# On the Determinant of Symplectic Matrices*

D. Steven Mackey[†] Niloufer Mackey[†]

February 22, 2003

**Abstract**

A collection of new and old proofs showing that the determinant of any symplectic matrix is +1 is presented. Structured factorizations of symplectic matrices play a key role in several arguments. A constructive derivation of the symplectic analogue of the Cartan-Dieudonné theorem is one of the new proofs in this essay.

**Key words.** symplectic, determinant, bilinear form, skew-symmetric, structure-preserving factorizations, symmetries, transvections, $\mathbb{G}$-reflector, Pfaffian.

**AMS subject classification.** 15A21, 15A23, 15A57, 15A63, 65F30

## 1 Introduction

This essay gathers together several proofs — some previously known, and some original — showing that the determinant of a symplectic matrix is always +1. The proofs will be presented only for real and complex symplectics, although the result is true for symplectic matrices with entries from any field. While some of the arguments can be adapted to general fields, only the proof using Pfaffians in Section 5 holds as written for any field.

**Definition 1.1.** A $2n \times 2n$ matrix $S$ with entries in the field $\mathbb{K}$ is said to be *symplectic* if $S^T J S = J$, where $J \stackrel{\text{def}}{=} \left[\begin{smallmatrix} 0 & I_n \\ -I_n & 0 \end{smallmatrix}\right]$. The set of all symplectic matrices over $\mathbb{K}$ is denoted by $Sp(2n, \mathbb{K})$.

Symplectic matrices can also be viewed as automorphisms of the bilinear form determined by the matrix $J$, that is $\langle x, y \rangle \stackrel{\text{def}}{=} x^T J y$. Recall that an automorphism of a bilinear form on $\mathbb{K}^n$ is a matrix $A$ such that $\langle Ax, Ay \rangle = \langle x, y \rangle$ for all $x, y \in \mathbb{K}^n$. It follows that the set of all automorphisms of any fixed non-degenerate form[1] is a multiplicative group. Other examples of automorphism groups arising from bilinear forms are the orthogonal and pseudo-orthogonal groups $O(n, \mathbb{K})$ and $O(p, q, \mathbb{K})$.

*Numerical Analysis Report No. 422, Manchester Centre for Computational Mathematics, Manchester, England, February 2003. This work was supported by Engineering and Physical Sciences Research Council Visiting Fellowship GR/S15563/01.

[†]Department of Mathematics, Western Michigan University, Kalamazoo MI 49008, USA. email: steve.mackey@wmich.edu, nil.mackey@wmich.edu

[1]A bilinear form $\langle x, y \rangle$ is non-degenerate iff $\{\langle x, y \rangle = 0, \forall\, y \Rightarrow x = 0\}$ and $\{\langle x, y \rangle = 0, \forall\, x \Rightarrow y = 0\}$ .

It is easy to see directly from Definition 1.1 that the determinant of any symplectic matrix (or more generally, the determinant of any automorphism of any non-degenerate bilinear form) has to be either $+1$ or $-1$.

$$\det(S^T J S) = \det J \quad \Rightarrow \quad (\det S)^2 \det J = \det J \quad \Rightarrow \quad \det S = \pm 1. \tag{1.1}$$

What is not obvious is why a determinant of $-1$ is never realized by any symplectic matrix, no matter what the field! This is especially surprising in view of the situation for matrices in other automorphism groups like $O(n, \mathbb{K})$ and $O(p, q, \mathbb{K})$, where both $+1$ and $-1$ determinants are easy to find. The aim of this essay is to shine some light on this unexpected result from various angles, hoping to demystify it to some degree. Our focus then, is the following theorem, with the field $\mathbb{K}$ restricted to $\mathbb{R}$ or $\mathbb{C}$.

**Theorem 1.2.** *Let $S \in Sp(2n, \mathbb{K})$. Then $\det S = +1$.*

A strategy common to many of the proofs is to factor a general symplectic matrix into a finite product of simpler symplectic matrices, each sufficiently simple that one can easily see that their determinants are all $+1$. Indeed, this essay might well have been entitled "Structured Factorizations of Symplectic Matrices" instead, with very little inaccuracy.

# 2 Two Proofs Using Structured Polar Decomposition

If $S = QP$ is the polar decomposition of a symplectic matrix $S$, then by the structured polar decomposition theorem in [8], both $Q$ and $P$ are necessarily symplectic. Hence by (1.1), $\det Q = \pm 1$ and $\det P = \pm 1$. But $\det P$ must be $+1$, since $P$ is positive definite. The issue therefore reduces to proving the following proposition.

**Proposition 2.1.** *Let $Q$ be any real symplectic orthogonal or complex symplectic unitary matrix. Then $\det Q = +1$.*

Two ways of proving Proposition 2.1 are presented when $Q \in SpO(2n, \mathbb{R})$, the group of real $2n \times 2n$ symplectic orthogonal matrices. As far as we know these methods do not generalize to complex symplectic unitary matrices, and a proof for this case is postponed until Section 3.1. We begin by observing that any $Q \in SpO(2n, \mathbb{R})$ commutes with $J$, and hence has the block form $Q = \left[\begin{smallmatrix} A & B \\ -B & A \end{smallmatrix}\right]$. Observe that it suffices to show $\det Q > 0$ in order to conclude $\det Q = +1$.

