Innovations in Incidence Geometry
Volume 1 (2005), Pages 67–131
ISSN 1781-6475

Algebraic structure of the perfect Ree-Tits generalized octagons

Kris Coolsaet

Abstract

We provide an algebraic description of the perfect Ree-Tits generalized octagons, i.e., an explicit embedding of octagons of this type in a 25-dimensional projective space. The construction is derived from the interplay between the 52-dimensional module of the Chevalley algebra of type F_4 over a field of even characteristic and its 26-dimensional submodule. We define a quadratic duality operator that interchanges special sets of (totally) isotropic elements in those modules and establish the points of the octagon as absolute points of this duality. We introduce many algebraic operations that can be used in the study of the generalized octagon. We also prove that the Ree group acts as expected on points and pairs of points.

Keywords: Ree-Tits generalized octagon, Chevalley algebra of type F4

MSC 2000: 51E12, 17B25

Contents

1 Introduction and overview 68
2 The Chevalley algebra of type F_4 71
3 Isotropic and totally isotropic elements 84
4 Automorphisms 89
5 The quotient algebra $\mathbf{Q}$ 97
6 The operator $Q(\cdot)$ 99
7 Octagonality 107
8 The Ree group ${}^2\widehat{\mathsf{F}_4}(K)$ 109
9 Pairs of octagonal elements 116
10 The Ree-Tits octagon 122

1 Introduction and overview

Of all generalized polygons related to simple groups [5] the generalized octagons have probably been studied the least. Perhaps one of the reasons is the fact that only a single (infinite) family of examples is known and that these *Ree-Tits* octagons, and their embeddings in a projective space, are not easily constructed.

The standard way to define a Ree-Tits octagon is as a coset geometry of a Ree group. This group is itself constructed by 'twisting' the Chevalley group of type F_4 over a suitable field K of characteristic 2. Another construction technique makes use of a special coordinatization (introduced by H. Van Maldeghem [5]) which is itself based on the properties of the Ree groups, in particular the commutation relations displayed in [4].

What is missing for the Ree-Tits octagon, and exists for all other classical generalized polygons, is an explicit embedding into some projective?space. In Section 10 of this text we shall construct such an embedding into a 25-dimensional space, where the points (lines) of the Ree-Tits octagon are a subset of the points (lines) of the projective space. We will provide explicit 'formulae' which can be applied to the projective coordinates of a point to determine whether or not it is a point of the octagon. And likewise, we shall provide a means to determine from the projective coordinates of two points of the octagon, whether they are collinear in the octagon, and more generally, what is their mutual distance.

These 'formulae' are not so elementary as in the case of the other generalized polygons. For example, we shall prove that a point Ke belongs to the Ree-Tits octagon if and only if both $e^2 = 0$ and $q([e, \mathbf{W}], e) = 0$. The first expression is shorthand for a system of 26 quadratic equations in 26 variables (the 26 coordinates of e). The second expression corresponds to a system of 676 polynomials of degree 3 in 26 variables, additionally involving a field automorphism σ of K. (For contrast, compare this to the formulae needed to define the points of the split Cayley hexagon : a single quadratic equation in 7 variables suffices.)

This text looks at the octagon (and the metasymplectic space to which its points and lines belong) from three different perspectives : groups, geometries and algebras, concentrating on the latter.

We start with the well-known 52-dimensional Chevalley algebra $\mathbf{F}$ of type F_4 and its 26-dimensional module $\mathbf{W}$ in the special case of characteristic 2. In this particular case $\mathbf{W}$ is a subalgebra of $\mathbf{F}$. (Many of the definitions and propositions in the first few sections of this text can also be extended to fields with different characteristics.)

$\mathbf{W}$ is isomorphic to the exceptional (quadratic) 26-dimensional Jordan algebra over the field K of which $\mathbf{F}$ is the algebra of derivations. (We have chosen to work independently of this fact, primarily because Jordan algebras, especially in characteristic 2, are not so easily accessible and we do not want to assume prior knowledge of them in this paper.)

Not all elements of $\mathbf{W}$ and $\mathbf{F}$ are of the same interest to us. Instead we single out a set of so-called *isotropic* elements of $\mathbf{W}$ and a set of *totally isotropic* elements of $\mathbf{F}$. These special elements allow us to attach geometrical meaning to the algebras $\mathbf{W}$ and $\mathbf{F}$: isotropic elements (or more correctly 1-dimensional subspaces of isotropic elements) turn out to be *points* of the associated metasymplectic space, and totally isotropic elements correspond to *symplecta*.

Isotropic and totally isotropic elements also allow us to define a Chevalley group $\widehat{\mathsf{F}_4}(K)$ of type F_4 and its action on $\mathbf{W}$ and $\mathbf{F}$. For each isotropic element e there is a corresponding group element $x(e)$ and each totally isotropic element E corresponds to a group element $x(E)$. The Chevalley group is generated by all these elements. Not suprisingly the (non-zero) isotropic elements form a single orbit under this action, and so do the (non-zero) totally isotropic elements.

To establish the various relations between different isotropic and totally isotropic elements we need to define a large number of algebraic operations on these elements. Some of these are new, others bear a direct relation to the operators used in the construction of the 27-dimensional module of the Chevalley algebra of type E_6 [1, 3]. In many cases ours are specific versions of operators that are more generally known, but were simplified for the special context of this paper in order to make it more self-contained and intelligible.

These operations also have a geometric interpretation. For example, two points e and f are collinear in the metasymplectic space if and only $e \cdot f = 0$, $[e, f] = 0$ and $e * f = 0$. Two symplecta E, F have trivial intersection if and only if $[E, F] \neq 0$. We regret that it was beyond the scope of this text to establish these connections in more detail, at least in the case of the metasymplectic space. We do give a more complete treatment for the octagon.

Because K has characteristic 2, not only $\mathbf{W}$ is a subalgebra of $\mathbf{F}$, but it is also an ideal and moreover isomorphic to its quotient $\mathbf{Q} = \mathbf{F}/\mathbf{W}$. Many of the operations defined on $\mathbf{F}$ turn out to be well-defined on $\mathbf{Q}$ and the algebra isomorphism μ is 'compatible' with many of them (e.g., $[\mu(e), \mu(f)] = \mu([e, f])$ and $\langle \mu(e), \mu(f) \rangle = e \cdot f$).

We intend to use μ to 'twist' both the Chevalley group and the metasymplectic geometry. At least one hurdle needs to be overcome : in geometric terms we would like to map points onto symplecta in such a way that (symmetric) incidence is preserved. In algebraic terms this has two consequences. Isotropic elements should be mapped to totally isotropic elements, and the algebraic operation that indicates incidence should be preserved.

The first problem is that totally isotropic elements belong to $\mathbf{F}$ and not to $\mathbf{Q}$. (However, it can be proved that no two totally isotropic elements can differ only in an element of $\mathbf{W}$, hence this is not really an issue). Secondly, and unfortunately, the algebraic operation that we use to check incidence is an operation on $\mathbf{W} \times \mathbf{F}$ and not on $\mathbf{W} \times \mathbf{Q}$, and there seems no easy way to make it so.

It turns out that we need yet another new operation $Q(\cdot)$ to solve this problem (defined by means of μ). This 'duality' operation Q is quadratic and not linear like μ, Q maps isotropic elements of $\mathbf{W}$ to totally isotropic elements of $\mathbf{F}$ (and not of $\mathbf{Q}$), and most importantly, Q transforms the 'incidence operation' into something very much like it, sufficiently so for incidence to be preserved. (As far as we are aware, this paper is the first to give an explicit algebraic definition of a duality operation of this kind.)

Note that applying first Q and then μ^{-1} brings us 'almost' back to the original : we only need an extra application of the Frobenius automorphism. Because of this supplementary automorphism we shall be forced in Section 7 to introduce a Tits automorphism σ (and restrict ourselves to fields for which such σ exists).

Essentially, the points of the octagon $\mathcal{O}$ can now be defined as 'absolute' elements of the duality Q. The last sections of this text simply prove that $\mathcal{O}$ turns out to be what we expect of it. We use Q to define the Ree group ${}^2\widehat{\mathsf{F}_4}(K)$ and prove that it has the desired properties : it is transitive on the points of $\mathcal{O}$ and on all pairs of points at a given distance. The corresponding elations and root groups have properties that make $\mathcal{O}$ a Moufang octagon. We also indicate how the different operators that have been defined in the text can be used in various ways to determine the distance between two given points and to compute the coordinates of the points on the shortest path between them (when they are not opposite).

We strongly regret that we had to restrict ourselves to the case of a field K that is *perfect*. From section 7 onwards we frequently use the fact that the Tits morphism σ has an inverse σ^{-1} and we occasionally have to extract a square root of an element of K. We currently do not see how to avoid this.

Finally, a word on notation and proofs.

In this text we had to introduce several new operations and choose an appropriate notation for them. For some, like the Lie bracket, it was obvious how to do this, for others, there were some difficulties. In particular, it would have been nice to choose the same notation for operations that later turn out to be 'compatible' with respect to μ and Q. Hence $E \cdot F$ could have been preferred to $\langle E, F\rangle$. However, although in general E, F stand for elements of $\mathbf{F}$, they can also belong to $\mathbf{W}$ which is a subspace of $\mathbf{F}$, and then this notation would be ambiguous. For the same reason we cannot call totally isotropic elements of $\mathbf{F}$

isotropic, although both notions are clearly dual with respect to Q. We hope that the reader does not get too confused.

Many of the lemmas, propositions and theorems in this text have proofs that are rather technical. In many cases these proofs can be (and have been) verified by computer using a symbolic computer algebra system. There is only one proposition (Proposition 2.2) which we do not prove 'by hand' because this would have taken numerous extra pages. Although the property itself is very relevant to the rest of the text, we did not think that the proof would contribute greatly to its understanding.

2 The Chevalley algebra of type F_4

In this section we shall review some general properties of the Chevalley algebra of type F_4. We restrict ourselves to the case where the base field K has characteristic 2. More information, also for the case of general characteristic, can be found in [2].

Consider a root system Φ of type F_4. The elements of Φ can be expressed as 4-tuples of real coordinates, in the following way :

1. There are 24 roots whose coordinates are permutations of 4-tuples of the form $(\pm 1, \pm 1, 0, 0)$.
2. There are 16 roots with coordinates of the form $(\pm\frac{1}{2}, \pm\frac{1}{2}, \pm\frac{1}{2}, \pm\frac{1}{2})$.
3. There are 8 roots whose coordinates are permutations of 4-tuples of the form $(\pm 1, 0, 0, 0)$.

We will denote coordinate quadruples by a shorthand notation of the form `1000`, $0\bar{1}\bar{1}0$, `+--+`, where $\bar{1}$ stands for -1, `+` for $\frac{1}{2}$ and `-` for $-\frac{1}{2}$.

Half of the roots have Euclidian length $\sqrt{2}$, and are called *long roots* (they correspond to the first case above). The other half have length 1, and are called *short* roots (the other two cases). We shall denote the sets of long (resp. short) roots by Φ_L (resp. Φ_S).

It is customary to express relations between the roots not in terms of the Euclidian inner product $r \cdot s$, but by means of the following binary product :

$$\langle r, s\rangle \stackrel{\text{def}}{=} 2\frac{r \cdot s}{r \cdot r}. \tag{1}$$

This operator has the advantage that for $r, s \in \Phi$, $\langle r, s\rangle$ is always an integer (in fact $\langle r, s\rangle = 0, \pm 1, \pm 2$). However, $\langle r, s\rangle$ is linear only in the second argument, and not in the first.

The value of $\langle r, s\rangle$ is closely related to whether $r+s, r-s$ are roots and of what kind, as summarised in the following tables :

When $r, s \in \Phi_S$

$\langle r, s\rangle =$	if and only if
2	$r = s$
1	$r - s \in \Phi_S$
0	both $r - s, r + s \in \Phi_L$
-1	$r + s \in \Phi_S$
-2	$r = -s$

When $r, s \in \Phi_L$

$\langle r, s\rangle =$	if and only if
2	$r = s$
1	$r - s \in \Phi_L$
0	$r - s, r + s \notin \Phi$
-1	$r + s \in \Phi_L$
-2	$r = -s$

(2)

When $r \in \Phi_S$, $s \in \Phi_L$

$\langle r, s\rangle =$	if and only if
2	$r - s \in \Phi_S$
0	$r - s, r + s \notin \Phi$
-2	$r + s \in \Phi_S$

When $r \in \Phi_L$, $s \in \Phi_S$

$\langle r, s\rangle =$	if and only if
1	$r - s \in \Phi_S$
0	$r - s, r + s \notin \Phi$
-1	$r + s \in \Phi_S$

With the root system Φ we may associate the *dual* root system Φ^* of roots r^* of the form

$$r^* \stackrel{\text{def}}{=} \frac{2r}{r \cdot r}, \qquad \text{with } r \in \Phi.$$

The element r^* is called the *co-root* corresponding to r. Note that $\langle r^*, s^*\rangle = \langle s, r\rangle$.

The roots

$$r_1 \stackrel{\text{def}}{=} \mathtt{1\bar{1}00}, \quad r_2 \stackrel{\text{def}}{=} \mathtt{01\bar{1}0}, \quad r_3 \stackrel{\text{def}}{=} \mathtt{0010}, \quad r_4 \stackrel{\text{def}}{=} \mathtt{---+}, \tag{3}$$

called *simple* roots, form a so-called *fundamental system*. Every root of Φ can be written as a linear combination of simple roots, with *integral* coefficients. If $r = \alpha_1 r_1 + \alpha_2 r_2 + \alpha_3 r_3 + \alpha_4 r_4$, then

$$\begin{aligned}
\alpha_1 &= 2\langle r_1, r\rangle + 3\langle r_2, r\rangle + 2\langle r_3, r\rangle + 1\langle r_4, r\rangle \\
\alpha_2 &= 3\langle r_1, r\rangle + 6\langle r_2, r\rangle + 4\langle r_3, r\rangle + 2\langle r_4, r\rangle \\
\alpha_3 &= 4\langle r_1, r\rangle + 8\langle r_2, r\rangle + 6\langle r_3, r\rangle + 3\langle r_4, r\rangle \\
\alpha_4 &= 2\langle r_1, r\rangle + 4\langle r_2, r\rangle + 3\langle r_3, r\rangle + 2\langle r_4, r\rangle
\end{aligned} \tag{4}$$

With these notations the co-roots satisfy

$$r^* = \begin{cases} 2\alpha_1 r_1^* + 2\alpha_2 r_2^* + \alpha_3 r_3^* + \alpha_4 r_4^*, & \text{when } r \in \Phi_S, \\ \alpha_1 r_1^* + \alpha_2 r_2^* + \frac{1}{2}\alpha_3 r_3^* + \frac{1}{2}\alpha_4 r_4^*, & \text{when } r \in \Phi_L. \end{cases} \tag{5}$$

From (4) it can be inferred that when $r \in \Phi_L$, both α_3 and α_4 are even integers. Hence the coefficients that occur in (5) are integral. Indeed, the co-roots $r_1^*, \ldots, r_4^*$ form a fundamental system for the dual root system Φ^*.

If we write $r^* = \alpha_1^* r_1^* + \cdots + \alpha_4^* r_4^*$ we obtain

$$\begin{aligned}
\alpha_1^* &= 2\langle r, r_1\rangle + 3\langle r, r_2\rangle + 4\langle r, r_3\rangle + 2\langle r, r_4\rangle \\
\alpha_2^* &= 3\langle r, r_1\rangle + 6\langle r, r_2\rangle + 8\langle r, r_3\rangle + 4\langle r, r_4\rangle \\
\alpha_3^* &= 2\langle r, r_1\rangle + 4\langle r, r_2\rangle + 6\langle r, r_3\rangle + 3\langle r, r_4\rangle \\
\alpha_4^* &= 1\langle r, r_1\rangle + 2\langle r, r_2\rangle + 3\langle r, r_3\rangle + 2\langle r, r_4\rangle
\end{aligned} \tag{6}$$

To every root r we associate the reflection w_r about the hyperplane orthogonal to r :

$$w_r(x) \stackrel{\text{def}}{=} x - \langle r, x\rangle r$$

It is a (defining) property of a root system that these reflections always map roots onto roots. The finite group W generated by all reflections w_r with $r \in \Phi$ is the *Weyl group* of type F_4. This group acts regularly on all fundamental systems of roots in Φ and it acts transitively on roots of the same type (short or long) and on pairs of roots with the same type and inner product. In other words, the 16 cases listed in (2) correspond exactly to the orbits of W on ordered pairs (r, s) of roots.

The Chevalley algebra $\mathbf{F}$ of type F_4 over a field K is a 52-dimensional vector space over K with a bilinear 'Lie bracket' operator $[\cdot,\cdot]$. $\mathbf{F}$ can be written as a direct sum of the following form :

$$\mathbf{F} = \mathbf{H} \oplus \bigoplus_{r \in \Phi} K e_r.$$

The elements e_r, one for every $r \in \Phi$, are the *root vectors* of $\mathbf{F}$, and $\mathbf{H}$ is the Cartan subalgebra of dimension 4, generated by the elements $h_r \stackrel{\text{def}}{=} [e_r, e_{-r}]$ for $r \in \Phi$.

Not all elements h_r, $r \in \Phi$ are linearly independent. In fact, the elements h_r can be expressed as linear combinations of the 4 elements $h_{r_1}, \ldots, h_{r_4}$, in a manner very similar to (5) :

$$h_r = \begin{cases} 2\alpha_1 h_{r_1} + 2\alpha_2 h_{r_2} + \alpha_3 h_{r_3} + \alpha_4 h_{r_4}, & \text{when } r \in \Phi_S, \\ \alpha_1 h_{r_1} + \alpha_2 h_{r_2} + \frac{1}{2}\alpha_3 h_{r_3} + \frac{1}{2}\alpha_4 h_{r_4}, & \text{when } r \in \Phi_L. \end{cases} \tag{7}$$

The coefficients should now be interpreted as elements of the base field K. (In our particular case of characteristic 2 we could therefore have left out the terms involving $2\alpha_1$ and $2\alpha_2$.) Note that, because the characteristic of K is small, it may happen that $h_r = h_{r'}$ although $r \neq r'$.

The 4 elements h_{r_i} together with the 48 root vectors e_r, form a *Chevalley basis* for $\mathbf{F}$. As $\operatorname{char} K = 2$, the Chevalley basis elements can be chosen to satisfy

$$\begin{aligned}
&[h_r, h_s] &&= 0, \\
&[e_s, h_r] = [h_r, e_s] &&= \langle r, s\rangle e_s, \\
&[e_r, e_{-r}] &&= h_r, \\
&[e_r, e_s] &&= e_{r+s}, && \text{when } r+s \in \Phi \text{ and } \langle r,s\rangle \neq 0, \\
&[e_r, e_s] &&= 0, && \text{otherwise.}
\end{aligned} \tag{8}$$

Note that in particular $[e_r, e_s] = 0$ whenever $\langle r, s\rangle \geq 0$, by (2).

The second equality in (8) needs some extra attention. First note that $\langle r, s\rangle$ should be interpreted as an element of K, which is possible because $\langle r, s\rangle$ is always an integer. Secondly, note that this equality holds for all $r \in \Phi$ and not only for simple roots r_i, even though the h_r are not linearly independent and the equality is not linear in r.

We shall often need to express an element $A \in \mathbf{F}$ as a linear combination of Chevalley base elements. The corresponding coordinates shall be denoted by $A[r]$ for $r \in \Phi$ and $A[i]$ for $i = 1, 2, 3, 4$. More precisely, we have

$$A = \sum_{r\in\Phi} A[r]e_r + \sum_{i=1}^{4} A[i]h_{r_i}. \tag{9}$$

The values for $h_r[i]$ can be derived from (6) reduced modulo 2, since (7) shows that the elements h_r behave like the co-roots r^* and hence that $h_r[i] = \alpha_i^* \bmod 2$.

$$\begin{aligned}
h_r[1] &= \langle r, r_2\rangle, & h_r[2] &= \langle r, r_1\rangle, \\
h_r[3] &= \langle r, r_4\rangle, & h_r[4] &= \langle r, r_3\rangle + \langle r, r_1\rangle
\end{aligned} \tag{10}$$

In terms of coordinates (8) can be expressed as follows : let $A, B, C \in \mathbf{F}$ with $C = [A, B]$, then

$$\begin{aligned}
C[t] &= \sum_{\substack{r,s\in\Phi \\ r+s=t, \langle r,s\rangle\neq 0}} A[r]B[s] + \sum_{i=1}^{4} \langle r_i, t\rangle (A[t]B[i] + A[i]B[t]), && \text{for } t \in \Phi, \\
C[i] &= \sum_{r\in\Phi} A[r]B[-r]h_r[i], && \text{for } i = 1, \dots, 4.
\end{aligned} \tag{11}$$

Because $\operatorname{char} K = 2$, the Lie bracket $[\cdot, \cdot]$ is a *symmetric* bilinear operation on $\mathbf{F}$. We associate a quadratic operator $\cdot^2$ to this bilinear operator by choosing values on the base vectors of $\mathbf{F}$ as follows :

$$e_r^2 = 0, \qquad h_{r_i}^2 = h_{r_i} \qquad \text{for } r \in \Phi,\ i = 1, \dots, 4 \tag{12}$$

and setting

$$(A + kB)^2 = A^2 + k[A, B] + k^2 B^2, \qquad \text{for every } A, B \in \mathbf{F},\ k \in K. \tag{13}$$

(Properties (12–13) uniquely determine the action of $\cdot^2$ on $\mathbf{F}$.)

In terms of coordinates we have :

$$\begin{aligned} A^2[t] &= \sum_{\substack{\{r,s\}\subset\Phi \\ r+s=t, \langle r,s\rangle\neq 0}} A[r]A[s] + \sum_{i=1}^{4} \langle r_i, t\rangle A[i]A[t], && \text{for } t \in \Phi, \\ A^2[i] &= A[i]^2 + \sum_{\{r,-r\}\subset\Phi} A[r]A[-r]h_r[i], && \text{for } i = 1, \ldots, 4. \end{aligned} \tag{14}$$

(The first sum in the first equation treats each pair exactly once. The sum in the last equation consists of 24 terms, one for each pair $\{r, -r\}$ of roots. Equations (14) could serve as an alternative definition of $\cdot^2$.)

Note that $h_r^2 = h_r$ for all $r \in \Phi$ (and not only when $r = r_i$ is a simple root).

Proposition 2.1. *Let $X, A \in \mathbf{F}$, then*

$$[[X, A], A] = [X, A^2]. \tag{15}$$

Proof. The Jacobi identity proves that $[[X, A], B] + [[X, B], A] = [X, [A, B]]$, hence by (13) it is sufficient to prove (15) when A is a Chevalley base element.

Let $r \in \Phi$. We express $[X, e_r]$ in terms of coordinates of X :

$$[X, e_r] = \sum_{\substack{s\in\Phi \\ r+s\in\Phi, \langle r,s\rangle\neq 0}} X[s]e_{r+s} + X[-r]h_r + \sum_i X[i]\langle r_i, r\rangle e_r,$$

We have $\langle r, r+s\rangle = \langle r, r\rangle + \langle r, s\rangle = 2 + \langle r, s\rangle \geq 0$, and hence $[e_{r+s}, e_r] = 0$. Also $[h_r, e_r] = 0$ and $[e_r, e_r] = 0$, and therefore $[[X, e_r], e_r] = 0 = [X, e_r^2]$.

Similarly, $[X, h_r] = \sum_{s\in\Phi} X[s]\langle r, s\rangle e_s$, so $[[X, h_r], h_r] = \sum_{s\in\Phi} X[s]\langle r, s\rangle^2 e_s$, which is the same as $[X, h_r]$ because $\langle r, s\rangle^2 = \langle r, s\rangle$ in characteristic 2. ☐

Let $\mathbf{W}$ denote the subspace of $\mathbf{F}$ generated by the elements h_r, e_r, restricted to the short roots $r \in \Phi_S$:

$$\mathbf{W} \stackrel{\text{def}}{=} \mathbf{I} \oplus \bigoplus_{r\in\Phi_S} Ke_r, \tag{16}$$

where $\mathbf{I}$ is the subspace of $\mathbf{H}$ generated by all h_r with $r \in \Phi_S$. In other words : $A \in \mathbf{W}$ if and only if its coordinates satisfy $A[r] = 0$, for all $r \in \Phi_L$, and $A[1] = A[2] = 0$. Hence $\mathbf{W}$ has dimension 26.

