ELSEVIER Journal of Geometry and Physics 19 (1996) 99–122

JOURNAL OF GEOMETRY AND PHYSICS

Real transformations with polynomial invariants

E.N. Dancer [a], J.F. Toland [b,*]

[a] School of Mathematics and Statistics, University of Sydney, Sydney, NSW 2006, Australia
[b] School of Mathematical Sciences, University of Bath, Claverton Down, Bath BA2 7AY, United Kingdom

Received 6 October 1994; revised 18 March 1995

Abstract

This paper seeks to generalise one aspect of classical Krein theory for linear Hamiltonian systems by examining how the existence of a non-trivial, homogeneous, polynomial W of degree $m \geq 2$ with $\langle Ax, \nabla W(x)\rangle = 0, x \in \mathbb{R}^N$, affects the spectrum of a real linear transformation A on $\mathbb{R}^N$. Amongst other things it is shown that (i) such a W exists if, and only if, the spectrum of A is linearly dependent over the natural numbers, and (ii) there exists such a W which is non-degenerate if, and only if, all the eigenvalues of A are imaginary and semi-simple. In classical Krein theory W is quadratic. Our enquiry is motivated by a theory of topological invariants for dynamical systems which have a first integral. Degenerate Hamiltonian systems are a special class where the present considerations are relevant.

Keywords: Krein theory; Polynomial invariants; Degenerate Hamiltonian systems; Homotopy invariant; Flows with a first integral
1991 MSC: 11C99, 15A18, 15A57, 34A30, 34A34, 34C35

1. Introduction

Let B denote a real, symmetric, non-singular matrix on $\mathbb{R}^{2n}$. From the classical theory of linear Hamiltonian systems it is well-known (see [5,8,9]) that if A is any non-singular, real matrix and $\langle Ax, Bx\rangle = 0$ for all $x \in \mathbb{R}^{2n}$, then $\tilde{J} = AB^{-1}$ is skew-symmetric and the spectrum $\sigma(A)$ of $A = \tilde{J}B$ is closed under multiplication by -1. Note that the condition $\langle Ax, Bx\rangle = 0, x \in \mathbb{R}^{2n}$, means that the quadratic polynomial $V(x) = \langle Bx, x\rangle$ is an invariant for the flow of a Hamiltonian differential equation $\dot{x} = Ax$. The purpose here is to generalise these classical results on the spectrum of A by considering systems $\dot{x} = Ax$ with polynomial invariants of degree higher than two in $\mathbb{R}^N$. Such systems need not be

* Corresponding author. E-mail: jft@maths.bath.ac.uk.

SSDI 0393-0440(95)00026-7

Hamiltonian, indeed N need not be even, though some of our results are new even when they are. If $f : \mathbb{R}^N \to \mathbb{R}^N$ and $W : \mathbb{R}^N \to \mathbb{R}$, $N \geq 2$, are smooth functions with the property that

$$\langle \nabla W(x), f(x) \rangle = 0, \quad x \in \mathbb{R}^N, \tag{1.1}$$

tl n W, which is constant on solutions of the ordinary differential equation

$$\frac{dx}{dt}(t) = f(x(t)), \quad t \in \mathbb{R}, \ x(t) \in \mathbb{R}^N, \tag{1.2}$$

is called a first integral of the flow defined by (1.2) or, equivalently, an *invariant* of f. When 0 is an equilibrium of the flow and the Hessian B of W is non-singular, the effect on the spectrum of the non-singular matrix $A = f'(0)$ is covered by the classical theory because $\langle Ax, Bx \rangle = 0$ for all $x \in \mathbb{R}^N$ and N must be even. However, in natural cases the Hessian is not invertible (see [4]). We therefore ask how inferences can be drawn from the existence of a more general function W satisfying (1.1). An example is the following result, part (a) of which is well-known (see [1]) and part (b) of which is proved in [4, Section 2.2].

Proposition 1.1. *Suppose $f(0) = 0, A = f'(0) : \mathbb{R}^N \to \mathbb{R}^N$ is invertible and (1.1) holds.*
(a) *If $W \neq 0$ on a deleted neighbourhood of 0 in $\mathbb{R}^N$ then all the eigenvalues of A are purely imaginary.*
(b) *If $\nabla W \neq 0$ on a deleted neighbourhood of 0 in $\mathbb{R}^N$ then N is even. If $f'(0)$ has no imaginary eigenvalues, then it has the same number of eigenvalues with positive real part as with negative real part.*

2. The main results

The focus is on the case $f = A$, when A is a real, linear transformation and W is a homogeneous polynomial of arbitrary degree. It is instructive to point out how this special case relates to the problem for general polynomial and real-analytic invariants, before discussing its theory in detail.

Suppose in (1.1) that $f(0) = 0$ and that $f'(0) : \mathbb{R}^N \to \mathbb{R}^N$ has transformation matrix A. Then it follows from (1.1) that

$$0 = \lim_{t \downarrow 0} \langle \nabla W(tx), t^{-1} f(tx) \rangle = \langle \nabla W(0), Ax \rangle, \quad x \in \mathbb{R}^N.$$

If A is invertible it follows that $\nabla W(0) = 0$. Without supposing that A is invertible, suppose henceforth that $\nabla W(0) = 0$. It follows from (1.1) that

$$0 = \lim_{t \downarrow 0} \langle t^{-1} \nabla W(tx), t^{-1} f(tx) \rangle = D^2 W(0)(x, Ax) = \langle Bx, Ax \rangle, \quad x \in \mathbb{R}^N, \tag{2.1}$$

where $D^m W(0)$, the mth derivative of W at 0, is a real, symmetric, m-linear form on $\mathbb{R}^N$ and the symmetric matrix B is the Hessian of W at 0. A differentiation with respect to x gives

$$\langle Bx, Ay \rangle + \langle By, Ax \rangle = 0, \quad x, y \in \mathbb{R}^N. \tag{2.2}$$