**First Approach** In [6], relationships between a number of classes of structured complex matrices and doubly-structured real matrices are described, together with correspondences between their canonical forms. One part of this story is the following connection between the group of $n \times n$ complex unitary matrices $U(n)$, and $SpO(2n, \mathbb{R})$. Let $A + iB$ with $A, B \in \mathbb{R}^{n \times n}$ denote an $n \times n$ complex unitary matrix. Then the map

$$\begin{array}{rcl} U(n) & \longrightarrow & SpO(2n, \mathbb{R}) \\ A + iB & \longmapsto & \left[\begin{smallmatrix} A & B \\ -B & A \end{smallmatrix}\right] \end{array}$$

is a group isomorphism [6]. Now any unitary matrix $A + iB$ is normal, and hence unitarily similar to $D_1 + iD_2$, where $D_1$ and $D_2$ are real $n \times n$ diagonal matrices. Using this isomorphism we may now conclude that any real symplectic orthogonal matrix $Q = \left[\begin{smallmatrix} A & B \\ -B & A \end{smallmatrix}\right]$ is similar via a real symplectic orthogonal similarity to $D = \left[\begin{smallmatrix} D_1 & D_2 \\ -D_2 & D_1 \end{smallmatrix}\right]$. But $D$ is permutation similar to a direct sum of $2 \times 2$ matrices of the form $\left[\begin{smallmatrix} c & d \\ -d & c \end{smallmatrix}\right]$, with determinant $c^2 + d^2 > 0$. Thus

$$\det Q \;=\; \det \begin{bmatrix} A & B \\ -B & A \end{bmatrix} \;=\; \det \begin{bmatrix} D_1 & D_2 \\ -D_2 & D_1 \end{bmatrix} \;>\; 0. \qquad \square$$

**Second Approach** This argument can be found in [11]. Imitating the complex diagonalization of a real $2\times2$ matrix $\left[\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right]$, the real $2n\times2n$ matrix $\left[\begin{smallmatrix} A & B \\ -B & A \end{smallmatrix}\right]$ can be block-diagonalized to a complex $2n \times 2n$ matrix as follows:

$$\frac{1}{\sqrt{2}} \begin{bmatrix} I_n & -iI_n \\ -iI_n & I_n \end{bmatrix} \begin{bmatrix} A & B \\ -B & A \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} I_n & iI_n \\ iI_n & I_n \end{bmatrix} \;=\; \begin{bmatrix} A+iB & 0 \\ 0 & A-iB \end{bmatrix}.$$

Thus we have

$$\begin{aligned} \det Q \;=\; \det \begin{bmatrix} A & B \\ -B & A \end{bmatrix} \;=\; \det \begin{bmatrix} A+iB & 0 \\ 0 & A-iB \end{bmatrix} \;&=\; \det(A+iB)\ \det(A-iB) \\ &=\; \det(A+iB)\ \overline{\det(A+iB)} \;>\; 0. \end{aligned}$$

Note that $A - iB = \overline{A+iB}$ because $A$ and $B$ are real matrices. $\square$

## 3 Proof Using Structured $\mathbf{Q\widehat{R}}$ Decomposition

A constructive proof of a symplectic QR-like decomposition for any $2n \times 2n$ real or complex symplectic matrix is now used to prove Theorem 1.2.

**Proposition 3.1.** *For any $S \in Sp(2n, \mathbb{K})$ there exists a factorization $S = Q\widehat{R}$, where $Q$ is symplectic and unitary with $\det Q = +1$, $\widehat{R} = \left[\begin{smallmatrix} R & Z \\ 0 & R^{-T} \end{smallmatrix}\right]$ is symplectic, and $R$ is $n \times n$ upper triangular. If $S$ is real, then $Q$ and $R$ can also be chosen to be real.*

Clearly $\det \widehat{R} = +1$, and since $\det Q$ will be $+1$ by construction, $\det S$ is forced to be $+1$, thus establishing Theorem 1.2. For conjugate symplectic matrices, that is, $S \in \mathbb{C}^{2n\times2n}$ such that $S^*JS = J$, a QR-like decomposition similar to the one given here can be found in [2], [3]. However, the determinant of a conjugate symplectic matrix can be any number on the unit circle in the complex plane; this can be seen by considering the conjugate symplectic matrices $e^{i\theta}I_{2n}$, where $\theta \in \mathbb{R}$. So the result of Theorem 1.2 does not extend to conjugate symplectic matrices.

The construction is presented only for complex symplectic matrices $S$; we leave it to the reader to check that the argument goes through in the real case in a similar fashion.

As usual in a $QR$-like decomposition algorithm, we start by reducing the first column $\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]$ of $S$ to a scalar multiple of $e_1$. (Here and in the following, $x, y, z$ and $w$ will denote vectors in $\mathbb{C}^n$.) It is important that we preserve the symplectic structure, so we do this reduction using only tools that are symplectic as well as unitary[2]. This can be done in the following three-step process:

$$\begin{bmatrix} x \\ y \end{bmatrix} \xrightarrow[(a)]{H_1} \begin{bmatrix} z \\ \beta e_1 \end{bmatrix} \xrightarrow[(b)]{G_1} \begin{bmatrix} w \\ 0 \end{bmatrix} \xrightarrow[(c)]{K_1} \begin{bmatrix} \alpha e_1 \\ 0 \end{bmatrix}. \tag{3.1}$$

[2]See [10] for further details on symplectic and symplectic unitary tools.

**Step(a):** Premultiply by $H_1 = \left[\begin{smallmatrix} \overline{U}_1 & 0 \\ 0 & U_1 \end{smallmatrix}\right]$, where $U_1$ is any $n \times n$ unitary reflector that maps $y$ to $\beta e_1 \in \mathbb{C}^n$. It does not matter what the polar angle of $\beta$ is, so there is no need to make any special choices to keep control of it. The matrix $H_1$ is symplectic and unitary, and $\det H_1 = \det \overline{U}_1 \det U_1 = e^{-i\theta} e^{i\theta} = +1$.

