= q^{\gamma+\delta} \begin{bmatrix} \rho \\ 1 \end{bmatrix}, \end{aligned}$$

as desired.

Case $(\rho', \gamma', \delta') = (\rho + 1, \gamma, \delta)$. By Lemma 15.3, Λ consists of those elements in x_y^+ that cover x, i.e. the atoms of the poset x_y^+. Now $|\Lambda|$ is given in (36.18). Eliminating γ^t in (36.18) using $\gamma^t = a - \rho - \delta$, we obtain (37.3), as desired.

Case $(\rho', \gamma', \delta') = (\rho, \gamma, \delta - 1)$. By Lemma 14.3, Λ consists of the elements in $[x \star y, x]$ that are covered by x (in P), i.e., the atoms of the poset $[x \star y, x]^*$. Now $|\Lambda|$ is given in (36.17). Eliminating γ^t in (36.17) as in the previous case, we obtain (37.7), as desired.

Case $(\rho', \gamma', \delta') = (\rho, \gamma - 1, \delta)$. In this case Λ consists of the elements in $[x \wedge y, x]$ that are covered by x but are not in $[x \star y, x]$. If $\delta \geq 1$, then by the previous case,

$$\begin{aligned} |\Lambda| &= \begin{bmatrix} \gamma + \delta \\ 1 \end{bmatrix} - \frac{t_{D-a+\delta-1}}{t_{D-a-\gamma+\delta-1}} \begin{bmatrix} \delta \\ 1 \end{bmatrix} \\ &= q^\delta \frac{t_{D-a-\gamma-1}}{t_{D-a-\gamma+\delta-1}} \begin{bmatrix} \gamma \\ 1 \end{bmatrix}, \end{aligned}$$

as desired. If $\delta = 0$ then $x \star y = x$, so

$$|\Lambda| = \begin{bmatrix} \gamma \\ 1 \end{bmatrix}$$

by Lemma 14.2(iii) and Lemma 27.5.

Case $(\rho', \gamma', \delta') = (\rho, \gamma + 1, \delta)$. We show $|\Lambda|$ equals the expression in (37.6) by induction on

$$\begin{aligned} \gamma^t &:= \gamma(y, x) \\ &= a - \rho - \delta. \end{aligned}$$

First consider the case $\gamma^t = 0$. Here x is relatively close to y in the sense of Lemma 20.2, so by Theorem 20.4,

$$\Lambda = \{z \in P \mid z \text{ covers } x\}.$$

In particular by Theorem 33.4,

$$\begin{aligned} |\Lambda| &= \eta_i \qquad (i = \text{rank}(x)), \\ &= \eta_{\rho+\gamma+\delta}, \end{aligned}$$

which is what (37.6) reduces to in this case. Now assume $\gamma^t > 0$. Set

$$\begin{aligned} \Omega &:= \{u \in x_y^+ \mid u \text{ covers } x\} \\ &= \{u \in P \mid u \geq x, \ \sigma(u) = (\rho + 1, \gamma, \delta)\}, \end{aligned}$$

and observe by (37.3) that

$$|\Omega| = \frac{t_{D-\rho-\gamma-1}}{t_{D-\rho-\gamma-\delta-1}} \begin{bmatrix} a - \rho - \delta \\ 1 \end{bmatrix}. \tag{37.10}$$

Set

$$\Psi := \{v \in P \mid v \geq x,\ \ \sigma(v) = (\rho+1, \gamma+1, \delta)\}.$$

Observe by induction and (37.6) (with $(a', \rho', \gamma', \delta') = (a, \rho+1, \gamma, \delta)$) that each element in Ω is covered by exactly

$$q^{a-\rho-\delta-1} \frac{t_{D-\rho-\gamma-2}}{t_{D-\rho-\gamma-\delta-2}} \frac{t_{D-a-\gamma-1}}{t_{D-a-\gamma+\delta-1}} \eta_{a+\gamma}$$

elements in Ψ. Also observe by Theorem 17.2(i),(iv) that each element $v \in \Psi$ covers a unique element in Ω, i.e., $x \vee (v \wedge y)$. From our above observations

$$|\Psi| = |\Omega| q^{a-\rho-\delta-1} \frac{t_{D-\rho-\gamma-2}}{t_{D-\rho-\gamma-\delta-2}} \frac{t_{D-a-\gamma-1}}{t_{D-a-\gamma+\delta-1}} \eta_{a+\gamma}. \tag{37.11}$$

Observe by (37.3) (with $(a', \rho', \gamma', \delta') = (a, \rho, \gamma+1, \delta)$) that each element in Λ is covered by exactly

$$\frac{t_{D-\rho-\gamma-2}}{t_{D-\rho-\gamma-\delta-2}} \begin{bmatrix} a-\rho-\delta \\ 1 \end{bmatrix}$$

elements in Ψ. Also observe by Theorem 17.3(iii),(iv) that each element in Ψ covers exactly q elements in Λ. From the above two observations

$$|\Lambda| \frac{t_{D-\rho-\gamma-2}}{t_{D-\rho-\gamma-\delta-2}} \begin{bmatrix} a-\rho-\delta \\ 1 \end{bmatrix} = |\Psi| q. \tag{37.12}$$

Combining (37.10)-(37.12), we find $|\Lambda|$ equals the expression (37.6), as desired.

Case $(\rho', \gamma', \delta') = (\rho, \gamma, \delta+1)$. First assume $a+\gamma < D$. In this case Λ consists of all the elements in P that cover x, but are not counted in (37.3) or (37.6). By (37.3), (37.6), and Theorem 33.4, we find

$$\begin{aligned} |\Lambda| \ = \ \eta_{\rho+\gamma+\delta} \ &- \ \frac{t_{D-\rho-\gamma-1}}{t_{D-\rho-\gamma-\delta-1}} \begin{bmatrix} a-\rho-\delta \\ 1 \end{bmatrix} \\ &- \ q^{a-\rho-\delta} \frac{t_{D-\rho-\gamma-1}}{t_{D-\rho-\gamma-\delta-1}} \frac{t_{D-a-\gamma-1}}{t_{D-a-\gamma+\delta-1}} \eta_{a+\gamma}. \end{aligned}$$

Evaluating this using (27.3), (27.4), (33.1), (33.8), we obtain (37.8).

Now assume $a+\gamma = D$. In this case, Λ consists of all the elements in P that cover x, but are not counted in (37.3). Proceeding as above, we find

$$|\Lambda| \ = \ \eta_{\rho+\gamma+\delta} \ - \ \frac{t_{D-\rho-\gamma-1}}{t_{D-\rho-\gamma-\delta-1}} \begin{bmatrix} D-\gamma-\rho-\delta \\ 1 \end{bmatrix}.$$

Evaluating this using (27.3), (27.4), (33.1), (33.8), we obtain (37.9), as desired.

In any other case, the expression (37.1) equals 0 by the staircase theorem. This proves Theorem 37.1.

