5 Unitary groups

In this section we analyse the unitary groups in a similar way to the treatment of the symplectic groups in the last section. Note that the treatment here applies only to the isometry groups of Hermitian forms which are not anisotropic. So in particular, the Lie groups $\mathrm{SU}(n)$ over the complex numbers are not included.

Let V be a vector space over F, σ an automorphism of F of order 2, and B a non-degenerate σ-Hermitian form on V (that is, $B(y,x) = B(x,y)^\sigma$ for all $x,y \in V$). It is often convenient to denote c^σ by $\overline{c}$, for any element $c \in F$.

Let F_0 denote the fixed field of F. There are two important maps from F to F_0 associated with σ, the *trace* and *norm* maps, defined by

$$\begin{aligned} \mathrm{Tr}(c) &= c + \overline{c}, \\ N(c) &= c \cdot \overline{c}. \end{aligned}$$

Now Tr is an additive homomorphism (indeed, an F_0-linear map), and N is a multiplicative homomorphism. As we have seen, the image of Tr is F_0; the kernel is the set of c such that $c^\sigma = -c$ (which is equal to F_0 if the characteristic is 2 but not otherwise).

Suppose that F is finite. Then the order of F is a square, say $F = \mathrm{GF}(q^2)$, and $F_0 = \mathrm{GF}(q)$. Since the multiplicative group of F_0 has order $q-1$, a non-zero element $c \in F$ lies in F_0 if and only if $c^{q-1} = 1$. This holds if and only if $c = a^{q+1}$ for some $a \in F$ (as the multiplicative group of F is cyclic), in other words, $c = a \cdot \overline{a} = N(a)$. So the image of N is the multiplicative group of F_0, and its kernel is the set of $(q+1)$st roots of 1. Also, the kernel of Tr consists of zero and the set of $(q-1)$st roots of -1, the latter being a coset of $F_0^\times$ in $F^\times$.

The Hermitian form on a hyperbolic plane has the form

$$B(x,y) = x_1\overline{y_2} + y_1\overline{x_2}.$$

An arbitrary Hermitian formed space is the orthogonal direct sum of r hyperbolic planes and an anisotropic space. We have seen that, up to scalar multiplication, the following hold:

(a) over $\mathbb{C}$, an anisotropic space is positive definite, and the form can be taken to be
$$B(x,y) = x_1\overline{y_1} + \cdots + x_s\overline{y_s};$$

(b) over a finite field, an anisotropic space has dimension at most one; if non-zero, the form can be taken to be
$$B(x,y) = x\overline{y}.$$

5.1 The unitary groups

Let A be the matrix associated with a non-degenerate Hermitian form B. Then $A = \overline{A}^\top$, and the isometry group of B (the *unitary group* $\mathrm{U}(V,B)$ consists of all invertible matrices P which satisfy $\overline{P}^\top AP = A$.

Since A is invertible, we see that

$$N(\det(P)) = \det(\overline{P}^\top)\det(P) = 1.$$

So $\det(P) \in F_0$. Moreover, a scalar matrix cI lies in the unitary group if and only if $N(c) = c\overline{c} = 1$.

The *special unitary group* $\mathrm{SU}(V,B)$ consists of all elements of the unitary group which have determinant 1 (that is, $\mathrm{SU}(V,B) = \mathrm{U}(V,B) \cap \mathrm{SL}(V)$), and the *projective special unitary group* is the factor group $\mathrm{SU}(V,B)/SU(V,B) \cap Z$, where Z is the group of scalar matrices.

In the case where $F = \mathrm{GF}(q^2)$ is finite, we can unambiguously write $\mathrm{SU}(n,q)$ and $\mathrm{PSU}(n,q)$, since up to scalar multiplication there is a unique Hermitian form on $\mathrm{GF}(q^2)^n$ (with rank $\lfloor n/2 \rfloor$ and germ of dimension 0 or 1 according as n is even or odd). (It would be more logical to write $\mathrm{SU}(n,q^2)$ and $\mathrm{PSU}(n,q^2)$ for these groups; we have used the standard group-theoretic convention.

Proposition 5.1 *(a)* $|\mathrm{U}(n,q) = q^{n(n-1)/2}\prod_{i=1}^{n}(q^i - (-1)^i)$.

(b) $|\mathrm{SU}(n,q)| = |\mathrm{U}(n,q)|/(q+1)$.

(c) $|\mathrm{PSU}(n,q)| = |\mathrm{SU}(n,q)|/d$*, where* $d = (n,q+1)$.