**Step(b):** Design a $2 \times 2$ unitary Givens rotation, $G = \left[\begin{smallmatrix} c & s \\ -\overline{s} & \overline{c} \end{smallmatrix}\right]$, where $\det G = |c|^2 + |s|^2 = 1$, to map $\left[\begin{smallmatrix} z_1 \\ \beta \end{smallmatrix}\right] \in \mathbb{C}^2$ to $\left[\begin{smallmatrix} w_1 \\ 0 \end{smallmatrix}\right]$. Symplectically embed $G$ in rows $1$, $n+1$ of $I_{2n}$ to obtain a symplectic unitary matrix $G_1$ with $\det G_1 = +1$. Once again, we do not need to control the polar angle of $w_1$.

**Step(c):** Premultiply by $K_1 = \left[\begin{smallmatrix} U_2 & 0 \\ 0 & \overline{U}_2 \end{smallmatrix}\right]$, where $U_2$ is an $n \times n$ unitary reflector such that $U_2 w = \alpha e_1$. Once again, we do not need to keep control of the polar angle of $\alpha$. Observe that $K_1$ is symplectic and unitary with $\det K_1 = +1$.

We now pause to make a crucial observation on the consequences of having preserved the symplectic structure. It is worth pointing out that the following lemma applies to automorphisms of any scalar product[3], not just those in the symplectic groups. We use $A^\star$ to denote the adjoint of $A$ with respect to the scalar product under discussion. Recall that $A^\star$ is the unique matrix such that $\langle Ax, y\rangle = \langle x, A^\star y\rangle$ for all vectors $x$, $y$. For more on adjoints and automorphisms see [7] and [10].

**Lemma 3.2.** *Suppose $A$ is an automorphism of a scalar product $\langle\cdot,\cdot\rangle$ on $\mathbb{K}^n$, and the first column of $A$ is $\alpha e_1$ for some nonzero $\alpha \in \mathbb{K}$. Then the first column of the adjoint $A^\star$ with respect to $\langle\cdot,\cdot\rangle$ is also a scalar multiple of $e_1$.*

*Proof.* It follows from the definition of adjoint that $A^\star A = I$ when $A$ is an automorphism. Thus

$$Ae_1 = \alpha e_1 \quad \Rightarrow \quad A^\star A e_1 = \alpha A^\star e_1 \quad \Rightarrow \quad e_1 = \alpha A^\star e_1 \quad \Rightarrow \quad A^\star e_1 = \alpha^{-1} e_1. \qquad \square$$

*Remark* 3.1. This proof can be used, mutatis mutandis, to show that if the $j$th column (row) of an automorphism $A$ is a scalar multiple of $e_k$ ($e_k^T$), then the $k$th column (row) of $A^\star$ is a scalar multiple of $e_j$ ($e_j^T$).

For symplectic matrices, the adjoint can be expressed in block form: if $A = \left[\begin{smallmatrix} E & F \\ G & H \end{smallmatrix}\right]$ where $E$, $F$, $G$, $H \in \mathbb{C}^{n\times n}$, then $A^\star = \left[\begin{smallmatrix} H^T & -F^T \\ -G^T & E^T \end{smallmatrix}\right]$. Now if the first column of $A$ is $\alpha e_1$, then by Lemma 3.2 the first column of $H^T$ is $e_1/\alpha$ and the first column of $G^T$ is 0. Equivalently, in $A$ the first row of $G$ is 0 and the first row of $H$ is $e_1^T/\alpha$. Thus our structure-preserving (i.e. symplectic) three-step reduction of the first column of $S$ results in a matrix of the form

$$Q_1 S \;=\; (K_1 G_1 H_1) S \;=\; \begin{bmatrix} \alpha & * & * & * \\ 0 & \boxed{L} & * & \boxed{M} \\ 0 & 0 & \alpha^{-1} & 0 \\ 0 & \boxed{N} & * & \boxed{P} \end{bmatrix}, \quad L, M, N, P \in \mathbb{C}^{(n-1)\times(n-1)} \tag{3.2}$$

with many more zeroes than were directly targeted by the reduction process. Furthermore it can be shown that the submatrix $\widetilde{S} = \left[\begin{smallmatrix} L & M \\ N & P \end{smallmatrix}\right]$ forms a $(2n-2) \times (2n-2)$ *symplectic*

[3]By a scalar product we mean any non-degenerate bilinear or sesquilinear form on $\mathbb{K}^n$.

matrix, and the result now follows by induction: the inductive hypothesis provides a $(2n-2)\times(2n-2)$ symplectic unitary matrix $\widetilde{Q}_2$ such that $\det\widetilde{Q}_2=+1$ and $\widetilde{Q}_2\widetilde{S}=\left[\begin{smallmatrix}\widetilde{R} & \widetilde{Z}\\ 0 & \widetilde{R}^{-T}\end{smallmatrix}\right]$, where $\widetilde{R}$ is upper triangular. Then with $Q_2$ defined to be the symplectic embedding of $\widetilde{Q}_2$ into rows and columns 2 through $n$, and $n+2$ through $2n$ of $I_{2n}$, we have $\det Q_2=1$, and

$$Q_2\,Q_1S \;=\; \begin{bmatrix} \alpha & * & * & * \\ 0 & \widetilde{R} & * & \widetilde{Z} \\ 0 & 0 & \alpha^{-1} & 0 \\ 0 & 0 & * & \widetilde{R}^{-T} \end{bmatrix} \;=\; \begin{bmatrix} R & Z \\ 0 & R^{-T} \end{bmatrix} \;=\; \widehat{R},$$

giving the desired factorization $S=Q\widehat{R}$, where $Q=Q_1^*\,Q_2^*$ is symplectic unitary with $\det Q=+1$. □

*Remark* 3.2. A modification of Step (c) can ensure that the diagonal entries of the upper triangular matrix $R$ in Proposition 3.1 are all positive. The unitary reflector $U_2$ can always be designed so that $U_2w=\alpha e_1$, with $\alpha>0$. For details, see [5] or [7, Section 8.2].

