The Suzuki groups

Robert A. Wilson

University of Manchester, 19th February 2013

What's purple and twisted? — A Suzuki grape!

Introduction

There are ten families of finite simple exceptional groups of Lie type, of which two were known to L. E. Dickson (1901–5), and the rest were discovered around 1955–60. The last three of these were the Suzuki groups and two families of Ree groups, which have remained somewhat mysterious to this day. Over the past few years I have discovered some new ways to construct them, which I feel are considerably simpler than previous methods.

Suzuki constructed his groups as groups of 4×4 matrices over fields of characteristic 2. The small Ree groups can be written as 7×7 matrices over fields of characteristic 3, and the large Ree groups as 26×26 matrices over fields of characteristic 2. (There is a sense in which the last of these can be thought of as 13×13 matrices over extensions of $GF(4)$ with a field automorphism: in this sense the three case then have projective dimensions 3, 6 and 12 respectively, over F^{2k+1}, where $F = GF(2), GF(3), GF(4)$ respectively.)

The standard constructions are often expressed in the language of *buildings* (due to Tits), but we shall go back to basics and adopt a more rural, specifically viticultural, metaphor, to describe the way these groups (Suzuki grapes?!) are grown (groan!). Much of the terminology is already agricultural: everything is done over a *field*, and arises from a *root system*. The novelty of my approach is a new *product* which is added to the process. Otherwise the process of growing grapes and making wine is much the same as it always has been, since the days of Suzuki, Ree and Tits, 50 years ago.

The talk is divided into several sections:

1. Ploughing the **field**.
2. Grading the **roots**.
3. Imposing a **linear** structure, applying **products**, and **pruning**.

4. **Growing** the groups.
5. **Harvesting** the crop.
6. **Fermenting** the juice.
7. **Drinking** the wine.
8. Conclusion: distilling the **spirit**.

1 Ploughing 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 $2 = 0$.

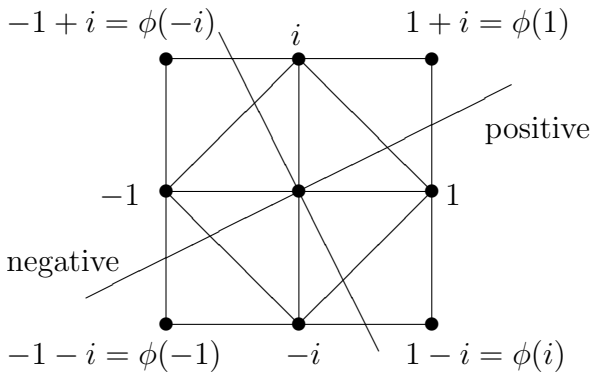
This automorphism has order $2k+1$, since the multiplicative group of F has order $2^{2k+1} - 1$, so $x^{2^{2k+1}-1} = 1$ for all non-zero $x \in F$, that is $x^{2^{2k+1}} = x$ for all $x \in F$.

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 $\tau = \sigma^{-1}$.

2 Examining and grading the roots

The root system. Our roots form a *root system*, which is embedded in a (complex) plane (plain?). 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. Linear maps from short roots to long are of two types: $z \mapsto (\pm 1 \pm i)z$ and $z \mapsto (\pm 1 \pm i)\overline{z}$. The latter square to $z \mapsto 2z$ and have eigenvalues $\pm\sqrt{2}$. We choose

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

Notice that $\phi^2 = 2$. Indeed, ϕ has two eigenvalues, $\pm\sqrt{2}$, and the corresponding eigenspaces are drawn over the frame in the picture above.

Short roots for long. 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

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

Grading the roots. 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.

3 Linearity, products, and pruning

We need to put various *products* and *structures* on the roots in order to help them grow. We shall use the structure of the root system to define these on a 4-dimensional vector space W over the field F.

The vector space. Each root r grows into a stem e_r. These four stems, together with *linear combinations* over the *field*, give us a vector space. More formally, 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 (indeed, alternating) inner product on W by defining

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

and *extending bilinearly over the field*(!). Thus the inner product is a bilinear map $W \times W \to F$.