It follows from (8) that $[\mathbf{W}, \mathbf{F}] \leq \mathbf{W}$ and hence that $\mathbf{W}$ is an ideal and a subalgebra of $\mathbf{F}$. Also note that $\mathbf{W}^2 \leq \mathbf{W}$ by (12).

Below we repeat (8) for the special case that both $r, s \in \Phi_S$. These could be considered the defining relations for $\mathbf{W}$:

$$
\begin{aligned}
[h_r, h_s] &= 0, \\
[h_r, e_s] &= \langle r, s\rangle e_s, \\
[e_r, e_{-r}] &= h_r, \\
[e_r, e_s] &= 0, && \text{when } r+s \neq 0,\ r+s \notin \Phi_S, \\
[e_r, e_s] &= e_{r+s}, && \text{when } r+s \in \Phi_S.
\end{aligned}
\tag{17}
$$

As before, $[e_r, e_s] = 0$ whenever $\langle r, s\rangle \geq 0$.

It turns out (although we do not need this here) that $\mathbf{W}$ can be extended to a 27-dimensional module of the Chevalley algebra of type E_6. There is a well-known construction of this module [3] that uses a bilinear operator $\times$ and a related quadratic operator $\#$. These operators are closely related (but not identical) to the Lie bracket and the square operator we use here.

When restricted to $r, s \in \Phi_S$ the product $\langle r, s\rangle$ is bilinear, symmetric and integral. This allows us to define the following symmetric bilinear dot product on $\mathbf{W}$, with values in K :

$$
\begin{aligned}
e_r \cdot e_s &\stackrel{\text{def}}{=} \begin{cases} 1 & \text{when } r = -s, \\ 0 & \text{otherwise.} \end{cases} \\
e_r \cdot h_s = h_s \cdot e_r &\stackrel{\text{def}}{=} 0, \\
h_r \cdot h_s &\stackrel{\text{def}}{=} \langle r, s\rangle, && \text{for } r, s \in \Phi_S.
\end{aligned}
\tag{18}
$$

(This dot product is related to the product $(\cdot,\cdot)$ used in the construction of the 27-dimensional E_6-module mentioned above.)

Consider $a \in \mathbf{W}$. In terms of coordinates with respect to the Chevalley base elements we find

$$
a \cdot e_r = a[-r], \quad a \cdot h_{r_3} = a[4], \quad a \cdot h_{r_4} = a[3],
$$

and hence, for general $a, b \in \mathbf{W}$:

$$
a \cdot b = \sum_{r \in \Phi_S} a[r]b[-r] + a[3]b[4] + a[4]b[3]. \tag{19}
$$

Note that $a \cdot a = 0$ for all $a \in \mathbf{W}$. The dot product is also nondegenerate on $\mathbf{W}$, in other words, if $a \cdot x = 0$ for all $x \in \mathbf{W}$, then necessarily $a = 0$.

Let $A \in \mathbf{F}$, $r, s \in \Phi_S$. Then we easily compute the following :

$$\begin{aligned} [e_r, A] \cdot e_s &= \begin{cases} \sum_{i=1}^4 \langle r_i, r\rangle A[i], & \text{when } r = -s, \\ A[-r-s], & \text{when } \langle r, s\rangle = 0 \text{ or } -1, \\ 0, & \text{otherwise.} \end{cases} \\ [h_s, A] \cdot e_r &= \langle s, r\rangle A[-r], \\ [e_r, A] \cdot h_s &= \langle s, r\rangle A[-r], \\ [h_r, A] \cdot h_s &= 0. \end{aligned} \tag{20}$$

It follows that $[\mathbf{W}, A] = 0$ for $A \in \mathbf{F}$ only when $A = 0$.

As an immediate consequence of (20) we find

$$[a, C] \cdot a = 0, \quad [a, C] \cdot b = [b, C] \cdot a, \quad \text{for } a, b \in \mathbf{W} \text{ and } C \in \mathbf{F}. \tag{21}$$

When $C = c$ belongs to $\mathbf{W}$ we may interchange the roles of a and c in the above, to obtain

$$[a, b] \cdot c = [b, c] \cdot a = [c, a] \cdot b, \qquad \text{for all } a, b, c \in \mathbf{W}. \tag{22}$$

Taking the dot product of d with the Jacobi identity $[[a, b], c]+[[b, c], a]+[[c, a], b] = 0$, and applying (22) yields

$$[a, b] \cdot [c, d] + [a, c] \cdot [b, d] + [a, d] \cdot [b, c] = 0, \qquad \text{for all } a, b, c, d \in \mathbf{W}. \tag{23}$$

To the symmetric bilinear dot product we associate a quadratic *norm* function $N(\cdot)$ as follows :

$$N(e_r) \stackrel{\text{def}}{=} 0, \qquad N(h_{r_i}) \stackrel{\text{def}}{=} 1, \qquad \text{for } r \in \Phi_S,\ i = 3, 4. \tag{24}$$

with

$$N(a + kb) = N(a) + ka \cdot b + k^2 N(b), \qquad \text{for } a, b \in \mathbf{W},\ k \in K. \tag{25}$$

In terms of coordinates :

$$N(a) = \sum_{\{r,-r\}\subset\Phi_S} a[r]a[-r] + a[3]^2 + a[3]a[4] + a[4]^2. \tag{26}$$

where the sum is taken such that every pair $\{r, -r\}$ occurs exactly once. Note that $N(h_r) = 1$ for all $r \in \Phi_S$ (and not only when $r = r_i$ is a simple root).

The following is a basic structural 'axiom' for the algebra $\mathbf{W}$:

Proposition 2.2. *Let $a \in \mathbf{W}$, then*

$$(a^2)^2 = N(a)a^2 + (a^2 \cdot a)a \tag{27}$$

Sketch of the proof. The fastest way to prove this proposition is by means of a computer. When expressed in terms of coordinates (27) can be interpreted as a set of 26 polynomial identities of the fourth degree in 26 different coordinate variables $a[r]$, $a[i]$, with $r \in \Phi_S$, $i = 3, 4$. These identities can then be verified by computer.

Without the use of a computer, a technical proof consists of proving (27) and its corollaries (28–31) in the special case that a, b, c, d are Chevalley base elements. There are however many very different cases to consider.

Finally, note that (27) is a consequence of a similar identity $(a^\#)^\# = D(a)a$ which is known [1, 3] to hold in the 27-dimensional E_6-module over K of which $\mathbf{W}$ is a subspace of co-dimension 1. □

From (27) we may derive many other identities. For example, substituting $a + kb$ for a and grouping the terms according to the degree in k, one obtains, for every $a, b \in \mathbf{W}$:

$$[[a,b],a^2] = (a\cdot b)a^2 + N(a)[a,b] + (a^2\cdot b)a + (a^2\cdot a)b \tag{28}$$

$$[a,b]^2 + [a^2,b^2] = N(a)b^2 + (a\cdot b)[a,b] + N(b)a^2 + (a^2\cdot b)b + (a\cdot b^2)a \tag{29}$$

Substituting $b + c$ for b in the above, yields, for $a, b, c \in \mathbf{W}$,

$$\begin{aligned}[[a,b],[a,c]] + [a^2,[b,c]] = {} & N(a)[b,c] + (a\cdot b)[a,c] + (a\cdot c)[a,b] \\ & + (b\cdot c)a^2 + (a^2\cdot b)c + (a^2\cdot c)b + ([a,b]\cdot c)a\end{aligned} \tag{30}$$

and finally, substituting $a + d$ for a, gives, for $a, b, c, d \in \mathbf{W}$,

$$\begin{aligned}&[[a,b],[c,d]] \;+\; [[a,c],[b,d]] \;+\; [[a,d],[b,c]] \\ &= (a\cdot b)[c,d] \;+\; (a\cdot c)[b,d] \;+\; (a\cdot d)[b,c] \;+\; (b\cdot c)[a,d] \;+\; (b\cdot d)[a,c] \\ &\;+\; (c\cdot d)[a,b] + \;([a,b]\cdot c)d \;+\; ([b,c]\cdot d)a \;+\; ([c,d]\cdot a)b \;+\; ([d,a]\cdot b)c\end{aligned} \tag{31}$$

Lemma 2.3. *Let $H : \Phi_S \to K$ satisfy*

$$H(r+s) = H(r) + H(s) \tag{32}$$

for all $r, s \in \Phi_S$ such that $r+s \in \Phi_S$. Then there exist four constants $H_1, \ldots, H_4 \in K$ with the property

$$H(r) = \sum_{i=1}^{4} \langle r_i, r\rangle H_i, \qquad \textit{for all } r \in \Phi_S. \tag{33}$$

Proof. Consider the 5 roots $s_0 = \texttt{++++}$, $s_1 = \bar{1}000$, $s_2 = 0\bar{1}00$, $s_3 = 00\bar{1}0$ and $s_4 = 000\bar{1}$. Note that every short root of the form $(\pm\frac{1}{2}, \pm\frac{1}{2}, \pm\frac{1}{2}, \pm\frac{1}{2})$ can be written as $s_0 + s_i$, $s_0 + s_i + s_j$, $s_0 + s_i + s_j + s_k$ or $s_0 + s_1 + s_2 + s_3 + s_4$, with $i, j, k \in \{1, \ldots, 4\}$. Also `1000` is the sum of `+---` and `++++` and likewise for `0100`, `0010` and `0001`.

By (32) we may conclude that for every $r \in \Phi_S$, $H(r)$ can be written as a linear combination of $H(s_0)$,..., $H(s_4)$. However, these five values are not independent. Indeed, we have

$$\begin{array}{rcccl} r & = & \texttt{---+} & = & ((s_0 + s_1) + s_2) + s_3, \\ s & = & \texttt{++--} & = & ((s_0 + s_3) + s_4), \\ r + s & = & \texttt{00}\bar{\texttt{1}}\texttt{0} & = & s_3, \end{array}$$

and hence $H(r) = H(s_0) + H(s_1) + H(s_2) + H(s_3)$, $H(s) = H(s_0) + H(s_3) + H(s_4)$, and $H(r+s) = H(s_3)$. Therefore (32) implies $H(s_1) + H(s_2) + H(s_3) + H(s_4) = 0$.

We have

$$\begin{array}{rcrrrr} s_0 & = & r_1 & +2r_2 & +3r_3 & +r_4 \\ s_1 & = & -r_1 & -r_2 & -r_3 & \\ s_2 & = & & -r_2 & -r_3 & \\ s_3 & = & & & -r_3 & \\ s_4 & = & -r_1 & -2r_2 & -3r_3 & -2r_4. \end{array} \tag{34}$$

Setting

$$H_1 \stackrel{\text{def}}{=} H(s_2)+H(s_3),\ H_2 \stackrel{\text{def}}{=} H(s_1)+H(s_2),\ H_3 \stackrel{\text{def}}{=} H(s_0)+H(s_4),\ H_4 \stackrel{\text{def}}{=} H(s_3),$$

we also easily verify that

$$\begin{array}{rcrrrr} H(s_0) & = & H_2 & & +H_4 & +H_3 \\ H(s_1) & = & H_2 & +H_1 & +H_4 & \\ H(s_2) & = & & H_1 & +H_4 & \\ H(s_3) & = & & & H_4 & \\ H(s_4) & = & H_2 & & +H_4. & \end{array} \tag{35}$$

Comparing (34) and (35) and using the fact that $\langle r_i, r_j \rangle$ is even unless $\{i, j\} = \{1, 2\}$ or $\{3, 4\}$, we see that (33) is satisfied for $r = s_0, \ldots, s_4$. The lemma then follows because (33) is linear in r. $\square$

Proposition 2.4. *Let $\tau \in \mathrm{Hom}(\mathbf{W}, \mathbf{W})$ be a linear transformation that satisfies*

$$\tau(a^2) = [\tau(a), a], \qquad \text{for all } a \in \mathbf{W}, \tag{36}$$

or equivalently

$$\tau(a^2) \cdot b = \tau(a) \cdot [a, b], \qquad \text{for all } a, b \in \mathbf{W}, \tag{37}$$

then there exists a unique element $T \in \mathbf{F}$ such that $\tau(a) = [a, T]$ for every $a \in \mathbf{W}$.

Proof. Let $r, s, t, u \in \Phi_S$. We shall compute $\tau(e_r) \cdot e_s$. Setting $a = e_r, h_r$ and $b = e_s, h_s$ in (37) we easily obtain that τ must satisfy

$$\begin{aligned} \tau(e_r) \cdot h_r &= 0, & & \\ \tau(e_r) \cdot e_r &= 0, & & \\ \tau(e_r) \cdot e_s &= 0, & & \text{when } \langle r, s\rangle = 1, \\ \tau(h_r) \cdot e_s &= 0, & & \text{when } \langle r, s\rangle \neq \pm 1, \\ \tau(h_r) \cdot h_s &= 0. & & \end{aligned} \tag{38}$$

(The third identity is obtained by setting $b = e_{r-s}$.)

Substituting c for b and $a + b$ for a in (37) we find that τ must satisfy

$$\tau([a, b]) \cdot c = \tau(a) \cdot [b, c] + \tau(b) \cdot [a, c], \qquad \text{for all } a, b, c \in \mathbf{W}. \tag{39}$$

Setting $a = e_r$, $b = h_u$ and $c = e_s$ in this equation, we find

$$\tau(h_u) \cdot [e_r, e_s] = \langle u, r + s\rangle \tau(e_r) \cdot e_s,$$

Hence, when $\langle r, t\rangle = 1$ we may take $t = r+s$ and find $\tau(h_u) \cdot e_t = \langle u, t\rangle \tau(e_r) \cdot e_s$. Because we can always find u such that $\langle u, t\rangle = 1$ we see that $\tau(e_r) \cdot e_s$ only depends on the value of $t = r + s$. We shall denote this value by $T_{(-t)}$. Combining this with the fourth identity of (38) we obtain :

$$\begin{aligned} \tau(e_r) \cdot e_s &= T_{(-r-s)}, & & \text{when } \langle r, s\rangle = -1, \\ \tau(h_s) \cdot e_r &= \langle s, r\rangle T_{(-r)} & & \end{aligned} \tag{40}$$

(For the second equation we have renamed the variables u, t to s, r, respectively.)

Now consider $t \in \Phi_L$ and $r, s, r', s' \in \Phi_S$ with $t = r + s = r' + s'$ but $\{r, s\} \neq \{r', s'\}$. We have $\langle r, t\rangle = \langle r, r + s\rangle = \langle r, r\rangle + \langle r, s\rangle = 2$ and hence also $\langle r, t\rangle = \langle r, r'\rangle + \langle r, s'\rangle = 2$. This implies $\langle r, r'\rangle = \langle r, s'\rangle = 1$ and therefore $r' - r \in \Phi_S$. Setting $a = e_r$, $b = e_{r'-r}$ and $c = e_{s'}$ in (39), we find

$$\tau(e_{r'}) \cdot e_{s'} = \tau(e_r) \cdot [e_{r'-r}, e_{s'}] + \tau(e_{r'-r}) \cdot [e_r, e_{s'}] = \tau(e_r) \cdot e_s.$$

It follows that the value of $\tau(e_r) \cdot e_s$ depends only on $t = r + s$. As before, we shall denote this value by $T_{(-t)}$ (but now t is a long root), obtaining

$$\tau(e_r) \cdot e_s = T_{(-r-s)}, \qquad \text{when } \langle r, s\rangle = 0. \tag{41}$$

Setting $a = e_r$, $b = e_s$, $c = e_{-s}$ in (39) we find

$$\tau(e_r) \cdot h_s = \tau([e_r, e_s]) \cdot e_{-s} + \tau([e_r, e_{-s}]) \cdot e_s,$$

and by distinguishing different cases according to the value of $\langle r, s\rangle$, we obtain

$$\tau(e_r)\cdot h_s = \langle s, r\rangle T_{(-r)} \tag{42}$$

Finally, consider $r, s \in \Phi_S$ such that $\langle r, s\rangle = -1$ and set $a = e_r$, $b = e_s$ and $c = e_{-r-s}$ to obtain

$$\tau(e_{r+s})\cdot e_{-r-s} = \tau(e_r)\cdot e_{-r} + \tau(e_s)\cdot e_{-s}.$$

We may now apply Lemma 2.3 with $H(r) \stackrel{\text{def}}{=} \tau(e_r)\cdot e_{-r}$ to prove that there exist constants $H_1, \dots, H_4$ such that

$$\tau(e_r)\cdot e_{-r} = \sum_{i=1}^{4}\langle r_i, r\rangle H_i. \tag{43}$$

Comparing equations (38–43) with (20) we see that τ behaves exactly like the elements $T \in \mathbf{F}$ with coordinates $T[r] = T_{(r)}$ and $T[i] = H_i$. □

Corollary 2.5. *Let $b, c \in \mathbf{W}$ be such that $b \cdot c = 0$ and $[b, c] = 0$. Then there exists a unique element $b * c \in \mathbf{F}$ such that*

$$[a, b * c] = [[a, b], c] + (a\cdot b)c + (a\cdot c)b, \quad \textit{for all } a \in \mathbf{W}. \tag{44}$$

Proof. We apply Proposition 2.4 to the map which maps a onto the right hand side of (44). We verify (36) :

$$\begin{aligned}
&[[a^2, b], c] + (a^2\cdot b)c + (a^2\cdot c)b + [[a, b], c], a] + (a\cdot b)[a, c] + (a\cdot c)[a, b]\\
&\quad = [[a, b], a], c] + [[a, b], c], a] + (a^2\cdot b)c + (a^2\cdot c)b + (a\cdot b)[a, c] + (a\cdot c)[a, b]\\
&\quad = [[a, b], [a, c]] + (a^2\cdot b)c + (a^2\cdot c)b + (a\cdot b)[a, c] + (a\cdot c)[a, b],
\end{aligned}$$

applying the Jacobi identity to the first two terms. By (30) this reduces to

$$[a^2, [b, c]] + N(a)[b, c] + (b\cdot c)a^2 + ([a, b]\cdot c)a$$

which is zero, as $[b, c] = 0$ and $b \cdot c = 0$. (Note that, by (22), the last term is also equal to $(a\cdot [b, c])a$.) □

It should be noted that the $*$-product is only partially defined on $\mathbf{W}\times\mathbf{W}$. (Its definition could be extended to all possible $b, c \in \mathbf{W}$ if we would allow its value to belong to a Chevalley algebra larger than $\mathbf{F}$, of type E_6. However, we do not need this for this text.)

We realize that the reader might perceive the introduction of $*$ as the definition of 'yet another operator'. However, it will play an important role in the definition of totally isotropic elements (in the next section) and in this way help to provide a geometric interpretation of the algebraic theory which is being developed.

The following lemma provides some simple examples of products of this kind.

Lemma 2.6. *Let $r, s \in \Phi_S$ be such that $\langle r, s\rangle = 0$. Then $r + s \in \Phi_L$ and*

$$e_r * e_s = e_{r+s}, \qquad (e_r + e_{-s}) * (e_s + e_{-r}) = e_{r+s} + e_{-r-s} + h_{r+s}. \tag{45}$$

When $r, s \in \Phi_S$ such that $\langle r, s\rangle = 1$ then $e_r * e_s = 0$.

Proof. Write $t = r + s$.

1. Let $u \in \Phi_S$. Then $[e_u, e_r * e_s] = [[e_u, e_r], e_s] + (e_u \cdot e_r)e_s + (e_u \cdot e_s)e_r$ is non-zero only in the following cases :

   – When $u = -r$ and then it is equal to $[h_r, e_s] + e_s = \langle r, s\rangle e_s + e_s = e_s$.
   – When $u = -s$ and then it is equal to e_r.
   – When $u \neq s$, $\langle u, r\rangle = -1$ and then it is equal to $[e_{u+r}, e_s]$, which is e_{u+t} provided that $\langle u + r, s\rangle = -1$

   By (2) $t + u \in \Phi_S$ if and only if $-2 = \langle u, t\rangle = \langle u, r\rangle + \langle u, s\rangle$ if and only if $u = -r$, $u = -s$ or $\langle u, r\rangle = \langle u, s\rangle = -1$. In other words, $[e_u, e_r * e_s]$ is non-zero exactly for those $u \in \Phi_S$ for which $t + u \in \Phi_S$, and then the value is e_{u+t}. This is exactly the same behavior as that of e_t.

   Likewise $[h_u, e_r * e_s] = \langle u, r\rangle[e_r, e_s] = 0$ and also $[h_u, e_t] = 0$, for all $u \in \Phi_S$.

2. Consider the expression

$$[[e_u, e_r], e_{-r}] + (e_u \cdot e_r)e_{-r} + (e_u \cdot e_{-r})e_r. \tag{46}$$

   The value of this expression is non-zero in the following cases :

   – When $u = -r$ and then it is equal to $[h_r, e_{-r}] + e_{-r} = e_{-r}$.
   – When $u = r$ and then it is equal to e_r.
   – When $\langle u, r\rangle = -1$ and then it is equal to $[e_{u+r}, e_{-r}] = e_u$.

   We may obtain a similar result for the expression

$$[[e_u, e_{-s}], e_s] + (e_u \cdot e_s)e_{-s} + (e_u \cdot e_{-s})e_s, \tag{47}$$

   and combining these results we obtain the following value for the sum of (46) and (47) : it is equal to e_u in those cases where $u = r, -r, s, -s$ or when either $\langle u, r\rangle = -1$ or $\langle u, -s\rangle = -1$ (but not both), and otherwise the value is 0. As in the first part of this proof, the identity $\langle u, t\rangle = \langle u, r\rangle + \langle u, s\rangle$ may be used to show that these conditions are equivalent to $\langle u, t\rangle = \pm 2$ or $\langle t, u\rangle = \pm 1$. Hence the value for the sum of (46) and (47) is $\langle t, u\rangle e_u = [h_t, e_u]$. In combination with the first part of the lemma, this proves the right hand identity in (45). (We leave it to the reader to verify the identity when applied to h_u instead of e_u.)

3. When $\langle r, s\rangle = 1$ we find that $[e_u, e_r * e_s] = [[e_u, e_r], e_s] + (e_u \cdot e_r)e_s + (e_u \cdot e_s)e_r$ is always zero except possibly in the following cases :

   – When $u = -r$, but then it is equal to $[h_r, e_s] + e_s = \langle r, s\rangle e_s + e_s = 0$.
   – When $u = -s$, but then it is equal to $[e_{r-s}, e_s] + e_r = e_r + e_r = 0$.
   – When $\langle u, r\rangle = -1$ and then it is equal to $[e_{u+r}, e_s]$, which again is 0, for now $\langle s, u + r\rangle = \langle s, u\rangle + \langle s, r\rangle = \langle s, u\rangle + 1 \geq 0$. □

Lemma 2.7. *Let $b, c \in \mathbf{W}$ be such that $[b, c] = 0$, $b \cdot c = 0$. Let $a \in \mathbf{W}$. Then*

1. $b * b = b^2$,
2. $b * c = c * b$,
3. $[b, b * c] = 0$ *and hence* $[a, b * c] \cdot b = 0$,
4. $[[a, b * c], b] = [[a, b], b * c] = [[a, b^2], c] = [[a, c^2], b]$,
5. $[a, (b * c)^2] = [[a, b^2], c^2] = [[a, c^2], b^2]$.