Proof (a) We use Theorem 3.17, with either $n = 2r$, $\varepsilon = -\frac{1}{2}$, or $n = 2r+1$, $\varepsilon = \frac{1}{2}$, and with q replaced by q^2, noting that, in the latter case, $|G_0| = q+1$. It happens that both cases can be expressed by the same formula! On the same theme, note that, if we replace $(-1)^i$ by 1 (and $q+1$ by $q-1$ in parts (b) and (c) of the theorem), we obtain the orders of $\mathrm{GL}(n,q)$, $\mathrm{SL}(n,q)$, and $\mathrm{PSL}(n,q)$ instead.

(b) As we noted, det is a homomorphism from $\mathrm{U}(n,q)$ onto the group of $(q+1)$st roots of unity in $\mathrm{GF}(q^2)^\times$, whose kernel is $\mathrm{SU}(n,q)$.

(c) A scalar cI belongs to $\mathrm{U}(n,q)$ if $c^{q+1} = 1$, and to $\mathrm{SL}(n,q^2)$ if $c^n = 1$. So $|Z \cap \mathrm{SL}(n,q^2)| = d$, as required.

We conclude this section by considering unitary transvections, those which preserve a Hermitian form. Accordingly, let $T : x \mapsto x + (xf)a$ be a transvection, where $af = 0$. We have

$$\begin{array}{rcl} B(xT, yT) & = & B(x + (xf)a, y + (yf)a) \\ & = & B(x,y) + (xf)\overline{B(y,a)} + \overline{(yf)}B(x,a) + (xf)\overline{(yf)}B(a,a). \end{array}$$

So T is unitary if and only if the last three terms vanish for all x, y. Putting $y = a$ we see that $(xf)\overline{B(a,a)} = 0$ for all x, whence (since $f \neq 0$) we must have $B(a,a) = 0$. Now choosing y such that $B(y,a) = 1$ and setting $\lambda = \overline{(yf)}$, we have $xf = \lambda B(x,a)$ for all x. So a unitary transvection has the form

$$x \mapsto x + \lambda B(x,a)a,$$

where $B(a,a) = 0$. In particular, an anisotropic space admits no unitary transvections. Also, choosing x and y such that $B(x,a) = B(y,a) = 1$, we find that $\mathrm{Tr}(\lambda) = 0$. Conversely, for any $\lambda \in \ker(\mathrm{Tr})$ and any a with $B(a,a) = 0$, the above formula defines a unitary transvection.

5.2 Hyperbolic planes

In this section only, we use the convention that $\mathrm{U}(2, F_0)$ means the unitary group associated with a hyperbolic plane over F, and σ is the associated field automorphism, having fixed field F_0.

Theorem 5.2 $\mathrm{SU}(2, F_0) \cong \mathrm{SL}(2, F_0)$.

Proof We will show, moreover, that the actions of the unitary group on the polar space and that of the special linear group on the projective space correspond, and that unitary transvections correspond to transvections in $\mathrm{SL}(2, F_0)$. Let $K = \{c \in F : c + \overline{c} = 0\}$ be the kernel of the trace map; recall that the image of the trace map is F_0.

With the standard hyperbolic form, we find that a unitary matrix

$$P = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

must satisfy $\overline{P}^\top AP = A$, where

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Hence

$$a\overline{c}+\overline{a}c=0,\quad b\overline{c}+\overline{a}d=1,\quad b\overline{d}+\overline{b}d=0.$$

In addition, we require, that $\det(P)=1$, that is, $ad-bc=1$.

From these equations we deduce that $b+\overline{b}=c+\overline{c}=0$, that is, $b,c\in K$, while $a-\overline{a}=d-\overline{d}=0$, that is, $a,d\in F_0$.

Choose a fixed element $u\in K$. Then $\lambda\in K$ if and only if $u\lambda\in F_0$. Also, $u^{-1}\in K$. Hence the matrix

$$P^{\dagger}=\begin{pmatrix} a & ub \\ u^{-1}c & d \end{pmatrix}$$

belongs to $\mathrm{SL}(2,F_0)$. Conversely, any matrix in $\mathrm{SL}(2,F_0)$ gives rise to a matrix in $\mathrm{SU}(2,F_0)$ by the inverse map. So we have a bijection between the two groups. It is now routine to check that the map is an isomorphism.

Represent the points of the projective line over F by $F\cup\{\infty\}$ as usual. Recall that ∞ is the point (rank 1 subspace) spanned by $(0,1)$, while c is the point spanned by $(1,c)$. We see that ∞ is flat, while c is flat if and only if $c+\overline{c}=0$, that is, $c\in K$. So the map $x\mapsto x$ takes the polar space for the unitary group onto the projective line over F_0. It is readily checked that this map takes the action of the unitary group to that of the special linear group.