Corollary 37.2. *Let P denote a regular quantum matroid with parameters (D, q, α, β). Pick any $x \in P$, pick any $y \in \mathrm{top}(P)$, and set*

$$\begin{aligned} \rho &:= \rho(x, y), \\ \delta &:= \delta(x, y) \\ &= \mathrm{rank}(x) - \rho. \end{aligned}$$

Then

$$\text{(i)} \quad \left|\{z \in P \mid x \textit{ covers } z,\ \partial(z,y) = \partial(x,y) + 1\}\right| = q^{\delta} \begin{bmatrix} \rho \\ 1 \end{bmatrix},$$

$$\text{(ii)} \quad \left|\{z \in P \mid x \textit{ covers } z,\ \partial(z,y) = \partial(x,y) - 1\}\right| = \begin{bmatrix} \delta \\ 1 \end{bmatrix}.$$

Suppose $x \notin \mathrm{top}(P)$. Then

(iii)

$$\left|\{z \in P \mid z \textit{ covers } x,\ \partial(z,y) = \partial(x,y) - 1\}\right| = \frac{t_{D-\rho-1}}{t_{D-\rho-\delta-1}} \begin{bmatrix} D-\rho-\delta \\ 1 \end{bmatrix},$$

(iv)

$$\left|\{z \in P \mid z \textit{ covers } x,\ \partial(z,y) = \partial(x,y) + 1\}\right| = \frac{q^{D-\rho-\delta-1}}{t_{D-\rho-\delta-1}} \left(\beta - \alpha \begin{bmatrix} \delta \\ 1 \end{bmatrix}\right) \begin{bmatrix} D-\rho-\delta \\ 1 \end{bmatrix}.$$

Proof. This is the case $a = D$ in Theorem 37.1. With the notation of that theorem, observe $\mathcal{D} = \mathcal{D}(D, 0)$, so by Definition 21.2, $\gamma = 0$ for all $(\rho, \gamma, \delta) \in V\mathcal{D}$. Setting $\Delta\gamma = 0$ in Theorem 13.5, we find the sets (i)–(iv) above equal the sets of vertices adjacent to x and contained in $\sigma^{-1}(\rho - 1, 0, \delta)$, $\sigma^{-1}(\rho, 0, \delta - 1)$, $\sigma^{-1}(\rho + 1, 0, \delta)$, $\sigma^{-1}(\rho, 0, \delta + 1)$, resp. The cardinality of these sets is given in (37.2), (37.7), (37.3), (37.9), respectively, (where $\gamma = 0$, $a = D$). This proves Corollary 37.2.

§38. The graph on top(P) is distance-regular

Definition 38.1. A finite, connected, undirected graph $\Gamma = (V\Gamma, E\Gamma)$ of diameter d is said to be *distance-regular*, with *intersection numbers* c_i, b_i $(0 \le i \le d)$, whenever for all integers i $(0 \le i \le d)$, and all $x, y \in V\Gamma$ at distance $\partial_\Gamma(x,y) = i$, the scalars

$$c_i := \left|\{z \in V\Gamma \mid xz \in E\Gamma,\ \ \partial_\Gamma(y,z) = i-1\}\right|,$$
$$b_i := \left|\{z \in V\Gamma \mid xz \in E\Gamma,\ \ \partial_\Gamma(y,z) = i+1\}\right|,$$

are independent of x, y.

Theorem 38.2. *Let P denote a nontrivial regular quantum matroid with parameters (D, q, α, β). Then* (i)–(iii) *hold below.*

(i) *The graph on* top(P) *is distance-regular, with intersection numbers*

(38.1)
$$c_i = \begin{bmatrix} i \\ 1 \end{bmatrix} (1 + \alpha \begin{bmatrix} i-1 \\ 1 \end{bmatrix}) \qquad\qquad (0 \le i \le d),$$

(38.2)
$$b_i = (\begin{bmatrix} D \\ 1 \end{bmatrix} - \begin{bmatrix} i \\ 1 \end{bmatrix})(\beta - \alpha \begin{bmatrix} i \\ 1 \end{bmatrix}) \qquad\qquad (0 \le i \le d),$$

where $d := \operatorname{diam}_{\text{top}}(P)$.

(ii) *Suppose* $d < D$. *Then*

(38.3)
$$\beta = \alpha \begin{bmatrix} d \\ 1 \end{bmatrix}.$$

(iii) *Assume* $d = D$. *Then the graph on* top(P) *has classical parameters* (D, q, α, β) *in the sense of Brouwer, Cohen, Neumaier* [B-C-N].

Proof. (i) Routine application of Corollary 37.2 using line (22.1).
(ii) Immediate from (i) and the observation $b_d = 0$.
(iii) Immediate from [B-C-N, p194].

§39. The classification of the regular quantum matroids with rank at least 4

In this section, we classify the nontrivial regular quantum matroids with rank at least 4. We do this as follows. Let P denote a nontrivial regular quantum matroid with parameters (D, q, α, β), and assume $D \geq 4$. First, we show $\alpha \in \{0, q-1, q\}$. In each case, we invoke a result from the literature to identify P, giving us our classification. Our main result is Theorem 39.6.

Lemma 39.1. *Let P denote a nontrivial regular quantum matroid with parameters (D, q, α, β), and assume $\alpha \neq q$. Then* (i)–(iii) *hold below.*

(i) $(q-1-\alpha)t_{D-i}^{-1} \geq 0 \qquad (2 \leq i \leq D)$.
(ii) $\Delta_i \neq \emptyset \qquad (2 \leq i \leq D)$.
(iii) $(q-1-\alpha)t_{D-i}^{-1} \in \mathbb{Z} \qquad (2 \leq i \leq D)$.

Proof. (i) Recall $\alpha \leq q$ by (28.4), and we assume $\alpha \neq q$, so $\alpha \leq q-1$. Also t_{D-i} is positive by Lemma 33.2(ii).

(ii) By Lemma 28.2(iii), it suffices to find $x, y \in A_P$
assume P is not a modular atomic lattice, there exists $x \in A_P$ and there exists $u \in \text{top}(P)$ such that $x \not\leq u$. We show there exists an atom $y \in \text{Shadow}(u)$ such that $x \vee y$ does not exist. To this end, recall by (15.2) and Theorem 19.3(i) that for all $y \in \text{Shadow}(u)$, $x \vee y$ exists if and only if $y \in [0, u \star x]$. Hence, it suffices to show $\text{Shadow}(u) \backslash [0, u \star x]$ is not empty. Observe by (27.5),

$$|\text{Shadow}(u)| = \begin{bmatrix} D \\ 1 \end{bmatrix}. \tag{39.1}$$

To compute the number of atoms in $[0, u \star x]$, observe by Definitions 12.1, 13.1 that $\rho(u,x) = 0$, $\gamma(u,x) = D-1$, $\delta(u,x) = 1$, $\gamma(x,u) = 0$. By Theorem 36.3(i) (with $\rho = 0$, $\gamma = D-1$, $\delta = 1$, $\gamma^t = 0$), the number of atoms in $[0, u \star x]$ equals

$$\frac{t_{D-1}}{t_{D-2}} \begin{bmatrix} D-1 \\ 1 \end{bmatrix}. \tag{39.2}$$

Observe

$$\begin{aligned} \begin{bmatrix} D \\ 1 \end{bmatrix} - \frac{t_{D-1}}{t_{D-2}} \begin{bmatrix} D-1 \\ 1 \end{bmatrix} &= 1 + \frac{q-1-\alpha}{t_{D-2}} \begin{bmatrix} D-1 \\ 1 \end{bmatrix} \\ &\geq 1 \end{aligned}$$

by (33.1) and (i) above, so $\text{Shadow}(u) \backslash [0, u \star x]$ is not empty, as desired.

(iii) Let the integer i be given. There exists $xy \in \Delta_i$ by part (ii) above, so by Corollary 34.2,

$$\frac{q-1-\alpha}{t_{D-i}} = q - \text{zig-zag}(x, y)$$

is an integer.

Theorem 39.2. *Let P denote a nontrivial q-line regular, α-zig-zag regular quantum matroid with rank $D \geq 3$.*

(i) *Suppose $D \geq 4$. Then $\alpha \in \{0, q-1, q\}$.*
(ii) *Suppose $D = 3$. Then $\alpha = q$ or $1+\alpha$ divides q.*

Proof. (i) We assume $\alpha \neq 0$, $\alpha \neq q-1$, $\alpha \neq q$, and get a contradiction. Observe $\alpha \geq 1$ by Lemma 28.6 and our assumptions, so P is dual-line regular by Theorem 31.1. It follows P is regular by Definition 32.1. Now on one hand, by Lemma 39.1(i),(iii) (with $i = D-2$), we find $(q-1-\alpha)t_2^{-1}$ is a positive integer. On the other hand, by (27.2), (33.1),

$$\begin{aligned}\frac{q-1-\alpha}{t_2} &= \frac{q-1-\alpha}{1+\alpha(q+1)}\\ &\leq \frac{q-2}{q+2}\\ &< 1,\end{aligned}$$

a contradiction.