Proof. The first four statements are immediate consequences of the definition and the fact that $b \cdot b = 0$ for all $b \in \mathbf{W}$. For the last statement, we find

$$\begin{aligned}[a, (b * c)^2] &= [[a, b * c], b * c] \\ &= [[[a, b * c], b], c] + ([a, b * c] \cdot b)c + ([a, b * c] \cdot c)b = [[a, b^2], c^2],\end{aligned}$$

using the first statements of the lemma. □

Lemma 2.8. *Let $b, c \in \mathbf{W}$, $A \in \mathbf{F}$ be such that $[b, c] = [b, [c, A]] = 0$, $b \cdot c = b \cdot [c, A] = 0$ and hence $[c, [b, A]] = 0$ and $c \cdot [b, A] = 0$. Then*

$$[A, b * c] = [b, A] * c + [c, A] * b. \tag{48}$$

Proof. Consider $x \in \mathbf{W}$. Then

$$\begin{aligned}[x, [b, A] * c] + [x, [c, A] * b] = [[x, c], [b, A]] + (x \cdot c)[b, A] + (x \cdot [b, A])c + \\ [[x, b], [c, A]] + (x \cdot b)[c, A] + (x \cdot [c, A])b.\end{aligned}$$

Also

$$\begin{aligned}[x,[b*c,A]] &= [[x,b*c],A]+[[x,A],b*c] \\ &= [[[x,b],c],A]+(x\cdot b)[c,A]+(x\cdot c)[b,A]+ \\ &\qquad [[[x,A],b],c]+([x,A]\cdot b)c+([x,A]\cdot c)b.\end{aligned}$$

Finally, several applications of the Jacobi identity yield

$$\begin{aligned}&[[x,b],[c,A]]+[[[x,b],c],A]+[[[x,A],b],c]+[[x,c],[b,A]] \\ &\quad = [[[x,b],A],c]+[[[x,A],b],c]+[[x,c],[b,A]] \\ &\quad = [[[b,A],x],c]+[[x,c],[b,A]] = [[[b,A],c],x] = 0.\end{aligned}$$

Adding these results, we find that $[x,[A,b*c]] = [x,[b,A]*c]+[x,[c,A]*b]$, for all $x \in \mathbf{W}$. □

3 Isotropic and totally isotropic elements

An element $a \in \mathbf{W}$ shall be called *isotropic* if and only if $a^2 = 0$. The elements e_r with $r \in \Phi_S$ serve as typical examples of isotropic elements. Another example is provided by the element $e_r + e_{-r} + h_r$. Also $e_r + e_s$ is isotropic when $r, s \in \Phi_S$ and $\langle r, s\rangle = 0$ or 1.

A subspace S of $\mathbf{W}$ is called *totally isotropic* when all its elements are isotropic. By (13) this means that S is totally isotropic if and only if every base $a_1, \ldots, a_d$ of S satisfies $a_i^2 = [a_i, a_j] = 0$, for $1 \le i, j \le d$.

Clearly Ke is totally isotropic if and only if $e^2 = 0$, and $Ke + Kf$ is totally isotropic if and only if $e^2 = f^2 = [e, f] = 0$.

The one-dimensional totally isotropic subspaces Ke will later serve as points of a geometry (which is metasymplectic [5], although we shall not expand on this) and the operators defined on $\mathbf{W}$, when applied to these points, will enable us to distinguish between various relations of pairs of such points (for example, Ke and Kf shall turn out to be collinear, or equal, if and only if $e \cdot f = 0$, $[e, f] = 0$ and $e * f = 0$).

The following lemma lists some simple properties of isotropic elements of $\mathbf{W}$.

Lemma 3.1. *Let e be an isotropic element of $\mathbf{W}$. Then*

1. *$N(e) = 0$,*
2. *$[a, e]^2 = (a \cdot e)[a, e] + (a^2 \cdot e)e$ for all $a \in \mathbf{W}$,*

3. *$[[[a,e],b],e] = [[a,e],[b,e]] = (a\cdot e)[b,e] + (b\cdot e)[a,e] + ([a,b]\cdot e)e$, for all $a,b\in\mathbf{W}$.*

4. *$e * [a,e] = (a\cdot e)e$, for all $a\in\mathbf{W}$.*

Proof. 1. Applying (28) to $a = e$, we find that $N(e)[e,b] = 0$ for all $b\in\mathbf{W}$. Hence $N(e)e = 0$, and therefore $N(e) = 0$.

2,3. These follow immediately from (29, 30) and the above.

4. Note that $e\cdot[a,e] = 0$ and $[e,[a,e]] = [a,e^2] = 0$, hence $e * [a,e]$ is well-defined. Now, for every $b\in\mathbf{W}$ we have

$$[b, e * [a,e]] = [[b,e],[a,e]] + (b\cdot[a,e])e + (b\cdot e)[a,e] = (a\cdot e)[b,e],$$

by the previous statement. □

An element $E\in\mathbf{F}$ of the form $E = e * f$ for some $e,f\in\mathbf{W}$ such that $e^2 = [e,f] = f^2 = 0$ and $e\cdot f = 0$, shall be called *totally isotropic*. We shall call $\{e,f\}$ a *defining pair* for E. (A defining pair need not be unique.)

The base elements e_r with $r\in\Phi_L$ are typical examples of totally isotropic elements. Indeed, we may always find $s,t\in\Phi_S$ such that $r = s+t$ and $\langle s,t\rangle = 0$, and then $e_r = e_s * e_t$. By (45) also the element $e_r + e_{-r} + h_r$ is totally isotropic.

(In terms of metasymplectic spaces, when $E = e * f$ is totally isotropic and non-zero, the one-dimensional subspace KE corresponds to a symplecton, the unique symplecton incident with both Ke and Kf.)

Lemma 3.2. *Let E be a totally isotropic element of $\mathbf{F}$ with defining pair $\{e,f\}$. Let $a,b\in\mathbf{W}$. Then*

1. *$[e,E] = [E,f] = 0$,*
2. *$[a,E]\cdot f = e\cdot[E,a] = 0$,*
3. *$[[a,E],f] = [e,[E,a]] = 0$,*
4. *$E^2 = 0$.*
5. *$[a,E] * f = (a\cdot f)E$ and, by symmetry, $e * [F,a] = (e\cdot a)F$,*
6. *$[a,E]^2 = 0$ and hence $[[a,E],[b,E]] = 0$.*
7. *$N([a,E]) = 0$ and hence $[a,E]\cdot[b,E] = 0$.*
8. *$[a,E] * [b,E] = ([a,E]\cdot b)E$.*

Proof. 1–4. These statements are immediate applications of Lemma 2.7.

5. As a consequence of statements 2 and 3 of this lemma, the $*$-products are well-defined. We have

$$[a, E] * f = [a, e * f] * f = [[a, e], f] * f + (a \cdot e) f * f + (a \cdot f) e * f,$$

where again each term on the right hand side is well-defined. Now, $f * f = 0$ by Lemma 2.7, hence it remains to be proved that $[[a, e], f] * f = 0$.

Write b for $[a, e]$ and let $x \in \mathbf{W}$. Then

$$[x, [b, f] * f] = [[x, [b, f]], f] + (x \cdot f)[b, f] + (x \cdot [b, f]) f = (b \cdot f)[x, f]$$

by application of Lemma 3.1–3. Hence $[b, f] * f = (b \cdot f) f = 0$, because $b \cdot f = [a, e] \cdot f = [e, f] \cdot a = 0$.

6. We have

$$\begin{aligned} [a, E]^2 &= ([[a, e], f] + (a \cdot f) e + (a \cdot e) f)^2 \\ &= [[a, e], f]^2 + (a \cdot f)^2 e^2 + (a \cdot e)^2 f^2 \\ &\quad + (a \cdot f)[[[a, e], f], e] + (a \cdot e)[[[a, e], f], f] + (a \cdot e)(a \cdot f)[e, f] \\ &= [[a, e], f]^2 \\ &= ([a, e] \cdot f)[a, f] + ([a, e]^2 \cdot f) f = ([a, e]^2 \cdot f) f \end{aligned}$$

by Lemma 3.1–2. Also, by the same lemma,

$$[a, e]^2 \cdot f = (a \cdot e)[a, e] \cdot f + (a^2 \cdot e)(e \cdot f) = 0.$$

7. Follows from the above and from Lemma 3.1–1.

8. Set $e' = [a, E]$ and $E' = e' * f$. Then by the above $E' = (a \cdot f) E$. Set $f' = [b, E']$ and $E'' = e' * f'$. Then by the above $E'' = (b \cdot e') E' = (a \cdot f)([a, E] \cdot b) E$. Also $e' * f' = (a \cdot f)[a, E] * [b, E]$.

   It follows that the result holds whenever $a \cdot f \neq 0$, and because the identity is linear in a, it will also hold when $a \cdot f = 0$. □

Statement 6 of this lemma justifies the name 'totally isotropic' for E : it proves that $[\mathbf{W}, E]$ is a totally isotropic subspace of $\mathbf{W}$.

Lemma 3.3. *Let E be a totally isotropic element of* $\mathbf{F}$ *with defining pair $\{e, f\}$. Let $A \in \mathbf{F}$. Define*

$$\langle E, A \rangle \stackrel{\text{def}}{=} [e, A] \cdot f. \tag{49}$$

Then

1. $[[e, A], E] = \langle E, A \rangle e$, and

2. $[[[a,E],A],E] = [[a,E],[A,E]] = \langle E,A\rangle[a,E]$, *for every* $a \in \mathbf{W}$.
3. $[E,A]^2 = \langle A,E\rangle[E,A] + \langle A^2,E\rangle E$.

Proof. 1. We find

$$\begin{aligned} [[e,A],E] = [[e,A],e*f] &= [[[e,A],e],f] + ([e,A]\cdot f)e + ([e,A]\cdot e)f \\ &= [[A,e^2],f] + ([e,A]\cdot f)e = \langle E,A\rangle e. \end{aligned}$$

2. Apply this result to $e' = [a,E]$ and $E' = e' * f$. By Lemma 3.2 we have $E' = (a\cdot f)E$. We find

$$(a\cdot f)[[[a,E],A],E] = [[e',A],E'] = ([e',A]\cdot f)e' = ([e',A]\cdot f)[a,E].$$

Also

$$\begin{aligned} [e',A]\cdot f &= [[a,E],A]\cdot f = [[a,e*f],A]\cdot f \\ &= [[[a,e],f],A]\cdot f + (a\cdot f)([e,A]\cdot f) + (a\cdot e)([f,A]\cdot f) \\ &= [[[a,e],A],f]\cdot f + [[[a,e],[f,A]]\cdot f + (a\cdot f)([e,A]\cdot f) \\ &= [f,[f,A]]\cdot[a,e] + (a\cdot f)([e,A]\cdot f) = (a\cdot f)([e,A]\cdot f) \end{aligned}$$

Hence, whenever $a\cdot f \neq 0$, we have $[[[a,E],A],E] = ([e,A]\cdot f)[a,E]$, and because this identity is linear in a, it will also hold when $a\cdot f = 0$.

3. Take $a \in \mathbf{W}$. We find

$$\begin{aligned} [a,[E,A]^2] &= [[a,[E,A]],[E,A]] \\ &= [[[a,[E,A]],E],A] + [[[a,[E,A]],A],E] \\ &= [[[a,E],A],E],A] + [[a,A],E],E],A] \\ &\qquad + [[[a,E],A],A],E] + [[a,A],E],A],E] \\ &= \langle A,E\rangle[[a,E],A] + \langle A^2,E\rangle[a,E] + \langle A,E\rangle[[a,A],E] \\ &= \langle A,E\rangle[a,[E,A]] + \langle A^2,E\rangle[a,E], \end{aligned}$$

and this is true for any $a \in \mathbf{W}$. □

Note that statement 2 of this lemma proves that the value of $\langle E,A\rangle$ is independent of the choice of the defining pair.

Lemma 3.4. *The operation* $\langle E,A\rangle$ *as defined in (49) can be extended in a unique way to a bilinear operator* $\langle\cdot,\cdot\rangle$ *defined over all elements of* $\mathbf{F}$.

This operator is symmetric and satisfies

$$\langle \mathbf{F},\mathbf{W}\rangle = \langle \mathbf{W},\mathbf{F}\rangle = 0, \tag{50}$$

$$\begin{aligned}
\langle e_r, e_{-r}\rangle &= 1, && \textit{when } r \in \Phi_L \\
\langle e_r, e_s\rangle &= 0, && \textit{when } r, s \in \Phi_L,\ r \neq -s \\
\langle h_r, e_s\rangle = \langle e_s, h_r\rangle &= 0, && \\
\langle h_r, h_s\rangle &= \langle r, s\rangle, && \textit{when } r, s \in \Phi_L
\end{aligned} \tag{51}$$

Proof. Note that by definition $\langle E, A\rangle$ is linear in its second argument. We need to prove that it can be 'made' linear in its first argument : in other words, if a linear combination $\sum_i k_i E_i$ of totally isotropic elements E_i turns out to be zero, then $\sum_i k_i\langle E_i, A\rangle = 0$ should be zero for every $A \in \mathbf{F}$.

Because e_r and $e_r + e_{-r} + h_r$ are totally isotropic for every $r \in \Phi_L$, A can always be written as a a linear combination of totally isotropic elements and elements of $\mathbf{W}$. Hence it is sufficient to prove the linearization property above in the case that A is totally isotropic and in the case that A belongs to $\mathbf{W}$.

When A is totally isotropic, we claim that $\langle E, A\rangle = \langle A, E\rangle$. Indeed, let $\{e, f\}$ be a defining pair for E and $\{e', f'\}$ for A, then

$$\begin{aligned}
\langle E, A\rangle = [e, e' * f'] \cdot f &= [[e, e'], f'] \cdot f + (e \cdot f')(e' \cdot f) + (e \cdot e')(f \cdot f') \\
&= [e, e'] \cdot [f, f'] + (e \cdot f')(e' \cdot f) + (e \cdot e')(f \cdot f'),
\end{aligned}$$

and this identity remains unchanged when interchanging e with e' and f with f'.

So in this case,

$$\sum_i k_i\langle E_i, A\rangle = \sum_i k_i\langle A, E_i\rangle = \langle A, \sum_i k_i E_i\rangle = 0,$$

as $\langle\cdot,\cdot\rangle$ is known to be linear in its second argument.

When A belongs to $\mathbf{W}$, we have $\langle E, A\rangle = [e, A] \cdot f = [e, f] \cdot A = 0$, by (21). Hence $\sum_i k_i\langle E_i, A\rangle$ is trivially zero.

This proves that $\langle\cdot,\cdot\rangle$ can be extended to a bilinear operator that is fully defined over $\mathbf{F}$. We shall now establish that the formulae (50) and (51) are the only ones possible. Because every element of $\mathbf{F}$ can be written as a linear combination of an element of $\mathbf{W}$ and elements of the form e_r, h_r, with $r \in \Phi_L$, this will prove that the bilinear product is unique.

Consider the special case $E = e_r$ with $r \in \Phi_L$. By (45) we may write $e_r = e_s * e_t$ with $s, t \in \Phi_S$ such that $s + t = r$. By (49) and (20) we have $\langle e_r, A\rangle = [e_s, A] \cdot e_t = A[-s-t] = A[-r]$. This proves $\langle e_r, h_s\rangle = 0$ and $\langle e_r, e_s\rangle = 0$ for every $s \in \Phi$, except the case $\langle e_r, e_{-r}\rangle = 1$.

As a second special case, consider $E = e_r + e_{-r} + h_r$. With s, t as before, we find

$$\begin{aligned}\langle e_r + e_{-r} + h_r, A\rangle &= [e_s + e_{-t}, A] \cdot (e_t + e_{-s}) \\ &= A[-s-t] + \sum_{i=1}^{4} \langle r_i, s\rangle A[i] + \sum_{i=1}^{4} \langle r_i, t\rangle A[i] + A[s+t] \\ &= A[-r] + \sum_{i=1}^{4} \langle r_i, r\rangle A[i] + A[r]\end{aligned}$$

Hence $\langle h_r, A\rangle = \sum_{i=1}^{4} \langle r_i, r\rangle A[i]$ and then $\langle h_r, e_s\rangle = 0$ and $\langle h_r, h_{r_j}\rangle = \langle r_j, r\rangle$, and hence $\langle h_r, h_s\rangle = \langle s, r\rangle$. □

This lemma also provides us with a tool to compute $\langle A, B\rangle$ in terms of coordinates :

$$\langle A, B\rangle = \sum_{r\in\Phi_L} A[r]B[-r] + A[1]B[2] + B[2]A[1]. \tag{52}$$

Every totally isotropic element E satisfies $E^2 = 0$, but this is not a sufficient condition. Indeed

Lemma 3.5. $\mathbf{W} - \{0\}$ *does not contain any totally isotropic elements.*

Proof. Let $e \in \mathbf{W}$ be a totally isotropic element of $\mathbf{F}$. Then $e^2 = 0$ and $[a, e]^2 = 0$ for all $a \in \mathbf{W}$. By Lemma 3.1–2 we then have $(a \cdot e)[a, e] = 0$ for every $a \in \mathbf{W}$. Hence $[a, e] = 0$ whenever $a \cdot e \neq 0$, and because the condition is linear in a, also when $a \cdot e = 0$. So $[\mathbf{W}, e] = 0$ and hence $e = 0$. □

4 Automorphisms

A nonsingular semi-linear transformation $g : \mathbf{W} \to \mathbf{W} : a \mapsto a^g$ shall be called an *automorphism* of $\mathbf{W}$ if it satisfies

$$(a^2)^g = (a^g)^2, \qquad \text{for all } a \in \mathbf{W}. \tag{53}$$

As an immediate consequence of (13) we find that also $[a^g, b^g] = [a, b]^g$. Moreover

Lemma 4.1. *Let g be an automorphism of $\mathbf{W}$ which is semi-linear with corresponding field automorphism σ. Then*

$$\begin{aligned} N(a^g) &= N(a)^\sigma, && \textit{for all } a \in \mathbf{W}, \\ a^g \cdot b^g &= (a \cdot b)^\sigma && \textit{for all } a, b \in \mathbf{W}. \end{aligned} \tag{54}$$

Proof. Note that the second equation follows from the first, by (25).

Let $a \in \mathbf{W}$. Write $b = a^g$. Applying g to equation (27) we obtain

$$(b^2)^2 = N(a)^\sigma b^2 + (a^2 \cdot a)^\sigma b,$$

because g is a semi-linear operation. But also by (27) we have $(b^2)^2 = N(b)b^2 + (b^2 \cdot b)b$. Hence

$$(N(b) - N(a)^\sigma)b^2 = ((a^2 \cdot a)^\sigma - b^2 \cdot b)b. \tag{55}$$

When a and a^2 are linearly independent, and hence b and b^2 are linearly independent, this implies $N(b) = N(a)^\sigma$ (and $(a^2 \cdot a)^\sigma = b^2 \cdot b$).

When $a^2 = ka$ for some $k \in K - \{0\}$, and hence $b^2 = k^\sigma b$, we find that $a^2 \cdot a = b^2 \cdot b = 0$. Hence the right hand side of (55) is zero, and again $N(b) = N(a)^\sigma$, provided $b^2 \neq 0$ (and hence $a^2 \neq 0$).

Finally, when $a^2 = 0$, we have $N(a) = 0$, by Lemma 3.1, but then also $b^2 = 0$ and $N(b) = 0$. □

Lemma 4.2. *Every automorphism g of $\mathbf{W}$ can be extended in a unique way to a nonsingular semi-linear transformation of $\mathbf{F}$ that satisfies*

$$(A^2)^g = (A^g)^2, \qquad \textit{for all } A \in \mathbf{F}, \tag{56}$$

and then

$$\begin{aligned} [A,B]^g &= [A^g, B^g], && \textit{for all } A, B \in \mathbf{F}, \\ (a * b)^g &= a^g * b^g, && \textit{for all } a, b \in \mathbf{W}. \end{aligned} \tag{57}$$

Proof. Note that (57) is an immediate consequence of (56). Also, if such an extension exists, it should satisfy

$$[a^g, A^g] = [a, A]^g, \qquad \text{for all } a \in \mathbf{W},\ A \in \mathbf{F}. \tag{58}$$

Consider the linear transformation τ_A that maps $a \in \mathbf{W}$ onto $[a^{g^{-1}}, A]^g$. We shall verify that τ_A satisfies the conditions of Proposition 2.4. Indeed, write $b = a^{g^{-1}}$, then

$$\tau_A(a^2) = [(a^2)^{g^{-1}}, A]^g = [b^2, A]^g = [[b,A],b]^g = [[b,A]^g, b^g] = [\tau_A(a), a].$$

As a consequence, we may define A^g to be the unique element of $\mathbf{F}$ for which $[a, A^g] = \tau_A(a) = [a^{g^{-1}}, A]^g$, for all $a \in \mathbf{W}$. Note that this definition of A^g satisfies (58).

It remains to be proved that (56) is satisfied for all $A \in \mathbf{F}$. It is sufficient to prove that $[a, (A^2)^g] = [a, (A^g)^2]$ for all $a \in \mathbf{W}$. Again write $b = a^{g^{-1}}$, then

$$[a, (A^2)^g] = [b, A^2]^g = [[b,A],A]^g = [[a,A^g],A^g],$$

by (58). □

A nonsingular semi-linear transformation g of $\mathbf{F}$ that satisfies (56) shall be called an *automorphism of* $\mathbf{F}$. The lemma above proves that there is a 1–1 relation between g and its restriction to $\mathbf{W}$ (which is an automorphism of $\mathbf{W}$). We shall therefore often drop the distinction between both types of automorphism.

The lemma also shows that to prove that two automorphisms are equal over $\mathbf{F}$, it is sufficient to prove that they are equal over $\mathbf{W}$.

Every automorphism σ of the field K can be extended to an automorphism of $\mathbf{F}$ simply by letting it act on the coordinates with respect to the Chevalley basis. An automorphism of this kind shall be called a *field* automorphism.

The following proposition provides another type of automorphism.

Proposition 4.3. *Let e be an isotropic element of $\mathbf{W}$. Define the linear map*

$$x(e) : \mathbf{F} \to \mathbf{F} : A \mapsto A^{x(e)} \stackrel{\text{def}}{=} A + [e, A] + e * [A, e]. \tag{59}$$

Then $x(e)$ is an automorphism of $\mathbf{F}$. Its restriction to $\mathbf{W}$ satisfies

$$a^{x(e)} = a + [e, a] + (a \cdot e)e, \qquad \textit{for all } a \in \mathbf{W}. \tag{60}$$

For $k, k' \in K$ we have $x(ke)x(k'e) = x((k + k')e)$ and in particular $x(e)^2 = 1$.

Proof. We first verify (56).

$$\begin{aligned}
&(A^{x(e)})^2 \\
&= (A + [e, A] + e * [A, e])^2 \\
&= A^2 + [e, A]^2 + (e * [A, e])^2 + [A, [e, A]] + [A, e * [A, e]] + [[e, A], e * [A, e]]
\end{aligned}$$

Set $b = [e, A]$ and $c = e$ and apply Lemma 2.7 to find that $[[e, A], e * [A, e]] = 0$ and $(e * [A, e])^2 = 0$. As $[b, [c, A]] = 0$ and $b \cdot [c, A] = 0$ we may apply Lemma 2.8 to obtain

$$[A, e * [A, e]] = [[A, [A, e]], e] + [[A, e], A] * e + [e, A] * [e, A] = [e, A^2] * e + [e, A]^2.$$

Combining these results, we find

$$(A^{x(e)})^2 = A^2 + [e, A^2] + (A^2)^e = (A^2)^{x(e)}.$$

By Lemma 3.1–4, (59) reduces to (60) when restricted to $\mathbf{W}$.

For the final statement it is sufficient to prove that $a^{x(ke)x(k'e)} = a^{x((k+k')e)}$ for all $a \in \mathbf{W}$. By (60) we have

$$\begin{aligned}
(a^{x(ke)})^{x(k'e)} &= (a + k[e, a] + k^2(e \cdot a)e)^{x(k'e)} \\
&= a + k[e, a] + k^2(e \cdot a)e + k'[e, a] + k'^2(e \cdot a)e \\
&= a + (k + k')[e, a] + (k + k')^2(e \cdot a)e.
\end{aligned}$$

Note that $x(e)^2 = 1$ implies that $x(e)$ is nonsingular. ☐

The group generated by all $x(e)$ with e an isotropic element of $\mathbf{W}$ shall de denoted by $\widehat{\mathsf{F}_4}(K)$ and is called a *Chevalley group* of type F_4. Note that this is a subgroup of the general linear group of $\mathbf{F}$.

Proposition 4.4. *Let E be a totally isotropic element of $\mathbf{F}$. Define the linear map*

$$x(E) : \mathbf{F} \to \mathbf{F} : A \mapsto A^{x(E)} \stackrel{\text{def}}{=} A + [E, A] + \langle A, E\rangle E. \tag{61}$$

Then $x(E)$ is an automorphism of $\mathbf{F}$. Its restriction to $\mathbf{W}$ satisfies

$$a^{x(E)} = a + [a, E], \qquad \textit{for all } a \in \mathbf{W}. \tag{62}$$

For $k, k' \in K$ we have $x(kE)x(k'E) = x((k + k')E)$ and in particular $x(E)^2 = 1$.

