RELATIONS IN HYPERREFLECTION GROUPS

ERICH W. ELLERS

Abstract. Hyperreflection groups \( G_m \) are generalizations of groups generated by reflections. A hyperreflection group is generated by hyperreflections if \( \dim V \) is finite. A hyperreflection is a simple mapping \( \sigma \) such that \( \det \sigma = \gamma \), where \( \gamma^m = 1 \). If the field of scalars is commutative, the order of \( \sigma \) is \( m \). Our main result states that every relation between hyperreflections and their inverses is a consequence of relations of lengths 2, 4, and \( m \). The most interesting special case occurs for \( m = 2 \). Then our result refers to relations between reflections.

1. Introduction. Let \( V \) be a vector space over a commutative field \( K \) with \( \text{char} \ K \neq 2 \) and \( G_2 \) the subgroup of the general linear group \( \text{GL}(V) \), consisting of all transformations \( \pi \) such that \( (\det \pi)^2 = 1 \). The group \( G_2 \) is generated by reflections, i.e. involutory dilatations. We shall show that every relation between reflections is a consequence of relations of lengths 2 and 4.

Similar results for other groups have been obtained by a number of authors (cf. list of references). The task of finding short relations such that all other relations are consequences of them is known as the relation problem (cf. F. Bachmann [2, p. 341]). Its solution provides the basis for the characterization of groups (cf. [2]).

In fact, we shall solve the relation problem in a more general setting. We allow the vector space to be infinite dimensional and the field \( K \) to be a skewfield. Our main result (Theorem 5) concerns hyperreflection groups \( G_m \) (§3), generated by \( \gamma \) and \( \gamma^{-1} \)-dilatations. Hyperreflection groups are natural generalizations of \( G_2 \). They were introduced in [7]. The results obtained there have been applied to equiaffinities [7], affinities [9], and projectivities [11].

The intuitive idea in the proof of Theorem 5 is to gradually reduce the length of a relation using \( n \)-relations for \( n = 2, 4, m \). In the process we find it necessary first to lengthen the word by inserting suitable mappings \( \sigma \) and their inverses in order to be able to apply Lemma 2. The purpose of §2 is solely to establish the existence of these mappings. The assumptions of Lemma 1 simply reflect the requirements that occur in the course of the proof of Theorem 5.

2. Existence of special \( \epsilon \)-dilatations. Let \( V \) be a left vector space of arbitrary dimension over the field \( K \) where \( K \) is not necessarily commutative. The dual space of \( V \) will be denoted by \( V^* \).

For our investigations it will often be important to consider the following two subspaces which are attached to each element \( \pi \) in the general linear group \( \text{GL}(V) \);
namely $B(\sigma) = \{x^\sigma - x; x \in V\}$, called the *path* of $\sigma$, and $F(\pi) = \{x \in V; x^\pi = x\}$, called the *fix* of $\pi$. A transformation $\pi$ is called *simple* if $\text{codim} \ F(\pi) = 1$.

Let $\varepsilon \in K \setminus \{0\}$. A simple transformation $\rho$ is called an *$\varepsilon$-dilatation* if there is some $r \in V \setminus \{0\}$ and some $\psi \in V^*$ such that $\rho: x \mapsto x + x^\psi r$ for all $x \in V$ and $1 + r^\psi = \varepsilon$. Here, $\rho$ determines the field element $\varepsilon$ only up to conjugates. Two cases of special interest: $1$-dilatations are transvections and $(-1)$-dilatations are reflections.

We are now already in a position where we can prove a technical-looking lemma which will play a key role in the proof of our main result on relations.

**Lemma 1.** Let $\dim V > 3$ and $|K| > 3$, or $\dim V = 2$ and $|K| > 5$, and let $\varepsilon \in K \setminus \{0\}$.

Assume $\tau$, $\omega$, $\rho$ are simple transformations in $\text{GL}(V)$ such that

\[
\text{codim}(F(\tau) \cap F(\omega) \cap F(\rho)) > 2
\]

and

\[
v \in V \setminus (F(\tau) \cup F(\omega) \cup F(\rho) \cup F(\omega)^{r^{-1}} \cup F(\omega)^{\rho}).
\]

Then there is an $\varepsilon$-dilatation $\sigma$ such that $F(\sigma) \supseteq F(\tau) \cap F(\omega) \cap F(\rho)$; $F(\sigma) \neq F(\tau)$, $F(\rho)$, and $B(\sigma) \neq B(\tau)$, $B(\rho)$. Furthermore, $v \notin F(\sigma)$, $v \in F(\omega)^{\sigma}$, $v^{\sigma} \notin K\nu + F(\tau\sigma)$, $v^{\sigma} \notin K\nu + F(\sigma^{-1}\rho)$.