(ii) Assume $\alpha \neq 0$, $\alpha \neq q$; otherwise the result is trivial. Then as in (i) above, P is regular. By Lemma 39.1(iii) (with $i=2$, $D=3$), we find $(q-1-\alpha)t_1^{-1}$ is an integer. By (33.1),

$$\begin{aligned}\frac{q-1-\alpha}{t_1} &= \frac{q-1-\alpha}{1+\alpha}\\ &= \frac{q}{1+\alpha} - 1,\end{aligned}$$

so $1+\alpha$ divides q, as desired. This proves Theorem 39.2.

Suppose P is a nontrivial regular quantum matroid with parameters (D, q, α, β), and assume $D \geq 4$. In each of the three cases in Theorem 39.2(i), we can identify P. If $\alpha = 0$, then P is known by Theorems 30.3, 30.4, 30.5. In each of the other cases $\alpha = q-1$, $\alpha = q$, there is a result in the literature of diagram geometries that identifies P. We quote these results below, translated into the language of quantum matroids via Theorem 23.1. First, we eliminate the easy case $q = 1$.

Theorem 39.3. *Let D denote an integer at least 2, and let P denote a poset. Then the following are equivalent.*

(i) *P is a nontrivial 1-line regular, 1-zig-zag regular quantum matroid with rank D.*

(ii) *There exists an integer* $N > D$ *such that* P *is isomorphic to the truncated Boolean algebra* $B(D, N)$.

Suppose (i), (ii) *hold. Then* P *is* $(N - D)$*-dual-line regular.* (See Example 40.1(1).)

Proof. Routine.

Theorem 39.4 ([Bu1, Theorem 8]). *Let* D *denote an integer at least* 3, *let* q *denote an integer at least* 2, *and let* P *denote a poset. Then the following are equivalent.*

(i) P *is a nontrivial* q*-line regular,* q*-zig-zag regular quantum matroid with rank* D.

(ii) q *is a prime power, and there exists an integer* $N > D$ *such that* P *is isomorphic to the truncated projective geometry* $L_q(D, N)$.

Suppose (i), (ii) *hold. Then* P *is* β*-dual-line regular, where*

$$\beta = q\frac{q^{N-D} - 1}{q - 1}.$$

See Example 40.1(3).

Theorem 39.5 ([Sp1, Theorem 3]). *Let* D *denote an integer at least* 3, *let* q *denote an integer at least* 2, *and let* P *denote a poset. Then the following are equivalent.*

(i) P *is a nontrivial* q*-line regular,* $(q-1)$*-zig-zag regular quantum matroid with rank* D.

(ii) q *is a prime power, and there exists an integer* $N > D$ *such that* P *is isomorphic to the attenuated space* $A_q(D, N)$.

Suppose (i), (ii) *hold. Then* P *is* $(q^{N-D} - 1)$*-dual-line regular.*

(See Example 40.1(4)).

We now arrive at the central theorem of this paper.

Theorem 39.6. *Let* D *denote an integer at least* 4, *and let* P *denote a poset. Then the following are equivalent.*

(i) P *is a nontrivial regular quantum matroid with rank* D.

(ii) P *is isomorphic to one of the following:*

(iia) *A truncated Boolean algebra* $B(D, N)$, $(D < N)$.

(iib) *A Hamming matroid* $H(D, N)$, $(2 \leq N)$.

(iic) *A truncated projective geometry* $L_q(D, N)$, $(D < N)$.

(iid) *An attenuated space* $A_q(D, N)$, $(D < N)$.

(iie) *A classical polar space of rank* D.

Proof. (i) $\to$ (ii). Let (D, q, α, β) denote the parameters of P. Then

$$q \geq 1 \tag{39.3}$$

by Lemma 27.2,

$$\alpha \in \{0, q-1, q\} \tag{39.4}$$

by Theorem 39.2, and

$$\beta \geq 1 \tag{39.5}$$

by Lemma 27.9, Lemma 27.10, and Definition 32.2.

First assume $\alpha = 0$. In this case $\mathrm{Rad}(P) = 0$ by Theorem 29.3(i),(iv) and (39.5), so P is nondegenerate by Definition 26.1. Now P is a Tits polar space of rank D by Theorem 30.3. If $q = 1$ then by Theorem 30.4 and (39.5), P is isomorphic to $H(D, N)$, where $N = \beta + 1 \geq 2$. If $q \geq 2$ then by Theorem 30.5, P is isomorphic to a classical polar space of rank D.

Next assume $\alpha = q - 1$. In this case we may assume $q \geq 2$; otherwise $\alpha = 0$ by (39.3), and our previous remarks apply. Now by Theorem 39.5, q is a prime power, and P is isomorphic to $A_q(D, N)$ for some integer $N > D$.

Finally assume $\alpha = q$. If $q = 1$, then by Theorem 39.3, P is isomorphic to $B(D, N)$ for some integer $N > D$. If $q \geq 2$, then by Theorem 39.4, q is a prime power, and P is isomorphic to $L_q(D, N)$ for some integer $N > D$.

(ii) $\to$ (i). Assume P is isomorphic to $B(D, N)$, for some integer $N > D$. Then P is a nontrivial regular quantum matroid of rank D by Theorem 39.3.

Assume P is isomorphic to $H(D, N)$, for some integer $N \geq 2$. Then P is a Tits polar space by Theorem 30.4, so P is a nondegenerate quantum matroid of rank D by Theorem 30.3. In particular P is nontrivial. P is 1-line regular and $(N-1)$-dual-line regular by Theorem 30.4, and 0-zig-zag regular by Theorem 30.3.

Assume P is isomorphic to $L_q(D, N)$, for some integer $N > D$. Then P is a nontrivial, regular quantum matroid of rank D by Theorem 39.4.

Assume P is isomorphic to $A_q(D, N)$, for some integer $N > D$. Then P is a nontrivial, regular quantum matroid of rank D by Theorem 39.5.

Assume P is a classical polar space of rank D. Then P is a Tits polar space of rank D by Theorem 30.5, so P is a nondegenerate quantum matroid of rank D by Theorem 30.3. In particular P is nontrivial. P is line-regular by Theorem 30.5, 0-zig-zag regular by Theorem 30.3, and dual-line regular by Theorem 30.5. This proves Theorem 39.6.

Corollary 39.7. *Let D denote an integer at least 4. Then the following are equivalent statements concerning a finite, undirected graph Γ.*

(i) *Γ is isomorphic to the graph on $\operatorname{top}(P)$, where P is a nontrivial regular quantum matroid with rank D.*

(ii) *Γ is isomorphic to one of the following:*
  - (iia) *The Johnson graph $J(D,N)$, $(D<N)$.*
  - (iib) *The Hamming graph $H(D,N)$, $(2\le N)$.*
  - (iic) *The q-Johnson graph $J_q(D,N)$, $(D<N)$.*
  - (iid) *The bilinear forms graph $H_q(D,N)$, $(D<N)$.*
  - (iie) *A dual polar space graph of diameter D.*

(See [B-I, p300] for the definitions of these graphs).

Proof. Immediate from Theorem 39.6.

Corollary 39.8. *Let P denote a regular quantum matroid with rank $D\ge 4$. Then P is embeddable (in the sense of Definition 6.3).*

Proof. Concerning the examples in Theorem 39.6, observe P is an A-matroid in cases (iia), (iib), and a V-matroid in cases (iic)–(iie).

§40. The examples of regular quantum matroids

Example 40.1. Let D denote an integer at least 2. In each of the following cases 1–5, P is a nontrivial regular quantum matroid of rank D. In each case, the parameters q, α, β are given. (See Definition 32.1). By Theorem 39.6, there are no other nontrivial regular quantum matroids with rank $D\ge 4$.

