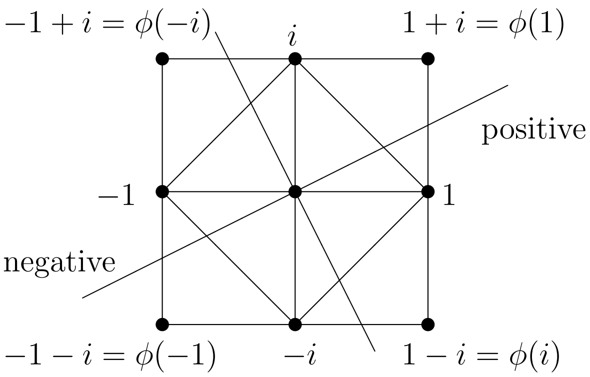
The Suzuki groups

Robert A. Wilson

University of Birmingham, 23rd March 2011

1 The roots

The square. We begin by studying the symmetry group of the square.



The dihedral group of order 8. In the complex plane, the symmetry group is generated by

- rotation $r : z \mapsto iz$, and
- reflection $s : z \mapsto \overline{z}$.

These give rise to the other

- rotations $z \mapsto -z$, $z \mapsto -iz$, $z \mapsto z$, and
- reflections $z \mapsto i\overline{z}$, $z \mapsto -\overline{z}$, $z \mapsto -i\overline{z}$.

Short and long roots. These symmetries can be thought of as permutations of either

- the set $\mathcal{S} = \{\pm 1, \pm i\}$ of four *short roots*, or
- the set $\mathcal{L} = \{\pm 1 \pm i\}$ of four *long roots*.

Explicitly, we have

- $r = (1, i, -1, -i) = (1+i, -1+i, -1-i, 1-i)$,
- $s = (i, -i) = (1+i, 1-i)(-1+i, -1-i)$.

Swapping short and long. In other words we can think of the symmetries of either the *square* or the *diamond*. Reflecting in a suitable line swaps the two (on a suitable scale):

$$z \mapsto \frac{1+i}{\sqrt{2}}\overline{z}.$$

Scaling up or down by a factor of $\sqrt{2}$ gives:

- $\phi : z \mapsto (1+i)\overline{z} : \mathcal{S} \to \mathcal{L}$;
- $\psi : z \mapsto \frac{1}{2}(1+i)\overline{z} : \mathcal{L} \to \mathcal{S}$.

A new addition on short roots. If we add together two perpendicular short roots, we get a long root. The map ψ then takes this result to a short root. This gives us a kind of 'addition', $x \oplus y = \psi(x+y)$ where this is defined, and $x \oplus y = 0$ otherwise. Specifically, we have

$$\begin{array}{rcl} 1 \oplus i & = & 1 \\ 1 \oplus (-i) & = & i \\ -1 \oplus i & = & -i \\ -1 \oplus (-i) & = & -1 \end{array}$$

Grading. If we grade the short (and/or long) roots by their projections onto the axis of the reflection $z \mapsto (1+i)\overline{z}/\sqrt{2}$, that is in the order $-1 < -i < i < 1$, then we find that $\oplus$ respects this grading in the sense that if $y < z$ then $x \oplus y < x \oplus z$ whenever this is defined. This is obvious, because both ordinary addition of roots, and ψ, preserve this grading.

2 The trunk

We shall use the structure of the roots to define some structures on a 4-dimensional vector space W over a suitable field F of characteristic 2. We first investigate the field.

The field. Our field must have order $q = 2^{2k+1}$, for some non-negative integer k, so is built as a suitable quotient $\mathbb{Z}_2[X]/(f)$, where f is an irreducible polynomial of degree $2k+1$ over $\mathbb{Z}_2$, the integers modulo 2.

Field automorphisms. The *Frobenius map* $x \mapsto x^2$ is an automorphism because

- $(xy)^2 = x^2y^2$ because multiplication is commutative, and
- $(x+y)^2 = x^2 + y^2$ because the cross term $2xy$ is zero, since $2 = 0$.