*Remark* 3.3. A matrix $\widehat{R}=\left[\begin{smallmatrix} R & Z\\ 0 & R^{-T}\end{smallmatrix}\right]$ where $R$ is $n\times n$ upper triangular will be said to be *quasi-upper triangular*, or quasi-triangular for short.

### 3.1 $\mathbf{Q\widehat{R}}$ Decomposition of Complex Symplectic Unitary Matrices

Suppose that we apply the construction described in Section 3 to reduce a matrix $S$ that is both complex symplectic and unitary to quasi-upper triangular form. Because the transformations used are symplectic and unitary, the matrix iterates will remain symplectic and unitary throughout the reduction process.

Consider the result of reducing the first column of $S$ to $\alpha e_1$. Because the reduced matrix is unitary, we must have $|\alpha|=1$, and so the first row is forced to be $\alpha e_1^T$. But because the reduced matrix is also symplectic, it must have the form shown in (3.2). This in turn forces the $(n+1)$th column to be $\alpha^{-1}e_{n+1}$. Thus the double-structure-preserving reduction of the first column of $S$ results in

$$Q_1S \;=\; \begin{bmatrix} e^{i\theta} & 0 & 0 & 0 \\ 0 & \boxed{L} & 0 & \boxed{M} \\ 0 & 0 & e^{-i\theta} & 0 \\ 0 & \boxed{N} & 0 & \boxed{P} \end{bmatrix}. \tag{3.3}$$

Inductively continuing the reduction process on the symplectic unitary submatrix $\left[\begin{smallmatrix} L & M\\ N & P\end{smallmatrix}\right]$ now leads to $\widehat{R}=\left[\begin{smallmatrix} D & 0\\ 0 & D^{-1}\end{smallmatrix}\right]$ that is *diagonal* rather than merely quasi-triangular, and $\det\widehat{R}=+1$ is even more obvious than before.

Thus we now have a proof of Proposition 2.1 for complex symplectic unitary matrices: any matrix of this type can be factored into a product of symplectic double Householders, embedded symplectic Givens, and a diagonal symplectic unitary, and each of these factors has determinant $+1$. This now completes the polar decomposition proof of Section 2 for the complex case.

## 4 Proofs Using Symplectic $\mathbb{G}$-reflectors

In this section we exploit the mapping properties of symplectic $\mathbb{G}$-reflectors developed in [7], to give two proofs that the determinant of any real or complex symplectic matrix is $+1$. Symplectic $\mathbb{G}$-reflectors (called symplectic transvections in [1], [4]) are elementary transformations, i.e. rank-one modifications of the identity, that are also symplectic. In a certain sense they are the simplest kind of symplectic transformation, since they act as the identity on a hyperplane. Indeed, one may equivalently define symplectic $\mathbb{G}$-reflectors to be those $2n \times 2n$ symplectic matrices that have a $(2n-1)$-dimensional fixed-point subspace. In [7] it is shown that any symplectic $\mathbb{G}$-reflector can be expressed in the form

$$G = I + \beta uu^T J, \quad \text{for some } \ 0 \neq \beta \in \mathbb{K}\,, \ \ 0 \neq u \in \mathbb{K}^{2n}, \tag{4.1}$$

and conversely, any $G$ given by (4.1) is always a symplectic $\mathbb{G}$-reflector.

The first step is to show that the determinant of any symplectic $\mathbb{G}$-reflector is $+1$. Three proofs of this fact are given in Section 4.1. The mapping capabilities of symplectic $\mathbb{G}$-reflectors are next developed in Section 4.2. Then in Section 4.3, $\mathbb{G}$-reflectors replace the double Householders and embedded Givens transformations used in the algorithm of Section 3 to reduce a symplectic matrix to quasi-triangular form. This reduction results in another proof of the determinant result very much in the spirit of the proof in Section 3.

Finally, in Section 4.4 a constructive argument shows that symplectic $\mathbb{G}$-reflectors are building blocks for the entire symplectic group — every $S \in Sp(2n, \mathbb{K})$ can be expressed as a finite product of symplectic $\mathbb{G}$-reflectors. Since every symplectic $\mathbb{G}$-reflector has $+1$ determinant, it follows that the same must be true for a general symplectic matrix.

### 4.1 Determinant of Symplectic $\mathbb{G}$-reflectors

We give three proofs that the determinant of any symplectic $\mathbb{G}$-reflector is $+1$. Two of these proofs use the notion of isotropic vector: a nonzero vector $x \in \mathbb{K}^n$ is *isotropic* with respect to some scalar product $\langle\cdot,\cdot\rangle$ if $\langle x, x\rangle = 0$. In this case, because $J$ is skew-symmetric and the scalar product is bilinear, $\langle x, x\rangle \overset{\text{def}}{=} x^T J x \equiv 0$, so every nonzero vector is isotropic.