1. The truncated Boolean algebra $B(D,N)\,(D<N)$ **[Bu1], [Te].**

Let A denote a set of cardinality N.

$$P=\{\,x\subseteq A\mid\ |x|\le D\},$$

$$x\le y \text{ whenever } x \text{ is a subset of } y\ \ (x,y\in P),$$

$$\operatorname{rank}(x)=|x|\ \ (x\in P),$$

$$q=1,\qquad \alpha=1,\qquad \beta=N-D.$$

2. The Hamming matroid $H(D,N)$ $(2 \le N)$ **[De], [Te].**
Set

$$A = A_1 \cup A_2 \cup \cdots \cup A_D \qquad \text{(disjoint union)},$$

where $|A_i| = N (1 \le i \le D)$.

$$P = \{x \subseteq A \mid \quad |x \cap A_i| \le 1 \quad \text{for all } i \quad (1 \le i \le D)\},$$

$$x \le y \;\text{ whenever }\; x \;\text{ is a subset of }\; y \; (x,\, y \in P),$$

$$\operatorname{rank}(x) = |x| \;\; (x \in P),$$

$$q = 1, \qquad \alpha = 0, \qquad \beta = N - 1.$$

3. The truncated projective geometry $L_q(D,N)$ $(D < N)$ **[Bu1], [Sta], [Te].**

Let V denote a vector space of dimension N over the field $GF(q)$.

$$P = \{x \mid x \;\text{ is a subspace of }\; V, \quad \dim(x) \le D\},$$

$$x \le y \;\text{ whenever }\; x \;\text{ is a subspace of }\; y \; (x,\, y \in P),$$

$$\operatorname{rank}(x) = \dim(x) \;\; (x \in P),$$

$$\alpha = q, \qquad \beta = q\frac{q^{N-D} - 1}{q - 1}.$$

4. The attenuated space $A_q(D,N)$ $(D < N)$ **[De], [Hu], [Sp1], [Sta], [Te].**

Let V denote a vector space of dimension N over the field $GF(q)$, and fix a subspace $w \subseteq V$ of dimension $N - D$.

$$P = \{x \mid x \;\text{ is a subspace of }\; V, \quad x \cap w = 0\},$$

$$x \le y \;\text{ whenever }\; x \;\text{ is a subspace of }\; y \; (x, y \in P),$$

$$\operatorname{rank}(x) = \dim(x)\;, (x \in P),$$

$$\alpha = q - 1, \qquad \beta = q^{N-D} - 1.$$

5. The classical polar spaces of rank D **over** $GF(q)$ **[C-J-P], [Ca2], [Mu].**

Let V denote a vector space over the field $GF(q)$, and assume V possesses one of the following nondegenerate forms:

name	**dim V**	**form**	ε
$B_D(q)$	$2D+1$	quadratic	0
$C_D(q)$	$2D$	alternating	0
$D_D(q)$	$2D$	quadratic (Witt index D)	-1
${}^2D_{D+1}(q)$	$2D+2$	quadratic (Witt index D)	1
${}^2A_{2D}(r)$	$2D+1$	Hermitean ($q=r^2$)	$\frac{1}{2}$
${}^2A_{2D-1}(r)$	$2D$	Hermitean ($q=r^2$)	$-\frac{1}{2}$.

We call a subspace of V *totally isotropic* whenever the form vanishes completely on that subspace. In each of the above cases, the dimension of any maximal isotropic subspace is D.

$$P = \{x \mid x \text{ is an isotropic subspace of } V\},$$

$$x \leq y \text{ whenever } x \text{ is a subspace of } y \ (x,\, y \in P),$$

$$\operatorname{rank}(x) = \dim(x) \ (x \in P),$$

$$\alpha = 0, \qquad \beta = q^{1+\varepsilon}.$$

§41. Directions for further research

In this section we give some conjectures and open problems concerning quantum matroids and related topics. See also Problem 4.4, Conjecture 6.5, and Conjecture 7.13 in the text.

Conjecture 41.1. *Let P denote a quantum matroid with rank $D \geq 2$. Let us say P is* thick line connected *whenever for all distinct atoms $x, y \in A_P$, there exists an integer $d \geq 1$ and a sequence $x = x_0, x_1, \ldots, x_d = y$ $(x_0, x_1, \ldots, x_d \in A_P)$ such that $x_i \vee x_{i+1}$ exists and is a thick line for all i $(0 \leq i < d)$. We conjecture that if P is thick line connected and if D is sufficiently large, then every line in P is thick.*

Problem 41.2. Let us say a finite, undirected graph Γ is a *quantum matroid graph* whenever there exists a quantum matroid P such that Γ is isomorphic to the graph on $\operatorname{top}(P)$. Find a simple combinatorial property that characterizes the quantum matroid graphs among all the finite undirected graphs. See Corollary 39.7.

Problem 41.3. Let P denote a classical polar space (Example 40.1(5)). What quantum matroid is dual to P in the sense of Definition 3.2?

Problem 41.4. For a classical matroid (Definition 1.1), one has the dependency axioms, the hyperplane axioms, the circuit axioms, the bond axioms, etc. See for example [Wh, Chapter 2]. To what extent are there analogous axioms for the $\mathcal{P}$-matroids, where $\mathcal{P}$ is any modular atomic lattice ?

Conjecture 41.5. *Let P denote a quantum matroid. Pick any $x, y \in P$, and let G denote the minimal geodesically closed subposet of P containing x, y. We conjecture*

$$rank(G) = \gamma(x, y) + \gamma(y, x) + \delta(x, y).$$

(See Definitions 12.1, 13.1).

Problem 41.6. Which quantum matroids are Cohen-Macaulay? (See[B-G-S]).

Problem 41.7. Let P denote a quantum matroid of rank D. Let us call P *weakly zig-zag regular* whenever for all integers i $(1 \leq i \leq D-1)$, and for all $x, y \in P$ such that $\text{rank}(x) = i$, $\text{rank}(y) = i$, the number of paths in P with endpoints x, y and shape $(i, i-1, i, i+1, i)$ equals the number of paths in P with endpoints x, y and shape $(i, i+1, i, i-1, i)$. If P is regular then P is weakly zig-zag regular. Classify the weakly zig-zag regular quantum matroids.

Problem 41.8. Let D denote an integer at least 3, and let q denote an integer at least 2. Find a short, direct proof, not involving Theorem 39.5, that any nontrivial q-line regular, $(q-1)$-zig-zag regular quantum matroid is embeddable. (See Definition 6.3.)

Problem 41.9. Let N denote an integer at least 3, let q denote a prime power, and let V denote an N dimensional vector space over the field $GF(q)$. Pick an integer D $(2 \leq D < N)$, and let P denote a nontrivial $(q-1)$-zig-zag regular V-matroid of rank D that spans V. Find a short, direct proof, not involving Theorem 39.5, that there exists a subspace $w \subseteq V$ such that $\dim(w) = N - D$ and such that

$$P = \{\, x \mid x \text{ is a subspace of } V, \quad x \cap w = 0\}.$$

(See Example 40.1(4).)

Problem 41.10. Let $\mathcal{P}$ denote a modular atomic lattice. Classify the uniform $\mathcal{P}$-basis systems. Give a short, direct proof, that does not refer to Theorems 35.3, 39.6. (See Definition 35.1.)

Problem 41.11. Let P denote a quantum matroid, and pick any $x, y \in P$. What can be said about $\mathrm{Rad}(x_y^+)$? Under what conditions is u_v^+ nondegenerate for all $u, v \in P$ such that $u \vee v$ does not exist? (See line (15.1) and Definition 26.1.)

References

[Ar] E. Artin, "Geometric Algebra", Interscience, New York, 1957.

[Ba] R. Baer, "Linear algebra and projective geometry", Academic Press, New York, 1952.

[B-I] E. Bannai and T. Ito, "Algebraic Combinatorics I: Association Schemes", Lecture Note 58, Benjamin-Cummings, Menlo Park, 1984.

[Bi] G. Birkhoff, "Lattice Theory", 3rd ed., Amer. Math. Soc. Colloq. Publ., vol. 25, Amer. Math. Soc., Providence, R.I., 1967.