Proof. We first verify (56).

$$\begin{aligned}
&(A^{x(E)})^2 \\
&= (A + [E, A] + \langle A, E\rangle E)^2 \\
&= A^2 + [E, A]^2 + \langle A, E\rangle^2 E^2 + [A, [E, A]] + \langle A, E\rangle[A, E] + \langle A, E\rangle[[A, E], E] \\
&= A^2 + [E, A]^2 + [E, A^2] + \langle A, E\rangle[A, E] \\
&= A^2 + [E, A^2] + \langle A^2, E\rangle E,
\end{aligned}$$

by Lemma 3.3. And this is the same as $(A^{x(E)})^2$. By (50) , (61) reduces to (62) when restricted to $\mathbf{W}$.

For the final statement it is sufficient to prove that $a^{x(kE)x(k'E)} = a^{x((k+k')E)}$ for all $a \in \mathbf{W}$. By (60) we have

$$\begin{aligned}
(a^{x(kE)})^{x(k'E)} &= (a + k[E, a])^{x(k'E)} \\
&= a + k[E, a] + k'[E, a] + kk'[E, [E, a]] = a + (k + k')[E, a].
\end{aligned}$$

Note that $x(E)^2 = 1$ implies that $x(E)$ is nonsingular. $\square$

In the special case of elements ke_r with $k \in K$, $r \in \Phi_S$ or $r \in \Phi_L$, we use the notation $x_r(k) \stackrel{\text{def}}{=} x(ke_r)$. The following identities can easily be verified

$$\begin{aligned}
e_{-r}^{x_r(k)} &= e_{-r} + kh_r + k^2 e_r, && \\
e_s^{x_r(k)} &= e_s, && \text{when } \langle r, s\rangle \geq 0, \\
e_s^{x_r(k)} &= e_s + ke_{r+s}, && \text{when } \langle r, s\rangle = -1, \\
e_s^{x_r(k)} &= e_s + ke_{r+s} + k^2 e_{2r+s}, && \text{when } \langle r, s\rangle = -2,\ s \neq -r, \\
h_s^{x_r(k)} &= h_s + \langle s, r\rangle ke_r &&
\end{aligned} \tag{63}$$

It can be proved that the set of all elements $x_r(k)$ with $k \in K$, $r \in \Phi$ generate the group $\widehat{\mathsf{F}_4}(K)$. This is often taken as the definition of $\widehat{\mathsf{F}_4}(K)$ (cf. [2]).

When g, h are group elements, we shall write $g^h \stackrel{\text{def}}{=} h^{-1}gh$. We have

Lemma 4.5. *Let e be an isotropic element of $\mathbf{W}$, E a totally isotropic element of $\mathbf{F}$ and g an automorphism of $\mathbf{F}$. Then*

$$x(e)^g = x(e^g), \qquad x(E)^g = x(E^g). \tag{64}$$

Proof. Let $A \in \mathbf{F}$. We have

$$(A^{x(e)})^g = (A + [e,A] + [e,A] * e)^g = A^g + [e^g, A^g] + [e^g, A^g] * e^g = (A^g)^{x(e^g)},$$

hence $x(e)g = gx(e^g)$. The second equation is proved in a similar way. □

Proposition 4.6. *Let e, f be isotropic elements of $\mathbf{W}$ such that $e \cdot f = 0$. Let E, F be totally isotropic elements of $\mathbf{F}$ such that $\langle E, F\rangle = 0$. Then*

$$\begin{aligned} x(e)x(f) &= x(f)x(e)x([e,f]), \\ x(e)x(E) &= x(E)x(e)x([e,E])x(e * [e,E]), \\ x(E)x(F) &= x(F)x(E)x([E,F]). \end{aligned} \tag{65}$$

If moreover $[e,f] = 0$ then

$$x(e * f) = x(e)x(f)x(e+f). \tag{66}$$

Proof. First consider the case $[e,f] = 0$. Note that in that case $f^{x(e)} = f + [e,f] + (e \cdot f)e = f$. For $a \in \mathbf{W}$ we find

$$\begin{aligned} a^{x(f)x(e)} &= (a + [f,a] + (a \cdot f)f)^{x(e)} \\ &= a^{x(e)} + [f^{x(e)}, a^{x(e)}] + (a \cdot f)f^{x(e)} \\ &= a + [e,a] + (a \cdot e)e + [f,a] + [f,[e,a]] + (a \cdot e)[e,f] + (a \cdot f)f \\ &= a + [e,a] + [f,a] + [[a,e],f] + (a \cdot e)e + (a \cdot f)f, \end{aligned}$$

and then

$$a^{x(f)x(e)x(e*f)} = a^{x(f)x(e)} + [a^{x(f)x(e)}, e * f].$$

Using Lemma 2.7 to discard most of the resulting terms, we find

$$[a^{x(f)x(e)}, e * f] = [a, e * f]$$

and we finally obtain

$$\begin{aligned} a^{x(f)x(e)x(e*f)} &= a + [e,a] + [f,a] + (a \cdot e)f + (a \cdot f)e + (a \cdot e)e + (a \cdot f)f \\ &= a + [e+f, a] + (e+f, a) \cdot (e+f) = a^{x(e+f)}. \end{aligned}$$

Rearranging terms yields (66).

Now, by Lemma 4.5, we have

$$x(f)x(e)x(f) = x(e^{x(f)}) = x(e + [e,f] + (e \cdot f)f) = x(e + [e,f]).$$

As $e \cdot [e,f] = 0$ and $[e,[e,f]] = 0$, we may apply (66) to see that $x(e + [e,f]) = x(e)x([e,f])x(e * [e,f]) = x(e)x([e,f])$, the first equation of (65).

Similarly,

$$x(E)x(e)x(E) = x(e + [e,E]) = x(e)x([e,E])x(e * [e,E]),$$

because $e \cdot [e,E] = 0$ and $[e,[e,E]] = 0$. This yields the second identity of (65).

Finally, for $a \in \mathbf{W}$ we find

$$a^{x(F)x(E)} = (a + [a,F])^{x(E)} = a + [a,F] + [a,E] + [[a,F],E]$$

and then

$$\begin{aligned} & a^{x(F)x(E)x(F)} \\ & \quad = a + [a,F] + [a,E] + [[a,F],E] + [a,F] + [a,E],F] + [[[a,F],E],F] \\ & \quad = a + [a,E] + [a,[E,F]], \end{aligned}$$

for $[[[a,F],E],F] = \langle E,F\rangle [a,F] = 0$, by Lemma 3.3–2. Similarly

$$a^{x(E)x([E,F])} = (a + [a,E])^{x([E,F])} = a + [a,E] + [a,[E,F]] + [[a,E],[E,F]].$$

And again $[[a,E],[E,F]] = \langle E,F\rangle [a,E] = 0$, by Lemma 3.3–2. □

This proposition does not apply when $e \cdot f \neq 0$ or $\langle E,F\rangle \neq 0$. In the special case $e \cdot f = 1$ and $\langle E,F\rangle = 1$ we may consider the elements

$$\begin{aligned} n(e,f) &\stackrel{\text{def}}{=} x(e)x(f)x(e) = x(e + f + [e,f]), \\ n(E,F) &\stackrel{\text{def}}{=} x(E)x(F)x(E) = x(E + F + [E,F]) \end{aligned} \tag{67}$$

We leave it to the reader to verify that these elements have the following interesting property :

$$e^{n(e,f)} = f, \quad f^{n(e,f)} = e, \quad E^{n(E,F)} = F, \quad F^{n(E,F)} = E. \tag{68}$$

In the special case of elements ke_r with $k \in K$,$k \neq 0$, $r \in \Phi$, we write $n_r(k) \stackrel{\text{def}}{=} n(ke_r, k^{-1}e_{-r})$. Also of interest are the elements $h_r(k) \stackrel{\text{def}}{=} n_r(k)n_r(1)$.

The following identities can easily be verified

$$\begin{aligned} e_s^{n_r(1)} &= e_{w_r(s)}, \\ h_s^{n_r(1)} &= h_{w_r(s)}, \qquad \text{for all } s \in \Phi. \end{aligned} \tag{69}$$

and

$$\begin{aligned} e_s^{h_r(k)} &= k^{\langle r,s\rangle} e_s, \\ h_s^{h_r(k)} &= h_s, \qquad\qquad \text{for all } s \in \Phi. \end{aligned} \tag{70}$$

Note that the group generated by all $n_r(1)$, $r \in \Phi$ acts on the root spaces Ke_s in the same way as the Weyl group W acts on the roots $s \in \Phi$. If $w \in W$ we shall denote the corresponding element of $\widehat{\mathsf{F}_4}(K)$ by $n(w)$. We have

$$\begin{aligned} e_s^{n(w)} &= e_{w(s)}, \\ h_s^{n(w)} &= h_{w(s)}, \qquad \text{for all } s \in \Phi. \end{aligned} \tag{71}$$

The elements $n(e,f)$ and $n(E,F)$ prove useful in the following

Theorem 4.7. *The group $\widehat{\mathsf{F}_4}(K)$ acts transitively on the isotropic elements of $\mathbf{W} - \{0\}$ and on the totally isotropic elements of $\mathbf{F} - \{0\}$.*

Proof. 1. Because the Weyl group acts transitively on the short roots, we may use elements $n_r(1)$ to map e_r onto e_s whenever $r, s \in \Phi_S$. If an isotropic element f has a coordinate $f[r] \neq 0$, then we may use $n(f, e_{-r}/f[r])$ to map f onto $f[r]e_r$ and then we may use $h_s(f[r])$ with $\langle s, r\rangle = -1$ to map this further onto e_r. Hence, it is sufficient to prove that $f \neq 0$ always has at least one coordinate $f[r]$ different from 0.

If this were not the case, then $f \in \mathbf{I}$ and hence $f = f[3]h_{r_3} + f[4]h_{r_4}$. Then $f^2 = f[3]^2 h_{r_3} + f[4]^2 h_{r_4}$, and this must be 0 because f is isotropic. Hence $f[3] = f[4] = 0$ and therefore $f = 0$.

2. We can use the same argument on totally isotropic elements F in $\mathbf{F} - \{0\}$. It remains to be proved that we can always find $r \in \Phi_L$ such that $F[r] \neq 0$.

Assume the contrary : let $F \in \mathbf{W} + \mathbf{H}$. Setting $F^2 = 0$ and computing coordinates, we easily prove that $F[1] = F[2] = 0$. Hence F must be an element of $\mathbf{W}$ and hence $F = 0$ by Lemma 3.5. □

It is also possible to determine all orbits of $\widehat{\mathsf{F}_4}(K)$ on *pairs* of isotropic or totally isotropic elements, but as was already mentioned in the introduction, this would lead us too far. In fact, the algebraic framework which we have established so far can serve as a valuable tool in the construction of the meta-symplectic space associated with $\mathbf{W}$, $\mathbf{F}$ and $\widehat{\mathsf{F}_4}(K)$. This is done using essentially the same techniques as we will use in Sections 8–10 of this text to construct the Ree-Tits octagon.

Lemma 4.8. *Let $E \in \mathbf{F} - \{0\}$ be totally isotropic. Then $[\mathbf{W}, E]$ is a totally isotropic subspace of $\mathbf{W}$ with the following properties :*

1. $\dim[\mathbf{W}, E] = 6$.
2. *If* $v, w \in [\mathbf{W}, E]$*, then* $v * w \in K\,E$ *and the bilinear form* $v * w/E$ *makes* $[\mathbf{W}, E]$ *into a symplectic space.*
3. *Every element* e *of* $[\mathbf{W}, E]$ *can be written as* $e = [w, E]$ *such that* w *is an isotropic element of* $\mathbf{W}$.
4. *Let* $e \neq 0$*. Then* e *belongs to* $[\mathbf{W}, E]$ *if and only if* E *has a defining pair containing* e.
5. *An element* $e \in \mathbf{W}$ *belongs to* $[\mathbf{W}, E]$ *if and only if* $[[e, \mathbf{W}], E] = 0$.

Proof. By Theorem 4.7 we only need to prove this for a single non-zero isotropic element of $\mathbf{W}$, say $E = e_{1\bar{1}00}$.

1. It is easily verified that in this case $[\mathbf{W}, E]$ is generated by the following six elements :

$$e_{1000},\ e_{0\bar{1}00},\ e_{+-++},\ e_{+---},\ e_{+-+-},\ e_{+--+}. \tag{72}$$

2. It can be inferred from the inner products of the corresponding roots that $e * f = 0$ is zero for any pair of elements of (72) except for the following

$$e_{1000} * e_{0\bar{1}00} = e_{+-++} * e_{+---} = e_{+-+-} * e_{+--+} = e_{1\bar{1}00} = E. \tag{73}$$

   The value of $u*v/E$ is then equal to $u_1v_2+u_2v_1+u_3v_4+u_4v_3+u_5v_6+u_6v_5$, when u_j, v_j are the coordinates of u, v with respect to the basis listed in (72).

3. The element u with coordinates $u_1, \ldots, u_6$ can be written as $[w, E]$ with

$$w = u_1e_{0100} + u_2e_{\bar{1}000} + u_3e_{-+++} + u_4e_{-+--} + u_5e_{-++-} + u_6e_{-+-+}.$$

   Note that w belongs to $[\mathbf{W}, e_{\bar{1}100}]$ and therefore is isotropic.

4. If e has a non-zero coordinate with respect to the base above, say $e_1 \neq 0$, then $\{e, e_{0\bar{1}00}/e_1\}$ is a defining pair for E, by (73).

   Conversely, applying Lemma 3.3–1 we find $e = [w, E]$ with $w = [e, e_{\bar{1}100}]$.

5. Clearly, if $e = [w, E]$ then $[[e, \mathbf{W}], E] = [[[w, E], \mathbf{W}], E] = 0$, by Lemma 3.3–2 and (50).

Conversely, $[[e, \mathbf{W}], E] = 0$ if and only if $[[[e, \mathbf{W}], E] \cdot \mathbf{W} = 0$ and this is equivalent to $[\mathbf{W}, E] \cdot [e, \mathbf{W}] = e \cdot [[\mathbf{W}, E], \mathbf{W}] = 0$. The space $[[\mathbf{W}, E], \mathbf{W}]$ is the sum of the spaces $[e_r, \mathbf{W}]$ where e_r is one of the six base vectors from (72), and from this we may compute that $[E, \mathbf{W}]$ is generated by $\mathbf{I}$ together with all vectors e_s where $s \in \Phi_S$ *except* when s is in the following list :

$$\bar{1}000,\ 0100,\ {-}{+}{-}{-},\ {-}{+}{+}{+},\ {-}{+}{-}{+},\ {-}{+}{+}{-}. \tag{74}$$

Note that for each root s in this list e_{-s} is one of the six base vectors of $[\mathbf{W}, E]$. Hence $e \cdot [[\mathbf{W}, E], \mathbf{W}] = 0$ if and only if $e \in [\mathbf{W}, E]$. □

5 The quotient algebra Q

Just like $\mathbf{W}$ is a subalgebra of $\mathbf{F}$ related to the short roots, there is another well-known subalgebra of $\mathbf{F}$ related to the long roots. This is the Chevalley algebra of type D_4 which we denote by $\mathbf{D}$:

$$\mathbf{D} \stackrel{\text{def}}{=} \mathbf{H} \oplus \bigoplus_{r \in \Phi_L} K e_r.$$

This algebra has dimension 28. The defining relations for $\mathbf{D}$ can be derived from (8) for the special case that both $r, s \in \Phi_L$:

$$\begin{aligned}
[h_r, h_s] &= 0, \\
[h_r, e_s] &= \langle r, s \rangle e_s, \\
[e_r, e_{-r}] &= h_r, \\
[e_r, e_s] &= 0, && \text{when } r + s \neq 0,\ r + s \notin \Phi_L, \\
[e_r, e_s] &= e_{r+s}, && \text{when } r + s \in \Phi_L.
\end{aligned} \tag{75}$$

Note that also $\mathbf{D}^2 \leq \mathbf{D}$.

The similarity between (17) and (75) suggests that by somehow mapping short roots to long roots in a way that preserves their properties we could establish an isomorphism between $\mathbf{W}$ and $\mathbf{D}$. Because $\dim \mathbf{W} \neq \dim \mathbf{D}$ this is not immediately possible. However, $\mathbf{D}$ contains $\mathbf{I}$ as an ideal, and the quotient $\mathbf{Q} \stackrel{\text{def}}{=} \mathbf{D}/\mathbf{I} \simeq \mathbf{F}/\mathbf{W}$ is another Lie algebra of dimension 26.

Consider the following linear transformation γ on the 4-dimensional real vector space generated by Φ :

$$\gamma : r_1 \mapsto r_4^* = 2r_4, \quad r_2 \mapsto r_3^* = 2r_3, \quad r_3 \mapsto r_2^* = r_2, \quad r_4 \mapsto r_1^* = r_1$$

This map satisfies $\gamma(r) \cdot \gamma(s) = 2r \cdot s$ and has the property that it maps the root system Φ onto its dual Φ^*.

We use γ to define a map $\overline{\cdot}$ on Φ that interchanges short and long roots :

$$\overline{r} \stackrel{\text{def}}{=} \gamma(r)^* = \begin{cases} \gamma(r), & \text{when } r \in \Phi_S, \\ \frac{1}{2}\gamma(r), & \text{when } r \in \Phi_L. \end{cases}$$

It can easily be proved that $\langle \overline{r}, \overline{s} \rangle = \langle s, r \rangle$.

Lemma 5.1. *Let w be an element of the Weyl group W. Define $\overline{w} : \Phi \to \Phi : x \mapsto \overline{w}(x) \stackrel{\text{def}}{=} \overline{w(\overline{x})}$. Then*

$$\overline{w_r} = w_{\overline{r}}, \qquad \text{for all } r \in \Phi, \tag{76}$$

and hence $\overline{W} = W$.

Proof. Consider $x \in \Phi, r \in \Phi$. We have $\overline{w_r}(x) = \overline{w_r(\overline{x})}$. Note that W preserves the length of a root, and that $\overline{\cdot}$ interchanges short and long roots. It follows that $\overline{w_r(\overline{x})} = \frac{1}{2}\gamma(w_r(\gamma(x))$. Hence

$$\begin{aligned} \overline{w_r}(x) = \frac{1}{2}\gamma(w_r(\gamma(x))) &= \frac{1}{2}\gamma(\gamma(x) - \langle r, \gamma(x) \rangle r) \\ &= \frac{1}{2}\gamma(\gamma(x)) - \frac{1}{2}\langle r, \gamma(x) \rangle \gamma(r) \\ &= x - \frac{r \cdot \gamma(x)}{r \cdot r}\gamma(r) \\ &= x - \frac{\gamma(r) \cdot \gamma(\gamma(x))}{\gamma(r) \cdot \gamma(r)}\gamma(r) \\ &= x - \frac{1}{2}\langle \gamma(r), 2x \rangle \gamma(r) \\ &= x - \langle \gamma(r), x \rangle \gamma(r) = w_{\gamma(r)}(x), \end{aligned} \tag{77}$$

and because the reflection operator $w_{\gamma(r)}(x)$ does not change when we multiply $\gamma(r)$ by a scalar, this is equal to $w_{\overline{r}}(x)$. $\square$

We may use the map $\overline{\cdot}$ to construct an isomorphism between $\mathbf{W}$ and $\mathbf{Q} \simeq \mathbf{F}/\mathbf{W}$. Define a linear transformation $\mu : \mathbf{W} \to \mathbf{Q}$ as follows :

$$\mu(e_r) \stackrel{\text{def}}{=} e_{\overline{r}} + \mathbf{W}, \quad \mu(h_r) \stackrel{\text{def}}{=} h_{\overline{r}} + \mathbf{W}, \qquad \text{for every } r \in \Phi_S.$$

That μ is an isomorphism of Lie algebras is an immediate consequence of (17), (75) and the properties of γ. We only need to prove that it is well-defined, for it potentially may occur that $h_r = h_s$ although $r \neq s$. However, it is easily verified that in that case also $h_{\overline{r}} = h_{\overline{s}} \bmod \mathbf{W}$.

Many of the operations on $\mathbf{F}$ are also well-defined on $\mathbf{Q}$. For $A, B \in \mathbf{F}$ we easily prove

$$\begin{aligned} [A + \mathbf{W}, B + \mathbf{W}] &= [A, B] + \mathbf{W} \\ (A + \mathbf{W})^2 &= A^2 + \mathbf{W} \\ \langle A + \mathbf{W}, B + \mathbf{W} \rangle &= \langle A, B \rangle \end{aligned}$$

For $a, b \in \mathbf{W}$, we find

$$\begin{aligned} [\mu(a), \mu(b)] &= \mu([a, b]), \\ \mu(a)^2 &= \mu(a^2), \\ \langle \mu(a), \mu(b) \rangle &= a \cdot b. \end{aligned} \tag{78}$$

(The last identity is obtained by comparing (18) and (50)).

6 The operator $Q(\cdot)$

It turns out that μ is not the best candidate for a *duality* operator which would enable us to 'twist' $\widehat{\mathsf{F}_4}(K)$ and obtain a meaningful definition of the Ree-Tits octagon. Such an operator should map isotropic elements onto totally isotropic elements (points onto symplecta) in such a way that (symmetric) incidence is preserved. (Ke and KE are 'incident' if and only if $e \in [\mathbf{W}, E]$.) Unfortunately $\mathbf{Q}$ does not (yet) have a notion of total isotropicity, and moreover, there is no obvious way to associate $\mu([\mathbf{W}, E])$ with $\mu^{-1}(E)$, or for that matter, to give a useful definition of $[\mathbf{W}, E]$ when $E \in \mathbf{Q}$.

Instead, in this section we shall introduce a duality operator $Q(\cdot)$ with images in $\mathbf{F}$ (and not $\mathbf{Q}$). This operator is quadratic (and not linear like μ) and is defined only on isotropic elements. (There seems to be no elegant way to extend it to all of $\mathbf{W}$.)

Let e be an isotropic element of $\mathbf{W}$, let $w \in \mathbf{W}$ and $A, B \in \mathbf{F}$. Then $[e, w] \cdot [e, A] = w \cdot [e, [e, A]] = 0$ by (21). Hence the value of $[e, A + \mathbf{W}] \cdot [e, B + \mathbf{W}]$ is well-defined.

Proposition 6.1. *Let e be an isotropic element of $\mathbf{W}$. Then there is a unique element $Q(e)$ of $\mathbf{F}$ satisfying*

$$[a, Q(e)] \cdot b = [e, \mu(a)] \cdot [e, \mu(b)], \qquad \textit{for all } a, b \in \mathbf{W}. \tag{79}$$

We have

$$\mu([a, Q(e)]) = [e, \mu(a)] * e \bmod \mathbf{W}, \qquad \textit{for all } a \in \mathbf{W}. \tag{80}$$

Proof. Consider the unique linear map $\tau_e \in \mathrm{Hom}(\mathbf{W}, \mathbf{W})$ that satisfies $\tau_e(a) \cdot b = [e, \mu(a)] \cdot [e, \mu(b)]$, for all $a, b \in \mathbf{W}$. It is sufficient to prove that τ_e satisfies the conditions of proposition 2.4. Indeed, let $A = \mu(a)$, $B = \mu(b)$, then several applications of (21) and (22) yield

$$\begin{aligned}\tau_e(a)\cdot[a,b] &= [e,\mu(a)]\cdot[e,\mu([a,b])]\\ &= [e,A]\cdot[e,[A,B]]\\ &= [e,A]\cdot[[e,A],B]+[e,A]\cdot[[e,B],A]\\ &= [e,A]\cdot[[e,B],A]\\ &= [[[e,B],A],A]\cdot e = [[e,B],A^2]\cdot e = [e,A^2]\cdot[e,B]\end{aligned}$$

and this is equal to $\tau_e(a^2)\cdot b$.