**Lemma 4.1.** *Suppose $G \in Sp(2n, \mathbb{K})$ is a $\mathbb{G}$-reflector. Then* $\det G = +1$.

**Proof by continuity:** Let $G = I + \beta uu^T J$ be an arbitrary symplectic $\mathbb{G}$-reflector. Consider the continuous path of matrices given by $G(t) = I + (1-t)\beta uu^T J$, with $0 \leq t \leq 1$. Note that $G(0) = G$. Now $G(t)$ is a symplectic $\mathbb{G}$-reflector for $0 \leq t < 1$, so $\det G(t) = \pm 1$ for all $t < 1$. But $\lim_{t\to 1} G(t) = I$, so by continuity $\det G(t) = +1$ for all $t$, in particular for $t = 0$. $\square$

**Proof by squaring:** This argument can be found in Artin [1]. Observe that any symplectic $\mathbb{G}$-reflector is the *square* of another symplectic $\mathbb{G}$-reflector. In particular, since $u^T J u = 0$ for all $u \in \mathbb{K}^{2n}$, we have

$$G \ = \ I + \beta uu^T J \ = \ (I + \tfrac{1}{2}\,\beta uu^T J)^2 \ = \ S^2.$$

Thus $\det G = \det(S^2) = (\det S)^2 = (\pm 1)^2 = 1$. $\square$

**Proof by eigenvalues:** In [7, Proposition 6.2] it is shown quite generally that whenever the vector $u$ in the formula for a $\mathbb{G}$-reflector (from any automorphism group) is isotropic, then the $\mathbb{G}$-reflector is non-diagonalizable. For a symplectic $\mathbb{G}$-reflector $G = I + \beta uu^T J$, the vector $u \in \mathbb{K}^{2n}$ is always isotropic, and thus $G$ is always non-diagonalizable. Now since $G$ is a $\mathbb{G}$-reflector, it acts as the identity on a hyperplane, so $G$ has eigenvalue $\lambda = 1$ with geometric multiplicity at least $2n - 1$. If $G$ had any eigenvalue other than $\lambda = 1$, then $G$ would be diagonalizable. Thus a symplectic $\mathbb{G}$-reflector $G$ has only the eigenvalue $+1$, and hence $\det G = +1$. $\square$

### 4.2 Mapping Properties of Symplectic $\mathbb{G}$-reflectors

The next theorem is a special case of a more general result proved in [7, Theorem 8.2] for automorphism groups arising from a large class of scalar products. Its statement has been specialized here for the case of the real and complex symplectic groups, where $\langle x, y\rangle = x^T Jy$.

**Theorem 4.2 (Symplectic $\mathbb{G}$-reflector mapping theorem).**
*Let $x$, $y$ be distinct nonzero vectors in $\mathbb{K}^{2n}$. Then there exists a symplectic $\mathbb{G}$-reflector $G$ such that $Gx = y$ if and only if $\langle y, x\rangle \neq 0$. Furthermore, if $G$ exists then it is unique, and can be expressed as*

$$G = I + \frac{1}{\langle y, x\rangle} uu^T J \quad \text{where} \quad u = y - x. \tag{4.2}$$

For the factorizations in Section 4.3 and Section 4.4, we will need to map a vector $x$ to $e_1$ by symplectic $\mathbb{G}$-reflectors. Building on Theorem 4.2, the next lemma shows that this can always be done, although it may sometimes require two symplectic $\mathbb{G}$-reflectors to accomplish the task.

**Lemma 4.3 (Symplectic two reflector mapping property).**
*Any nonzero $x \in \mathbb{K}^{2n}$ can be mapped to $e_1$ by a product of at most two symplectic $\mathbb{G}$-reflectors.*

*Proof.* Let $0 \neq x \in \mathbb{K}^{2n}$. Since $\langle e_1, x\rangle = e_1^T Jx = x_{n+1}$, we conclude from Theorem 4.2 that $x$ can be mapped to $e_1$ by a single symplectic $\mathbb{G}$-reflector if $x_{n+1} \neq 0$. On the other hand, if $x_{n+1} = 0$, then we can get to $e_1$ by a composition of two symplectic $\mathbb{G}$-reflectors: send $x$ to some vector $y$ with $y_{n+1} \neq 0$, and follow by mapping $y$ to $e_1$. There are several cases to consider.

**Case 1.** $(x_{n+1} = 0,\ x_1 \neq 0)$
In this case we can map $x$ to $y = e_{n+1}$, since $\langle y, x\rangle = e_{n+1}^T Jx = -x_1 \neq 0$.

**Case 2.** $(x_{n+1} = 0,\ x_1 = 0,$ and $x_j \neq 0$, for some $j$ with $2 \leq j \leq n)$
Here we can map $x$ to $y = e_{n+1} + e_{n+j}$, since $\langle y, x\rangle = (e_{n+1} + e_{n+j})^T Jx = -x_j \neq 0$.

**Case 3:** $(x_{n+1} = 0,\ x_1 = 0,$ and $x_{n+j} \neq 0$, for some $j$ with $2 \leq j \leq n)$
In this case we can map $x$ to $y = e_j + e_{n+1}$, since $\langle y, x\rangle = (e_j + e_{n+1})^T Jx = x_{n+j} \neq 0$.

Thus we have shown that any nonzero $x \in \mathbb{K}^{2n}$ with $x_{n+1} = 0$ can be mapped by a symplectic $\mathbb{G}$-reflector to $y$ with $y_{n+1} \neq 0$ (in fact $y_{n+1} = 1$); then by Theorem 4.2, $y$ can be mapped to $e_1$ by a second symplectic $\mathbb{G}$-reflector. $\square$

*Remark* 4.1. Lemma 4.3 is a special case of a general result concerning the mapping capabilities of $\mathbb{G}$-reflectors in a large class of scalar product spaces. Suppose $\mathbb{G}$ is the automorphism group of a scalar product that is symmetric or skew-symmetric bilinear, or Hermitian or skew-Hermitian sesquilinear. It is easy to show that $\langle x, x\rangle = \langle y, y\rangle$ is a necessary condition on $x, y \in \mathbb{K}^n$ in order for there to exist some $G \in \mathbb{G}$ such that $Gx = y$. In [9] it is shown, by a nonconstructive argument, that for any nonzero $x, y \in \mathbb{K}^n$ such that $\langle x, x\rangle = \langle y, y\rangle$ there is a $G \in \mathbb{G}$ such that $Gx = y$, where $G$ is the product of at most two $\mathbb{G}$-reflectors. This is the general "Two $\mathbb{G}$-reflector Mapping Theorem". By contrast, it should be noted that the proof given in Lemma 4.3 for the special symplectic case of the Two $\mathbb{G}$-reflector Mapping Theorem is completely constructive.