[B-G-S] A. Bjorner, A. M. Garsia, R. P. Stanley, An introduction to Cohen-Macauley partially ordered sets,, in "Ordered Sets", (I. Rival, ed.), 1982, pp. 583–615.

[Bo] R. C. Bose, Strongly regular graphs, partial geometries, and partially balanced designs, Pacific J. Math., **13** (1963), 389–419.

[B-C-N] A. E. Brouwer, A. M. Cohen, A. Neumaier, "Distance-Regular Graphs", Springer, Berlin-Heidelberg, 1989.

[Bu1] F. Buekenhout, Diagrams for geometries and groups, J. Combin. Theory Ser. A, **27** (1979), 121–151.

[Bu2] F. Buekenhout, The basic diagram of a geometry, in "Geometries and Groups, Lecture Notes 893", Springer, 1981, pp. 1–29.

[Bu-C] F. Buekenhout and A. Cohen, "Diagram Geometry", to appear.

[Ca1] P. Cameron, Dual polar spaces, Geom. Dedicata, **12** (1982), 75–85.

[Ca2] P. Cameron, "Projective and Polar Spaces", QMW Maths Notes 13.

[C-J-P] H. Cypers, P. Johnson, A. Pasini, On the embeddability of polar spaces, Geom. Dedicata, to appear.

[De] P. Delsarte, Association schemes and t-designs in regular semi-lattices, J. Combin. Theory Ser. A, **20** (1976), 230–243.

[Di] J. Dieudonné, "La géométrie des groupes classiques", Springer, Berlin, 1963.

[Fa] V. Faigle, Geometries on partially ordered sets, J. Combin. Theory Ser. B, **28** (1980), 26-51.

[Hu] T. Huang, A characterization of the association schemes of bilinear forms, European J. Combin., **8** (1987), 159–173.

[K-L-S] B. Korte, L. Lovasz, R. Schrader, "Greedoids", Springer Verlag, Berlin, 1991.

[Mu] A. Munemasa, The geometry of orthogonal groups over finite fields, preprint.

[Ox] J. Oxley, "Matroid Theory", Oxford University Press, Oxford, 1992.

[Pa] A. Pasini, "Diagram Geometries", Oxford University Press, 1994.

[Sh] E. Shult, Characterization of the Lie incidence geometries,, in "Surveys in Cominatorics", London Math. Soc. Lecture Note 82, 1983.

[Sp1] A. Sprague, Incidence structures whose planes are Nets, European J. Combin., **2** (1981), 193–204.

[Sp2] A. Sprague, Pasch's axiom and projective spaces, Discrete Math., **33** (1981), 79–87.

[St] R. Stanley, "Enumerative Combinatorics I", Wadsworth, Belmont, 1986.

[Sta] D. Stanton, Harmonics on posets, J. Combin. Theory Ser. A, **40** (1985), 136–149.

[Ta] D. E. Taylor, "The Geometry of the Classical Groups", Heldermann Verlag, Berlin, 1992.

[Te] P. Terwilliger, The incidence algebra of a uniform poset, in "Coding Theory and Design Theory part I: Coding Theory", IMA volumes in Mathematics and its applications Vol. 20, Springer, New York, 1990, pp. 193–212.

[Ti] J. Tits, "Buildings of Spherical Type and Finite BN-pairs", Lecture Notes in Math. 386, Springer, Berlin, 1974.

[We] D. Welsh, "Matroid Theory", Academic Press, London, 1976.

[Wh] N. White (ed.), "Theory of Matroids", Encyclopedia Math. Appl., Cambridge U. Press, 1992.

[V-Y] O. Veblen, J. W. Young, "Projective Geometry", Boston, 1919.

Department of Mathematics
University of Wisconsin
480 Lincoln Dr., Madison
WI 53706, U.S.A.

Advanced Studies in Pure Mathematics 24, 1996
Progress in Algebraic Combinatorics
pp. 443–453

Geometry of Matrices

Zhe-xian Wan

***In Memory of Professor L. K. Hua* (1910–1985)**

§1. Introduction

The study of the geometry of matrices was initiated by L. K. Hua in the mid forties [5–10]. At first, relating to his study of the theory of functions of several complex variables, he began studying four types of geometry of matrices over the complex field, i.e., geometries of rectangular matrices, symmetric matrices, skew-symmetric matrices, and hermitian matrices. In 1949, he [11] extended his result on the geometry of symmetric matrices over the complex field to any field of characteristic not 2, and in 1951 he [12] extended his result on the geometry of rectangular matrices to any division ring distinct from $\mathbb{F}_2$ and applied it to problems in algebra and geometry. Then the study of the geometry of matrices was succeeded by many mathematicians. In recent years it has also been applied to graph theory.

To explain the problems of the geometry of matrices we are interested in, it is better to start with the Erlangen Program which was formulated by F. Klein in 1872. It says: "A geometry is the set of properties of figures which are invariant under the nonsingular linear transformations of some group". There F. Klein pointed out the intimate relationship between geometry, group, and invariants. Then a fundamental problem in a geometry in the sense of Erlangen Program is to characterize the transformation group of the geometry by as few geometric invariants as possible. The answer to this problem is often called the fundamental theorem of the geometry.

In a geometry of matrices, the points of the associated space are a certain kind of matrices of the same size, and there is a transformation

Received February 28, 1995.
Revised May 16, 1995.
This paper was presented at the International Conference on Algebraic Combinatorics, Fukuoka, Japan, November 22–26, 1993.

group acting on this space. Take the geometry of rectangular matrices as an example. Let D be a division ring, and m and n be integers ≥ 2. The *space* of the geometry of rectangular matrices over D consists of all $m \times n$ matrices over D and is denoted by $\mathcal{M}_{m\times n}(D)$. The elements of $\mathcal{M}_{m\times n}(D)$ are called the *points* of the space. $\mathcal{M}_{m\times n}(D)$ admits transformations of the following form

$$\begin{array}{rcl} \mathcal{M}_{m\times n}(D) & \to & \mathcal{M}_{m\times n}(D) \\ X & \mapsto & PXQ+R, \end{array} \tag{1}$$

where $P \in GL_m(D)$, $Q \in GL_n(D)$, and $R \in \mathcal{M}_{m\times n}(D)$. All these transformations form a *transformation group* of $\mathcal{M}_{m\times n}(D)$, which is denoted by $G_{m\times n}(D)$. Then the geometry of rectangular matrices aims at the study of the invariants of its geometric figures (or subsets) under $G_{m\times n}(D)$. For instance, for the figure formed by two $m\times n$ matrices X_1 and X_2 over D, rank $(X_1 - X_2)$ is an invariant under $G_{m\times n}(D)$. If rank $(X_1 - X_2) = 1$, X_1 and X_2 are called *adjacent.* L. K. Hua proved that the invariant "adjacency" alone is "almost" sufficient to characterize the transformation group $G_{m\times n}(D)$ of $\mathcal{M}_{m\times n}(D)$, which will be explained in detail in the next section.

§2. Geometry of rectangular matrices

Fundamental Theorem of the Geometry of Rectangular Matrices. Let D be a division ring, m and n integers ≥ 2, $\mathcal{A}$ a bijective map from $\mathcal{M}_{m\times n}(D)$ to itself. Assume that both $\mathcal{A}$ and $\mathcal{A}^{-1}$ preserve the adjacency, i.e., for any two points X_1 and X_2 of $\mathcal{M}_{m\times n}(D)$, X_1 and X_2 are adjacent if and only if $\mathcal{A}(X_1)$ and $\mathcal{A}(X_2)$ are adjacent. Then, when $m \neq n$, $\mathcal{A}$ is of the form

$$\mathcal{A}(X) = PX^{\sigma}Q + R \ \text{ for all } X \in \mathcal{M}_{m\times n}(D), \tag{2}$$

where $P \in GL_m(D)$, $Q \in GL_n(D)$, $R \in \mathcal{M}_{m\times n}(D)$, σ is an automorphism of D, and X^{σ} is the matrix obtained from X by applying σ to all its entries. When $m = n$, besides (1) $\mathcal{A}$ can also be of the form

$$\mathcal{A}(X) = P\,{}^t(X^{\tau})\,Q + R \ \text{ for all } X \in \mathcal{M}_{m\times n}(D), \tag{3}$$

where P, Q, and R have the same meaning as above, and τ is an anti-automorphism of D. Conversely, both maps (2) and (3) are bijections, and they and their inverses preserve the adjacency. Q.E.D.