This automorphism has order $2k+1$, since the multiplicative group of F has order $2^{2k+1} - 1$, so $x^{2^{2k+1}-1} = 1$ for all non-zero $x \in F$, that is $x^{2^{2k+1}} = x$ for all $x \in F$.

Conversely, every automorphism is determined by where it takes $[X]$: it must go to one of the $2k+1$ roots of f. Thus the automorphism group of this field is cyclic of order $2k+1$.

The square root of two. Consider the map σ on F defined by $x^\sigma = x^{2^{k+1}}$. Then σ squares to the map $x \mapsto x^{2^{2k+2}} = x^2$, that is to the Frobenius map. Informally, we have $x^{\sigma^2} = x^2$ so we can write $\sigma = \sqrt{2}$. For simplicity later on, write $\tau = \sigma^{-1}$.

The vector space. We define W to be the 4-dimensional vector space over F with basis $\{e_{-1}, e_{-i}, e_i, e_1\}$, that is e_x for $x \in \mathcal{S}$.

The inner product. We put a symmetric inner product on W by defining

- $e_x.e_{-x} = 1$, and
- $e_x.e_y = 0$ otherwise,

and extending bilinearly. Thus the inner product is a bilinear map $W \times W \to F$.

The outer product. We define a symmetric outer product $\star : W \times W \to W$, by

- $e_x \star e_y = e_{x \oplus y}$ whenever $x \oplus y$ is defined, and
- $e_x \star e_y = 0$ otherwise,

and extending semi-linearly via

$$u \star (v + \lambda w) = u \star v + \lambda^\tau (u \star w).$$

The restricted outer product. The bilinear inner product $. : W \times W \to F$ corresponds naturally to a linear map $W \otimes W \to F$, defined by $\sum_i u_i \otimes v_i \mapsto \sum_i u_i.v_i$. By rank–nullity, this latter map has kernel (null-space) of codimension 1 in $W \times W$. Call this kernel K.

Similarly, the outer product $\star : W \times W \to W$ corresponds to a map $W \otimes W \to W$ defined by $\sum_i u_i \otimes v_i \mapsto \sum_i u_i \star v_i$. From now on we only consider the outer product restricted to the subspace K of codimension 1.

3 The branches

We shall be interested in the group of all linear maps on W which preserve both the inner product, and the restricted outer product. First we construct a few elements and subgroups of this group.

Coordinate permutations. The coordinate permutation $e_x \mapsto e_{-x}$ (extended linearly to the map $\sum_x \lambda_x e_x \mapsto \sum_x \lambda_x e_{-x}$) is clearly a symmetry of the whole construction. This element generates a group of order 2 which I shall call the *Weyl group.*

Diagonal matrices. Consider the map

$$d_\lambda : \sum_x \alpha_x e_x \mapsto \sum_x \lambda_x \alpha_x e_x,$$

where $\lambda_1 = \lambda$, $\lambda_i = \lambda^{\sigma-1}$, and $\lambda_{-x} = \lambda_x^{-1}$. That is, d_λ is the diagonal map $\mathrm{diag}(\lambda^{-1}, \lambda^{1-\sigma}, \lambda^{\sigma-1}, \lambda)$. Since this map is inverted by conjugation by the non-trivial element of the Weyl group, in order to check that it preserves the two products it is sufficient to check:

- $(\lambda_x e_x).(\lambda_{-x} e_{-x}) = \lambda_x \lambda_{-x} (e_x.e_{-x}) = 1$,
- $(\lambda_1 e_1) \star (\lambda_i e_i) = (\lambda^{1+\sigma-1})^\tau e_1 = \lambda_1 e_1$,
- and, since $\lambda = \lambda_i^{\sigma+1}$, we have $(\lambda_1 e_1) \star (\lambda_{-i} e_{-i}) = (\lambda_i^\sigma)^\tau e_i = \lambda_i e_i$,

as the other products are all zero.

Thus we obtain a cyclic group of order $q-1$, as λ ranges over all the non-zero elements of the field. I will call this group the *maximal torus* T. Since it is normalised by W, we see that $N = TW$ is a dihedral group of order $2(q-1)$.