Finally, applying (78), we find

$$\langle\mu([a,Q(e)],\mu(b)\rangle = [a,Q(e)]\cdot b = [e\cdot\mu(a)]\cdot[e,\mu(b)] = \langle[e,\mu(a)]*e,\mu(b)\rangle,$$

and as any element of $\mathbf{Q}$ can be written as $\mu(b)$ for some $b\in\mathbf{W}$, we obtain (80). □

Proposition 6.2. *Let $e,f\in\mathbf{W}$ such that $e^2=f^2=0$. Then*

1. *$Q(e)$ is totally isotropic (or zero).*
2. *If $e\cdot f=0$, then $Q([e,f])=[Q(e),Q(f]$.*
3. *If $[e,f]=0$ then $Q(e+f)=Q(e)+Q(f)+\mu^{-1}(e*f)$.*
4. *If $A\in\mathbf{F}$ is totally isotropic and $a=\mu^{-1}(A)$, then $Q([e,A])=a*[Q(e),a]$.*

Proof. Let $a,b,c,d\in\mathbf{W}$. For ease of notation we introduce the following abbreviations :

$$E\stackrel{\text{def}}{=}Q(e),\ F\stackrel{\text{def}}{=}Q(f),\ x\stackrel{\text{def}}{=}[a,Q(e)],\ y\stackrel{\text{def}}{=}[b,Q(E)].$$

We choose $A,B,C,D,X,Y\in\mathbf{F}$ be such that

$$A=\mu(a),\ B=\mu(b),\ C=\mu(c),\ D=\mu(d)\bmod\mathbf{W},$$

and

$$X=\mu(x)=[e,A]*e,\ Y=\mu(y)=[e,B]*e\bmod\mathbf{W}.$$

We shall also abbreviate $[u,E]\cdot v$ to E_{uv} for general $u,v\in\mathbf{W}$. Note that $[e,X]=[e,Y]=0$ and

$$[e,[C,X]]=[[e,C],X]+[[e,X],C]=[[e,C],e*[e,A]]=([e,C]\cdot[e,A])e,$$

and hence

$$[e,[C,X]]=E_{ac}e,\ [e,[D,X]]=E_{ad}e,\ [e,[C,Y]]=E_{bc}e,\ [e,[D,Y]]=E_{bd}e.$$

Also $[[e,A],Y]=[[e,B],X]=E_{ab}e$.

1. We compute

$$\begin{aligned}&[[c,[a,Q(e)]],[b,Q(e)]]\cdot d\\&=[[c,x],y]\cdot d=[c,x]\cdot[d,y]=x\cdot[c,[d,y]]\\&=[e,A]\cdot[e,\mu([c,[d,y]])]\\&=[e,A]\cdot[e,[C,[D,Y]]]\\&=[e,A]\cdot[C,[e,[D,Y]]]+[e,A]\cdot[[e,C],[D,Y]]\\&=[e,A]\cdot[C,[e,[D,Y]]]+[e,A]\cdot[[[e,C],Y],D]+[e,A]\cdot[[[e,C],D],Y].\end{aligned}$$

   Now,

$$\begin{aligned}&[e,A]\cdot[C,[e,[D,Y]]]=[e,A]\cdot[C,E_{bd}e]=E_{ac}E_{bd}\\&[e,A]\cdot[[[e,C],Y],D]=[e,A]\cdot[E_{bc}e,D]=E_{bc}E_{ad}\\&[e,A]\cdot[[[e,C],D],Y]=[[e,A],Y]\cdot[[e,C],D]=E_{ab}e\cdot[[e,C],D]=E_{ab}E_{cd}.\end{aligned}$$

   This proves

$$[[c,x],y]\cdot d=E_{ac}E_{bd}+E_{bc}E_{ad}+E_{ab}E_{cd}. \tag{81}$$

   Adding to this expression the expression obtained by interchanging the roles of a and b (and hence those of x and y) yields $[c,[x,y]]\cdot d=0$, for all $c,d\in\mathbf{W}$. Hence $[x,y]=0$.

   Setting $a=b$ in (81) we get $[c,x^2]\cdot d=0$, for all $c,d\in\mathbf{W}$. Therefore $x^2=0$ and by symmetry $y^2=0$.

   Also

$$x\cdot y=[a,Q(e)]\cdot y=[e,A]\cdot[e,Y]=e\cdot[[e,A],Y]=E_{ab}e\cdot e=0.$$

   This proves that $x*y$ is well-defined. We find

$$\begin{aligned}[c,x*y]\cdot d&=[[c,x],y]\cdot d+(c\cdot y)(x\cdot d)+(c\cdot x)(y\cdot d)\\&=[[c,x],y]\cdot d+E_{bc}E_{ad}+E_{ac}E_{bd},\end{aligned}$$

   and by (81) we find that this is equal to $E_{ab}E_{cd}=E_{ab}[c,Q(e)]\cdot d$, for all $c,d\in\mathbf{W}$. Hence $x*y=([a,Q(e)]\cdot b)Q(e)$. If $Q(e)\neq 0$ we may always find a,b such that $[a,Q(e)]\cdot b=1$, and then $Q(e)=x*y$. This proves that $Q(e)$ is totally isotropic with defining pair $\{x,y\}$.

2. Let $e\cdot f=0$. By Lemma 3.1–2 we have $[e,f]^2=0$ and therefore $[e,f]$ is

isotropic and $Q([e,f])$ is well-defined. We compute

$$
\begin{aligned}
&[[a,E],F]\cdot b = [x,Q(f)]\cdot b = [f,X]\cdot[f,B]\\
&= [f, e*[e,A]]\cdot[f,B]\\
&= [[e,f],[e,A]]\cdot[f,B] + ([e,A]\cdot f)(e\cdot[f,B])\\
&= [[f,B],[e,A]]\cdot[e,f] + ([e,A]\cdot f)(e\cdot[f,B])\\
&= [[[f,B],e],A]\cdot[e,f] + [[[f,B],A],e]\cdot[e,f] + ([e,A]\cdot f)(e\cdot[f,B])\\
&= [[[f,B],e],A]\cdot[e,f] + [[f,B],A]\cdot[[e,f],e] + ([e,A]\cdot f)(e\cdot[f,B])\\
&= [[[f,B],e],A]\cdot[e,f] + ([e,A]\cdot f)(e\cdot[f,B])
\end{aligned}
$$

and hence, adding the same formula with the roles of e and f interchanged,

$$
\begin{aligned}
[a,[E,F]] &= [[[f,B],e],A]\cdot[e,f] + [[[e,B],f],A]\cdot[e,f]\\
&= [[[e,f],A],B]\cdot[e,f]\\
&= [[e,f],A]\cdot[[e,f],B] = [a,Q([e,f]]\cdot b.
\end{aligned}
$$

This is true for all $a,b\in\mathbf{W}$, hence $[Q(e),Q(f)] = Q([e,f])$.

3. Note that $e{+}f$ is isotropic and hence $Q(e{+}f)$ is well-defined. Let $a,b\in\mathbf{W}$. Applying the definition of $Q(\cdot)$, we find

$$
\begin{aligned}
&[a,Q(e+f)]\cdot b - [a,Q(e)]\cdot b - [a,Q(f)]\cdot b\\
&\quad= [e+f,A]\cdot[e+f,B] - [e,A]\cdot[e,B] - [f,A]\cdot[f,B]\\
&\quad= [e,A]\cdot[f,B] + [e,B]\cdot[f,A]\\
&\quad= [[f,B],A]\cdot e + e\cdot[[B,[f,A]]\\
&\quad= e\cdot[f,[A,B]] = \langle e*f,[A,B]\rangle\\
&\quad= \mu^{-1}(e*f)\cdot[a,b] = [a,\mu^{-1}(e*f)]\cdot b.
\end{aligned}
$$

This is true for all $a,b\in\mathbf{W}$, hence $Q(e+f)-Q(e)-Q(f) = \mu^{-1}(e*f)$.

4. Let $c,d\in\mathbf{W}$. Then

$$
[c,Q([e,A])]\cdot d = [[e,A],C]\cdot[[e,A],D] \tag{82}
$$

Also

$$
\begin{aligned}
[c,a*[a,Q(e)]]\cdot d &= [c,a*x]\cdot d\\
&= [c,a]\cdot[x,d] + (c\cdot x)(a\cdot d) + (c\cdot a)(x\cdot d)\\
&= [[a,c],d]\cdot x + E_{ac}a\cdot d + E_{ad}a\cdot c
\end{aligned}
$$

Now

$$\begin{aligned}[][a,c],d]\cdot x &= [e,[[A,C],D]]\cdot[e,A]\\ &= e\cdot[[e,A],[D,[A,C]]]\\ &= e\cdot[[[e,A],D],[A,C]] + e\cdot[[[e,A],[A,C]],D]\end{aligned} \tag{83}$$

Because A is totally isotropic, we may apply Lemma 3.3–2 to reduce the second term of this result to

$$\begin{aligned}e\cdot[[[e,A],[A,C]],D] &= e\cdot\langle A,C\rangle[[e,A],D]\\ &= (a\cdot c)[e,D]\cdot[e,A] = E_{ad}a\cdot c.\end{aligned} \tag{84}$$

The first term in the result of (83) can be simplified as follows :

$$\begin{aligned}e\cdot[[[e,A],D],[A,C]] &= [e,[A,C]]\cdot[[e,A],D]\\ &= [[e,C],A]\cdot[[e,A],D] + [[e,A],C]\cdot[[e,A],D]\\ &= [e,C]\cdot[[[e,A],D],A] + [[e,A],C]\cdot[[e,A],D]\\ &= E_{ac}a\cdot d + [[e,A],C]\cdot[[e,A],D]\end{aligned}$$

Adding (84) yields the same result as (82). ☐

Lemma 6.3. *Let e be an isotropic element of $\mathbf{W}$. Then the coordinates of $Q = Q(e)$ are given by*

$$\begin{aligned}
&Q[t] = e[\overline{t}]^2, \text{ when } t\in\Phi_L, && Q[t] = \sum_{\substack{\{u,v\}\subset\Phi_S\\ u+v=\overline{t}}} e[u]e[v], \text{ when } t\in\Phi_S,\\
&Q[1] = e[4]^2, && Q[3] = \sum_{\substack{u\in\Phi_S\\ \langle r_1,u\rangle=-1}} e[u]e[-u],\\
&Q[2] = e[3]^2, && Q[4] = \sum_{\substack{u\in\Phi_S\\ \langle r_2,u\rangle=-1}} e[u]e[-u].
\end{aligned} \tag{85}$$

Proof. Let $r,s\in\Phi_S$. We have

$$[e,e_{\overline{r}}] = \sum_{\substack{u\in\Phi_S\\ \langle\overline{r},u\rangle=-1}} e[u]e_{\overline{r}+u}, \qquad [e,e_{\overline{s}}] = \sum_{\substack{v\in\Phi_S\\ \langle\overline{s},v\rangle=-1}} e[v]e_{\overline{s}+v},$$

and hence, by (79),

$$[e_r,Q(e)]\cdot e_s = [e,e_{\overline{r}}]\cdot[e,e_{\overline{s}}] = \sum_{u,v}{}' e[u]e[v],$$

where the sum ranges over all pairs $u, v \in \Phi_S$ such that $\langle \overline{r}, u\rangle = \langle \overline{s}, v\rangle = -1$ and $u + v + \overline{r} + \overline{s} = 0$. We consider different cases, according to the value of $\langle r, s\rangle = \langle \overline{r}, \overline{s}\rangle$.

1. $\langle r, s\rangle = \langle \overline{r}, \overline{s}\rangle = 0$ and then $t = -r - s \in \Phi_L$. Without loss of generality (because of the transitivity properties of the Weyl group W) we may choose $\overline{r} = \mathtt{1100}$, $\overline{s} = \mathtt{0011}$. The following table lists the corresponding values of $u, v \in \Phi_S$ for which $\langle \overline{r}, u\rangle = \langle \overline{s}, v\rangle = -1$.

u	$\bar{1}000$	$0\bar{1}00$	----	---+	--+-	--++
v	$00\bar{1}0$	$000\bar{1}$	----	+---	-+--	++--

   Of these, the only pair satisfying $u + v + \overline{r} + \overline{s} = 0$ is $u = v =$ ----, i.e., $u = v = -\frac{1}{2}(\overline{r} + \overline{s})$.

   This is equivalent to the condition $u = v = \overline{t}$. Applying (20) then yields the left hand formula for $Q[t]$.

2. $\langle r, s\rangle = \langle \overline{r}, \overline{s}\rangle = -1$ and then $t = -r - s \in \Phi_S$. Choose $\overline{r} = \mathtt{1100}$, $s = \bar{1}010$ (and hence $\overline{t} = -\overline{r} - \overline{s} = 0\bar{1}\bar{1}0$), yielding the following table of values for u, v :

u	$\bar{1}000$	$0\bar{1}00$	----	---+	--+-	--++
v	1000	$00\bar{1}0$	+---	++--	+--+	++-+

   Of these, there are three pairs (u, v) satisfying $u + v + \overline{r} + \overline{s} = 0$:

   $$(\texttt{----}, \texttt{+--+}), \quad (\texttt{---+}, \texttt{+---}), \quad (0\bar{1}00, 00\bar{1}0),$$

   and taken together with the corresponding pairs (v, u) we find exactly the six pairs of short roots that add up to the long root $\overline{t}$. This yields the right hand formula for $Q[t]$.

3. When $s = -r$ the conditions are satisfied whenever $u = -v$ and $\langle \overline{r}, u\rangle = -1$. We find
   $$[e_r, Q(e)] \cdot e_{-r} = \sum_{\substack{u \in \Phi_S \\ \langle \overline{r}, u\rangle = -1}} e[u]e[-u].$$

   We shall abbreviate this expression to $Q_{(r)}$. By (20) $Q_{(r)}$ is also equal to $\sum_i \langle r_i, r\rangle Q[i]$.

   Because $\langle r_3, r_j\rangle = 0$, except when $j = 4$ (and then it equals 1), we find $Q_{(r_3)} = Q[4]$, and likewise $Q[3] = Q_{(r_4)}$, giving the values listed in (85).

   Finally, take r, s such that $r + s = t \in \Phi_L$. Then $Q_{(r)} + Q_{(s)} = \sum_i \langle r_i, t\rangle Q[i]$. By choosing $t = r_1$ and r_2 we may use this equality to obtain values for $Q[2]$ and $Q[1]$.

By the above, we have

$$Q_{(r)} + Q_{(s)} = \sum_u{}' e[u]e[-u],$$

where the sum ranges over all $\{u, -u\}$ such that either $\langle \overline{r}, \pm u\rangle = \pm 1$ or $\langle \overline{s}, \pm u\rangle = \pm 1$, but not both. Note that $\overline{r} + \overline{s} = 2\overline{t}$, and hence $2\langle \overline{t}, u\rangle = 2\langle u, \overline{t}\rangle = \langle u, \overline{r}\rangle + \langle u, \overline{s}\rangle$, and therefore $\langle \overline{t}, u\rangle = \langle \overline{r}, u\rangle + \langle \overline{s}, u\rangle$. It follows that the sum ranges exactly over all $\{u, -u\}$ for which $\langle \overline{t}, \pm u\rangle = \pm 1$. Hence

$$Q_{(r)} + Q_{(s)} = \sum_{\substack{u\in\Phi_S \\ \langle u,\overline{t}\rangle=-1}} e[u]e[-u] = \sum_{\{u,-u\}\subset\Phi_S} \langle u, \overline{t}\rangle e[u]e[-u]. \qquad (86)$$

Now, $e^2 = 0$ and in particular $e^2[3] = e^2[4] = 0$. Using (14) this translates to

$$0 = e[3]^2 + \sum_{\{u,-u\}\subset\Phi_S} \langle u, r_4\rangle e[u]e[-u],$$

and similarly for $e[4]^2$.

Comparing this result with (86) applied to $t = r_1$ and $t = r_2$, we obtain the formulae for $Q[1]$ and $Q[2]$. □

Let $r, s \in \Phi_s$. We may use (85) to compute the following example values of $Q(e)$:

$$\begin{aligned}
&Q(e_r) &&= e_{\overline{r}}, && \\
&Q(e_r + e_s) &&= e_{\overline{r}} + e_{\overline{s}}, &&\text{when } \langle r, s\rangle = 1, \\
&Q(e_r + e_s) &&= e_{\overline{r}} + e_{\overline{s}} + e_{\overline{r+s}}, &&\text{when } \langle r, s\rangle = 0, \\
&Q(e_r + e_{-r} + h_r) &&= e_{\overline{r}} + e_{-\overline{r}} + h_{\overline{r}}, &&
\end{aligned} \qquad (87)$$

We extend the Frobenius morphism $k \mapsto k^{\mathrm{frob}} \stackrel{\mathrm{def}}{=} k^2$ over K to $\mathbf{F}$ by applying it to the coordinates with respect to the Chevalley basis. When K is a perfect field this map is nonsingular and then $\cdot^{\mathrm{frob}}$ is a field automorphism of $\mathbf{F}$.

The three left hand equations of (85) imply

$$Q(e) = \mu(e^{\mathrm{frob}}) \bmod \mathbf{W}, \qquad \text{for isotropic } e \in \mathbf{W}. \qquad (88)$$

Proposition 6.4. *Let $e, f \in \mathbf{W}$ be such that $e^2 = f^2 = 0$. Then*

1. $\langle Q(e), Q(f)\rangle = (e \cdot f)^2$.
2. $Q([e, Q(f)]) = f^{\mathrm{frob}} * [Q(e), f^{\mathrm{frob}}]$.

3. $[e, Q(f)] = 0$ *if and only if* $[f^{\text{frob}}, Q(e)] = 0$.
4. $e \in [\mathbf{W}, Q(f)]$ *if and only if* $f^{\text{frob}} \in [\mathbf{W}, Q(e)]$.

Proof. 1. Using (78) and (88) we obtain

$$\langle Q(e), Q(f)\rangle = \langle \mu(e^{\text{frob}}), \mu(f^{\text{frob}})\rangle = e^{\text{frob}} \cdot f^{\text{frob}} = (e \cdot f)^{\text{frob}} = (e \cdot f)^2.$$

2. Apply Proposition 6.2–4 to $A = Q(f)$ and then $a = f^{\text{frob}}$, by (88).
3. By the above, $[Q(e), f^{\text{frob}}] = 0$ clearly implies $Q([e, Q(f)]) = 0$ and hence $[e, Q(f)] = 0$. Substituting f^{frob} for e and e for f in this result, we obtain that $[Q(f), e]^{\text{frob}} = 0$ implies $[f^{\text{frob}}, Q(e)] = 0$.
4. Let $a \in \mathbf{W}$ be such that $e = [w, Q(f)]$. By Lemma 4.8 we can always choose w to be isotropic. Then by the above $\{f^{\text{frob}}, [f^{\text{frob}}, Q(w)]\}$ is a defining pair for $Q(e)$, and then by the same lemma $f^{\text{frob}} \in [\mathbf{W}, e]$. The converse can be obtained by substituting f^{frob} for e and e for f in this result. □

Proposition 6.5. *Let g be an automorphism of $\mathbf{W}$ (and $\mathbf{F}$). Define $\overline{g} \overset{\text{def}}{=} \mu g \mu^{-1}$. Then $\overline{g}$ is an automorphism of $\mathbf{W}$ (and $\mathbf{F}$).*

When e is an isotropic element of $\mathbf{W}$, we have

$$\overline{x(e)} = x(Q(e)), \qquad \overline{x(Q(e))} = x(e^{\text{frob}}). \tag{89}$$

Proof. We check the definition of automorphism. Let $a \in \mathbf{W}$ and choose $A \in \mathbf{F}$ such that $A = \mu(a) \bmod \mathbf{W}$. Then $\overline{g}$ maps a^2 onto

$$\mu^{-1}((A^2)^g) = \mu^{-1}((A^g)^2) = (\mu^{-1}(A^g))^2,$$

and hence $(a^2)^{\overline{g}} = (a^{\overline{g}})^2$. Also note that $\overline{g^{-1}} = \mu g^{-1} \mu^{-1}$ is the inverse of $\overline{g}$, hence $\overline{g}$ is nonsingular.

Let e be as stated. Then $\overline{x(e)}$ maps a onto $\mu^{-1}(A + e * [A, e])$. (Note that $[A, e] = 0 \bmod \mathbf{W}$.) And this is equal to $a + [a, Q(e)]$, by (80). Similarly $\overline{x(Q(e))}$ maps a to $\mu^{-1}(A + [A, Q(e)]) = a + [a, e^{\text{frob}}]$, by (88). □

As a particular case of (89) we find

$$\overline{x_r(k)} = \begin{cases} x_{\overline{r}}(k^2), & \text{when } r \in \Phi_S, \\ x_{\overline{r}}(k), & \text{when } r \in \Phi_L. \end{cases} \tag{90}$$

Note that by definition $\overline{gh} = \overline{g}\,\overline{h}$ for any two automorphisms g and h of $\mathbf{W}$. Also (89) implies

$$\overline{\overline{g}} = g^{\text{frob}}, \qquad \text{for } g \in \widehat{\mathsf{F}_4}(K), \tag{91}$$

and hence $\overline{\widehat{\mathsf{F}_4}(K)} = \widehat{\mathsf{F}_4}(K)$ when K is a perfect field. In that case the map $g \mapsto \overline{g}$ is an automorphism of the group $\widehat{\mathsf{F}_4}(K)$ (called a *graph* automorphism).

Let e, f be isotropic elements of $\mathbf{W}$. From Proposition 6.4 it follows that $e \cdot f = 1$ if and only if $\langle Q(e), Q(f)\rangle = 1$. Hence $n(Q(e), Q(f))$ is well-defined whenever $n(e, f)$ is. Proposition 6.5 then implies

$$\overline{n(e,f)} = n(Q(e), Q(f)), \qquad \overline{N(Q(e), Q(f))} = n(e,f)^{\mathrm{frob}}.$$

In particular

$$\overline{n_r(k)} = \begin{cases} n_{\overline{r}}(k^2), & \text{when } r \in \Phi_S, \\ n_{\overline{r}}(k), & \text{when } r \in \Phi_L, \end{cases} \tag{92}$$

and then

$$\overline{h_r(k)} = \begin{cases} h_{\overline{r}}(k^2), & \text{when } r \in \Phi_S, \\ h_{\overline{r}}(k), & \text{when } r \in \Phi_L. \end{cases} \tag{93}$$

Using Lemma 5.1 the case $k = 1$ of (92) can easily be generalized to

$$\overline{n(w)} = n(\overline{w}), \qquad \text{for all } w \in W. \tag{94}$$

Proposition 6.6. *Let e be an isotropic element of $\mathbf{W}$. Let g be an automorphism of $\mathbf{W}$ (and $\mathbf{F}$). Then*

$$Q(e^g) = Q(e)^{\overline{g}}. \tag{95}$$

Proof. Let $a, b \in \mathbf{W}$, $g \in \widehat{\mathsf{F}_4}(K)$. We have

$$\begin{aligned} [a^{\overline{g}}, Q(e)^{\overline{g}}] \cdot b^{\overline{g}} &= [a, Q(e)] \cdot b \\ &= [e, \mu(a)] \cdot [e, \mu(b)] \\ &= [e^g, \mu(a)^g] \cdot [e^g, \mu(b)^g] \\ &= [e^g, \mu(a^{\overline{g}})] \cdot [e^g, \mu(b^{\overline{g}})] = [a^{\overline{g}}, Q(e^g)] \cdot b^{\overline{g}} \end{aligned}$$

□

7 Octagonality

From now on we shall require that K has a *Tits* automorphism, i.e., a field automorphism σ with the property

$$(k^\sigma)^\sigma = k^2, \quad \text{for every } k \in K. \tag{96}$$

Note that we require σ to have an inverse σ^{-1}, and consequently, that K is a perfect field.

For finite fields a Tits automorphism exists if and only if the order of the field is an odd power 2^{2m+1} of 2, and then $k^\sigma \stackrel{\text{def}}{=} k^{2^{m+1}}$.