When $D \neq \mathbb{F}_2$, the theorem was proved by L. K. Hua [12] in 1951. The proof for the case $D = \mathbb{F}_2$ was supplemented by Z. Wan and Y.

Wang [24] in 1962. The key tool to prove this theorem is the maximal set introduced by L. K. Hua. A *maximal set* in $\mathcal{M}_{m\times n}(D)$ is a maximal set of points such that any two of them are adjacent. Thus the concept of a maximal set is actually the concept of a maximal clique appeared in graph theory twenty years later. Clearly a bijective map $\mathcal{A}$ for which both $\mathcal{A}$ and $\mathcal{A}^{-1}$ preserve the adjacency carries maximal sets into maximal sets. The main steps Hua used to prove the above theorem is as follows: First he determined the normal forms of maximal sets under $G_{m\times n}(D)$. They are

$$\left\{ \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix} \middle| \, x_{11}, x_{12}, \dots, x_{1n} \in D \right\} \tag{4}$$

and

$$\left\{ \begin{pmatrix} x_{11} & 0 & \dots & 0 \\ x_{21} & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ x_{m1} & 0 & \dots & 0 \end{pmatrix} \middle| \, x_{11}, x_{21}, \dots, x_{m1} \in D \right\}. \tag{5}$$

Then by defining the intersection of two maximal sets which contain two adjacent points in common to be a *line* in any one of them, he proved that $\mathcal{A}$ induces bijective maps on maximal sets, which carries lines into lines and that a line in the maximal set (4) is of the form

$$\left\{ \begin{pmatrix} ta_{11}+b_{11} & ta_{12}+b_{12} & \dots & ta_{1n}+b_{1n} \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix} \middle| \, t \in D \right\}, \tag{6}$$

where $a_{11}, a_{12}, \dots, a_{1n}, b_{11}, b_{12}, \dots, b_{1n} \in D$. When $D \neq \mathbb{F}_2$, by the fundamental theorem of affine geometry, after subjecting $\mathcal{A}$ to a bijective map of the form (2) or (3) (which will be needed only when $m = n$), it can be assumed that $\mathcal{A}$ leaves both the maximal sets (4) and (5) pointwise fixed. Finally it can be proved that $\mathcal{A}$ leaves every point of $\mathcal{M}_{m\times n}(D)$ fixed.

In [12], from the above theorem L. K. Hua deduced the explicit forms of automorphisms, semi-automorphisms, Jordan automorphisms, and Lie automorphisms of the total matrix ring $\mathcal{M}_n(D) (n \geq 2)$ over D. For Jordan automorphisms it is assumed that the characteristic of D is not 2, and for Lie automorphisms it is assumed that the characteristic of D is not 2 and 3. He also deduced the fundamental theorem of the projective geometry of rectangular matrices over D (for detailed proof, cf. [17]). When D is a field, the latter was proved by W. L. Chow [2] in 1949. In 1965, S. Deng and Q. Li [3] deduced the fundamental theorem of the geometry of rectangular matrices over a field from Chow's result.

Call the points of $\mathcal{M}_{m\times n}(D)$ the *vertices* and define two vertices *adjacent* if they are adjacent points. Then a *graph* is obtained. Denote this graph by $\Gamma(\mathcal{M}_{m\times n}(D))$. Naturally, the fundamental theorem of the geometry of rectangular matrices can be interpreted as a theorem on graph automorphisms of $\Gamma(\mathcal{M}_{m\times n}(D))$ [1].

§3. Geometry of alternate matrices

In this section we assume that F is a field and n is an integer ≥ 2. Let A be an $n\times n$ matrix over F. If ${}^tA=-A$ and all entries along the main diagonal of A are 0's, then A is called an $n\times n$ *alternate matrix* over F. Denote by $\mathcal{K}_n(F)$ the set of all $n\times n$ alternate matrices over F, and call it the *space* of the geometry of $n\times n$ alternate matrices and its elements the *points*. Transformations of $\mathcal{K}_n(F)$ to itself of the following form

$$\begin{array}{rcl} \mathcal{K}_n(F) & \to & \mathcal{K}_n(F) \\ X & \mapsto & {}^tPXP+K, \end{array} \tag{7}$$

where $P\in GL_n(F)$ and $K\in\mathcal{K}_n(F)$, form a *transformation group* of $\mathcal{K}_n(F)$, denoted by $GK_n(F)$. Let X_1 and $X_2\in\mathcal{K}_n(F)$. If rank $(X_1-X_2)=2$, then X_1 and X_2 are said to be *adjacent*. Clearly, the adjacency is an invariant under $GK_n(F)$. Conversely, we have

Fundamental Theorem of the Geometry of Alternate Matrices. Let F be a field of any characteristic, n an integer ≥ 4, and $\mathcal{A}$ a bijective map from $\mathcal{K}_n(F)$ to itself. Assume that both $\mathcal{A}$ and $\mathcal{A}^{-1}$ preserve the adjacency. Then, when $n>4$, $\mathcal{A}$ is of the form

$$\mathcal{A}(X)=a\,{}^tPX^{\sigma}P+K \text{ for all } X\in\mathcal{K}_n(F), \tag{8}$$

where $a\in F^*$, $P\in GL_n(F)$, $K\in\mathcal{K}_n(F)$, and σ is an automorphism of F. When $n=4$, $\mathcal{A}$ is of the form

$$\mathcal{A}(X)=a\,{}^tP(X^*)^{\sigma}P+K \text{ for all } X\in\mathcal{K}_4(F), \tag{9}$$

where a, P, K, and σ have the same meaning as above, and $X \to X^*$ is either the identity map of $\mathcal{K}_4(F)$ or the following map

$$(10)\qquad \begin{pmatrix} 0 & x_{12} & x_{13} & x_{14} \\ -x_{12} & 0 & x_{23} & x_{24} \\ -x_{13} & -x_{23} & 0 & x_{34} \\ -x_{14} & -x_{24} & -x_{34} & 0 \end{pmatrix} \mapsto \begin{pmatrix} 0 & x_{12} & x_{13} & x_{23} \\ -x_{12} & 0 & x_{14} & x_{24} \\ -x_{13} & -x_{14} & 0 & x_{34} \\ -x_{23} & -x_{24} & -x_{34} & 0 \end{pmatrix}.$$

Conversely, both maps (8) and (9) are bijective, and they and their inverses preserve the adjacency. Q.E.D.

The above theorem was proved by M. Liu [16] in 1966, the proof relies also on the concept of *maximal sets.* When $F = \mathbb{C}$ and $\mathcal{A}$ satisfies further conditions, it was proved by L. K. Hua [5] in 1945. The map (10) was also discovered by L. K. Hua [5] in 1945.

This theorem has also applications to algebra and geometry [16], and can also be interpreted as a theorem on graph automorphisms [1].