By transitivity, it is enough to consider the unitary transvections $x\mapsto x+\lambda B(x,a)a$, where $a=(0,1)$. In matrix form, these are

$$P=\begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix},$$

with $\lambda\in K$. Then

$$P^{\dagger}=\begin{pmatrix} 1 & u\lambda \\ 0 & 1 \end{pmatrix},$$

which is a transvection in $\mathrm{SL}(2,F_0)$, as required. ■

In particular, we see that $\mathrm{PSU}(2,F_0)$ is simple if $|F_0|>3$.

5.3 Generation and simplicity

We follow the now-familiar pattern. First we treat two exceptional finite groups, then we show that unitary groups are generated by unitary transvections and that most are simple. By the preceding section, we may assume that the rank is at least 3.

The finite unitary group $\mathrm{PSU}(3,q)$ is a 2-transitive group of permutations of the q^3+1 points of the corresponding polar space (since any two such points are spanned by a hyperbolic pair) and has order $(q^3+1)q^3(q^2-1)/d$, where $d = (3,q+1)$. Moreover, any two points span a line containing $q+1$ points of the polar space. The corresponding geometry is called a *unital*.

For $q=2$, the group has order 72, and so is soluble. In fact, it is *sharply 2-transitive*: a unique group element carries any pair of points to any other.

Exercise 5.1 (a) Show that the unital associated with $\mathrm{PSU}(3,2)$ is isomorphic to the *affine plane* over $\mathrm{GF}(3)$, defined as follows: the points are the vectors in a vector space V of rank 2 over $\mathrm{GF}(3)$, and the lines are the cosets of rank 1 subspaces of V (which, over the field $\mathrm{GF}(3)$, means the triples of vectors with sum 0).

(b) Show that the automorphism group of the unital has the structure $3^2:\mathrm{GL}(2,3)$, where 3^2 denotes an elementary abelian group of this order (the translation group of V) and : denotes semidirect product.

(c) Show that $\mathrm{PSU}(3,2)$ is isomorphic to $3^2:Q_8$, where Q_8 is the quaternion group of order 8.

(d) Show that $\mathrm{PSU}(3,2)$ is not generated by unitary transvections.

We next consider the group $\mathrm{PSU}(4,2)$, and outline the proof of the following theorem:

Theorem 5.3 $\mathrm{PSU}(4,2)\cong\mathrm{PSp}(4,3)$.

Proof Observe first that both these groups have order 25 920. We will construct a geometry for the group $\mathrm{PSU}(4,2)$, and use the technical results of Section 4.4 to identify it with the generalised quadrangle for $\mathrm{PSp}(4,3)$. Now it has index 2 in the full automorphism group of this geometry, as also does $\mathrm{PSp}(4,3)$, which is simple; so these two groups must coincide.

The geometry is constructed as follows. Let V be a vector space of rank 4 over $\mathrm{GF}(4)$ carrying a Hermitian form of polar rank 2. The projective space $\mathrm{PG}(3,4)$ derived from V has $(4^4-1)/(4-1)=85$ points, of which $(4^2-1)(4^{3/2}+1)/(4-1)=45$ are points of the polar space, and the remaining 40 are points on which the form does not vanish (spanned by vectors x with $B(x,x)=1$). Note that $40=(3^4-1)/(3-1)$ is equal to the number of points of the symplectic generalised quadrangle over $\mathrm{GF}(3)$. Let Ω denote this set of 40 points.

Define an F-line to be a set of four points of Ω spanned by the vectors of an orthonormal basis for V (a set of four vectors x_1, x_2, x_3, x_4 with $B(x_i, x_i) = 1$ and $B(x_i, x_j) = 0$ for $i \neq j$). Note that two orthogonal points p, q of Ω span a non-degenerate 2-space, which is a line containing five points of the projective space of which three are flat and the other two belong to Ω. Then $\{p, q\}^\perp$ is also a non-degenerate 2-space containing two points of Ω, which complete $\{p, q\}$ to an F-line. Thus, two orthogonal points lie on a unique F-line, while two non-orthogonal points lie on no F-line. It is readily checked that, if $L = \{p_1, p_2, p_3, p_4\}$ is an F-line and q is another point of Ω, then p has three non-zero coordinates in the orthonormal basis corresponding to L, so q is orthogonal to a unique point of L. Thus, the points of Ω and the F-lines satisfy condition (a) of Section 4.4; that is, they form a generalised quadrangle.