### 4.3 Quasi-triangular Reduction

We now reduce $S \in Sp(2n, \mathbb{K})$ to quasi-triangular form using symplectic $\mathbb{G}$-reflectors, rather than the symplectic unitary tools described in Section 3. Once again, the proof is constructive.

**Proposition 4.4.** *For any $S \in Sp(2n, \mathbb{K})$ there exist symplectic $\mathbb{G}$-reflectors $G_1, G_2, \ldots, G_m$ such that*

$$G_m G_{m-1} \cdots G_1 S \;=\; \widehat{R}\,,$$

*where $\widehat{R} = \begin{bmatrix} R & Z \\ 0 & R^{-T} \end{bmatrix}$ is symplectic, $R$ is $n \times n$ upper triangular with only ones on the diagonal, and $m \leq 2n$.*

*Proof.* The procedure begins by using Lemma 4.3 to map the first column of $S$ to $e_1$. This requires a product $T_1$ of at most two $\mathbb{G}$-reflectors, $T_1 = G_1$ or $T_1 = G_2 G_1$, and by Lemma 3.2 we get

$$T_1 S \;=\; \begin{bmatrix} 1 & * & * & * \\ 0 & \boxed{L} & * & \boxed{M} \\ 0 & 0 & 1 & 0 \\ 0 & \boxed{N} & * & \boxed{P} \end{bmatrix}. \tag{4.3}$$

Then one continues inductively to reduce $\begin{bmatrix} L & M \\ N & P \end{bmatrix} \in Sp(2n-2, \mathbb{K})$ to quasi-triangular form. Since it takes at most two $\mathbb{G}$-reflectors per column, we see that a product of at most $2n$ $\mathbb{G}$-reflectors suffices to reduce any $S \in Sp(2n, \mathbb{K})$ to quasi-triangular form.

The only issue remaining is to see why any $(2n-2) \times (2n-2)$ symplectic $\mathbb{G}$-reflector $\widetilde{T}_2$ used in the inductive step can also be viewed as a $\mathbb{G}$-reflector in $Sp(2n, \mathbb{K})$. First observe that any $(2n-2) \times (2n-2)$ symplectic matrix $\widetilde{T}_2 = \begin{bmatrix} E & F \\ H & K \end{bmatrix}$, whether it is a $\mathbb{G}$-reflector or not, when embedded into $I_{2n}$ as

$$T_2 \;=\; \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \boxed{E} & 0 & \boxed{F} \\ 0 & 0 & 1 & 0 \\ 0 & \boxed{H} & 0 & \boxed{K} \end{bmatrix} \tag{4.4}$$

will be an element of $Sp(2n, \mathbb{K})$.

Now $\widetilde{T}_2$, being a $\mathbb{G}$-reflector, has a basis $\widetilde{v}_1, \widetilde{v}_2, \ldots, \widetilde{v}_{2n-3}$ for its fixed hyperplane in $\mathbb{K}^{2n-2}$. For each of these $\widetilde{v}_i = \begin{bmatrix} w_i \\ z_i \end{bmatrix} \in \mathbb{K}^{2n-2}$, $w_i,\, z_i \in \mathbb{K}^{n-1}$, define $v_i = \begin{bmatrix} 0 \\ w_i \\ 0 \\ z_i \end{bmatrix} \in \mathbb{K}^{2n}$. Then every $v_i$ will be fixed by $T_2$, as will $e_1$ and $e_{n+1}$. Thus $\mathcal{H} \stackrel{\text{def}}{=} \operatorname{span}\{v_1, v_2, \ldots, v_{2n-3}, e_1, e_{n+1}\}$ is a fixed $(2n-1)$-dimensional hyperplane for $T_2$, which proves that $T_2$ is indeed a $\mathbb{G}$-reflector in $Sp(2n, \mathbb{K})$. $\square$

### 4.4 $\mathbb{G}$-reflectors Generate the Symplectic Groups

It is possible to use symplectic $\mathbb{G}$-reflectors to take the reduction of a symplectic matrix even further than quasi-triangular form. In fact one can reduce all the way down to the identity.

**Proposition 4.5.** *For any $S \in Sp(2n, \mathbb{K})$ there exist symplectic $\mathbb{G}$-reflectors $G_1, G_2, \ldots G_m$ such that $G_m G_{m-1} \cdots G_1 S = I_{2n}$, where $m \leq 4n$.*

Before proving this proposition, we observe that the inverse of any $\mathbb{G}$-reflector is also a $\mathbb{G}$-reflector. See [7] for a geometric proof of this fact for $\mathbb{G}$-reflectors in any automorphism group $\mathbb{G}$, or simply observe that $G = I + \beta uu^T J \;\Rightarrow\; G^{-1} = I - \beta uu^T J$. Thus Proposition 4.5 immediately yields a constructive proof of the following factorization result. For nonconstructive proofs using somewhat more abstract methods, the reader is directed to [1] or [4, pp. 373–374].

**Theorem 4.6 (Symplectic $\mathbb{G}$-reflectors generate $Sp(2n, \mathbb{K})$).**
*Every $S \in Sp(2n, \mathbb{K})$ can be expressed as a product of at most $4n$ symplectic $\mathbb{G}$-reflectors.*

Since by Lemma 4.1 the determinant of any symplectic $\mathbb{G}$-reflector is $+1$, this factorization provides yet another proof that the determinant of any symplectic matrix is $+1$.