In particular, if $y \in \ker(B)$ then $Ay \in (\text{range}(B))^{\perp} = \ker(B)$ and so A is a linear transformation on $\ker(B)$. Now suppose $m \geq 3$ is the smallest natural number such that

$$D^m W(0)(x, x, \ldots, x, \cdot) \neq 0 \in (\ker(B))^*$$

for some $x \in \ker(B)$. This is equivalent to $D^m W(0)$ being non-zero on $\ker(B)$. (Here $(\ker(B))^*$ denotes the dual space of $\ker(B)$.) Then for all $x \in \ker(B)$,

$$\begin{aligned} 0 &= \lim_{t \downarrow 0} \langle t^{-m+1} \nabla W(tx), t^{-1} f(tx) \rangle = \lim_{t \downarrow 0} t^{-m+1} DW(tx)(t^{-1} f(tx)) \\ &= \frac{1}{(m-1)!} D^m W(x, x, \ldots, x, Ax). \end{aligned}$$

Therefore if V is defined by $V(x) = D^m W(x, x, \ldots, x)$, $x \in \ker(B)$, then $A : \ker(B) \to \ker(B)$ and $\langle \nabla V(x), Ax \rangle = 0$ for all $x \in \ker(B)$. Thus the study of the general condition (1.1) leads naturally to the particular case when f is linear and W is a homogeneous polynomial of degree m. At this stage it is appropriate to give some definitions [6].

By a *polynomial* on $\mathbb{R}^N$ is meant a function $W : \mathbb{R}^N \to \mathbb{R}$ with the property that

$$W(x + ty) = \sum_{\ell=0}^{m} w_\ell(x, y) t^\ell, \quad t \in \mathbb{R}, \ x, y \in \mathbb{R}^N, \tag{2.3}$$

where $w_\ell \in \mathbb{R}$ is independent of t. When $w_m \not\equiv 0$ the polynomial is said to have *degree m* and if

$$W(tx) = t^m W(x), \quad x \in \mathbb{R}^N, \ t \in \mathbb{R},$$

then W is said to be *homogeneous* of *degree m*. A homogeneous polynomial W of degree m is a *polynomial invariant* for an $N \times N$ matrix transformation A if, and only if,

$$\langle \nabla W(x), Ax \rangle = 0, \quad x \in \mathbb{R}^N. \tag{2.4}$$

Here $\langle \, , \rangle$ denotes the inner product on $\mathbb{R}^N$ relative to which ∇W is defined by the relation

$$\langle \nabla W(x), y \rangle = DW(x)(y), \quad x, y \in \mathbb{R}^N.$$

Suppose throughout that A is a fixed real linear transformation on $\mathbb{R}^N$. An element λ of $\sigma(A)$ is said to be semi-simple if its algebraic and geometric multiplicities coincide and simple if it has algebraic multiplicity 1. There follows a summary of our main conclusions.

The result of Theorem 4.2 is that A has a non-trivial, homogeneous, polynomial invariant of degree $m \geq 2$ if, and only if, there exist m eigenvalues, $\lambda_1, \ldots, \lambda_m$, of A (not necessarily distinct) with

$$\sum_{\ell=1}^{m} \lambda_\ell = 0. \tag{2.5}$$

In particular, $\sigma(A)$ is linearly independent over $\mathbb{N}$ (the natural numbers) if, and only if, A has no non-trivial homogeneous, polynomial invariant. This observation pertains to homogeneous, *polynomial* invariants and not to homogeneous invariants in general. As the following example shows, the smoothness assumption has more influence than might at first appear likely.

Example. Let $N = 2, A = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}$ and $W(x, y) = |x|^p|y|^q, p, q > 1$. Then W, which is positively homogeneous of degree $m = p + q$, is a polynomial if, and only if, p and q are natural numbers. Also W is an invariant for A if, and only if,

$$\alpha p + \beta q = 0. \tag{2.6}$$

Therefore if $\beta < 0, \alpha > 1$ the matrix A has homogeneous invariants with any given order of differentiability and invariants with any given degree of homogeneity. By composing one of these invariants with a smooth, real-valued function which is zero, along with all its derivatives, at 0 we obtain a smooth invariant for A. But only when $p, q \in \mathbb{N}$ is there a homogeneous, *polynomial* invariant of degree $m = p + q$. Note that when $p, q \in \mathbb{N}$, (2.6) is the precise form which (2.5) takes in this example.

Note also that if $\mathrm{i}\alpha \in \sigma(A), \alpha \in \mathbb{R}$, then a relation of the form (2.5) is obtained when m is even by putting $\lambda_\ell = (-1)^\ell \mathrm{i}\alpha, 1 \leq \ell \leq m$. Therefore, if A has an imaginary eigenvalue it has a non-trivial, homogeneous, polynomial invariant of every even degree.

For convenience with notation, the word polynomial will be used to mean 'homogeneous polynomial'; on a few occasions when this meaning is not intended we will refer to a 'general polynomial'; the adjective 'homogeneous' may be included elsewhere, but only for emphasis. A polynomial W is said to be *non-degenerate* if, and only if,

$$\nabla W(x) \neq 0, \quad x \in \mathbb{R}^N \backslash \{0\}.$$

(The notion of a non-degenerate first integral, which coincides with requiring the Hessian to be invertible when W is quadratic, is central in the topological degree theory of [2–4].) We show in Theorem 4.6 that if A is non-singular then all its non-degenerate, polynomial invariants have even degree and N is even. Also A has a non-degenerate, polynomial invariant of even degree $m \geq 3$ if, and only if, all its eigenvalues are imaginary and semi-simple (Theorem 4.5). This result is false in the classical case of quadratic invariants ($m = 2$). For example, if A is a real, $2n \times 2n$, non-singular symmetric matrix and J is the usual symplectic matrix $\begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$, then $W(x) = \langle Ax, x \rangle, x \in \mathbb{R}^{2n}$, defines a non-degenerate, quadratic invariant for the matrix JA, if A is non-singular. However, the eigenvalues of JA are not, in all cases, imaginary. Note also that $\widehat{W}(x) = (W(x))^2, x \in \mathbb{R}^N$, defines a polynomial invariant of JA of degree 4, which is non-degenerate if, and only if, A is positive- or negative-definite. Therefore our result contains the classical one that when A is positive- or negative-definite all the eigenvalues of JA are imaginary and semi-simple. Indeed, A has a non-degenerate polynomial invariant of *any* degree $m \geq 3$ only if all its eigenvalues are imaginary and semi-simple (Theorem 4.6).