The outer product. The outer product is only used for the Ree groups. It identically zero in the case of the Suzuki groups, so we need not concern ourselves with it.

The middle product. This is the new magic ingredient, which vastly simplifies the whole process of viticulture. We define a symmetric middle 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 non-zero products of basis vectors are

$$\begin{aligned} e_1 \star e_i &= e_1 \\ e_1 \star e_{-i} &= e_i \\ e_{-1} \star e_i &= e_{-i} \\ e_{-1} \star e_{-i} &= e_{-1} \end{aligned}$$

Pruning the middle product. (Apologies for mixing my metaphors.) The bilinear inner product $. : W \times W \to F$ corresponds naturally to a linear map $W \otimes W \to F$, or better, $W \wedge W \to W$, defined by $\sum_i u_i \wedge v_i \mapsto \sum_i u_i.v_i$. By rank–nullity, this latter map has kernel (null-space) of codimension 1 in $W \wedge W$. Call this kernel K.

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

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

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

4 The groups begin to grow

We shall be interested in the group of all linear maps on W which preserve both the inner product, and the restricted middle 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*. So far we have only a very tiny group: it is really only a bud.

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:

- $(\lambda_x e_x).(\lambda_{-x} e_{-x}) = \lambda_x \lambda_{-x} (e_x.e_{-x}) = 1$,
- $(\lambda_1 e_1) \star (\lambda_i e_i) = ((\lambda_1)^{1+\sigma-1})^\tau e_1 = \lambda_1 e_1$,
- 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 $2(q-1)$. Our group has blossomed into a flower.

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

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

(In fact, if we just specify $e_{-1} \mapsto e_{-1}$ and $e_i \mapsto e_i + e_{-i}$, then the star product determines the images of e_{-i} and e_1.) It is easy enough to check by eye that this preserves the inner product. To check the restricted middle 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*. Our groups are starting to bear fruit!

5 Harvesting the grapes

(Again, apologies for mixing my metaphors, and using 'grapes' for two different things.)

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 *grape*. Obviously, if v is a grape, then so is λv, for all $\lambda \neq 0$. Observe that $e_1 = e_1 \star e_i$, so e_1 is a grape. Therefore so is e_{-1}, its image under the Weyl group. Moreover, the Borel subgroup maps e_1 to $q^2(q-1)$ distinct grapes, coming in q^2 distinct 1-spaces (or *bunches of grapes*, as we might call them). Thus we get $(q^2+1)(q-1)$ grapes altogether so far, or q^2+1 *bunches*, up to scalar multiplication.

No more grapes. I claim that these are all the grapes. First note that the grading on the roots lifts to a grading on the bunches of grapes. That is, when we order the coordinates via the ordering on roots $-1 < -i < i < 1$, 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}. (So we only have two grades of grapes.) 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 grapes. The vines have been picked clean.

The harvest. In the standard terminology, what we have called a bunch of grapes is called a *point*, that is, a point is a 1-space $\langle v \rangle$ spanned by a grape v. We have shown that there are exactly $q^2 + 1$ points. (Of course, the bigger your field, the more bunches of grapes you get.) These points form the Suzuki–Tits *ovoid*, which therefore consists of our entire grape harvest. It's called an ovoid because it's considered to be egg-shaped, but it is clearly better to think of it as a big *fermenting bin* into which we put all our bunches of grapes.

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). In other words, the grapes are well mixed in the fermenting bin.

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 middle 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 $q - 1$.

6 Fermenting the juice

Some yeast from the group-theory cupboard. 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 $(1, 2, 4, 3)$.

7 The symposium

To summarise, we planted the *roots* in a *field* of characteristic 2 (and odd degree), and we put various *products* on them, and we watched as the plants produced fruit at particular *points* in the *space* over the field. We counted the bunches of grapes and deduced the *order* of the symmetry group. We showed that as long as the field is not *prime*, the Suzuki group is *perfect*, and using some group yeast deduced that it is *simple*.

Perfect grapes make vintage wine, and although it can be drunk now, I hope it will improve with age, and still be drunk with enjoyment many years from now.

And I think if you *distill* the *spirit* of my talk you will find that it is about *70 per cent proof*.