A unitriangular matrix. Consider next the matrix

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

interpreted as the linear map which fixes e_{-1} and maps

$$\begin{array}{rcl} e_{-i} & \mapsto & e_{-i} + e_{-1} \\ e_i & \mapsto & e_i + e_{-i} \\ e_1 & \mapsto & e_1 + e_i + e_{-i} + e_{-1} \end{array}$$

It is easy enough to check by eye that this preserves the inner product. To check the restricted outer product we have five things to check:

- $e_{-1} \star (e_{-i} + e_{-1}) = e_{-1} \star e_{-i}$;
- $e_{-1} \star (e_i + e_{-i}) = e_{-1} \star e_i + e_{-1} \star e_{-i} = e_{-i} + e_{-1}$;
- $(e_{-i} + e_{-1}) \star (e_1 + e_i + e_{-i} + e_{-1}) = (e_{-i} + e_{-1}) \star (e_1 + e_i) = ((e_{-i} \star e_1) + (e_{-1} \star e_i) + (e_{-i} \star e_i + e_{-1} \star e_1) = e_i + e_{-i}$;
- $(e_i + e_{-i}) \star (e_1 + e_i + e_{-i} + e_{-1}) = (e_i + e_{-i}) \star (e_1 + e_{-1}) = e_1 + e_{-1} + e_i + e_{-i}$,
- $e_{-1} \star (e_1 + e_i + e_{-i} + e_{-1}) + (e_{-i} + e_{-1}) \star (e_i + e_{-i}) = e_{-1} \star e_1 + e_{-i} \star e_i = 0$.

The lower unitriangular subgroup. It is easy to check that this matrix squares to the matrix

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

which has order 2. Conjugating this by d_λ gives a matrix with $\begin{pmatrix} \lambda^{-\sigma} & 0 \\ \lambda^{-2} & \lambda^{-\sigma} \end{pmatrix}$ in the bottom left-hand corner. Since $\lambda \mapsto \lambda^{-\sigma}$ is a bijection on the non-zero elements of the field, this gives us a group of order q. Similarly, if we conjugate the original unitriangular matrix by d_λ, we get a matrix with $\begin{pmatrix} 1 & 0 \\ \lambda^{\sigma-1} & 1 \end{pmatrix}$ in the top left-hand corner. This means that modulo the first group of order q, we get another group of order q, lifting to a group of order q^2 altogether.

This group of order q^2 is called the *unipotent* subgroup U, and the group $B = UT$ is a group of order $q^2(q-1)$, consisting of lower triangular matrices. I shall call it the *Borel subgroup*.

4 The leaves

Definition. Some of the non-zero vectors $v \in W$ have the special property that $v = v \star w$ for some w. Let us call such a vector a *leaf*. Obviously, if v is a leaf, then so is λv, for all $\lambda \neq 0$. Observe that $e_1 = e_1 \star e_i$, so e_1 is a leaf. Therefore so is e_{-1}, its image under the Weyl group. Moreover, the Borel subgroup maps e_1 to $q^2(q-1)$ distinct leaves, coming in q^2 distinct 1-spaces. Thus we get $(q^2+1)(q-1)$ leaves altogether so far, or $q^2 + 1$ up to scalar multiplication.

No more leaves. I claim that these are all the leaves. First note that when we order the coordinates via $-1 < -i < i < 1$, the leading term of v must appear on both sides of one of the defining equations $e_1 \star e_i = e_1$ etc., so must be e_1 or e_{-1}. We may also assume the leading coefficient is 1. If the leading term is e_{-1}, then $v = e_{-1}$.

If the leading term is e_1, then we can use the 'top part' of the unipotent subgroup U to clear out the term in e_i. After that, we can use the 'bottom part' of U to clear out the term in e_{-i}. Thus we can assume $v = e_1 + \lambda_{-1}e_{-1}$. Since $v \star v = 0$ we can assume w has no term in e_1, and the leading term of w is e_i in order to get the leading term correct in

$$(e_1 + \lambda_{-1}e_{-1}) \star (e_i + \cdots) = e_1 + \lambda_{-1}e_{-i} + \cdots .$$

But now we cannot get another term in e_{-i} from the lower terms in the product, so we must have $\lambda_{-1} = 0$. This completes the proof that there are no more leaves.