An invariant W is called *active* if

$$\operatorname{span}\{\nabla W(x) \colon x \in \mathbb{R}^N\} = \mathbb{R}^N.$$

If W is an active, polynomial invariant of degree $m \geq 2$ and $\lambda_1 \in \sigma(A)$, then there exist $\lambda_2, \ldots, \lambda_m \in \sigma(A)$, not necessarily distinct, such that (2.5) holds (Theorem 4.3). In the case $m = 2$, an invariant W is active if, and only if, its Hessian B at 0 is non-singular. This result therefore includes the classical one that if there exists a non-singular symmetric matrix B with $\langle Bx, Ax\rangle = 0$ for all $x \in \mathbb{R}^N$, then $\sigma(A)$ is closed under multiplication by -1.

When a non-trivial, polynomial invariant fails to be active on $\mathbb{R}^N$, it can control part of the spectrum of A, while having no influence on the rest. To see this, if W is a non-active, polynomial invariant for A let

$$\mathbb{R}^N = M \oplus M^{\perp} \quad \text{where } M = \operatorname{span}\{\nabla W(x)\colon\ x \in \mathbb{R}^N\},$$

and write $x = y + z \in M \oplus M^{\perp}, x \in \mathbb{R}^N$. Then for $t \in \mathbb{R}$, $y \in M$ and $z \in M^{\perp}$,

$$\frac{\mathrm{d}}{\mathrm{d}t} W(y + tz) = \langle \nabla W(y + tz), z\rangle = 0,$$

whence, if $x = y + z \in M \oplus M^{\perp}$,

$$W(x) = W(y).$$

If we write the vector $\nabla W(x) = \nabla W(y + z)$ as $(\nabla_y W(y + z), \nabla_z W(y + z))$, then

$$\nabla_z W(y + z) = 0 \quad \text{and} \quad \nabla_y W(y + z) = \nabla_y W(y).$$

Now with respect to the decomposition of $\mathbb{R}^N = M \oplus M^{\perp}$, let

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}.$$

Then (2.4), with $x = y + z$, has the form

$$0 = \langle A_{11}y + A_{12}z, \nabla_y W(y + z)\rangle = \langle A_{11}y + A_{12}z, \nabla_y W(y)\rangle$$

and

$$0 = \langle A_{21}y + A_{22}z, \nabla_z W(y + z)\rangle.$$

The second equality contains no information since $\nabla_z W = 0$. However, with $z = 0$ in the first, we find that

$$\langle A_{11}y, \nabla_y W(y)\rangle = 0, \quad y \in M.$$

In other words, $W|_M$ is a homogeneous, polynomial invariant for the matrix A_{11}, and W is active on M because $M = \operatorname{span}\{\nabla W(y)\colon\ y \in M\}$.

With $y = 0$, the first equation gives

$$\langle A_{12}z, \nabla_y W(y)\rangle = 0, \quad y \in M,\ z \in M^{\perp},$$

which implies that $A_{12} = 0$, since $M = \operatorname{span}\{\nabla_y W(y)\colon\ y \in M\}$. Hence

$$A = \begin{pmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{pmatrix},$$

from which it follows that $M^{\perp}$ is an invariant subspace of A and every eigenvalue of A_{11} is also an eigenvalue of A, with the same multiplicity. Note also that the degree of W restricted to M is the same as the degree of W on $\mathbb{R}^N$ (because $W(x) = W(y)$ where $x = y + z \in M \oplus M^{\perp}$). Hence every eigenvalue of A_{11} is also an eigenvalue of A and is involved in a relationship (2.5) with $(m-1)$ eigenvalues of A which are also eigenvalues of A_{11}.

Now we can regard the result following (2.2), that $\ker(B)$ is an invariant subspace for the operator A, as a special case of the present discussion in which $M = \text{range}(B)$ and $m = 2$. Thus an inductive procedure for deciding how a general polynomial invariant W influences the spectrum of A emerges. At the mth step, the theory of this paper shows how the mth derivative of W, behaving as an active, homogeneous polynomial invariant, controls a part of $\sigma(A)$ (the part which coincides with the spectrum of A_{11} in the above analysis) while simultaneously defining a new invariant subspace $M^{\perp}$ for A upon which higher derivatives of W influence the spectrum at the $(m+1)$st step. If W denotes a real-analytic invariant of A, the whole spectrum of A is constrained by W in this way, except when W is identically zero on a subspace of $\mathbb{R}^N$.

A polynomial is *non-vanishing* if $W(x) \neq 0, x \in \mathbb{R}^N \backslash \{0\}$. It is easy to see that, for homogeneous polynomials, non-vanishing $\Rightarrow$ non-degenerate $\Rightarrow$ active, and so the results outlined above strengthen Proposition 1.1 in the case when W is a homogeneous polynomial. Indeed, there is a scale of hypotheses, reflecting the decreasing influence of the invariant W on $\sigma(A)$, which can be described as follows:

$$\boldsymbol{H}_k: \qquad D^k W(x) \neq 0 \quad \text{for all } x \in \mathbb{R}^N \backslash \{0\}, \ k = 0, 1, \ldots, m-1.$$

Note that $\boldsymbol{H}_k$ with $k = 0, 1,$ and $m-1$ is equivalent to non-vanishing, non-degenerate and active, respectively and that $\boldsymbol{H}_k$ implies $\boldsymbol{H}_\ell$ if $k \leq \ell$.