§4. Geometry of symmetric matrices

In this section we assume again that F is a field and n is an integer ≥ 2. An $n \times n$ matrix S over F is called *symmetric* if ${}^tS = S$. Denote by $\mathcal{S}_n(F)$ the set of all $n \times n$ symmetric matrices over F, and call it the *space* of the geometry of $n \times n$ symmetric matrices and its elements the *points.* The set of all transformations of $\mathcal{S}_n(F)$ to itself of the form

$$(11)\qquad \begin{aligned} \mathcal{S}_n(F) &\to \mathcal{S}_n(F) \\ X &\mapsto {}^tPXP + S, \end{aligned}$$

where $P \in GL_n(F)$ and $S \in \mathcal{S}_n(F)$, forms a *transformation group* of $\mathcal{S}_n(F)$, denoted by $GS_n(F)$. Let $X_1, X_2 \in \mathcal{S}_n(F)$. When rank $(X_1 - X_2) = 1$, then X_1 and X_2 are said to be *adjacent.* Clearly, the adjacency of two points in $\mathcal{S}_n(F)$ is an invariant under $GS_n(F)$. Conversely, we have

Fundamental Theorem of the Geometry of Symmetric Matrices. Let F be a field of any characteristic and n be an integer ≥ 2; when F is of characteristic two and $F \neq \mathbb{F}_2$ we assume further that $n \geq 3$. Let $\mathcal{A}$ be a bijective map from $\mathcal{S}_n(F)$ to itself and assume that both $\mathcal{A}$ and $\mathcal{A}^{-1}$ preserve the adjacency. Then unless $n = 3$ and $F = \mathbb{F}_2$, $\mathcal{A}$ is of the form

$$(12)\qquad \mathcal{A}(X) = a\,{}^tPX^\sigma P + S \ \text{ for all } X \in \mathcal{S}_n(F),$$

where $a \in F^*$, $P \in GL_n(F)$, $S \in \mathcal{S}_n(F)$, and σ is an automorphism of F. When $n = 3$ and $F = \mathbb{F}_2$, the bijective map

$$(13)\begin{cases} \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{12} & x_{22} & 0 \\ x_{13} & 0 & x_{33} \end{pmatrix} \mapsto \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{12} & x_{22} & 0 \\ x_{13} & 0 & x_{33} \end{pmatrix} \\ \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{12} & x_{22} & 1 \\ x_{13} & 1 & x_{33} \end{pmatrix} \mapsto \begin{pmatrix} x_{11}+1 & x_{12}+1 & x_{13}+1 \\ x_{12}+1 & x_{22} & 1 \\ x_{13}+1 & 1 & x_{33} \end{pmatrix} \end{cases}$$

from $\mathcal{S}_3(\mathbb{F}_2)$ to itself preserves also the adjacency and $\mathcal{A}$ is a product of maps of the form (12) or (13). Q.E.D.

When $F = \mathbb{C}$ and $\mathcal{A}$ satisfies further conditions, the above theorem was first proved by L. K. Hua [5] in 1945. In 1949 he [11] proved the theorem for any field of characteristic not two by the method of constructing involutions. But there are some gaps in his paper [11] which the author could not fill in. Without any restriction on the characteristic of F the author [18,19] proved the above theorem. In the proof, besides the maximal sets which were defined in the same way as in the geometry of rectangular matrices and were called the *maximal sets of rank* 1 by the author, the *maximal sets of rank* 2 were also introduced. At first, the *distance* $d(X, Y)$ between two points X and Y of $\mathcal{S}_n(F)$ is defined to be the least integer d such that there is a sequence of $d + 1$ points

$$X_0 = X, X_1, X_2, \dots, X_d = Y$$

of $\mathcal{S}_n(F)$ for which any pair of consecutive points X_i and X_{i+1} $(i = 0, 1, 2, \dots, d-1)$ are adjacent. Assume that F is of characteristic not two. Then a subset $\mathcal{L}$ of $\mathcal{S}_n(F)$ is called a *maximal set of rank* 2 if (i) $\mathcal{L}$ contains a maximal set of rank 1, denoted by $\mathcal{M}$, (ii) for any $S \in \mathcal{L} \setminus \mathcal{M}$ and $M \in \mathcal{M}$, $d(S, M) = 2$, and (iii) for any $T \in \mathcal{S}_n(F)$, $d(T, M) = 2$ for all $M \in \mathcal{M}$ implies $T \in \mathcal{L}$. When F is characteristic two, the definition of maximal sets of rank 2 should be modified [19]. Clearly, if $\mathcal{A}$ is a bijective map of $\mathcal{S}_n(F)$ for which both $\mathcal{A}$ and $\mathcal{A}^{-1}$ preserve the adjacency, then $\mathcal{A}$ carries maximal sets of rank 1 into maximal sets of rank 1 and maximal sets of rank 2 into maximal sets of rank 2. The normal form of maximal sets of rank 1 under $GS_n(F)$ is

$$(14) \qquad \left\{ \begin{pmatrix} x & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix} \middle| \, x \in F \right\},$$

and the normal form of maximal sets of rank 2 under $GS_n(F)$ is

$$(15)\qquad \left\{ \left(\begin{array}{cccc} x_{11} & x_{12} & \dots & x_{1n} \\ x_{12} & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ x_{1n} & 0 & \dots & 0 \end{array} \right) \middle| \, x_{11}, x_{12}, \dots, x_{1n} \in F \right\}.$$

Then maximal sets of rank 2 are used in the proof of the above theorem instead of the maximal sets used in the proof of the fundamental theorem of the geometry of rectangular matrices. The case when $n = 2$, F is of characteristic two, and $F \neq \mathbb{F}_2$ still remains open.

When F is of characteristic not two, from the above theorem we can deduce the explicit form of the automorphisms of the Jordan ring of $n \times n$ symmetric matrices over F [18] and the fundamental theorem of the dual polar space of type C_n due to W. L. Chow [2] (cf. [15],[23]).

Call the points of $\mathcal{S}_n(F)$ *vertices*. Two vertices are said to be *adjacent* if they are adjacent as points. Then we obtain a *graph*, denoted by $\Gamma(\mathcal{S}_n(F))$. The fundamental theorem of the geometry of symmetric matrices can naturally be interpreted as a theorem on graph automorphisms of the graph $\Gamma(\mathcal{S}_n(F))$[18, 19].

It is interesting that when F is a finite field of characteristic not two and $n \geq 2$, and when F is a finite field of characteristic two and $n \geq 3$, besides $\Gamma(\mathcal{S}_3(\mathbb{F}_2))$, all $\Gamma(\mathcal{S}_n(F))$ are not distance-transitive. But the author [20] proved that $\Gamma(\mathcal{S}_3(\mathbb{F}_2))$ is distance-transitive, hence, distance-regular, and isomorphic to the graph of the folded 7-cube.

Now assume that F is of characteristic not two. Let $X_1, X_2 \in \mathcal{S}_n(F)$. When rank $(X_1 - X_2) = 1$ or 2, we say that X_1 and X_2 are adjacent. Then we obtain also a *graph*, denoted by $\Gamma^*(\mathcal{S}_n(F))$. From the fundamental theorem of the geometry of symmetric matrices we can deduce that the graph automorphisms of $\Gamma^*(\mathcal{S}_n(F))$ are of the form (12) (cf. [18]). When $F = \mathbb{F}_q$, the graph $\Gamma^*(\mathcal{S}_n(\mathbb{F}_q))$ was defined by Y. Egawa [4], who proved that it is distance-regular and computed its parameters.

§5. Geometry of Hermitian matrices

Let D be a division ring which possesses an involution. Denote the involution of D by $-$, i.e.,

$$(16)\qquad \begin{array}{rcl} - : D & \to & D \\ a & \mapsto & \bar{a}, \end{array}$$

is a bijective map which has the following properties: for any $a, b \in D$ we have

$$\overline{a+b} = \bar{a} + \bar{b}, \tag{17}$$

$$\overline{ab} = \bar{b}\bar{a}, \tag{18}$$

and

$$\bar{\bar{a}} = a. \tag{19}$$

Let

$$F = \{a \in D \mid \bar{a} = a\}. \tag{20}$$

Define the trace map

$$\begin{aligned} Tr : D &\rightarrow F \\ a &\mapsto a + \bar{a} \end{aligned} \tag{21}$$

and the norm map

$$\begin{aligned} N : D &\rightarrow F \\ a &\mapsto a\bar{a}. \end{aligned} \tag{22}$$

We make the following assumptions:

Assumption I F is a proper subfield of D and is contained in the center of D.

Assumption II The map Tr is surjective.

We remark that Assumption I excludes the case when D is a field and $-$ is the identity map.