As with the Frobenius automorphism, we shall extend σ to $\mathbf{F}$ by setting $e_r^\sigma \stackrel{\text{def}}{=} e_r$, $h_r^\sigma \stackrel{\text{def}}{=} h_r$, for $r \in \Phi$. We have $A^{\text{frob}} = (A^\sigma)^\sigma$ for every $A \in \mathbf{F}$. As before, when g is an automorphism of $\mathbf{W}$, we shall write $g^\sigma = \sigma^{-1} g \sigma$. Note that $\overline{\sigma} = \sigma$ and hence $\overline{g^\sigma} = \overline{g}^\sigma$. By Lemma 4.5 we have $x(e)^\sigma = x(e^\sigma)$ and $x(E)^\sigma = x(E^\sigma)$, when e (resp. E) is isotropic (resp. totally isotropic).

For $a, e \in \mathbf{W}$ such that e is isotropic we define the operator $q(\cdot,\cdot)$ as follows :

$$q(a,e) \stackrel{\text{def}}{=} [a, Q(e^{\sigma^{-1}})]. \tag{97}$$

We write $q(\mathbf{W}, e)$ for the set of elements $q(a,e)$ with $a \in \mathbf{W}$, i.e., $q(\mathbf{W},e) = [\mathbf{W}, Q(e^{\sigma^{-1}})]$.

(This is the last operator which we shall need to introduce in this text. It serves an important purpose in the definition of the Ree group and the Ree-Tits octagon and, as shall be shown in Theorem 10.1, also has an interesting geometric interpretation.)

An element $e \in \mathbf{W}$ shall be called *semi-octagonal* if and only if it is isotropic and $q(e,e) = 0$. An element $e \in \mathbf{W}$ shall be called *octagonal* if and only if it is isotropic and $e \in q(\mathbf{W}, e)$, or equivalently (by Lemma 4.8) if and only if $q([e,\mathbf{W}], e) = 0$. All octagonal elements are semi-octagonal, but not conversely.

Proposition 7.1. *Let e, f be isotropic elements of $\mathbf{W}$. Then*

1. $Q(q(e,f)^{\sigma^{-1}}) = f * q(f,e)$,
2. $q(e,f) = 0$ *if and only if* $q(f,e) = 0$,
3. $e \in q(\mathbf{W}, f)$ *if and only if* $f \in q(\mathbf{W}, e)$,
4. $q(e,e)$ *is octagonal and* $e \in q(\mathbf{W}, q(e,e))$.

Proof. Statements 1–3 are immediate consequences of Proposition 6.4 with $f^{\sigma^{-1}}$ substituted for f. Setting $f = e$ in the first statement of this proposition, we see that $\{e, q(e,e)\}$ is a defining pair for $Q(q(e,e)^{\sigma^{-1}})$. Hence both e and $q(e,e)$ belong to $[\mathbf{W}, Q(q(e,e)^{\sigma^{-1}})]$ and therefore $q(e,e)$ is octagonal. □

To study the (semi-)octagonal elements of $\mathbf{W}$, we first look at elements of the form e_r with $r \in \mathbf{W}$. Table 1 lists the 24 short roots $r \in \Phi_S$, the corresponding long roots $\overline{r}$ and the value of $r + \overline{r}$ when it belongs to Φ_S.

$\langle r, \overline{r}\rangle = 1$

r	$\overline{r}$
0100	0110
$0\bar{1}00$	$0\bar{1}\bar{1}0$
0001	1001
$000\bar{1}$	$\bar{1}00\bar{1}$
++++	0101
+−−+	$0\bar{1}01$
−++−	$010\bar{1}$
−−−−	$0\bar{1}0\bar{1}$

$\langle r, \overline{r}\rangle = 0$

r	$\overline{r}$
+++−	$\bar{1}100$
++−+	0011
+−++	$00\bar{1}1$
+−−−	$\bar{1}\bar{1}00$
−+++	1100
−+−−	$001\bar{1}$
−−+−	$00\bar{1}\bar{1}$
−−−+	$1\bar{1}00$

$\langle r, \overline{r}\rangle = -1$

r	$\overline{r}$	$r + \overline{r}$
1000	$\bar{1}001$	0001
$\bar{1}000$	$100\bar{1}$	$000\bar{1}$
0010	$01\bar{1}0$	0100
$00\bar{1}0$	$0\bar{1}10$	$0\bar{1}00$
++−−	$\bar{1}010$	−++−
+−+−	$\bar{1}0\bar{1}0$	−−−−
−+−+	1010	++++
−−++	$10\bar{1}0$	+−−+

Table 1: Short roots $r \in \Phi_S$ and the corresponding long roots $\overline{r}$.

The leftmost list corresponds to octagonal elements e_r, the middle list corresponds to those e_r that are semi-octagonal but not octagonal and for the rightmost list elements are not even semi-octagonal. In the last case the third column corresponds to the value of $q(e_r, e_r)$.

Lemma 7.2. *Consider Weyl group elements w, w' of the form $w = w_r w_{\overline{r}}$, where $r \in \Phi_S$ such that $\langle r, \overline{r}\rangle = 0$, and $w' = (w_s w_{\overline{s}})^2$ with $s \in \Phi_S$ such that $\langle s, \overline{s}\rangle = 1$.*

Then $\overline{w} = w$ and $\overline{w}' = w'$. The elements of this form generate a subgroup W' of W which is isomorphic to the dihedral group of order 16. This group has 3 orbits on Φ_S which correspond to the three lists of 8 roots given in Table 1.

Proof. When $\langle r, \overline{r}\rangle = 0$ the Weyl group elements w_r and $w_{\overline{r}}$ commute. Then $\overline{w} = \overline{w_r w_{\overline{r}}} = \overline{w_r}\,\overline{w_{\overline{r}}} = w_{\overline{r}} w_r = w$.

When $\langle s, \overline{s}\rangle = 1$ we have $w' = (w_s w_{\overline{s}})^2 = w_{\overline{s}}^{w_s} w_{\overline{s}} = w_{w_s(\overline{s})} w_{\overline{s}} = w_{s-\overline{s}} w_{\overline{s}}$. Also $\overline{s - \overline{s}} = \overline{s} - 2s$, hence $\overline{w'} = w_{\overline{s}-2s} w_s = w_s w_s w_{\overline{s}-2s} w_s = w_s w^{w_s}_{\overline{s}-2s} = w_s w_{\overline{s}-2s+s} = w_s w_{\overline{s}-s}$. But also $w' = (w_s w_{\overline{s}})^2 = w_s w_s^{w_{\overline{s}}} = w_s w_{s-\overline{s}}$.

For example, with $s =$ 0100 we obtain w' which maps roots (x, y, z, t) to $(x, -y, -z, t)$. With $r =$ +++− we obtain w, which maps (x, y, z, t) to $\frac{1}{2}(-x + y - z + t, x - y - z + t, -x - y + z + t, x + y + z + t)$. We leave it to the reader to compute W' and prove that it has the stated properties. □

8 The Ree group ${}^2\widehat{\mathsf{F}_4}(K)$

Proposition 8.1. *Let g be an automorphism of $\mathbf{W}$ satisfying $\overline{g} = g^\sigma$, i.e.,*

$$(a^g)^\sigma = (a^\sigma)^{\overline{g}}, \qquad \textit{for all } a \in \mathbf{W}. \tag{98}$$

Let e, f be isotropic elements of $\mathbf{W}$. Then

1. $q(e^g, f^g) = q(e, f)^g$.
2. $e \in q(\mathbf{W}, f)$ *if and only if* $e^g \in q(\mathbf{W}, f^g)$.

Therefore g maps octagonal elements of $\mathbf{W}$ onto octagonal elements of $\mathbf{W}$ and semi-octagonal elements onto semi-octagonal elements.

Proof. We have $q(e^g, f^g) = [e^g, Q(f^g)^{\sigma^{-1}}] = [e^g, Q(f)^{\overline{g}\sigma^{-1}}] = [e^g, Q(f)^{\sigma^{-1}g}] = q(e, f)^g$ by (95), proving the first statement. Also, if $e \in q(\mathbf{W}, f)$, then $e^g \in q(\mathbf{W}, f)^g = q(\mathbf{W}, f^g)$, by the above. $\square$

A first example of an element g that satisfies (98) is given by n_w where w is a Weyl group element such that $w = \overline{w}$. Examples of such w are provided by Lemma 7.2.

Also, when $h \in \widehat{\mathsf{F}_4}(K)$ is such that h^σ and $\overline{h}$ commute, then $g = h\overline{h^{\sigma^{-1}}}$ satisfies (98). Indeed

$$\overline{g} = \overline{h\overline{h^{\sigma^{-1}}}} = \overline{h}\,\overline{\overline{h^{\sigma^{-1}}}} = \overline{h}h^\sigma = h^\sigma\overline{h} = (h\overline{h^{\sigma^{-1}}})^\sigma = g^\sigma$$

For example, consider $h = h_r(k)$ as in (70), with $r \in \Phi_S$. Then $\overline{h} = h_{\overline{r}}(k^2)$ commutes with h^σ, and therefore $i_r(k) \stackrel{\text{def}}{=} h_r(k)h_{\overline{r}}(k^\sigma)$ satisfies (98). We have

$$\begin{aligned} e_s^{i_r(k)} &= k^{\langle r,s\rangle}k^{\langle \overline{r},s\rangle\sigma}e_s, & \\ h_s^{i_r(k)} &= h_s, & \text{for all } s \in \Phi. \end{aligned} \tag{99}$$

If $q(e, e) = 0$ then $x(e^\sigma)$ and $x(Q(e))$ commute and $y(e) = x(Q(e^{\sigma^{-1}}))x(e) = x(e)x(Q(e^{\sigma^{-1}}))$ satisfies (98). This is generalized in the following

Lemma 8.2. *Let e be an isotropic element of $\mathbf{W}$. Define*

$$y(e) \stackrel{\text{def}}{=} x(Q(e^{\sigma^{-1}}))x(e)x(q(e, e)). \tag{100}$$

Then $y(e)$ satisfies (98).

Proof. Write $f = q(e, e)$. Note that $Q(f) = e^\sigma * f^\sigma$ by Proposition 7.1–1. Then

$$\begin{aligned} x(e^\sigma)x(Q(e)) &= x(Q(e))x(e^\sigma)x([e^\sigma, Q(e)])x(e^\sigma * [e^\sigma, Q(e)]) \\ &= x(Q(e))x(e^\sigma)x(f^\sigma)x(Q(f)). \end{aligned} \tag{101}$$

We have

$$\begin{aligned}
\overline{y(e)} &= \overline{x(Q(e^{\sigma^{-1}}))}\,\overline{x(e)}\,\overline{x(f)} \\
&= x(e^{\sigma})x(Q(e))x(Q(f)), && \text{by (89)} \\
&= x(Q(e))x(e^{\sigma})x(f^{\sigma})x(Q(f))x(Q(f)), && \text{by (101)} \\
&= x(Q(e))x(e^{\sigma})x(f^{\sigma}) = y(e)^{\sigma}.
\end{aligned}$$

□

The group generated by all elements n_w with $\overline{w} = w$, all elements $i_r(k)$ and all elements $y(e)$ as defined in (100) shall be denoted by ${}^2\widehat{\mathsf{F}_4}(K)$. This is a so-called *Ree* group or *twisted Chevalley group* of type ${}^2\mathsf{F}_4$. For further details we refer to [2].

For $k \in K$, $r \in \Phi_S$, we write

$$y_r(k) \stackrel{\text{def}}{=} y(ke_r) = \begin{cases} x_{\overline{r}}(k^{\sigma})x_r(k), & \text{when } \langle r, \overline{r}\rangle \neq -1, \\ x_{\overline{r}}(k^{\sigma})x_r(k)x_{r+\overline{r}}(k^{1+\sigma}), & \text{when } \langle r, \overline{r}\rangle = -1. \end{cases} \tag{102}$$

It can be proved that the set of all elements $y_r(k)$ with $k \in K$, $r \in \Phi_S$ generate the group ${}^2\widehat{\mathsf{F}_4}(K)$. This is often taken as the definition of ${}^2\widehat{\mathsf{F}_4}(K)$ (cf. [2]).

Lemma 8.3. *Let e be an isotropic element of $\mathbf{W}$, let $k_1, k_2 \in K$. Then*

$$y(k_1 e)y(k_2 e) = y(k_1^{\sigma} k_2\, q(e,e))y((k_1 + k_2)e). \tag{103}$$

In particular $y(e)^2 = 1$ when e is semi-octagonal, and $y(e)^4 = 1$ when it is not.

Proof. Write $d = q(e,e)$, $D = Q(d^{\sigma^{-1}})$ and $E = Q(e^{\sigma^{-1}})$. Note that $[d,e] = [d,E] = 0$ and $[e,E] = d$. By Proposition 7.1–1 d is octagonal, implying $[d,D] = 0$, and $d * e = D$, implying $[e,D] = 0$. From (65) it then follows that $x(d)$ commutes with $x(e)$, $x(D)$ and $x(E)$, while $x(D)$ commutes with $x(d)$ and $x(e)$. From this we obtain

$$\begin{aligned}
y(k_1 e)y(k_2 e) &= x(k_1^{\sigma}E)x(k_1 e)x(k_1 k_1^{\sigma} d)x(k_2^{\sigma}E)x(k_2 e)x(k_2 k_2^{\sigma} d) \\
&= x(k_1^{\sigma}E)x(k_1 e)x(k_2^{\sigma}E)x(k_2 e)x(k_1 k_1^{\sigma} d)x(k_2 k_2^{\sigma} d).
\end{aligned}$$

Also by (65) we find $x(e)x(E) = x(E)x(e)x([e,E])x(e * [e,E])$ which is equal to $x(E)x(e)x(d)x(D)$. Hence $x(k_1 e)x(k_2^{\sigma}E) = x(k_2^{\sigma}E)x(k_1 e)x(k_1 k_2^{\sigma} d)x(k_1^2 k_2^{\sigma} D)$.

This yields

$$
\begin{aligned}
&y(k_1e)y(k_2e)\\
&= x(k_1^\sigma E)x(k_2^\sigma E)x(k_1e)x(k_1k_2^\sigma d)x(k_1^2k_2^\sigma D)x(k_2e)x(k_1k_1^\sigma d)x(k_2k_2^\sigma d)\\
&= x((k_1+k_2)^\sigma E)x((k_1+k_2)e)x(k_1k_2^\sigma + k_1k_1^\sigma d + k_2k_2^\sigma d)x(k_1^2k_2^\sigma D)\\
&= y((k_1+k_2)e)x((k_1+k_2)^{1+\sigma} + k_1k_2^\sigma + k_1k_1^\sigma d + k_2k_2^\sigma d)x(k_1^2k_2^\sigma D)\\
&= y((k_1+k_2)e)x(k_2k_1^\sigma d)x(k_1^2k_2^\sigma D)\\
&= y((k_1+k_2)e)y(k_2k_1^\sigma d).
\end{aligned}
$$

where we use Propositions 4.3 and 4.4 to simplify the products of the form $x(ke)x(k'e)$ and $x(kE)x(k'E)$. $\square$

Lemma 8.4. *Let e be an octagonal (isotropic) element of* $\mathbf{W}$. *Let $g \in {}^2\widehat{\mathsf{F}_4}(K)$. Then*

$$y(e^g) = y(e)^g. \tag{104}$$

Proof. We have $y(e^g) = x(Q(e^g)^{\sigma^{-1}})x(e^g)$. Using (98) we obtain $Q(e^g)^{\sigma^{-1}} = Q(e)^{\overline{g}\sigma^{-1}} = Q(e)^{\sigma^{-1}g}$ and hence $y(e^g) = x(Q(e^{\sigma^{-1}})^g)x(e^g)$, which by (64) is equal to $x(Q(e^{\sigma^{-1}}))^g x(e)^g = y(e)^g$. $\square$

The remainder of this section shall be devoted to show transitivity of ${}^2\widehat{\mathsf{F}_4}(K)$ on the set of octagonal elements of $\mathbf{W}$. In order to do this we shall make use of a 'reduction process' which can be summarized as follows : we choose an element of the form $y_t(k)$ to map a given $a \in \mathbf{W}$ onto $a' \in \mathbf{W}$ with the property that a' has one more coordinate that is 'known' to be zero. Successive reductions reduce the number of non-zero coordinates of a while staying in the same orbit of $\widehat{\mathsf{F}_4}(K)$, until the final image of a has a sufficiently simple structure.

In general, when it is known that $a[r] \neq 0$ for some $r \in \Phi_S$, we choose $s \in \Phi_S$ such that $\langle r, s\rangle = 1$ and then apply an element of the form $y_t(k)$ with $t = s - r$. From (63) and (102) it follows that the coordinate of the image a' of a at position s is now equal to $a'[s] = a[s] + ka[r]$. Hence, setting $k = a[s]/a[r]$ we obtain $a'[s] = 0$. When $\langle r, s\rangle = 0$ we obtain a similar effect when we use t such that $\overline{t} = s - r$. Now $a'[s] = a[s] + k^\sigma a[r]$ and we take $k = (a[s]/a[r])^{\sigma^{-1}}$.

This transformation does however also affect other coordinates $a[s]$ of a, and when performing successive reductions, we should make sure that earlier annihilations are not 'undone' by later actions.

Consider how the action of $x_r(k)$ affects coordinates $a[s]$ of a, with $s \in \Phi_S$. (Coordinates $a[3]$ and $a[4]$ will turn out not to be important). It follows from (63) that only those coordinates are affected that satisfy $\langle r, s\rangle > 0$. As a consequence, the action of $y_r(k)$ will only affect those coordinates for which $\langle r, s\rangle > 0$

or $\langle \overline{r}, s\rangle > 0$. (When $\langle r, \overline{r}\rangle = -1$, the extra condition $\langle r + \overline{r}, s\rangle$ does not impose an additional constraint.)

Let us now use these techniques to determine the various orbits of ${}^2\widehat{\mathsf{F}_4}(K)$ on the (non-zero) octagonal elements of $\mathbf{W}$. Consider such an element a. Because a is isotropic (and $\neq 0$) it cannot belong to $\mathbf{I}$ (cf. the proof of Theorem 4.7). Hence a has at least one non-zero coordinate $a[r]$ for $r \in \Phi_S$.

We shall first prove that we can choose $r = \mathtt{0100}$. Indeed, if a has coordinate $a[r] \neq 0$ for some r such that $\langle r, \overline{r}\rangle = 1$, then by Lemma 7.2 there exists $w \in \mathbf{W}$ such that $n(w) \in {}^2\widehat{\mathsf{F}_4}(K)$ and $r^w = \mathtt{0100}$. In that case $a' = a^{n(w)}$ is in the same orbit as a and $a'[\mathtt{0100}] \neq 0$. We may therefore use a' instead of a in the subsequent argument.

If $a[r] = 0$ for all r satisfying $\langle r, \overline{r}\rangle = 1$ but a has a non-zero coordinate for some $r \in \Phi_S$ with $\langle r, \overline{r}\rangle = 0$, then we proceed as follows. Using the same argument as above we may assume that $r = \mathtt{-+--}$. We then apply $y_{\mathtt{++++}}(1)$ to create a non-zero coordinate at position $\mathtt{0100}$. A similar technique can be used when $\langle r, \overline{r}\rangle = -1$, taking $r = \mathtt{++--}$ and applying $y_{\mathtt{-+++}}(1)$.

Henceforth we may assume that $a[r] = a[\mathtt{0100}] \neq 0$. We proceed in several stages.

Stage 1. Apply the reductions $y_t(k)$ in the top part of Table 2 (and in the order given) to annihilate subsequent coordinates $a[s]$.

s	$\overline{t}$	t	$t+\overline{t}$
++++	$00\bar{1}1$	+-++	
-+++	$10\bar{1}0$	--++	+--+
+++-	$\bar{1}0\bar{1}0$	+-+-	----
-++-	$00\bar{1}\bar{1}$	--+-	
++-+	$0\bar{1}01$	+--+	
-+-+	$1\bar{1}00$	---+	
++--	$\bar{1}\bar{1}00$	+---	
-+--	$0\bar{1}0\bar{1}$	----	
0010	$0\bar{1}10$	$00\bar{1}0$	$0\bar{1}00$
$00\bar{1}0$	$0\bar{1}\bar{1}0$	$0\bar{1}00$	

Table 2: Reductions used in the proof of Theorem 8.5.

We have listed s, t, $\overline{t}$ and $t + \overline{t}$ (when it belongs to Φ_S) for each step. The table is ordered in such a way that each group element $y_t(k)$ does not affect coordinates $a[s]$ for each s from the previous rows. We leave it to the reader to verify that indeed $\langle s, t\rangle \leq 0$ and $\langle s, \overline{t}\rangle \leq 0$ in each of these cases. Note that

$\langle r, s\rangle = 1$ and $t = s - r$ for all entries (in the top part), hence $\langle r, t\rangle = -1$ and therefore $y_t(k)$ never 'accidentally' makes $a[r]$ zero.

After applying the stated reductions, we end up with an image b of a for which $b[r] \neq 0$ and $b[s] = 0$ whenever $\langle r, s\rangle = 1$. Now, for such s we compute $b^2[s]$ using (14). Note that $u + v = s$ implies $\langle r, u\rangle + \langle r, v\rangle = 1$ and hence either $u = r$, $v = r$ or $\langle r, u\rangle = 1$ (and then $b[u] = 0$) or $\langle r, v\rangle = 1$ (and $b[v] = 0$). Hence $b^2[s] = b[r]b[s - r]$, and because b is isotropic and $b[r] \neq 0$, we find that $b[s-r] = 0$. When s runs through all short roots that satisfy $\langle r, s\rangle = 1$, $v = s-r$ runs through all short roots that satisfy $\langle r, v\rangle = -1$.

In other words, the eight remaining roots s for which a coordinate is (possibly) non-zero, are the 8 short roots s satisfying $\langle r, s\rangle = -2, 0$ or 2, which for our particular choice of r, correspond to the roots with integral coordinates, i.e., the cyclic permutations of 1000 and $\bar{1}$000.

Note that roots s, s' of this form satisfy $\langle s, s'\rangle = -2$, 0 or 2 and that also $\langle r_3, s\rangle = -2$, 0 or 2 for each s. This implies $b[4] = 0$, whenever $b^2 = 0$.

As a final step we may use the reductions in the bottom part of Table 2 to annihilate two more coordinates. (Note that this time $s = r + \bar{t}$).

Stage 2. Let e denote the image of a obtained at this point. We now have to consider several cases.

1. Assume that $e[0001] \neq 0$. Then we use the transformations on the right, this time with $r = 0001$ and $s = \bar{r} + t$, to end up with a an element of the following form :

s	$\bar{t}$	t	$\bar{t}+t$
1000	100$\bar{1}$	$\bar{1}$000	000$\bar{1}$
$\bar{1}$000	$\bar{1}$00$\bar{1}$	000$\bar{1}$	

$$e = k_1 e_{0100} + k_2 e_{0\bar{1}00} + k_3 e_{0001} + k_4 e_{000\bar{1}} + k_5 h_{r_3},$$

with $k_1 \neq 0$ and $k_3 \neq 0$. We easily compute that $e^2 = 0$ implies $k_5^2 = k_1 k_2 + k_3 k_4$.

We shall now express the fact that e is semi-octagonal in terms of $k_1, \ldots, k_5$. We use (85) to compute

$$\begin{aligned} Q(e) = k_1^2 e_{0110} &+ k_2^2 e_{0\bar{1}\bar{1}0} + k_3^2 e_{1001} + k_4^2 e_{\bar{1}00\bar{1}} \\ &+ k_1 k_3 e_{++++} + k_1 k_4 e_{-++-} + k_2 k_3 e_{+--+} + k_2 k_4 e_{----} \\ &+ k_5^2 h_{r_2} + k_1 k_2 h_{r_3} + k_3 k_4 h_{r_4}\,. \end{aligned}$$

When e is semi-octagonal, we have $q(e, e) = 0$ and hence

$$0 = q(e, e)^\sigma \cdot e_{0010} = [e^\sigma, Q(e)] \cdot e_{0010},$$

which by (20) is equal to

$$k_1^\sigma Q(e)[0\bar{1}\bar{1}0] + k_2^\sigma Q(e)[01\bar{1}0] + k_3^\sigma Q(e)[00\bar{1}\bar{1}] + k_4^\sigma Q(e)[00\bar{1}1] = k_1^\sigma k_2^2.$$

Because $k_1 \neq 0$, this implies $k_2 = 0$.

Similarly, computing $q(e,e)^\sigma \cdot e_{1000}$ we find $k_3^\sigma k_4^2 = 0$ and hence $k_4 = 0$, so $e = k_1 e_{0100} + k_3 e_{0001}$.