The following example illustrates how much weaker are the implications for a non-singular transformation A of the existence of an active, as opposed to a non-degenerate, first integral.

Example. Let

$$A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & 0 & -2 \end{pmatrix}$$

and let $W : \mathbb{R}^4 \to \mathbb{R}$ be the cubic polynomial defined by

$$W(\boldsymbol{x}) = x_2^2 x_4 - x_2^2 x_3 + x_1 x_2 x_4, \quad \boldsymbol{x} = (x_1, x_2, x_3, x_4) \in \mathbb{R}^4.$$

Then

$$\nabla W(\boldsymbol{x}) = (x_2 x_4, 2x_2(x_4 - x_3) + x_1 x_4, -x_2^2, x_2^2 + x_1 x_2)$$

and it is easily checked that $\langle \nabla W(\boldsymbol{x}), A\boldsymbol{x} \rangle = 0$, $\boldsymbol{x} \in \mathbb{R}^4$. Now, to check that V is active, suppose that for some $\boldsymbol{a} \in \mathbb{R}^4$,

$$\langle \nabla W(\boldsymbol{x}), \boldsymbol{a}\rangle = 0 \quad \text{for all } \boldsymbol{x} \in \mathbb{R}^4.$$

Then for all $(x_1, x_2, x_3, x_4) \in \mathbb{R}^4$,

$$0 = a_1x_2x_4 + 2a_2x_2(x_4 - x_3) + a_2x_1x_4 - a_3x_2^2 + a_4x_2^2 + a_4x_1x_2.$$

The x_1x_2 term gives $a_4 = 0$; the x_2^2 term then gives $a_3 = 0$; the x_1x_4 term gives $a_2 = 0$ and the x_2x_4 term gives $a_1 = 0$. Hence, $\text{span}\{\nabla W(\boldsymbol{x})\colon \boldsymbol{a} \in \mathbb{R}^4\} = \mathbb{R}^4$ and so V is active.

Hence this non-singular, real transformation A has an active, polynomial invariant W of odd degree, its eigenvectors are all real and none of them is semi-simple. If W were non-degenerate then its degree would be even and the eigenvalues of A would be imaginary and semi-simple. From Theorem 4.4 it is easy to see that in this example the degree of any polynomial invariant of A is a multiple of 3.

There follows a brief remark about the possibility of more than one polynomial invariant. If W_1 and W_2 are homogeneous polynomials, of possibly different degrees $m_1, m_2 \geq 2$, they are *strictly independent* if $\{\nabla W_1(x), \nabla W_2(x)\}$ is a linearly independent set for every $x \in \mathbb{R}^N\backslash\{0\}$. Note that if this is so then W_1 and W_2 are non-degenerate and m_1 and m_2 are both even, or both odd (Theorem 4.6). Hence if W_1 and W_2 are polynomial invariants with degrees $m \geq 3$ then all the eigenvalues of A are imaginary and semi-simple. We will prove amongst other things (Theorem 4.6) that if the eigenvalues of a non-singular, real transformation A with largest modulus are simple, then A does *not* have a strictly independent pair of polynomial invariants. This is an analogue of Lemma 1.1 of [3] which says that if A has a simple, imaginary eigenvalue it does not have a strictly independent pair of quadratic, polynomial invariants. It is easy to see [4, Lemma 3.3], that if $\alpha\nabla W_1(x_0) + \beta\nabla W_2(x_0) = 0$, $x_0 \in \mathbb{R}^N\backslash\{0\}$ and W_1 and W_2 are polynomial invariants of A, then $\alpha\nabla W_1(x(t)) + \beta\nabla W_2(x(t)) = 0$ for all $t \in \mathbb{R}$ when

$$\frac{\mathrm{d}x}{\mathrm{d}t}(t) = Ax(t), \quad t \in \mathbb{R}, \ x(0) = x_0.$$

Such remarks about the non-existence of strictly independent invariants is significant in the topological degree theory for flows with a first integral [2–4]. In the presence of a pair of strictly independent integrals the degree is trivial (see [3, p.570]).

3. Preliminaries

A function $V : (\mathbb{C}^N)^m \to \mathbb{C}$ is said to be a complex *m-linear form* on $\mathbb{C}^N$ if for each $\ell \in \{1, \ldots, m\}$ and $z_1, \ldots, z_{\ell-1}, z_{\ell+1}, \ldots, z_m$, the function

$$z \longmapsto V(z_1, \ldots, z_{\ell-1}, z, z_{\ell+1}, \ldots, z_m) \text{ is linear over } \mathbb{C},$$

and a complex m-linear form V is *symmetric* if

$$V(z_1, \ldots, z_m) = V(z_{\sigma(\ell)}, \ldots, z_{\sigma(m)}), \quad \sigma \in S_m,$$

where S_m denotes the group of permutations of m elements. A *complex polynomial* on $\mathbb{C}^N$ is defined by (2.3) with $t \in \mathbb{C}$, $x, y \in \mathbb{C}^N$. A polynomial or an m-linear form is said to be *real* if its value is real when its argument is real. Let C_m denote the complex space of all m-linear forms on $\mathbb{C}^N$ and let R_m be the real space of real m-linear forms. Let C_m^σ and R_m^σ denote the corresponding spaces of symmetric forms. In [6] it is shown that the linear operator Σ defined by

$$\Sigma V(z) = V(z, \dots, z), \quad V \in C_m, \ z \in \mathbb{C}^N, \tag{3.1}$$

is a surjection from C_m onto the space of complex, homogeneous polynomials of degree m, whose restriction to C_m^σ is a bijection with the same range. The analogous statement for R_m, R_m^σ and real, homogeneous polynomials follows by the same argument. Therefore, in considering real, homogeneous polynomial invariants W, there is no loss of generality in seeking W in the form ΣV for some V in R_m^σ.