A basic topological property of symplectic groups, closely tied to the determinant issue, now easily follows from Theorem 4.6.

**Theorem 4.7.** *$Sp(2n, \mathbb{K})$ is path-connected.*

*Proof.* Express $S \in Sp(2n, \mathbb{K})$ as a product of $\mathbb{G}$-reflectors, $S = G_1 G_2 \cdots G_m$. Then continuously deforming each $G_j = I + \beta_j u_j u_j^T J$ to the identity by

$$G_j(t) = I + (1-t)\beta_j u_j u_j^T J\,, \quad t \in [0,1]\,, \;\; 1 \leq i \leq m$$

gives us a continuous path $S(t) = G_1(t) G_2(t) \cdots G_m(t)$ from $S$ to $I_{2n}$ in $Sp(2n, \mathbb{K})$. $\square$

We now turn to the proof of Proposition 4.5. Begin the reduction of a general $S \in Sp(2n, \mathbb{K})$ as in Section 4.3. Use Lemma 4.3 to construct a matrix $T_1$ that maps the first column of $S$ to $e_1$, thus obtaining (4.3). Now comes the unconventional, but key step. Rather than working next on the second column — as one is accustomed to do in QR-like decompositions — proceed instead to the $(n+1)$st column. Our goal is to map the $(n+1)$st column to $e_{n+1}$, *without disturbing the first column*, which has been mapped to $e_1$. The following lemma tells us which symplectic $\mathbb{G}$-reflectors leave $e_1$ fixed.

**Lemma 4.8.** *Suppose $G$ is a symplectic $\mathbb{G}$-reflector such that $Gx = y$. Then $Ge_1 = e_1 \Leftrightarrow x_{n+1} = y_{n+1}$.*

*Proof.* $(\Rightarrow)$: $y_{n+1} = \langle e_1, y\rangle = \langle Ge_1, Gx\rangle = \langle e_1, x\rangle = x_{n+1}$.
$(\Leftarrow)$: From Theorem 4.2 we know that a $\mathbb{G}$-reflector mapping $x$ to $y$ is unique whenever it exists, and is specified by $G = I + \beta uu^T J$ with $u = y - x$ and $\beta = 1/\langle y, x\rangle$. Hence

$$Ge_1 \;=\; e_1 + \beta uu^T J e_1 \;=\; e_1 + \beta\langle u, e_1\rangle u \;=\; e_1 + \beta(x_{n+1} - y_{n+1})u \;=\; e_1. \qquad \square$$

For brevity, let $x \in \mathbb{K}^{2n}$ denote the $(n+1)$st column of $T_1 S$. By (4.3) we already have $x_{n+1} = 1$. If it is possible to map $x$ to $y = e_{n+1}$ by a symplectic $\mathbb{G}$-reflector $G$, then by Lemma 4.8, $G$ will automatically send $e_1$ to $e_1$. By Theorem 4.2, such a $G$ exists if and only if $\langle y, x \rangle = e_{n+1}^T J x = -x_1 \neq 0$.

Should $x_1 = 0$, then we can achieve our goal in two steps. First map $x$ to $z = e_1 + e_{n+1}$ by a symplectic $\mathbb{G}$-reflector $G_1$. This can be done since

$$\langle z, x \rangle = (e_1 + e_{n+1})^T J x = x_{n+1} - x_1 = x_{n+1} = 1 \neq 0.$$

Furthermore $x_{n+1} = z_{n+1} = 1$, so $G_1 e_1 = e_1$ by Lemma 4.8. Then $z$ can be mapped to $e_{n+1}$ by a second $\mathbb{G}$-reflector $G_2$, since $\langle e_{n+1}, z \rangle = -1 \neq 0$; again $G_2 e_1 = e_1$ by Lemma 4.8.

Thus we see that we can construct a symplectic matrix $T_2$ such that $T_2 x = e_{n+1}$, $T_2 e_1 = e_1$, and $T_2$ is the product of at most two symplectic $\mathbb{G}$-reflectors. (When $x_1 \neq 0$, $T_2 = G$, otherwise $T_2 = G_2 G_1$.) By Lemma 3.2 and its generalization in the accompanying remark, this gives us

$$T_2 T_1 S = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \boxed{A} & 0 & \boxed{B} \\ 0 & 0 & 1 & 0 \\ 0 & \boxed{C} & 0 & \boxed{D} \end{bmatrix},$$

where $T_2 T_1$ is the product of at most four symplectic $\mathbb{G}$-reflectors and $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right] \in Sp(2n-2, \mathbb{K})$. Note, though, that we can generically expect $T_2 T_1$ to be the product of just two $\mathbb{G}$-reflectors.

This process can be continued inductively on $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ until $S$ is reduced to the identity. The comments at the end of Section 4.3, showing how the $(2n-2) \times (2n-2)$ $\mathbb{G}$-reflectors from the induction step may be regarded as $2n \times 2n$ $\mathbb{G}$-reflectors, again apply. Thus we see that for any $S \in Sp(2n, \mathbb{K})$ there are symplectic $\mathbb{G}$-reflectors $G_1, G_2, \ldots, G_m$, $m \leq 4n$, such that $G_m G_{m-1} \cdots G_2 G_1 S = I_{2n}$, completing the proof of Proposition 4.5. Note that for a generic $S \in Sp(2n, \mathbb{K})$, we expect to have $m = 2n$.