Let n be an integer ≥ 2. An $n \times n$ matrix H over D is called *hermitian* if ${}^t\overline{H} = H$. The *space* of the geometry of hermitian matrices over D, denoted by $\mathcal{H}_n(D)$, is the set of all $n \times n$ hermitian matrices over D, whose elements are called the *points*. The set of transformations of $\mathcal{H}_n(D)$ to itself of the form

$$\begin{aligned} \mathcal{H}_n(D) &\rightarrow \mathcal{H}_n(D) \\ X &\mapsto {}^t\overline{P}XP + H, \end{aligned} \tag{23}$$

where $P \in GL_n(D)$ and $H \in \mathcal{H}_n(D)$, forms a *transformation group* of the space $\mathcal{H}_n(D)$, which is denoted by $GH_n(D)$. Let $X_1, X_2 \in \mathcal{H}_n(D)$. When rank $(X_1 - X_2) = 1$ then X_1 and X_2 are said to be *adjacent*. Clearly, the adjacency of two points is an invariant under $GH_n(D)$. Conversely, we have

Fundamental Theorem of the Geometry of Hermitian Matrices. Let D be a division ring which possesses an involution and assume that Assumptions I and II hold. Let n be an integer ≥ 2 and when $n = 2$ we assume that D is a field. Let $\mathcal{A}$ be a bijective map from $\mathcal{H}_n(D)$ to itself and assume that both $\mathcal{A}$ and $\mathcal{A}^{-1}$ preserve the adjacency. Then $\mathcal{A}$ is of the form

$$\mathcal{A}(X) = \alpha^t\overline{P}X^\sigma P + H \text{ for all } X \in \mathcal{H}_n(D), \tag{24}$$

where $\alpha \in F^*$, $P \in GL_n(D)$, $H \in \mathcal{H}_n(D)$, and σ is an automorphism of D which commutes with the involution $-$ of D. If we assume further that the norm map N is bijective, then we can assume that $\alpha = 1$. Q.E.D.

The above theorem was proved by the author [21,22] recently. In the proof, besides the maximal sets of rank 1 and rank 2, which were defined in a similar way as those in the geometry of symmetric matrices, the reduced maximal sets of rank 2 are also introduced. The normal form of maximal sets of rank 1 under $GH_n(D)$ is

$$\left\{ \left. \begin{pmatrix} x_{11} & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix} \right| x_{11} \in F \right\}, \tag{25}$$

and the normal form of maximal sets of rank 2 under $GH_n(D)$ is

$$\left\{ \left. \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ \overline{x_{12}} & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ \overline{x_{1n}} & 0 & \dots & 0 \end{pmatrix} \right| x_{11} \in F, x_{12}, \dots, x_{1n} \in D \right\}. \tag{26}$$

If $\mathcal{M}$ is a maximal set of rank 1, then there is a unique maximal set $\mathcal{L}$ of rank 2 containing $\mathcal{M}$. For any $\mathcal{M}$ containing the zero matrix 0, $\mathcal{L}$ has an additive group structure with respect to matrix addition, $\mathcal{M}$ is its subgroup, and the set of cosets of $\mathcal{L}$ relative to $\mathcal{M}$ is called a *reduced maximal set of rank* 2. Clearly, the reduced maximal set of rank 2 from $\mathcal{L}$ are all the maximal sets of rank 1 contained in $\mathcal{L}$. Hence, if we assume that $\mathcal{A}(0) = 0$, then $\mathcal{A}$ carries reduced maximal sets of rank 2 to reduced maximal sets of rank 2. The reduced maximal sets of rank 2 are used in the proof of the above theorem when $n \geq 3$ as the maximal sets in the proof of the fundamental theorem of the geometry of rectangular matrices. When $n = 2$ and D is a field, the theorem can be proved by studying three maximal sets of rank 1 which have a nonempty intersection [22]. The case when $n = 2$ and D is not a field still remains open.

When $D = \mathbb{C}$ and $\mathcal{A}$ satisfies some other conditions, the above theorem was proved by L. K. Hua [5] in 1945. When $D = \mathbb{F}_q$, it was proved by A. A. Ivanov and S. V. Shpectorov [14] in 1991.

The above theorem has also applications to algebra [22] and geometry [23], and can also be interpreted as a theorem on graph automorphisms [22].

References

[1] A. E. Brouwer, A. M. Cohen and A. Neumaier, Distance-regular Graphs, Springer-Verlag, 1989.

[2] W. L. Chow, On the geometry of algebraic homogeneous spaces, Ann. of Math. (2), **50** (1949), 32–67.

[3] S. Deng and Q. Li, On the affine geometry of algebraic homogeneous spaces, Acta Math. Sinica, **15** (1965), 651–663 (in Chinese). English translation: Chinese Mathematics, **7** (1965), 387–391.

[4] Y. Egawa, Association schemes of quadratic forms, J. Combin. Theory Ser. A, **38** (1985), 1–14.

[5] L. K. Hua, Geometries of matrices I. Generalizations of van Staudt's theorem, Trans. Amer. Math. Soc., **57** (1945), 441–481.

[6] L. K. Hua, Geometries of matrices I_1. Arithmetical construction, Trans. Amer. Math. Soc., **57** (1945), 482–490.

[7] L. K. Hua, Geometries of symmetric matrices over the real field I, Dokl. Akad. Nauk. SSSR(N.S.), **53** (1946), 95–97.

[8] L. K. Hua, Geometries of symmetric matrices over the real field II, Dokl. Akad. Nauk. SSSR(N.S.), **53** (1946), 195–196.

[9] L. K. Hua, Geometries of matrices II. Study of involutions in the geometry of symmetric matrices, Trans. Amer. Math. Soc., **61** (1947), 193–228.

[10] L. K. Hua, Geometries of matrices III. Fundamental theorems in the geometries of symmetric matrices, Trans. Amer. Math. Soc., **61** (1947), 229–255.

[11] L. K. Hua, Geometries of symmetric matrices over any field with characteristic other than two, Ann. of Math., **50** (1949), 8–31.

[12] L. K. Hua, A theorem on matrices over a sfield and its applications, Acta Math. Sinica, **1** (1951), 109–163.

[13] L. K. Hua and Z. Wan, Classical Groups, Shanghai Science and Technology Press, Shanghai, 1963 (in Chinese).

[14] A. A. Ivanov and S. V. Shpectorov, A characterization of the association schemes of Hermitian forms, J. Math. Soc. Japan, **43** (1991), 25–48.

[15] M. Liu, A proof of the fundamental theorem of projective geometry of symmetric matrices, Shuxue Jinzhan, **8** (1965), 283–292 (in Chinese).

[16] M. Liu, Geometry of alternate matrices, Acta Math. Sinica, **16** (1966), 104–135 (in Chinese). English translation: Chinese Mathematics, **8** (1966), 108–143.

[17] D. Pei, A proof of the fundamental theorem of the projective geometry of rectangular matrices, A Collection of Papers Celebrating the Fifth Anniversary of the Chinese University of Science and Technology, 99–110 (in Chinese).

[18] Z. Wan, Geometry of symmetric matrices and its applications I, Algebra Colloq., **1** (1994), 97–120.

[19] Z. Wan, Geometry of symmetric matrices and its applications II, Algebra Colloq., **1** (1994), 201–224.

[20] Z. Wan, The Graph of 3×3 binary symmetric matrices, Northeast. Math. J., **11** (1995), 1–2.

[21] Z. Wan, Geometry of hermitian matrices and its applications I, Algebra Colloq., **2** (1995), 167-192.

[22] Z. Wan, Geometry of hermitian matrices and its applications II, to appear in Algebra Colloq.

[23] Z. Wan, Geometry of Matrices, in preparation.

[24] Z. Wan and Y. Wang, Discussion on “A theorem on matrices over a sfield and its applications”, Shuxue Jinzhan, **5** (1962), 325–332 (in Chinese).

Institute of Systems Science
Chinese Academy of Science
Beijing, China
and
Department of Information Theory
Lund University
Lund, Sweden