The ovoid. For conformity with the standard terminology, define a *point* to be a 1-space $\langle v \rangle$ spanned by a leaf v. We have shown that there are exactly $q^2 + 1$ points. These points form the Suzuki–Tits *ovoid*.

Transitivity. We have seen that the symmetries given above generate a group which acts transitively on the $q^2 + 1$ points. Moreover, the stabilizer of the point $\langle e_{-1} \rangle$ acts transitively on the remaining q^2 points, so the group acts 2-transitively (and therefore primitively).

The stabilizer of two points. If we fix the points $\langle e_{-1} \rangle$ and $\langle e_1 \rangle$ then we must fix

- $\langle e_{-1} \rangle^{\perp} = \langle e_{-1}, e_{-i}, e_i \rangle$,
- $e_1 \star W = \langle e_1, e_i \rangle$, and
- their intersection $\langle e_i \rangle$; and
- similarly $\langle e_{-i} \rangle$.

So any such symmetry must be diagonal, say $e_x \mapsto \lambda_x e_x$. To preserve the inner product we must have

$$1 = e_x.e_{-x} = (\lambda_x e_x).(\lambda_{-x}e_{-x}) = \lambda_x\lambda_{-x}e_x.e_{-x} = \lambda_x\lambda_{-x}$$

so $\lambda_{-x} = \lambda_x^{-1}$.

To preserve the restricted outer product, we must have

$$\lambda_1 e_1 = (\lambda_1 e_1) \star (\lambda_i e_i) = (\lambda_1\lambda_i)^{\tau} e_1 \star e_i = (\lambda_1\lambda_i)^{\tau} e_1,$$

so $\lambda_1 = (\lambda_1\lambda_i)^{\tau}$, and therefore $(\lambda_1)^{\sigma} = \lambda_1\lambda_i$, so $\lambda_i = \lambda_1^{\sigma-1}$. Hence the group of diagonal symmetries is no bigger than the group of order $q-1$ already constructed.

We have shown that the stabilizer of two points has order $q - 1$.

5 The fruit

The Suzuki group. The group $Sz(F) = Sz(q)$ is defined as the set of linear maps on W which preserve the inner product and the restricted outer product. We have shown that $Sz(q)$ acts 2-transitively on the set of q^2+1 *points*, and that the two-point stabilizer has order $q-1$. Hence

$$|Sz(q)| = (q^2+1)q^2(q-1).$$

A group-theoretic lemma. If G is a permutation group which is primitive, and perfect, such that the point stabilizer is soluble, then G is simple.

Proof: Let H be the stabilizer of one of the points. Suppose, for a contradiction, that K is a normal subgroup of G, with $1 < K < G$. Then

- H is maximal, since G is primitive;
- K does not fix all the points, so does not fix any point, so $K \not\subseteq H$;
- hence $KH = G$;
- therefore $G/K = HK/K \cong H/H \cap K$ is soluble.

But this contradicts the assumption that G is perfect.

Simplicity. We assume that $q > 2$, and verify the hypotheses of this lemma:

- $Sz(q)$ is primitive on q^2+1 points, since it is 2-transitive.
- $Sz(q)$ is generated by conjugates of its maximal subgroup H, which is in turn generated by conjugates of T. But $T = (TW)'$ so $T \subseteq G'$, and therefore $Sz(q)$ is perfect.
- the point stabilizer has order $q^2(q-1)$, so is equal to B, which is lower-triangular, and therefore soluble.

We conclude that $Sz(q)$ is simple whenever $q > 2$.

The case $q = 2$. On the other hand, if $q = 2$, we have $|Sz(2)| = 5.4$, and $Sz(2)$ acts 2-transitively on the 5 points. Hence $Sz(2)$ is isomorphic to the Frobenius group of order 20, generated by the permutations $(0, 1, 2, 3, 4)$ and $(1, 2, 4, 3)$.