**Proof.** Put $F(\tau) \cap F(\omega) \cap F(\rho) = R$. If $\dim V > 3$, we can choose $w' \in F(\omega) \setminus (R \cup (K\nu + K\nu^\tau) \cup (K\nu + K\nu^{\rho^{-1}}))$; if $|K| > 5$, we can choose $w' \in F(\omega) \setminus R$.

Let $w \neq 0$ be a multiple of $w'$ such that $v - \nu w \notin F(\tau)$, $F(\rho)$. This can be done since $v \notin F(\tau)$, $F(\omega)$, $F(\rho)$ and $K$ contains more than three elements. If $K$ contains more than five elements, we can choose $w$ such that also $v - w \notin B(\tau)$, $B(\rho)$. Then $(K\nu + K\nu^\tau) \cap F(\tau) = C \subseteq R + K(v - \nu w)$. First we see that $C \subseteq R$; namely, if $C \subseteq R$, then $v^\tau \in F(\tau)$ or $w \in R + K\nu^\tau$. This is a contradiction. Now let $c \in C \setminus \{0\}$ and $C \subseteq R + K(v - \nu w)$. Then $c = r + \gamma(v - \nu w)$ with $\gamma \neq 0$ and $r \in R$, hence $v - \nu w \in F(\tau)$. This is a contradiction. Similarly, we get $(K\nu + K\nu^{\rho^{-1}}) \cap F(\rho) = D \subseteq R + K(v - \nu w)$.

Now we define $\sigma: x \mapsto x + x^\psi(v - w)$ where $\psi \in V^*$ such that $R^\psi = 0$, $v^\psi = \varepsilon$, and $w^\psi = 1$. This can be done by our choice of $w$ since $v \notin F(\omega)$. Then $(v - \nu w)^\psi = 0$. The space $F(\sigma)$ contains $R + K(v - \nu w)$, in fact, $F(\sigma) = R + K(v - \nu w)$ if codim $R = 2$. Then $C, D \not\subseteq F(\sigma)$. If codim $R = 3$, then $F(\sigma)$ is not yet completely determined by the conditions on $\psi$. But then we can choose $F(\sigma)$ such that $C, D \not\subseteq F(\sigma)$ since dim $C = \dim D = 1$ and $C, D \not\subseteq R + K(v - \nu w)$.

Now we get immediately that $v \notin F(\sigma)$, and $w^\sigma = v$, hence $v \in F(\omega)^\sigma$. Also, $F(\sigma) \neq F(\tau)$, $F(\rho)$, and $B(\sigma) \neq B(\tau)$, $B(\rho)$ by our choice of $w$. Therefore, $F(\tau\sigma) = F(\tau) \cap F(\sigma)$ and $F(\sigma^{-1}\rho) = F(\sigma) \cap F(\rho)$. Clearly, $F(\sigma) \supseteq R$.

If $v^{\sigma\tau} \in K\nu + F(\tau\sigma)$, then $v^\tau \in K\nu^{\sigma^{-1}} + F(\tau\sigma) = Kw + F(\tau\sigma)$, hence $K\nu^\tau + Kw \subseteq Kw + F(\tau\sigma)$. Now we get $C = (K\nu^{\sigma\tau} + Kw) \cap F(\tau\sigma) = (K\nu^\tau + Kw) \cap F(\tau\sigma) \subset F(\sigma)$ which is a contradiction.

\[\text{\footnotesize For elementary properties of } B(\pi) \text{ and } F(\pi), \text{ see e.g. } [6] \text{ or } [18].\]
If \( v^{\sigma^{-1} p} \in K v + F(\sigma^{-1} p) \), then \( w^p \in K v + F(\sigma^{-1} p) \) and consequently \( w \in K v^{\sigma^{-1}} + F(\sigma^{-1} p) \). Similar to the previous case, this yields \( D \subset F(\sigma) \). This is a contradiction again, and we have \( v^{\sigma^{-1} p} \notin K v + F(\sigma^{-1} p) \).