Lemma 3.1. *Suppose that W is a real, homogeneous polynomial of degree m. Then*
(a) *W is active if, and only if, $\boldsymbol{H}_{m-1}$ holds;*
(b) *W is non-degenerate if, and only if, $\boldsymbol{H}_1$ holds;*
(c) *W is non-vanishing if, and only if, $\boldsymbol{H}_0$ holds.*

Proof. Let $W = \Sigma V$, where $V \in R_m^\sigma$.

(a). Suppose W is not active. Then there exists $y \in \mathbb{R}^N \backslash \{0\}$ such that

$$\langle \nabla W(x), y \rangle = 0, \quad x \in \mathbb{R}^N. \tag{3.2}$$

Since $W = \Sigma V$ and V is symmetric, (3.2) may be re-written as

$$DW(x)(y) = V(x, x, \dots, x, y) = 0, \quad x \in \mathbb{R}^N.$$

After differentiating $(m-1)$ times and using the symmetry of V, we find

$$D^{m-1}W(y)(x_1, \dots, x_{m-1}) = V(y, x_1, \dots, x_{m-1}) = 0, \quad x_1, \dots, x_{m-1} \in \mathbb{R}^N. \tag{3.3}$$

Since $y \neq 0$, we have proved that $\boldsymbol{H}_{m-1}$ is false when W is not active. Conversely, if $\boldsymbol{H}_{m-1}$ is false then (3.3) holds for some non-zero $y \in \mathbb{R}^N$. Putting $x_1 = x_2 = \cdots = x_{m-1} = x$ we find that (3.2) holds and hence W is not active. This completes the proof of (a).
(b),(c). These are both immediate from the definitions. □

Remark. From (3.3) there follows the observation that $\boldsymbol{H}_k$ implies $\boldsymbol{H}_\ell$ when $k \leq \ell$.

Suppose that a basis $\{e_1, \dots, e_N\} \subset \mathbb{C}^N$ is *closed under conjugation* ($\{\bar{e}_1, \dots, \bar{e}_N\} = \{e_1, \dots, e_N\}$, where $\bar{}$ denotes the usual component-wise operation of complex conjugation in $\mathbb{C}^N$). Its dual basis $\{e_1^*, \dots, e_N^*\}$ for $(\mathbb{C}^N)^*$ is uniquely determined by the system of equations $e_i^*(e_j) = \delta_{ij}, 1 \leq i, j \leq N$. If $f \in (\mathbb{C}^N)^*$, let $\overline{f} \in (\mathbb{C}^N)^*$ be defined by $\overline{f}(z) = \overline{f(\overline{z})}, z \in \mathbb{C}^N$. Note that

$$\overline{e_i^*}(\overline{e_j}) = \overline{e_i^*(e_j)} = \delta_{ij}. \tag{3.4}$$

Hence $\{\overline{e_1^*}, \ldots, \overline{e_N^*}\}$ is the dual basis of $\{\overline{e}_1, \ldots, \overline{e}_N\}$ (which is also clearly a basis for $\mathbb{C}^N$). By uniqueness, $\{e_1^*, \ldots, e_N^*\}$ is also closed under the conjugation operation defined above on $(\mathbb{C}^N)^*$. Note that $\overline{e}_i^* = \overline{e_i^*}$.

Let Q denote the set of all functions q which map $\{1, \ldots, m\}$ into $\{1, \ldots, N\}$. If $q \in Q$, let $\overline{q} \in Q$ be defined by

$$e_{\overline{q}(j)} = \overline{e_{q(j)}}, \quad 1 \le j \le m. \tag{3.5}$$

Note that, for any i, j, the definition of $\overline{q}$ gives

$$e^*_{\overline{q}(i)}(e_j) = 1 \text{ if, and only if, } \overline{q}(i) = j, \text{ i.e. if, and only if, } \overline{e_{q(i)}} = e_j,$$

while

$$\overline{e^*_{q(i)}}(e_j) = \overline{e^*_{q(i)}(\overline{e}_j)} = 1 \text{ if, and only if, } \overline{e}_j = e_{q(i)}, \text{ i.e. if, and only if, } \overline{q}(i) = j.$$

Hence it follows that

$$e^*_{\overline{q}(i)} = \overline{e^*_{q(i)}}, \quad 1 \le i \le N. \tag{3.6}$$

If $q \in Q$, let $V_q \in C_m$ be defined by

$$V_q(z_1, \ldots, z_m) = \prod_{i=1}^{m} e^*_{q(i)}(z_i). \tag{3.7}$$

Then

$$\begin{aligned}\overline{V_q(z_1, \ldots, z_m)} &= \prod_{i=1}^{m} \overline{(e^*_{q(i)}(z_i))} \\ &= \prod_{i=1}^{m} \overline{e^*_{q(i)}}(\overline{z}_i), \quad \text{by definition of } \overline{e^*_{q(i)}} \\ &= \prod_{i=1}^{m} e^*_{\overline{q}(i)}(\overline{z}_i), \quad \text{by (3.6).}\end{aligned}$$

Hence for any $q \in Q$,

$$\overline{V_q(z_1, \ldots, z_m)} = V_{\overline{q}}(\overline{z}_1, \ldots, \overline{z}_m). \tag{3.8}$$

Theorem 3.2.