Now consider two points of Ω which are not orthogonal. The 2-space they span is degenerate, with a radical of rank 1. So of the five points of the corresponding projective line, four lie in Ω and one (the radical) is flat. Sets of four points of this type (which are obviously determined by any two of their members) will be the H-lines. It is readily checked that the H-lines do indeed arise in the manner described in Section 4.4, that is, as the sets of points of Ω orthogonal to two given non-orthogonal points. So condition (b) holds.

Now a point p of Ω lies in four F-lines, whose union consists of thirteen points. If q and r are two of these points which do not lie on an F-line with p, then q and r cannot be orthogonal, and so they lie in an H-line; since p and q are orthogonal to p, so are the remaining points of the H-line containing them. Thus we have condition (c). Now (d) is easily verified by counting, and the proof is complete. ∎

Exercise 5.2 (a) Give a detailed proof of the above isomorphism.

(b) If you are familiar with a computer algebra package, verify computationally that the above geometry for $\mathrm{PSU}(4, 2)$ is isomorphic to the symplectic generalised quadrangle for $\mathrm{PSp}(4, 3)$.

In our generation and simplicity results we treat the rank 3 case separately. In the rank 3 case, the unitary group is 2-transitive on the points of the unital.

Theorem 5.4 *Let (V, B) be a unitary formed space of Witt rank* 1*, with* $\mathrm{rk}(V) = 3$. *Assume that the field F is not* $\mathrm{GF}(2^2)$.

(a) $\mathrm{SU}(V, B)$ *is generated by unitary transvections.*

(b) $\mathrm{PSU}(V, B)$ *is simple.*

Proof We exclude the case of $\mathrm{PSU}(3,2)$ (with $F = \mathrm{GF}(2^2)$, considered earlier. Replacing the form by a scalar multiple if necessary, we assume that the germ contains vectors of norm 1. Take such a vector as second basis vector, where the first and third are a hyperbolic pair. That is, we assume that the form is

$$B((x_1,x_2,x_3),(y_1,y_2,y_3)) = x_1\overline{y_3} + x_2\overline{y_2} + x_3\overline{y_1},$$

so the isometry group is

$$\{P : \overline{P}^\top AP = A\}$$

where

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Now we check that the group

$$Q = \left\{ \begin{pmatrix} 1 & -\overline{a} & b \\ 0 & 1 & a \\ 0 & 0 & 1 \end{pmatrix} : N(a) + \mathrm{Tr}(b) = 0 \right\}$$

is a subgroup of $G = \mathrm{SU}(V,B)$, and its derived group consists of unitary transvections (the elements with $a = 0$).

Next we show that the subgroup T of V generated by the transvections in G is transitive on the set of vectors x such that $B(x,x) = 1$. Let x and y be two such vectors. Suppose first that $\langle x,y \rangle$ is nondegenerate. Then it is a hyperbolic line, and a calculation in $\mathrm{SU}(2,F_0)$ gives the result. Otherwise, there exists z such that $\langle x,z \rangle$ and $\langle y,z \rangle$ are nondegenerate, so we can get from x to y in two steps.

Now the stabiliser of such a vector in G is $\mathrm{SU}(x^\perp,B) = \mathrm{SU}(2,F_0)$, which is generated by transvections; and every coset of this stabiliser contains a transvection. So G is generated by transvections.

Now it follows that the transvections lie in G', and Iwasawa's Lemma (Theorem 2.7) shows that $\mathrm{PSL}(V,B)$ is simple. ■

Exercise 5.3 Complete the details in the above proof by showing

(a) the group $\mathrm{SU}(2,F_0)$ acts transitively on the set of vectors of norm 1 in the hyperbolic plane;

(b) given two vectors x,y of norm 1 in a rank 3 unitary space as in the proof, either $\langle x,y \rangle$ is a hyperbolic plane, or there exists z such that $\langle x,z \rangle$ and $\langle y,z \rangle$ are hyperbolic planes.

Theorem 5.5 *Let (V,B) be a unitary formed space with Witt rank at least 2. Then*

(a) $\mathrm{SU}(V,B)$ *is generated by unitary transvections.*

(b) $\mathrm{PSU}(V,B)$ *is simple.*

Proof We follow the usual pattern. The argument in the preceding theorem shows part (a) without change if $F \neq \mathrm{GF}(4)$. In the excluded case, we know that $\mathrm{PSU}(4,2) \cong \mathrm{PSp}(4,3)$ is simple, and so is generated by any conjugacy class (in particular, the images of the transvections of $\mathrm{SU}(4,2)$). Then induction shows the result for higher rank spaces over $\mathrm{GF}(4)$. Again, the argument in 3 dimensions shows that transvections are commutators; the action on the points of the polar space is primitive; and so Iwasawa's Lemma shows the simplicity. ∎