**Lemma 2.** Let \( e \in K \setminus \{0\} \) and \( \pi \in GL(V) \) with \( \dim B(\pi) = 2 \). If \( v \in V \) and \( v^\pi \notin K v + F(\pi) \), then there is an \( e \)-dilatation \( \sigma \) and a dilatation \( \rho \) in \( GL(V) \) such that \( \pi = \rho \sigma, v \in F(\rho), \) and \( F(\rho) \cap F(\sigma) = F(\pi) \).

**Proof.** We define \( \sigma: x \rightarrow x + x^\psi (v^\pi - v) \) where \( \psi \in V^* \) such that \( F(\psi) = 0, v^\psi = 1, \) and \( (v^\pi - v)^\psi = \varepsilon - 1 \). This is possible since \( v^\pi \notin K v + F(\pi) \). Then \( \sigma \) is an \( e \)-dilatation and \( \rho = \sigma^{-1} \in GL(V) \). Since \( v^\sigma = v^\pi \), we have \( v \in F(\rho) \) and \( F(\rho) = K v \oplus F(\pi), \) hence \( \rho \) is simple.

3. Hyperreflection groups. We shall now introduce hyperreflection groups over skewfields.

In [9], we have extended the definition of determinants\(^3\) to vector spaces of infinite dimension. If \( \pi \in GL(V) \) and \( \text{codim} F(\pi) \) is finite, then \( \det \pi = \det_{B(\pi)} \pi \). Let \( C(K^*) \) be the commutator subgroup of \( K^* \), the multiplicative group of \( K \). If \( \kappa \in K^*/C(K^*) \), then a simple transformation \( \rho \) with \( \det \rho = \kappa \) is called a \( \kappa \)-dilatation.

If \( e \) is a scalar distinct from zero such that \( e \cdot C(K^*) = \kappa \) and if \( \rho \) is an \( e \)-dilatation, then obviously \( \rho \) is also a \( \kappa \)-dilatation.

Let \( \Gamma \) be a cyclic subgroup of \( K^*/C(K^*) \). Assume that the order of \( \Gamma \) is \( m \neq 1 \) and \( \gamma \) is a generator of \( \Gamma \). Then a \( \gamma \)-dilatation is called a hyperreflection. We define \( G_m(V) = \{ \pi \in GL(V); \dim B(\pi) < \infty \text{ and } \det \pi \in \Gamma \} \) and call \( G_m(V) \) the hyperreflection group. The hyperreflections generate \( G_m(V) \) (cf. [9] and [7, Theorem 7]).

If our field \( K \) is commutative, then \( C(K^*) = \{1\} \), and \( \gamma \) can be taken as an element in \( K \). Since \( \gamma^m = 1 \), \( \gamma \) becomes a primitive \( m \)th root of unity. In order to define the hyperreflection group \( G_m(V) \), we have to assume that \( K \) contains an \( m \)th root of unity. If \( m = 2 \), then \( \gamma = -1 \) and a hyperreflection is a reflection.

Now we shall introduce a number of concepts pertaining specifically to solving the relation problem.

Let \( T \) be the set of all \( \gamma \) and \( \gamma^{-1} \)-dilatations in \( G_m(V) \). Then \( T \) generates \( G_m \), \( T^{-1} \subset T \), and \( T \) is normal in \( G_m \). If \( \sigma_i \in T \), then the \( n \)-tuple \( (\sigma_1, \ldots, \sigma_n) \) is a word in the free group \( F \) generated by the set \( T \). A word \( (\sigma_1, \ldots, \sigma_n) \) of length \( n \) is an \( n \)-relation if \( \sigma_1 \cdots \sigma_n = 1 \). Let \( S \) denote the normal subgroup of \( F \) that is generated by all \( n \)-relations with \( n = 2, 4, \) and \( n = m \) where \( F(\sigma_i) = F(\sigma_j) \) for all \( i, j \).