(a) *The set $\{V_q : q \in Q\}$ is a basis for C_m and if $V \in C_m$ then*

$$V = \sum_{q \in Q} \alpha_q V_q, \quad \textit{where } \alpha_q = V(e_{q(1)}, e_{q(2)}, \ldots, e_{q(m)}). \tag{3.9}$$

(b) *If $V \in C_m$ then $V \in R_m$ if, and only if, for all $(z_1, \ldots, z_m) \in (\mathbb{C}^N)^m$*

$$\overline{V(z_1, \ldots, z_m)} = V(\overline{z}_1, \ldots, \overline{z}_m). \tag{3.10}$$

(c) *If* $V \in C_m$ *then* (3.10) *holds, if and only if,*

$$\overline{\alpha_q} = \alpha_{\overline{q}} \quad \text{for all } q \in Q. \tag{3.11}$$

Proof

(a) It is clear that for each $q \in Q$ the function V_q is in C_m. Suppose that $q_i, 1 \leq i \leq r$, are distinct elements of Q and $\alpha_i \in \mathbb{C}, 1 \leq i \leq r$, are such that $\sum_{i=1}^{r} \alpha_i V_{q_i} = 0 \in C_m$. Then since

$$V_{q_i}(e_{q_j(1)}, e_{q_j(2)}, \dots, e_{q_j(m)}) \geq 0$$

with equality if, and only if, $i = j$, it follows immediately that $\alpha_i = 0$ for all $i, 1 \leq i \leq r$. Thus $\{V_q : q \in Q\}$ is linearly independent. Now if $V \in C_m$ and $z_i \in \mathbb{C}^N, 1 \leq i \leq m$, then $z_i = \sum_{j=1}^{N} e_j^*(z_i) e_j$, whence

$$V(z_1, \dots, z_m) = V\left(\sum_{j=1}^{N} e_j^*(z_1) e_j, \dots, \sum_{j=1}^{N} e_j^*(z_m) e_j\right) = \sum_{q \in Q} \alpha_q V_q(z_1, \dots, z_m).$$

Hence $\{V_q : q \in Q\}$ spans C_m. This proves (a).

(b) If $(z_1, \dots, z_m) \in (\mathbb{R}^N)^m$ and (3.10) holds then it is immediate that $V(z_1, \dots, z_m) \in \mathbb{R}$. Conversely, suppose that $V(z_1, \dots, z_m) \in \mathbb{R}$ wherever $(z_1, \dots, z_m) \in (\mathbb{R}^N)^m$. For this step only, let $\{e_1, \dots, e_N\}$ be the standard basis of $\mathbb{R}^N$ over $\mathbb{R}$ (which is also a basis for $\mathbb{C}^N$ over $\mathbb{C}$ and which is closed under conjugation). Then by (3.9), α_q is real for all $q, q = \overline{q}$ and

$$\begin{aligned}\overline{V(z_1, \dots, z_m)} &= \sum_{q \in Q} \alpha_q \overline{V_q(z_1, \dots, z_m)} \\ &= \sum_{q \in Q} \alpha_q V_{\overline{q}}(\overline{z}_1, \dots, \overline{z}_m), \quad \text{by (3.8)} \\ &= \sum_{q \in Q} \alpha_q V_q(\overline{z}_1, \dots, \overline{z}_m) = V(\overline{z}_1, \dots, \overline{z}_m).\end{aligned}$$

This proves (b).

(c) As in part (a), let $\{e_1, \dots, e_N\}$ be any basis of $\mathbb{C}^N$ which is closed under conjugation. If (3.10) holds then

$$\overline{\alpha_q} = \overline{V(e_{q(1)}, \dots, e_{q(m)})} = V(\overline{e_{q(1)}}, \dots, \overline{e_{q(m)}}) = V(e_{\overline{q}(1)}, \dots, e_{\overline{q}(m)}) = \alpha_{\overline{q}}.$$

Conversely, if (3.11) holds, then

$$\begin{aligned}V(z_1, \dots, z_m) &= \sum_{q \in Q} \alpha_q V_q(z_1, \dots, z_m) \\ &= \sum_{q \in Q} \overline{\alpha_{\overline{q}}}\, V_q(z_1, \dots, z_m), \quad \text{by (3.11)} \\ &= \sum_{q \in Q} \overline{\alpha_{\overline{q}} V_{\overline{q}}(\overline{z}_1, \dots, \overline{z}_m)}, \quad \text{by (3.8)}\end{aligned}$$

$$= \overline{\sum_{\overline{q}\in Q} \alpha_{\overline{q}}\, V_{\overline{q}}(\overline{z}_1, \ldots, \overline{z}_m)} = \overline{V(\overline{z}_1, \ldots, \overline{z}_m)}.$$

This proves (c) . □

If A is a real, linear transformation on $\mathbb{R}^N$ and W is a real, polynomial of degree m such that

$$\langle \nabla W(x), Ax\rangle = 0, \quad x \in \mathbb{R}^N, \tag{3.12}$$

suppose, without loss of generality, that $W = \Sigma V$, $V \in R_m^\sigma$. Because V is symmetrical, (3.12) may be re-written as

$$V(x, x, \ldots, x, Ax) = 0, \quad x \in \mathbb{R}^N, \tag{3.13}$$

which, after differentiating m times, gives

$$\sum_{\ell=1}^{m} V(x_1, x_2, \ldots, Ax_\ell, \ldots, x_m) = 0, \quad x_\ell \in \mathbb{R}^N, \ 1 \le \ell \le m. \tag{3.14}$$

Therefore, because V is symmetric, (3.13) and (3.14) are equivalent. But, when V is a general (not necessarily symmetric) element of R_m, it remains the case that (3.14) implies (3.12) when $W = \Sigma V$.

Hence *if* $V \in R_m$ *satisfies* (3.14) *and is non-trivial on the diagonal*, $\{(x, \ldots, x) : x \in \mathbb{R}^N\}$ *of* $(\mathbb{R}^N)^m$ *then* $W = \Sigma V$ *is a non-trivial, homogeneous, polynomial invariant for* A *of degree* m. *Also, all homogeneous, polynomial invariants* W *of degree* m *of* A *are in the form* $W = \Sigma V$, $V \in R_m^\sigma$, *where* V *satisfies* (3.14).