Finally, because e is octagonal, we must have $q([e, e_{+---}], e) = 0$. This expression evaluates to $k_3 k_1^\sigma e_{++++} = 0$, contradicting the assumptions $k_1 = e[0100] \neq 0$ and $k_3 = e[0001] \neq 0$.

2. Assume $e[000\bar{1}] \neq 0$. By Lemma 7.2 the group W' contains an element w that interchanges 0001 and $000\bar{1}$ and leaves 0100 invariant. The corresponding element $n(w) \in {}^2\widehat{\mathsf{F}_4}(K)$ then reduces this to the previous case.
3. If $e[0001] = e[000\bar{1}] = 0$ we have an element of the form

$$e = k_1 e_{0100} + k_2 e_{0\bar{1}00} + k_3 e_{1000} + k_4 e_{\bar{1}000} + k_5 h_{r_3},$$

with $k_1 \neq 0$. Again we easily compute that $e^2 = 0$ implies $k_5^2 = k_1 k_2 + k_3 k_4$. Now

$$\begin{aligned} Q(e) = k_1^2 e_{0110} &+ k_2^2 e_{0\bar{1}\bar{1}0} + k_3^2 e_{\bar{1}001} + k_4^2 e_{100\bar{1}} \\ &+ k_1 k_3 e_{-+++} + k_1 k_4 e_{+++-} + k_2 k_3 e_{---+} + k_2 k_4 e_{+---} \\ &+ k_5^2 h_{r_2} + k_5^2 h_{r_3} + k_3 k_4 h_{r_4} \end{aligned}$$

and

$$q(e,e)^\sigma \cdot e_{0010} = k_1^\sigma k_2^2\,, \qquad q(e,e)^\sigma \cdot e_{000\bar{1}} = k_3^\sigma k_3^2\,,$$
$$q(e,e)^\sigma \cdot e_{0001} = k_4^\sigma k_4^2\,.$$

As e is semi-octagonal, this implies $e = k_1 e_{0100}$ (and hence e is octagonal).

Stage 3. Finally we apply an element $i_s(k)$ with the property $\langle 0100, s\rangle = -1$ and $\langle 0100, \overline{s}\rangle = 0$. By (99) this element is mapped onto e_{0100}. We may take $s = $ --+- (and $\overline{s} = 00\bar{1}\bar{1}$).

This proves that every non-zero octagonal element can be mapped onto the same element by a suitable transformation of the group ${}^2\widehat{\mathsf{F}_4}(K)$.This completes the proof of

Theorem 8.5. *The group ${}^2\widehat{\mathsf{F}_4}(K)$ acts transitively on the non-zero octagonal elements of $\mathbf{W}$.*

9 Pairs of octagonal elements

The proof technique of Theorem 8.5 can be extended to provide more information on the orbit structure of ${}^2\widehat{\mathsf{F}_4}(K)$ on pairs of (non-zero) octagonal elements. The main result shall be stated in Theorem 9.5. The proof of that theorem is subdivided into several lemmas.

Lemma 9.1. *Let $k \in K - \{0\}$. Then ${}^2\widehat{\mathsf{F}_4}(K)$ acts transitively on all ordered pairs (e, f) of octagonal (isotropic) elements $e, f \in \mathbf{W}$ such that $e \cdot f = k$.*

Proof. By Theorem 8.5 we may choose $e = e_{\mathtt{0\bar{1}00}}$ without loss of generality. From $e \cdot f \neq 0$ it follows that the coordinate $f[\mathtt{0100}]$ is non-zero. We shall again apply the reductions $y_t(k)$ of Table 2 to map f onto an element of the form $k_1 e_{\mathtt{0100}}$. This time however, we are only allowed to use group elements $y_t(k)$ that leave e invariant.

From (102) it follows that $y_t(k)$ leaves e_r invariant (with $r \in \Phi_S$) whenever $\langle r, t\rangle \geq 0$ and $\langle r, \bar{t}\rangle \geq 0$. It is now easily verified that all reductions $y_t(k)$ from the table indeed satisfy $\langle t, \mathtt{0\bar{1}00}\rangle \geq 0$ and $\langle \bar{t}, \mathtt{0\bar{1}00}\rangle \geq 0$. Note that the argument of Stage 2 of Theorem 8.5 can again be applied to prove that the octagonality of f implies that Stage 1 will result in an element of the form $k_1 e_{\mathtt{0100}}$. (The reductions used in that argument need not leave e untouched.)

Finally, because the group ${}^2\widehat{\mathsf{F}_4}(K)$ leaves the dot product invariant, we must have $k_1 = k$. (Note that Stage 3 of Theorem 8.5 cannot be applied in this case.) $\square$

Lemma 9.2. *The group ${}^2\widehat{\mathsf{F}_4}(K)$ acts transitively on all ordered pairs (e, f) of octagonal (isotropic) elements $e, f \in \mathbf{W}$ such that $e \cdot f = 0$ and $[e, f] \neq 0$.*

Proof. This proof will use the same techniques as the proofs of Theorem 8.5 and Lemma 9.1. By Theorem 8.5 we may choose $e = e_{\text{\_\_\_\_}}$ without loss of generality.

The main argument of this proof will rely on the assumption that the coordinate $f[\mathtt{0100}]$ is non-zero. We shall therefore first prove that it is always possible to map f onto f' such that $f'[\mathtt{0100}] = 0$, by means of element of ${}^2\widehat{\mathsf{F}_4}(K)$ that leaves $e = e_{\text{\_\_\_\_}}$ invariant.

From $[e,f] = [e_{\texttt{----}}, f] \neq 0$ we easily deduce that either $f[3] + f[4] \neq 0$ or at least one of the coordinates $f[r] \neq 0$, with $r =$ 1000, 0100, 0010, 0001, -+++, +-++, ++-+ or +++-.

Below we give a list of group elements $y_t(1)$ which can be used to map f onto f' with the required properties, whenever $f[r] \neq 0$ for one of the given r

r	s	$\bar{t}$	t	$\bar{t}+t$
1000	+++-	$010\bar{1}$	-++-	
0100		none needed		
0010	+++-	$\bar{1}010$	++--	-++-
0001	0100	$010\bar{1}$	-++-	
-+++	0100	$\bar{1}010$	++--	-++-
+-++	+++-	$010\bar{1}$	-++-	
++-+	0100	$010\bar{1}$	-++-	
+++-	0100	$10\bar{1}0$	--++	

For example, the last row of the table tells us that, when $f[\texttt{0100}] = 0$ but $f[\texttt{+++-}] \neq 0$, we may apply $y_{\texttt{--++}}(1)$ to create a non-zero coordinate at position 0100. Similarly, the first row of the table indicates that when $f[\texttt{0100}] = f[\texttt{+++-}] = 0$ but $f[\texttt{1000}] \neq 0$ we may first apply $y_{\texttt{-++-}}(1)$ to obtain an image with a non-zero coordinate at position +++-, bringing us back to the first example. We leave it to the reader to verify that each of the listed group elements indeed leaves $e = e_{\texttt{----}}$ invariant.

Only the case where $f[r] = 0$ for each r in the table, remains. It turns out that now $f^2 = 0$ implies $f[3] + f[4] = 0$, contradicting $[e,f] = 0$. Indeed, we may use (14) to compute $f^2[3] + f^2[4] = f[3]^2 + f[4]^2 = (f[3] + f[4])^2$.

Henceforth we may assume that the coordinate $f[\texttt{0100}]$ is non-zero. We can map f onto $ke_{\texttt{0100}}$ using the reductions $y_t(k)$ of Table 2, *except* the first reduction (which corresponds to $s =$ ++++). We have no need for this reduction because $e \cdot f = e_{\texttt{----}} \cdot f = 0$ implies $f[\texttt{++++}] = 0$. It is easily verified that all other reductions in the table satisfy $\langle t, \texttt{----}\rangle \geq 0$ and $\langle \bar{t}, \texttt{----}\rangle \geq 0$ and therefore leave $e_{\texttt{----}}$ invariant.

This proves that f can be mapped onto an element of the form $ke_{\texttt{0100}}$. We get rid of the scalar k by applying $i_s(k)$ with $s =$ --++ and $\bar{s} = 10\bar{1}0$. Indeed

$$\langle \texttt{--++}, \texttt{0100}\rangle = -1, \quad \langle 10\bar{1}0, \texttt{0100}\rangle = 0, \quad \langle \texttt{--++}, \texttt{----}\rangle = 0, \quad \langle 10\bar{1}0, \texttt{----}\rangle = 0,$$

and hence $i_s(k)$ leaves e invariant and maps $ke_{\texttt{0100}}$ to $e_{\texttt{0100}}$. □

Lemma 9.3. *The group* $^2\widehat{\mathsf{F}_4}(K)$ *acts transitively on all ordered pairs* (e, f) *of octagonal (isotropic) elements* $e, f \in \mathbf{W}$ *such that* $e \cdot f = 0$*,* $[e, f] = 0$ *and* $e * f \neq 0$.

Proof. This proof will use the same techniques as the proofs of the previous lemmas in this section. Without loss of generality we may choose $e = e_{\mathtt{0001}}$.

As in the proof of Lemma 9.1 we first need to prove that f can be mapped onto an element f' for which $f'[\mathtt{0100}]$ is non-zero, using only group elements from $^2\widehat{\mathsf{F}_4}(K)$ that leave e invariant. (In this case this turns out to be slightly more complicated.)

Because $[e, f] = 0$ and $e \cdot f = 0$, we must have $f[r] = 0$ whenever $\langle \mathtt{0001}, r\rangle = -1$ or -2. By Lemma 3.1–4, f cannot belong to the set

$$S \stackrel{\text{def}}{=} \{[a, e] \mid a \in \mathbf{W}, a \cdot e = 0\}.$$

It is easily computed that S is precisely the subspace of $\mathbf{W}$ generated by all e_r such that $\langle \mathtt{0001}, r\rangle = 2$ or 1. Therefore either $f[3] \neq 0$, $f[4] \neq 0$ or $f[r] = 0$ for one of the 6 short roots r for which $\langle \mathtt{0001}, r\rangle = 0$. We list these roots below, together with the group elements $y_t(1)$ which can be used to map f onto f' with the required properties, whenever $f[r] \neq 0$. These group elements leave e invariant. We refer to the proof of Lemma 9.2 for an explanation of how this table should be interpreted.

r	s	$\bar{t}$	t	$\bar{t}+t$
0100		none needed		
0010	0100	$01\bar{1}0$	0010	0100
$\bar{1}000$	0100	1100	−+++	
$0\bar{1}00$	0100	0110	0100	
$00\bar{1}0$	0100	0110	0100	

Unfortunately, the short root $r = \mathtt{1000}$ cannot be handled in the same way. This means that we still need to treat the situation where $f[r] = 0$ for all roots r listed above, while $f[\mathtt{1000}] \neq 0$, $f[3] \neq 0$ or $f[4] \neq 0$.

Expanding $f^2[3] = f^2[4] = 0$ for this case results in $f[3] = f[4] = 0$. Likewise, $f^2[\mathtt{++++}] = f^2[\mathtt{++-+}] = f^2[\mathtt{+-++}] = f^2[\mathtt{+--+}] = 0$ lead to $f[\mathtt{-+++}] = f[\mathtt{-+++}] = f[\mathtt{-+++}] = f[\mathtt{-+++}] = 0$, and hence f must be of the following form

$$f = k_1 e_{\mathtt{1000}} + k_2 e_{\mathtt{0001}} + k_3 e_{\mathtt{++++}} + k_4 e_{\mathtt{+--+}} + k_5 e_{\mathtt{+-++}} + k_6 e_{\mathtt{++-+}}.$$

We may compute

$$Q(f) = k_1^2 e_{\bar{1}001} + k_2^2 e_{1001} + k_3^2 e_{0101} + k_4^2 e_{0\bar{1}01} + k_5^2 e_{00\bar{1}1} + k_6^2 e_{0011} + (k_1 k_2 + k_3 k_4 + k_5 k_6) e_{0001},$$

and from this result we obtain that $q(f,f)^\sigma = [f^\sigma, Q(f)]$ is equal to $k_1^\sigma k_1^2 e_{0001}$. Hence f cannot be (semi-)octagonal when $k_1 = f[1000] \neq 0$.

Henceforth we may assume that $f[0100] \neq 0$. We again apply the reductions of Table 2 except those that do not leave e invariant, i.e., those corresponding to $s =$ +++-, -++-, ++-- and -+--. Luckily, the corresponding coordinates $f[s]$ are already zero, because $[e,f] = 0$.

This proves that we can map f onto an element of the form e_{0100}. We use $i_s(k^{\sigma^{-1}})$ with $s = 00\bar{1}0$ and $\bar{s} = 0\bar{1}10$ to get rid of the final scalar k. Indeed,

$$\langle 00\bar{1}0, 0100\rangle = 0, \quad \langle 0\bar{1}10, 0100\rangle = -1, \quad \langle 00\bar{1}0, 0001\rangle = 0, \quad \langle 0\bar{1}10, 0001\rangle = 0,$$

and hence $i_s(k^{\sigma^{-1}})$ leaves e invariant and maps ke_{0100} onto e_{0100}. ☐

Lemma 9.4. *The group* $^2\widehat{\mathsf{F}_4}(K)$ *acts transitively on all ordered pairs* (e,f) *of octagonal (isotropic) elements* e, f *such that* $e \cdot f = 0$*,* $[e,f] = 0$*,* $e * f = 0$*, but* $f \notin Ke$*.*

Proof. This proof will use the same techniques as the proofs of the other lemmas in this section. Without loss of generality we may choose $e = e_{++++}$.

From the proof of the previous lemma we know that f must have $f[3] = f[4] = 0$ and $f[r] = 0$, whenever $\langle r, ++++\rangle \geq 0$, and because $f \notin Ke$ it follows that at least one coordinate $f[r] \neq 0$ with $\langle r, ++++\rangle = 1$. These are the same roots r as in the proof of Lemma 9.2, and in fact we can use the same transformations as in that lemma to ensure that we may choose $r = 0100$. (The transformations of Lemma 9.2 happen to also leave ++++ invariant, not entirely by coincidence.)

We may now apply the following reductions to f (taken from Table 2) in order to annihilate three of its coordinates :

s	$\bar{t}$	t	$t+\bar{t}$
++++	$00\bar{1}1$	+-++	
-+++	$10\bar{1}0$	--++	+--+
++-+	$0\bar{1}01$	+--+	

This reduces f to the form

$$f = k_1 e_{\mathsf{1000}} + k_2 e_{\mathsf{0100}} + k_3 e_{\mathsf{0010}} + k_4 e_{\mathsf{0001}} + k_5 e_{\mathsf{+-++}} + k_6 e_{\mathsf{+++-}}.$$

We compute

$$\begin{aligned} Q(f) = {} & k_1^2 e_{\mathsf{\bar{1}001}} + k_2^2 e_{\mathsf{0110}} + k_3^2 e_{\mathsf{01\bar{1}0}} + k_4^2 e_{\mathsf{1001}} + k_5^2 e_{\mathsf{00\bar{1}1}} + k_6^2 e_{\mathsf{\bar{1}100}} \\ & + k_1 k_2 e_{\mathsf{-+++}} + k_1 k_4 e_{\mathsf{0001}} + k_2 k_3 e_{\mathsf{0100}} + k_2 k_4 e_{\mathsf{++++}} + k_3 k_4 e_{\mathsf{++-+}} \\ & + (k_1 k_3 + k_5 k_6) e_{\mathsf{-+-+}}. \end{aligned}$$

Now, as f is octagonal, we have $0 = [[f^\sigma, a], Q(f)] \cdot b = [f^\sigma, a] \cdot [Q(f), b]$ for all $a, b \in \mathbf{W}$. Setting $a = e_{\mathsf{--+-}}$ and $b = e_{\mathsf{++--}}$ we obtain

$$\begin{aligned} [f^\sigma, e_{\mathsf{--+-}}] &= k_1^\sigma e_{\mathsf{+-+-}} + k_2^\sigma e_{\mathsf{-++-}} + k_4^\sigma e_{\mathsf{--++}}, \\ [Q(f), e_{\mathsf{++--}}] &= k_1^2 e_{\mathsf{-+-+}} + k_1 k_2 e_{\mathsf{0100}} + k_1 k_4 e_{\mathsf{++-+}}, \end{aligned}$$

and therefore $[f^\sigma, e_{\mathsf{--+-}}] \cdot [Q(f), e_{\mathsf{++--}}] = k_1^\sigma k_1^2$.

Likewise,

$$\begin{aligned} [f^\sigma, e_{\mathsf{+---}}] \cdot [Q(f), e_{\mathsf{--++}}] &= k_3^\sigma k_3^2, \\ [f^\sigma, e_{\mathsf{-++-}}] \cdot [Q(f), e_{\mathsf{000\bar{1}}}] &= k_5^\sigma k_5^2, \\ [f^\sigma, e_{\mathsf{+--+}}] \cdot [Q(f), e_{\mathsf{0\bar{1}00}}] &= k_6^\sigma k_6^2, \end{aligned}$$

proving that $k_1 = k_3 = k_5 = k_6 = 0$ when f is octagonal. Hence we are left with $f = k_2 e_{\mathsf{0100}} + k_4 e_{\mathsf{0001}}$, with $k_2 \neq 0$. In the proof of Theorem 8.5 (Stage 2) it was already shown that an element of this form can only be octagonal when $k_4 = 0$.

In a final step we get rid of the constant k_2 by applying $i_s(k_2)$ with $s = \mathsf{--++}$ and $\overline{s} = \mathsf{10\bar{1}0}$, the same element we used in the proof of Lemma 9.2. □

We combine Lemmas 9.1–9.4 into the following

Theorem 9.5. *The following is an exhaustive list of all orbits of* $^2\widehat{\mathsf{F}_4}(K)$ *on ordered pairs* (e, f) *of non-zero octagonal (isotropic) elements* $e, f \in \mathbf{W}$.

1. *For each* $k \in K - \{0\}$ *an orbit with representative* $(e_{\mathsf{0100}}, k e_{\mathsf{0\bar{1}00}})$.

2. *An orbit with representative* (e_{0100}, e_{----}).

3. *An orbit with representative* (e_{0100}, e_{0001}).

4. *An orbit with representative* (e_{0100}, e_{++++}).

5. *For each* $k \in K - \{0\}$ *an orbit with representative* (e_{0100}, ke_{0100}).

The following table lists several properties of the corresponding pairs (e, f).

	Case 1	*Case 2*	*Case 3*	*Case 4*	*Case 5*
$e \cdot f$	k	0	0	0	0
$[e, f]$	*not isotr.*	*semi-oct.*†	0	0	0
$q(e, f)$	*isotr.*‡	$\neq 0$, *oct.*	0	0	0
$f \in [e, \mathbf{F}]$	*no*	*no*	*yes*	*yes*	*yes*
$[e, \mathbf{W}] \cdot f$	$\neq 0$	$\neq 0$	0	0	0
$e + f$	*not isotr.*	*not isotr.*	*semi-oct.*†	*octag.*	*octag.*
$\mu^{-1}(e{*}f)^{\sigma^{-1}}$	*undef.*	*undef.*	$\neq 0$, *oct.*	0	0
$f \in [e, \mathbf{W}]$	*no*	*no*	*no*	*yes*	*yes*
$f \in q(\mathbf{W}, e)$	*no*	*no*	*no*	*yes*	*yes*
$[e, \mathbf{F}] \cdot f$	$\neq 0$	$\neq 0$	$\neq 0$	0	0
$f \in Ke$	*no*	*no*	*no*	*no*	$f = ke$
$q(\mathbf{W}, e) \cap q(\mathbf{W}, f)$	$\{0\}$	$\{0\}$	$K\mu^{-1}(e{*}f)^{\sigma^{-1}}$	$\dim = 3$	$q(\mathbf{W}, e) = q(\mathbf{W}, f)$

† *semi-octagonal, but not octagonal* ‡ *isotropic, but not semi-octagonal*

Proof. Only the table of properties remains to be checked. Note that all properties are 'invariant' for ${}^2\widehat{\mathsf{F}_4}(K)$, for example, when $g \in {}^2\widehat{\mathsf{F}_4}(K)$, then $[e, f]^g = [e^g, f^g]$, $q(e, f)^g = q(e^g, f^g)$, etc. The only invariant which has not yet been proved elsewhere in this text is $\mu^{-1}(e * f)^{\sigma^{-1}}$. We find

$$\mu^{-1}(e^g * f^g)^{\sigma^{-1}} = \mu^{-1}((e * f)^g)^{\sigma^{-1}} = \mu^{-1}(e * f)^{\overline{g}\sigma^{-1}} = \mu^{-1}(e * f)^{\sigma^{-1}g},$$

using (98) and the definition of $\overline{g}$.

As a consequence it is sufficient to check all listed properties only for a single (representative) pair in each orbit. We shall discuss the most difficult cases here and leave the rest to be verified by the reader.

Table 3 lists the relevant values for $f = e_r$ and $Q(f) = e_{\overline{r}}$ and the six corresponding base elements $e_{s_1}, \ldots, e_{s_6}$ of $q(\mathbf{W}, f)$. The last row corresponds to $f = e$. From this table we easily derive the results for $q(\mathbf{W}, e) \cap q(\mathbf{W}, f)$. Note

r	$\overline{r}$	s_1	s_2	s_3	s_4	s_5	s_6
$0\bar{1}00$	$0\bar{1}\bar{1}0$	$0\bar{1}00$	$00\bar{1}0$	$+--+$	$+---$	$---+$	$----$
$----$	$0\bar{1}0\bar{1}$	$0\bar{1}00$	$000\bar{1}$	$+-+-$	$+---$	$--+-$	$----$
0001	1001	1000	0001	$+--+$	$+-++$	$++-+$	$++++$
$++++$	0101	0100	0001	$++++$	$++-+$	$-+++$	$-+-+$
0100	0110	0100	0010	$++++$	$+++-$	$-+++$	$-++-$

Table 3: Auxiliary table for the proof of Theorem 9.5

that in the third case the intersection is generated by the single element $e_{\texttt{++++}}$. In that case $e * f = e_{\texttt{0101}}$ and indeed $\mu^{-1}(e * f)^{\sigma^{-1}} = e_{\texttt{++++}}$.

Consider the value of $e + f$. When $e = e_{\texttt{0100}}$ and $f = e_{\texttt{0001}}$ it was already shown in the proof of Theorem 8.5 (Stage 2) that $e + f$ is semi-octagonal but not octagonal. When $f = e_{\texttt{++++}}$ we obtain $Q(e + f) = Q(e_{\texttt{0100}} + e_{\texttt{++++}}) = e_{\texttt{0110}} + e_{\texttt{0101}}$, and then $[e_{000\bar{1}} + e_{\texttt{+--+}}, Q(e + f)] = (e + f)^\sigma$, proving that $e + f$ is octagonal. $\square$

10 The Ree-Tits octagon

Because scalar multiples of octagonal elements are again octagonal, it is natural to define a new geometry $\mathcal{O}$ whose *points* are the one dimensional subspaces Ke of $\mathbf{F}$ (i.e., the points of the 25-dimensional projective space associated with $\mathbf{W}$) for which e is octagonal (and isotropic).

Let $Ke, Kf \in \mathcal{O}$ be such that $Ke \neq Kf$. As an immediate consequence of the properties listed in Theorem 9.5, we find that the following are equivalent :

- Every 1-dimensional subspace $K(k_1 e + k_2 f)$ of $Ke + Kf$ belongs to $\mathcal{O}$.
- $f \in q(\mathbf{W}, e)$.
- $f \in [e, \mathbf{W}]$.
- $[e, \mathbf{F}] \cdot f = 0$.
- $\dim q(\mathbf{W}, e) \cap q(\mathbf{W}, f) = 3$.

- $e \cdot f = 0$, $[e, f] = 0$ and $e * f = 0$.

Two different points Ke, Kf of $\mathcal{O}$ that satisfy one of these statements shall be called *collinear*. In this case, the 2-dimensional subspace $Ke + Kf$ of $\mathcal{O}$ is called a *line* of $\mathcal{O}$.

We shall also introduce the following terminology, borrowed from [5]. Let Ke, Kf be different points of $\mathcal{O}$.