The word \( (\sigma_1', \ldots, \sigma_n') = (\pi') \) in \( F \) is derived from the word \( (\sigma_1, \ldots, \sigma_n) = (\pi) \) in \( F \) if \( (\pi') \equiv (\pi) \mod S \) and if \( \bigcap_{i=1}^n F(\sigma_i) \supset \bigcup_{i=1}^n F(\sigma_i) \).

Let \( v \in V \). The word \( (\sigma_1, \ldots, \sigma_n) \) in \( F \) has \( v \)-defect \( k \) if \( v \notin F(\sigma_i) \) for exactly \( k \) of the elements \( \sigma_i, i = 1, \ldots, n \); it is \( v \)-ordered if \( v \in F(\sigma_i) \) and \( v \notin F(\sigma_j) \) implies \( i < j \).

Lemma 3. If \((\pi) = (\sigma_1, \ldots, \sigma_n)\) is a relation with \(F(\sigma_i) = F(\sigma_j)\) for all \(i, j\), then the empty word \(\emptyset\) is derived from \((\pi)\).

Proof. Assume all \(\sigma_i\) are \(\gamma\)-dilatations; then \(n = km\).
If \(k = 1\), then \((\pi) \in \mathcal{S}\).
If \(k > 1\), then \(n > m + 2\) and \(\sigma_1 \cdots \sigma_{m-1} = \sigma\) is a \(\gamma^{-1}\)-dilatation.
The relation \((\sigma, \sigma_m, \ldots, \sigma_n)\) is derived from and shorter than \((\pi)\) if \(m > 2\). If \(m = 2\), then \(\sigma_1\sigma_2\sigma_3 = \sigma\) is a \(\gamma\)-dilatation and the relation \((\sigma, \sigma_4, \ldots, \sigma_n)\) is derived from and shorter than \((\pi)\).
The case that all \(\sigma_i\) are \(\gamma^{-1}\)-dilatations can be dealt with similarly.
Now assume there is at least one \(\gamma\) and one \(\gamma^{-1}\)-dilatation in \((\pi)\). If \(n = 2\), then \((\pi) \in \mathcal{S}\).
If \(n > 3\), then there are at least two dilatations with the same determinant and one with the inverse. Assume \(\det \sigma_i = \det \sigma_{i+1} = \gamma\) and \(\det \sigma_{i+2} = \gamma^{-1}\). Then \(\sigma_i\sigma_{i+1}\sigma_{i+2} = \sigma\) is a \(\gamma\)-dilatation and the relation \((\sigma_1, \ldots, \sigma_{i-1}, \sigma, \sigma_{i+3}, \ldots, \sigma_n)\) is derived from and shorter than \((\pi)\). Again, the case where there are two \(\gamma^{-1}\) and one \(\gamma\)-dilatations can be treated similarly.

For the next two proofs, the following remark will be helpful whenever we use Lemma 2. Let \(t, w \in \mathcal{T}\); then \(\det tw\) has one of the values \(1, \gamma^2, \gamma^{-2}\). If the assumptions of Lemma 2 are satisfied for \(\pi = tw\), we choose \(e\) such that \(e \cdot C(K^*) = \gamma\) or \(\gamma^{-1}\). Thus we see that there are dilatations \(\rho, \sigma \in \mathcal{T}\) such that \(tw = \rho\sigma\) and \(\det \rho = (\det \sigma)^{-1} = \gamma\) if \(\det tw = 1\), \(\det \rho = \det \sigma = \gamma\) if \(\det tw = \gamma^2\), and \(\det \rho = \det \sigma = \gamma^{-1}\) if \(\det tw = \gamma^{-2}\).

Lemma 4. Let \(\dim V > 3\) and \(|K| > 3\), or \(\dim V = 2\) and \(|K| > 5\). Assume \(\tau, \omega, \rho \in \mathcal{T}\) with at least two distinct fixes. If \((\tau, \omega, \rho)\) is a word with \(v\)-defect 3, then there is a word that is derived from \((\tau, \omega, \rho)\) and whose \(v\)-defect is at most 2.