Now suppose that $V \in R_m$ satisfies (3.14) and let $\{e_1, \ldots, e_N\}$ be the standard basis for $\mathbb{R}^N$ over $\mathbb{R}$ which is also a basis for $\mathbb{C}^N$ over $\mathbb{C}$. Then, for $q \in Q$, let $\alpha_q = V(e_{q(1)}, \ldots, e_{q(m)})$ and for $(z_1, \ldots, z_m) \in (\mathbb{C}^N)^m$ let

$$V(z_1, \ldots, z_m) = \sum_{q\in Q} \alpha_q V_q(z_1, \ldots, z_m). \tag{3.15}$$

If $x_1, \ldots, x_{m-1} \in \mathbb{R}^N$ and $z_m = x_m + \mathrm{i}y_m \in \mathbb{C}^N$, then

$$\sum_{\ell=1}^{m} V(x_1, \ldots, Ax_\ell, \ldots, x_{m-1}, z_m)$$
$$= \sum_{\ell=1}^{m} V(x_1, \ldots, Ax_\ell, \ldots, x_{m-1}, x_m) + \mathrm{i}V(x_1, \ldots, Ax_\ell, \ldots, x_{m-1}, y_m) = 0.$$

Now suppose that for some $k \le m - 1$,

$$\sum_{\ell=1}^{k} V(x_1, \ldots, Ax_\ell, \ldots, x_k, z_{k+1}, \ldots, z_m)$$
$$+ \sum_{\ell=k+1}^{m} V(x_1, \ldots, x_k, z_{k+1}, \ldots, Az_\ell, \ldots, z_m) = 0,$$

for all $(x_1, \ldots, x_k, z_{k+1}, \ldots, z_m) \in (\mathbb{R}^N)^k \times (\mathbb{C}^N)^{m-k}$. (We have just observed this when $k = m-1$.) Then

$$\begin{aligned}
&\sum_{\ell=1}^{k-1} V(x_1, \ldots, Ax_\ell, \ldots, x_{k-1}, z_k, \ldots, z_m) \\
&\quad + \sum_{\ell=k}^{m} V(x_1, \ldots, x_{k-1}, z_k, \ldots, Az_\ell, \ldots, z_m) \\
&\qquad = \sum_{\ell=1}^{k} V(x_1, \ldots, Ax_\ell, \ldots, x_k, \ldots, z_{k+1}, \ldots, z_m) \\
&\qquad + \sum_{\ell=k+1}^{m} V(x_1, \ldots, x_k, z_{k+1}, \ldots, Az_\ell, \ldots, z_m) \\
&\qquad + \mathrm{i}\left[\sum_{\ell=1}^{k} V(x_1, \ldots, Ax_\ell, \ldots, x_{k-1}, y_k, z_{k+1}, \ldots, z_m)\right. \\
&\qquad\quad \left. + \sum_{\ell=k+1}^{m} V(x_1, \ldots, x_{k-1}, y_k, z_{k+1}, \ldots, Az_\ell, \ldots, z_m)\right] = 0.
\end{aligned}$$

Hence, by induction, (3.14) implies that

$$\sum_{\ell=1}^{m} V(z_1, \ldots, Az_\ell, \ldots, z_m) = 0, \quad z_\ell \in \mathbb{C}^N, \ 1 \le \ell \le m. \tag{3.16}$$

Therefore any real V satisfying (3.14) can be extended to a complex V satisfying (3.16) and there is no loss of generality in considering (3.16) for elements V of R_m from the outset.

Now suppose that $V \in R_m$ is identically zero on the diagonal of $(\mathbb{R}^N)^m$. Then $V(x, x, \ldots, x) = 0, x \in \mathbb{R}^N$, and differentiation m times gives

$$\sum_{\sigma \in S_m} V(x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(m)}) = 0, \quad x_1, \ldots, x_n \in \mathbb{R}^N. \tag{3.17}$$

From this, it follows, by induction, that the extension of V as an element of C_m has the property that

$$V(x + \mathrm{i}y, x + \mathrm{i}y, \ldots, x + \mathrm{i}y) = 0, \quad x + \mathrm{i}y \in \mathbb{C}^N.$$

Hence if $V \in R_m$ is zero on the diagonal of $(\mathbb{R}^N)^m$, then it is zero on the diagonal of $(\mathbb{C}^N)^m$. Also the degree of ΣV is the same when regarded as a real or a complex polynomial.

4. Polynomial invariants of real transformations

Since A is real, $\sigma(A) = \{\lambda_1, \ldots, \lambda_n\}$ is closed under conjugation and $n \le N$. Now we introduce notation for a Jordan basis of A. For each $p, 1 \le p \le n$, let $\overline{p}$ be such that $\lambda_{\overline{p}} = \overline{\lambda_p}$ and let

$$\ker(\lambda_p I - A) = \operatorname{span}\{f_j^p : 1 \le j \le n(p)\}, \tag{4.1a}$$

where $\{f_j^p : 1 \le j \le n(p)\}$ is a linearly independent set with $f_j^{\overline{p}} = \overline{f_j^p}$ chosen as follows: for each $p \in \{1, \dots, n\}$, let $j \in \{1, \dots, n(p)\}$ and let there exist $\{e_{j,k}^p : 1 \le k \le m(j,p)\}$, the root vectors, satisfying

$$e_{j,1}^p = f_j^p, \quad Ae_{j,k+1}^p = \lambda_p e_{j,k+1}^p + e_{j,k}^p, \qquad 1 \le k \le m(j,p) - 1, \tag{4.1b}$$

$$e_{j,m(j,p)}^p \notin \operatorname{Range}(\lambda_p I - A), \tag{4.1c}$$

$$e_{j,k}^{\overline{p}} = \overline{e_{j,k}^p}, \quad p \in \{1, \dots, n\}, \quad j \in \{1, \dots, n(p)\}, \ k \in \{1, \dots, m(j,p)\}. \tag{4.1d}$$

Then $B = \{e_{j,k}^p : 1 \le p \le n, 1 \le j \le n(p), 1 \le k \le m(j,p)\}$ is a basis of $\mathbb{C}^N$ which is closed under conjugation relative to which A is in Jordan Normal Form. For convenience with notation later, let $e_{k,0}^p = 0$.