**Comparison with automorphism groups of symmetric bilinear forms:** Let $\mathbb{G}$ be the automorphism group of a symmetric bilinear form, e.g. $\mathbb{G} = O(n, \mathbb{K})$ or $\mathbb{G} = O(p, q, \mathbb{K})$. Then the Cartan-Dieudonné Theorem [1] [4, pp. 352–355], states that $\mathbb{G}$-reflectors (also known in this context as "symmetries") generate $\mathbb{G}$. Thus Theorem 4.6 can be viewed as the symplectic analogue of the Cartan-Dieudonné Theorem. So far the analogy between the automorphism groups of symmetric bilinear forms and the symplectic groups $Sp(2n, \mathbb{K})$, generated by a skew-symmetric bilinear form, is very close.

But when we look at the individual $\mathbb{G}$-reflectors we see a striking difference. For a symmetric bilinear form, *every* $\mathbb{G}$-reflector has determinant $-1$ rather than $+1$ (see [7] for a proof of this fact); thus both $+1$ and $-1$ determinants are realized in $\mathbb{G}$, depending only on whether the number of $\mathbb{G}$-reflectors used to generate an automorphism is even or odd. This very basic difference in the $\mathbb{G}$-reflectors in some sense "explains" why the determinants of general symplectic matrices behave differently than their counterparts in other automorphism groups.

As noted in Theorem 4.7, the symplectic groups are connected. This connectedness can be viewed as a strengthening of Theorem 1.2, since Theorem 1.2 follows from Theorem 4.7 but not conversely. On the other hand, any group $\mathbb{G}$ associated with a symmetric bilinear form must be disconnected, as a consequence of the existence of both $+1$ and $-1$ determinants in $\mathbb{G}$.

## 5 Proof using Pfaffians

The final proof presented in this essay comes from [1], and is the shortest, simplest, and most general proof of all. Its only drawback is that it relies on some rather non-obvious properties of Pfaffians, and thus does not go very far towards "demystifying" the symplectic determinant result.

While the notion of the Pfaffian is particular to skew-symmetric matrices, it is very general in the sense that these matrices may have entries from any commutative ring. The two fundamental results needed for the proof are stated below. For further details see [1] or [4].

- For any even integer $n \geq 2$, there is a polynomial in $n(n-1)/2$ variables with integer coefficients, denoted by Pf, with the following property. For *any* $n \times n$ skew-symmetric matrix $K$ (with entries in any commutative ring), the number $\mathrm{Pf}\ K$ obtained by evaluating the polynomial Pf at the upper triangular entries of $K$ (i.e. $K_{ij}$ for $i < j$) satisfies
$$\det K \;=\; (\mathrm{Pf}\ K)^2.$$
Modulo a certain normalizing condition, the polynomial Pf is unique. Note that $\mathrm{Pf}\, J \neq 0$, since $J$ is non-singular.

- Congruence transformations preserve skew-symmetry, and Pfaffians behave nicely with respect to congruences. For any $A \in \mathbb{K}^{n \times n}$ and any $n \times n$ skew-symmetric $K$, we have
$$\mathrm{Pf}(A^T K A) \;=\; \det A \cdot \mathrm{Pf}\, K. \tag{5.1}$$

It can now be shown very quickly that any symplectic matrix, with entries from any field, has determinant $+1$. Recall that $A \in Sp(2n, \mathbb{K}) \;\Rightarrow\; A^T J A = J$. Then equation (5.1) implies
$$\mathrm{Pf}\, J \;=\; \mathrm{Pf}(A^T J A) \;=\; \det A \cdot \mathrm{Pf}\, J.$$
Cancelling $\mathrm{Pf}\, J$ shows that $\det A = +1$. $\square$

**Acknowledgements**. This work was carried out during a sabbatical at the University of Manchester. The authors thank Nick Higham, Fran Tisseur and the Department of Mathematics at the University of Manchester for their support and hospitality. Thanks are also due to Nick and Fran for helpful comments on this paper.

## References

[1] E. Artin. *Geometric Algebra.* Interscience Tracts in Pure and Applied Mathematics. Interscience Publishers Inc., New York, 1957.

[2] Angelika Bunse-Gerstner. Matrix factorizations for symplectic QR-like methods. *Linear Algebra Appl.*, 83:49–77, 1986.

[3] Ralph Byers. *Hamiltonian and symplectic algorithms for the algebraic Riccati equation.* PhD thesis, Cornell University, Ithaca, N.Y., 1983.

[4] Nathan Jacobson. *Basic Algebra I.* W. H. Freeman and Company, San Francisco, 1974.

[5] R. B. Lehoucq. The computation of elementary unitary matrices. *ACM Trans. Math. Software*, 22(4):393–400, 1996.

[6] D. Steven Mackey and Niloufer Mackey. Symplectic orthogonal canonical forms for doubly structured matrices. In preparation.

[7] D. Steven Mackey, Niloufer Mackey, and Françoise Tisseur. $\mathbb{G}$-reflectors in scalar product spaces. Numerical Analysis Report No. 420, Manchester Centre for Computational Mathematics, Manchester, England, February 2003.

[8] D. Steven Mackey, Niloufer Mackey, and Françoise Tisseur. Structured factorizations for structured matrices. Numerical analysis report, Manchester Centre for Computational Mathematics, Manchester, England, 2003. In preparation.

[9] D. Steven Mackey, Niloufer Mackey, and Françoise Tisseur. Structured mapping problems. Numerical Analysis Report No. 421, Manchester Centre for Computational Mathematics, Manchester, England, 2003.

[10] D. Steven Mackey, Niloufer Mackey, and Françoise Tisseur. Structured tools for structured matrices. Numerical Analysis Report No. 419, Manchester Centre for Computational Mathematics, Manchester, England, February 2003.

[11] K. L. Meyer and G. R. Hall. *Introduction to Hamiltonian Dynamical Systems and the N-Body Problem*, volume 90 of *Applied Mathematical Sciences*. Springer, New York., 1991.