Proof. If \(v \in F(\omega)^{-1}\) or \(v \in F(\omega)^\rho\), then we can reduce the \(v\)-defect using \(\tau \omega = (\tau \omega \tau^{-1}) \tau \rho\) and \(\tau \omega = \tau \rho (\rho^{-1} \omega \rho)\), respectively. Now we assume \(v \notin F(\omega)^{-1} \cup F(\omega)^\rho\). By Lemma 1, there is some \(e\)-dilatation \(\sigma\) such that \(v \in F(\sigma^{-1} \omega \rho), e^\sigma \in K^v + F(\tau \sigma), \) and \(e^\sigma \rho \in K^v + F(\sigma^{-1} \rho)\). Therefore, we can apply Lemma 2 twice, for \(\pi = \tau \sigma\) and \(\pi = \sigma^{-1} \rho\), in order to reduce the \(v\)-defect.

Theorem 5. Let \(\dim V > 3\) and \(|K| > 3\), or \(\dim V = 2\) and \(|K| > 5\). Let \((\sigma_1, \ldots, \sigma_n) = (\pi)\) be a relation in \(F\). Then the empty word \(\emptyset\) is derived from \((\pi)\).

Proof. Let \(v \in V\). We can assume that \((\pi)\) is \(v\)-ordered since we can \(v\)-order by \([8, \text{Lemma 5}]\). We note that \(v\)-ordering does not increase the \(v\)-defect.

We shall now prove by induction on the \(v\)-defect of the word \((\sigma_1, \ldots, \sigma_n)\) that there is a derived word \(d(\sigma_1, \ldots, \sigma_n)\) which has \(v\)-defect zero.

Now assume \(v \in F(\sigma_i)\) and \(v \notin F(\sigma_{i+1})\) for some \(i < n\); otherwise the \(v\)-defect of \((\pi)\) is zero. If \(F(\sigma_{i+j}) = F(\sigma_{i+j})\) for all \(j > 1\), then \(\sigma_{i+1} \cdots \sigma_n = \Sigma\) is simple or \(\Sigma = \text{id}\). If \(\Sigma\) is simple, then \(F(\Sigma) = F(\sigma_{i+j})\). Since \(\sigma_1 \cdots \sigma_n \Sigma = 1\), we get \(v \in F(\Sigma) = F(\sigma_{i+j})\) and the \(v\)-defect is zero. If \(\Sigma = \text{id}\), then using Lemma 3 we also get a derived word whose \(v\)-defect is zero. Therefore, we can assume that \(F(\sigma_{i+j}) \neq F(\sigma_{i+k})\) for some \(j, k\).
Assume \( i = n - 2 \). Then \( \sigma_{i+1} = \sigma_{i+2} \) and \( \sigma_{i+k} = \sigma_{i+2} \). Also, \( v \in F(\sigma_{i+1}, \sigma_{i+2}) \).

Thus \( \sigma_{i+1}, \sigma_{i+2} \) is simple by [8, Lemma 6]. Now \( \sigma_{i+1}, \sigma_{i+2} \) can be decomposed into two \( \gamma \) or \( \gamma^{-1} \)-dilatations \( \sigma'_{i+1}, \sigma'_{i+2} \) as the case may require, such that \( F(\sigma'_{i+1}) = F(\sigma'_{i+2}) = F(\sigma_{i+1}, \sigma_{i+2}) \). The construction of \( \sigma'_{i+1} \) and \( \sigma'_{i+2} \) is similar to that in the proof of Lemma 2. Thus we have reduced the \( v \)-defect.

If \( i < n - 2 \), then there are three elements \( \sigma_{i+1}, \sigma_{i+k}, \sigma_{i+h} \) in \((\pi)\) with two distinct fixes. It is not difficult to see that we can assume \( \sigma_{i+1} = \sigma_{i+2}, \sigma_{i+k} = \sigma_{i+2}, \) and \( \sigma_{i+h} = \sigma_{i+3} \). Now we apply Lemma 4 to reduce the \( v \)-defect.

Since for every word \( (\sigma_1, \ldots, \sigma_n) \) in \( F \) the space \( \cap_{i=1}^n F(\sigma_i) \) has finite codimension, we can now use induction on this codimension to see that there is a derived word \( d(\sigma_1, \ldots, \sigma_n) \) that has \( v \)-defect zero for all \( v \in V \). But then, clearly, \( d(\sigma_1, \ldots, \sigma_n) = \emptyset \), the empty word.

**References**