- If Ke, Kf are not collinear, and $[e, f] = 0$, then Ke and Kf are said to be *cohyperlinear*.
- If $[e, f] \neq 0$ and $e \cdot f = 0$, then Ke and Kf are said to be *almost opposite*.
- If $e \cdot f \neq 0$, then Ke and Kf are said to be *opposite*.

The next and most important theorem of this text proves that the point-line geometry $\mathcal{O}$ satisfies the axioms of a generalized octagon, where 'collinear', 'cohyperlinear', 'almost opposite' and 'opposite' correspond to distance 2, 4, 6 and 8 in the incidence graph of $\mathcal{O}$.

Theorem 10.1. *Let Ke, Kf be different points of $\mathcal{O}$.*

1. *If Ke and Kf are collinear then every point of the line $Ke + Kf$ belongs to $\mathcal{O}$. Apart from the points of $Ke+Kf$ there are no other points of $\mathcal{O}$ collinear to both Ke and Kf.*
2. *If Ke and Kf are cohyperlinear, then there is a unique point of $\mathcal{O}$ that is collinear to both Ke and Kf. This point is $K\mu^{-1}(e * f)^{\sigma^{-1}}$. In the incidence graph of $\mathcal{O}$, Ke and Kf are at mutual distance 4.*
3. *If Ke and Kf are almost opposite then there are no points of $\mathcal{O}$ that are collinear to both Ke and Kf. There is a unique point of $\mathcal{O}$ that is collinear to Ke and cohyperlinear to Kf. This point is $Kq(f, e)$. In the incidence graph of $\mathcal{O}$, Ke and Kf are at mutual distance 6.*
4. *If Ke and Kf are opposite, then there are no points of $\mathcal{O}$ that are collinear to both Ke and Kf, or that are collinear to Ke and cohyperlinear to Kf (or vice versa). Every line of $\mathcal{O}$ through Kf contains exactly one point of $\mathcal{O}$ that is almost opposite to Ke. In the incidence graph of $\mathcal{O}$, Ke and Kf are at mutual distance 8.*

These are the only possible relations between different points of $\mathcal{O}$.

Proof. The different cases of this theorem correspond to 4 different cases of Theorem 9.5, but numbered the other way round. (The fifth case of Theorem

9.5 occurs when $Ke = Kf$.) Without loss of generality we may set $e = e_{0100}$ and set f to one of the representatives listed in Theorem 9.5.

1. (Case 4 of Theorem 9.5.) We take $f = e_{++++}$. If Kd is a point collinear to both Ke and Kf then $d \in q(\mathbf{W}, e) \cap q(\mathbf{W}, f)$. Using Table 3 we easily derive that d must be of the form
$$d = k_1 e_{0100} + k_2 e_{++++} + k_3 e_{-+++},$$
and then
$$Q(d) = k_1^2 e_{0110} + k_2^2 e_{0101} + k_3^2 e_{1100}.$$
From this we compute
$$[[d, e_{0\bar{1}00}], Q(d^{\sigma^{-1}})] = [k_1 h_{r_3} + k_2 e_{+-++} + k_3 e_{--++}, Q(d^{\sigma^{-1}})] = k_3 k_3^{\sigma} e_{++++}$$
and hence $k_3 = 0$ when d is octagonal, i.e, $d \in Ke + Kf$.
2. (Case 3 of Theorem 9.5.) A point collinear to both Ke and Kf must belong to the intersection $q(\mathbf{W}, e) \cap q(\mathbf{W}, f)$, which by Theorem 9.5 is exactly $K\mu^{-1}(e * f)^{\sigma^{-1}}$. Note also that $\mu^{-1}(e * f)^{\sigma^{-1}}$ is octagonal by the same theorem.
3. (Case 2 of Theorem 9.5.) We have $q(\mathbf{W}, e) \cap q(\mathbf{W}, f) = \{0\}$, hence there are no points collinear to both Ke and Kf.

   We take $f = e_{----}$. If $Kd \in \mathcal{O}$ is collinear to e and cohyperlinear to f, we must have $d \in q(\mathbf{W}, e)$ and $[d, f] = 0$. From Table 3 it easily follows that $d \in Ke_{-++-}$. Also $q(f, e) = [e_{----}, e_{0110}] = e_{-++-}$.
4. (Case 1 of Theorem 9.5.) We have $q(\mathbf{W}, e) \cap q(\mathbf{W}, f) = \{0\}$, hence there are no points collinear to both Ke and Kf.

   Take $f = e_{0\bar{1}00}$. As before, if $Kd \in \mathcal{O}$ is collinear to Ke and cohyperlinear to Kf, we must have $d \in q(\mathbf{W}, e)$ and $[d, f] = 0$. Table 3 than proves that $d \in Ke_{0010}$, but then d is not even semi-octagonal.

   Finally, consider the hyperplane of elements $w \in \mathbf{W}$ such that $e \cdot w = 0$. Because f does not belong to this hyperplane, every line of $\mathcal{O}$ through f intersects this hyperplane in a single point $Kw \in \mathcal{O}$. By definition $e \cdot w = 0$, hence Kw is either almost opposite to Ke, cohyperlinear to Ke, collinear to Ke or equal to Ke. The last possibility is easily ruled out, and the second and third possibility have been disproved above. □

The following theorem provides some insight into the 'local structure' of $\mathcal{O}$.

Theorem 10.2. *Let Ke be a point of $\mathcal{O}$. Then the points of $\mathcal{O}$ collinear with Ke form a cone with vertex Ke and a Suzuki-Tits ovoid as a base. The semi-octagonal elements of $q(\mathbf{W}, e)$ form a 5-dimensional subspace which is generated by the elements of this cone.*

Proof. By Theorem 9.5 all points collinear with Ke must lie in $q(\mathbf{W}, e)$. We may set $e = e_{0001}$ without loss of generality. Then $f \in q(\mathbf{W}, e)$ is of the form

$$f = k_1 e_{1000} + k_2 e_{0001} + k_3 e_{++++} + k_4 e_{+--+} + k_5 e_{+-++} + k_6 e_{++-+},$$

with $k_1, \ldots, k_6 \in K$.

In the proof of Lemma 9.3 we have already established that f is semi-octagonal if and only if $k_1 = 0$, and then

$$Q(f) = k_2^2 e_{1001} + k_3^2 e_{0101} + k_4^2 e_{0\bar{1}01} + k_5^2 e_{00\bar{1}1} + k_6^2 e_{0011} + k_*^2 e_{0001}$$

with $k_* \stackrel{\text{def}}{=} \sqrt{k_3 k_4 + k_5 k_6}$.

To find out in what cases f is octagonal we establish a set of generators for the 6-dimensional space $q(\mathbf{W}, f)$. Proposition 2.4 already shows that $e = e_{0010}$ belongs to this space because $f \in q(\mathbf{W}, e)$. Now, consider the following elements of $q(\mathbf{W}, f)$:

$$\begin{aligned} v_0 \stackrel{\text{def}}{=} q(e_{000\bar{1}}, f) = k_2^\sigma e_{1000} + k_3^\sigma e_{0100} + k_4^\sigma e_{0\bar{1}00} \\ + k_5^\sigma e_{00\bar{1}0} + k_6^\sigma e_{0010} + k_*^\sigma h_{0001}, \end{aligned}$$

$$\begin{aligned}
v_1 &\stackrel{\text{def}}{=} q(e_{++--}, f) = k_4^\sigma e_{+--+} + k_6^\sigma e_{++++} + k_*^\sigma e_{++-+}\,, \\
v_2 &\stackrel{\text{def}}{=} q(e_{+---}, f) = k_3^\sigma e_{++-+} + k_6^\sigma e_{+-++} + k_*^\sigma e_{+--+}\,, \\
v_3 &\stackrel{\text{def}}{=} q(e_{+++-}, f) = k_4^\sigma e_{+-++} + k_5^\sigma e_{++-+} + k_*^\sigma e_{++++}\,, \\
v_4 &\stackrel{\text{def}}{=} q(e_{+-+-}, f) = k_3^\sigma e_{++++} + k_5^\sigma e_{+--+} + k_*^\sigma e_{+-++}\,,
\end{aligned}$$

$$\begin{aligned}
v_5 &\stackrel{\text{def}}{=} q(e_{-+--}, f) = k_4^\sigma e_{---+} + k_6^\sigma e_{-+++} + k_*^\sigma e_{-+-+}\,, \\
v_6 &\stackrel{\text{def}}{=} q(e_{----}, f) = k_3^\sigma e_{-+-+} + k_6^\sigma e_{--++} + k_*^\sigma e_{---+}\,, \\
v_7 &\stackrel{\text{def}}{=} q(e_{-++-}, f) = k_4^\sigma e_{--++} + k_5^\sigma e_{-+-+} + k_*^\sigma e_{-+++}\,, \\
v_8 &\stackrel{\text{def}}{=} q(e_{--+-}, f) = k_3^\sigma e_{-+++} + k_5^\sigma e_{---+} + k_*^\sigma e_{--++}\,.
\end{aligned}$$

Note that $v_1, \ldots, v_4$ belong to the space generated by e_{++++}, e_{++-+}, e_{+-++} and e_{+--+}. Similarly, $v_5, \ldots, v_8$ belong to the space generated by e_{-+++}, e_{-+-+}, e_{--++} and e_{---+}.

Not all of these vectors are linearly independent. For example, we see that $k_3^\sigma v_1 = k_*^\sigma v_2 + k_6^\sigma v_4$ and $k_3^\sigma v_3 = k_5^\sigma v_2 + k_*^\sigma v_4$, and hence, when $k_3 \neq 0$ the vectors $v_1, \ldots, v_4$ generate a space of dimension 2. This is also the case when $k_4 \neq 0$, $k_5 \neq 0$ or $k_6 \neq 0$, and with a similar argument we can prove that also $v_5, \ldots, v_8$ generate a space of dimension 2. We may conclude that, unless $f \in Ke$, the elements $e, v_0, \ldots, v_8$ generate $q(\mathbf{W}, f)$.

To determine which of the elements $f \notin Ke$ is octagonal, i.e., whether $f \in q(\mathbf{W}, f)$, we need to consider two different cases :

Case 1. Assume $k_3 \neq 0$. By the above we must have that

$$f = k_6 e_{++-+} + k_4 e_{+--+} + k_3 e_{++++} + k_5 e_{+-++} + k_2 e_{0001}$$

is a linear combination of e_{0001} and

$$\begin{aligned} v_2 &= k_3^\sigma e_{++-+} + k_*^\sigma e_{+--+} + k_6^\sigma e_{+-++}, \\ v_4 &= k_5^\sigma e_{+--+} + k_3^\sigma e_{++++} + k_*^\sigma e_{+-++}. \end{aligned}$$

This yields no condition on k_2, proving that the resulting solutions form a cone with vertex $Ke = Ke_{0001}$. We obtain the following conditions on $k_3, \ldots, k_6$:

$$k_4 k_3^\sigma + k_6 k_*^\sigma + k_3 k_5^\sigma = 0, \qquad k_5 k_3^\sigma + k_6 k_6^\sigma + k_3 k_*^\sigma = 0.$$

Applying σ to these equations, and expanding k_*^2, yields :

$$k_4^\sigma k_3^2 + k_6^\sigma k_3 k_4 + k_6^\sigma k_5 k_6 + k_3^\sigma k_5^2 = 0, \qquad k_5^\sigma k_3^2 + k_6^\sigma k_6^2 + k_3^\sigma k_3 k_4 + k_3^\sigma k_5 k_6 = 0.$$

These equations are not independent and will be satisfied if and only if

$$k_4 = k_3^{-1-\sigma} k_6^{2+\sigma} + k_3^{-1} k_5 k_6 + k_3^{1-\sigma} k_5^\sigma. \tag{105}$$

Case 2. Let $k_3 = 0$. We find that $k_6 e_{++-+} + k_4 e_{+--+} + k_5 e_{+-++}$ now must belong to the space generated by

$$\begin{aligned} v_1 &= k_*^\sigma e_{++-+} + k_4^\sigma e_{+--+} + k_6^\sigma e_{++++} \\ v_2 &= k_*^\sigma e_{+--+} + k_6^\sigma e_{+-++} \\ v_3 &= k_5^\sigma e_{++-+} + k_*^\sigma e_{++++} + k_4^\sigma e_{+-++} \\ v_4 &= k_5^\sigma e_{+--+} + k_*^\sigma e_{+-++} \end{aligned}$$

with $k_*^2 = k_5 k_6$. This is equivalent to

$$k_3 = k_5 = k_6 = 0. \tag{106}$$

(105) and (106) together describe the well-known parameter equation of the *Suzuki-Tits ovoid* in a 3-dimensional projective space (cf. [5]). $\square$

We end this section by establishing the two types of root groups for $\mathcal{O}$.

Proposition 10.3. *Let Ke, Kf be points of $\mathcal{O}$ that are almost opposite. Then $y([e,f])$ fixes all elements (points and lines) on the shortest path joining Ke and Kf and all elements incident to (at least) one of those elements.*

Let L be any line of $\mathcal{O}$ through Ke and not on this path. Then the subgroup of ${}^2\widehat{\mathsf{F}_4}(K)$ of all $y(k\,[e,f])$ with $k \in K$ acts transitively on the points of L different from Ke.

Proof. By Lemma 8.4 and Theorem 9.5 we may choose $e = e_{\mathsf{0001}}$ and $f = e_{\mathsf{----}}$ without loss of generality. Then $[e,f]$ is the semi-octagonal element $e_{\mathsf{---+}}$ and $Q([e,f]^{\sigma^{-1}}) = e_{\mathsf{1\bar{1}00}}$.

The points on the shortest path joining Ke and Kf are Ke' and Kf' with $e' = q(f,e) = e_{\mathsf{+--+}}$ and $f' = q(e,f) = e_{\mathsf{0\bar{1}00}}$ and they are easily seen to be left invariant by $y([e,f])$. Hence also the lines on that path are left invariant, and all points on those lines.

Consider a line M through Ke'. M lies entirely inside the space $q(\mathbf{W}, e') = [\mathbf{W}, e_{\mathsf{0\bar{1}01}}]$ which is generated by the following vectors :

$$e_{\mathsf{0\bar{1}00}},\ e_{\mathsf{0001}},\ e_{\mathsf{+-++}},\ e_{\mathsf{+--+}},\ e_{\mathsf{--++}},\ e_{\mathsf{---+}}.$$

It is easily seen that each of these vectors is left invariant by $y([e,f])$, hence so is M. (By symmetry, the same holds for lines through f'.)

Finally, let L be a line through Ke, not equal to $Ke + Ke'$. From the proof of Theorem 10.2 we know that L lies on the cone with vertex Ke and a base belonging to the 4-dimensional space generated by $e_{\mathsf{++++}}, e_{\mathsf{+--+}}, e_{\mathsf{+-++}}, e_{\mathsf{++-+}}$. Of these four vectors, the only one not left invariant by $y(k\,[e,f])$ is $e_{\mathsf{++++}}$ which maps to $e_{\mathsf{++++}} + ke_{\mathsf{0001}} = e_{\mathsf{++++}} + ke$.

Now consider a point Kg of L. Using the notation of the proof of Theorem 10.2 we assign 'coordinates' $k_2, \ldots, k_5$ to the element g. It is easily seen that the case $k_3 = 0$ occurs only when $g \in Ke + Ke'$. Because k_3 is the 'coordinate' that corresponds to e_{++++}, we find that g is mapped by $y(k\,[e,f])$ to $g + kk_3e$, which lies on the same line L. Also, when k runs through K, the image of g encounters all points of the line L, except e.

Note that the elements $y(k\,[e,f])$ form a subgroup isomorphic to the additive group of K, by Lemma 8.3. □

The elements $y([e,f])$ of the proposition above are called *point-elations* of $\mathcal{O}$ and the group of elements $y(k\,[e,f])$ is a *root group* [5]. In [2, Proposition 13.6.3] this root group is listed as of 'type (v)'.

Proposition 10.4. *Let Ke, Kf be opposite points of $\mathcal{O}$ and let Kg be a point of $\mathcal{O}$ at distance 4 of both Ke and Kf. Then*

1. *The subset of elements c of $q(\mathbf{W}, g)$ that satisfy $[c,e] = [c,f] = 0$ is a 2-dimensional subspace of the form $Kg + Kd$ where d is isotropic but not semi-octagonal and $g = q(d,d)$.*
2. *Let $k, k', \ell, \ell' \in K$. Then*
$$y(kg)y(\ell d)y(k'g)y(\ell' d) = y((k + k' + \ell^\sigma \ell')\,g)y((\ell + \ell')d), \tag{107}$$
which makes $X \stackrel{\text{def}}{=} \{y(kg)y(\ell d) \mid k, \ell \in K\}$ a subgroup of ${}^2\widehat{\mathsf{F}_4}(K)$.
3. *X fixes all points and lines on the unique path of length 8 joining Ke to Kg and then to Kf.*
4. *X fixes all lines incident with any of the three 'middle' points of that path (i.e. Kg and the two points collinear to Kg).*
5. *X acts transitively on the set of lines through Ke which are not on this path.*

Proof. We shall first prove the proposition for the special case $e = e_{0001}$, $f = e_{000\bar{1}}$ and $g = e_{0\bar{1}00}$. Then $Q(g^{\sigma^{-1}}) = e_{0\bar{1}\bar{1}0}$ and it is easily computed that statement 1 of this proposition is satisfied by $d = e_{00\bar{1}0}$.

By (102) we have

$$y(kg) = x(k^\sigma e_{0\bar{1}\bar{1}0})x(ke_{0\bar{1}00}), \qquad y(\ell d) = x(\ell^\sigma e_{0\bar{1}10})x(\ell e_{00\bar{1}0})x(\ell\ell^\sigma e_{0\bar{1}00}).$$

From this we compute $d^{y(kg)} = d$ and hence, using (104),

$$y(\ell d)y(k'g) = y(k'g)y(\ell d)^{y(k'g)} = y(k'g)y(\ell d^{y(k'g)}) = y(k'g)y(\ell d). \tag{108}$$

Also, by (103),

$$y(kg)y(k'g) = y((k+k')g), \quad y(\ell d)y(\ell' d) = y(\ell^\sigma \ell' g)y((\ell+\ell')d). \tag{109}$$

Combining (108) and (109) yields (107).

The three 'middle' points on the unique path of length 8 joining Ke to Kg to Kf correspond to $e_{\mathtt{+--+}}$, $e_{0\bar{1}00}$ and $e_{\mathtt{----}}$. For each of these three vectors e_r, the following table lists the generators $e_r, e_{s_1}, \ldots, e_{s_4}$ of the corresponding 5-dimensional subspace of semi-octagonal elements of $q(\mathbf{W}, e_r)$. (These results were obtained by applying an appropriate element of W' to the example computed in the proof of Theorem 10.2.)

r	s_1	s_2	s_3	s_4
+−−+	0001	$0\bar{1}00$	−−−+	+−++
$0\bar{1}00$	+−−+	−−−−	+−−−	−−−+
−−−−	$0\bar{1}00$	$000\bar{1}$	−−+−	+−−−

Note that each of the listed vectors e_r, e_{s_i} is left invariant by $x(e_{0\bar{1}\bar{1}0})$, $x(e_{0\bar{1}10})$, $x(e_{0\bar{1}00})$ and $x(e_{00\bar{1}0})$, and hence also by $y(kg)y(\ell d)$, except for $e_{\mathtt{+-++}}$ and $e_{\mathtt{--+-}}$. In the these two cases we find

$$e_{\mathtt{+-++}}^{y(kg)y(\ell d)} = e_{\mathtt{+-++}} + \ell e_{\mathtt{+--+}}, \qquad e_{\mathtt{--+-}}^{y(kg)y(\ell d)} = e_{\mathtt{--+-}} + \ell e_{\mathtt{----}}.$$

This proves statements 3 and 4.

Now consider the line L joining Ke with $Ke_{\mathtt{++++}}$. The vector $e = e_{0001}$ is left invariant by $y(kg)y(\ell d)$, while $e_{\mathtt{++++}}$ is mapped to

$$\begin{aligned}
e_{\mathtt{++++}}^{y(kg)y(\ell d)} &= e_{\mathtt{++++}}^{y(\ell d)} + k^\sigma e_{\mathtt{+--+}}^{y(\ell d)} + k e_{\mathtt{+-++}}^{y(\ell d)} \\
&= e_{\mathtt{++++}} + \ell e_{\mathtt{++-+}} + \ell\ell^\sigma e_{\mathtt{+-++}} + \ell^2\ell^\sigma e_{\mathtt{+--+}} + k^\sigma e_{\mathtt{+--+}} + k e_{\mathtt{+-++}} + k\ell e_{\mathtt{+--+}} \\
&= e_{\mathtt{++++}} + (k^\sigma + k\ell + \ell^2\ell^\sigma) e_{\mathtt{+--+}} + (k + \ell\ell^\sigma) e_{\mathtt{+-++}} + \ell e_{\mathtt{++-+}}.
\end{aligned}$$

In terms of the 'coordinates' of the proof of Theorem 10.2 the resulting image has $k_3 = 1$, $k_4 = k^\sigma + k\ell + \ell^2\ell^\sigma$, $k_5 = k + \ell\ell^\sigma$ and $k_6 = \ell$. So, when (k, ℓ) runs through $K \times K$, the images of L range over all lines through Ke, except the line joining e and $e_{\mathtt{+--+}}$ (which corresponds to $k_3 = 0$). This proves statement 5.

We still need to prove that the proposition also holds for general e, f and g in the stated configuration. Denote by Ke' (resp. Kf') the unique points of $\mathcal{O}$

collinear to Ke and Kg (resp. Kf and Kg). By Lemma 8.4 and Theorem 9.5 we may choose $e' = e_{0001}$ and $f = e_{----}$ without loss of generality (e' and f are almost opposite). This implies $g = e_{+--+}$ and $f' = e_{0\bar{1}00}$. The special case used in the first part of this proof can now be applied to show that the line through Ke_{0001} and Ke can be chosen to be the line connecting Ke_{0001} and Ke_{++++}. Proposition 10.3 then further allows us to choose $Ke = Ke_{++++}$. Finally, an appropriate element of W' enables us to map the sequence of points e, e', g, f', f thus obtained, to the sequence e_{0001}, e_{+--+}, $e_{0\bar{1}00}$, e_{----}, $e_{000\bar{1}}$ of the first part of this proof. □

The group X defined in this proposition is again a *root group* and the elements $y(kg)y(\ell d)$ are called *line-elations* of $\mathcal{O}$. When $\ell = 0$ these are the *central elations* with center Kg. In [2, Proposition 13.6.3] X is listed as of 'type (*vi*)'.

Propositions 10.3 and 10.4 imply that $\mathcal{O}$ is a *Moufang polygon* [5] and therefore none other than the classical Ree-Tits generalized octagon $\mathsf{O}(K, \sigma)$.

Acknowledgements

I would very much like to thank H. Van Maldeghem without whom this text would never have been written (or at least have remained unfinished for another decade). I thank him for introducing the subject to me many years ago, for urging me on, for carefully reading the various drafts of this paper and for giving plenty of valuable advice.

I would also like to thank the referee for many useful comments and for his suggestion to include Propositions 10.3 and 10.4 as an appropriate conclusion to this article.

References

[1] **M. Aschbacher**, Some multilinear forms with large isometry groups, *Geometriae Dedicata*, **25** (1988), 417–465.

[2] **R. W. Carter**, *Simple groups of Lie type*, John Wiley & sons, London, 1972.

[3] **A. M. Cohen**, Point-line spaces associated with buildings, *Handbook of Incidence Geometry, Buildings and Foundations*, (ed. F. Buekenhout), Elsevier, Amsterdam, 1995, 647–737.

[4] **J. Tits**, Moufang octagons and the Ree groups of type ${}^2\mathsf{F}_4$, *Amer. J. Math.* **105** (1983), 539–594.

[5] **H. Van Maldeghem**, *Generalized Polygons*, Birkhäuser verlag, Basel, 1998.

Kris Coolsaet

DEPARTMENT OF APPLIED MATHEMATICS AND COMPUTER SCIENCE, GHENT UNIVERSITY, KRIJGSLAAN 281–S9, B–9000 GENT, BELGIUM

e-mail: `Kris.Coolsaet@UGent.be`