Now let $\mathcal{P}$ denote the set of all functions $P : \{1, \dots, m-1\} \to \{1, \dots, n\}$. If $P \in \mathcal{P}$ let $\overline{P} \in \mathcal{P}$ be defined by $\lambda_{\overline{P}(\ell)} = \overline{\lambda_{P(\ell)}}$, $1 \le \ell \le m-1$. If $P \in \mathcal{P}$ let $\mathcal{J}_P$ be the set of functions J on $\{1, \dots, m-1\}$ with $J(\ell) \in \{1, 2, \dots, n(P(\ell))\}$, $1 \le \ell \le m-1$.

If $P \in \mathcal{P}$ and $J \in \mathcal{J}_P$ let $\mathcal{K}_{J,P}$ denote those functions K on $\{1, \dots, m-1\}$ with $K(\ell) \in \{1, 2, \dots, m(J(\ell), P(\ell))\}$. Finally, if $K \in \mathcal{K}_{J,P}$ let K_ℓ be defined by

$$K_\ell(\ell') = \begin{cases} K(\ell') & \text{if } \ell \ne \ell', \\ K(\ell) - 1 & \text{if } \ell = \ell'. \end{cases}$$

(Note that K has range in $\mathbb{N}$ and K_ℓ has range in $\mathbb{N} \cup \{0\}$.) Let

$$|K| = \sum_{\ell=1}^{m-1} K(\ell), \quad K \in \mathcal{K}_{J,P} \text{ and } \mu_P = -\sum_{\ell=1}^{m-1} \lambda_{P(\ell)}. \tag{4.2}$$

Note that $\mu_{\overline{P}} = \overline{\mu_P}$ because of the definition of $\overline{P}$. Let $V \in R_m$ and for $P \in \mathcal{P}, J \in \mathcal{J}_P, K \in \mathcal{K}_{J,P}$ let $v_{J,K}^P \in (\mathbb{C}^N)^*$ be defined by

$$v_{J,K}^P(z) = V(e_{J(1),K(1)}^{P(1)}, e_{J(2),K(2)}^{P(2)}, \dots, e_{J(m-1),K(m-1)}^{P(m-1)}, z), \quad z \in \mathbb{C}^N. \tag{4.3}$$

Note that $v_{J,K_\ell}^P = 0$ if $K_\ell(\ell) = 0$. Since B is a basis for $\mathbb{C}^N$, the function V is known if all of the functionals $v_{J,K}^P$ are known. This is immediate by Theorem 3.2. Because of the discussion in Section 3, to find a non-trivial, polynomial invariant for A it is sufficient to find $V \in R_m$ such that V is non-trivial on the diagonal of $(\mathbb{R}^N)^m$ and

$$\sum_{i=1}^{m} V(z_1, z_2, z_{i-1}, Az_i, z_{i+1}, \dots, z_m) = 0 \quad \text{for all } (z_1, \dots, z_m) \in (\mathbb{C}^N)^m. \tag{4.4}$$

Let A^* denote the conjugate of A on $(\mathbb{C}^N)^*$ defined by $(Az^*)z = z^*(Az), z \in \mathbb{C}^N, z^* \in (\mathbb{C}^N)^*$. Note that relative to the basis for $(\mathbb{C}^N)^*$ dual to B, the transformation matrix for A^* is the transpose of the Jordan Normal Form of A.

Theorem 4.1. *Let $V \in R_m$, $P \in \mathcal{P}$, $J \in \mathcal{J}_P$ and $K \in \mathcal{K}_{J,P}$. If* (4.4) *holds then*

$$\text{(a)} \qquad v^{\overline{P}}_{J,K}(z) = \overline{v^P_{J,K}(\overline{z})} = \overline{v^P_{J,K}}(z), \tag{4.5}$$

$$\text{(b)} \qquad v^P_{J,K} \in \ker(\mu_P I - A^*)^{|K|+2-m}, \tag{4.6}$$

$$\text{(c)} \qquad (\mu_P I - A^*)v^P_{J,K} = \sum_{\ell=1}^{m-1} v^P_{J,K_\ell}. \tag{4.7}$$

Hence, by (b),

$$\text{(d)} \qquad v^P_{J,K} = 0 \quad \text{if } \mu_P \notin \sigma(A). \tag{4.8}$$

Proof.
(a). By the definition of $\overline{P}$ and (4.1)

$$\begin{aligned} v^{\overline{P}}_{J,K}(z) &= V(\overline{e^{P(1)}_{J(1),k(1)}}, \ldots, \overline{e^{P(m-1)}_{J(m-1),K(m-1)}}, z) \\ &= \overline{V(e^{P(1)}_{J(1),K(1)}, \ldots, e^{P(m-1)}_{J(m-1),K(m-1)}, \overline{z})}, \quad \text{by Theorem 3.2(b)} \\ &= \overline{v^P_{J,K}(\overline{z})}. \end{aligned}$$

(b), (c). If $P \in \mathcal{P}$, $J \in \mathcal{J}_P$, $K \in \mathcal{K}_{J,P}$ and $\ell \in \{1, 2, \ldots, m-1\}$ then

$$Ae^{P(\ell)}_{J(\ell),K(\ell)} = \lambda_{P(\ell)} e^{P(\ell)}_{J(\ell),K(\ell)} + e^{P(\ell)}_{J(\ell),K_\ell(\ell)}. \tag{4.9}$$

(Recall the convention that $e^P_{k,0} = 0$.) Therefore, from (4.4) with $z_m = z \in \mathbb{C}^N$ and $z_\ell = e^{P(\ell)}_{J(\ell),K(\ell)}$, $1 \le \ell \le m-1$, we obtain

$$\sum_{\ell=1}^{m-1} \lambda_{P(\ell)}\, v^P_{J,K}(z) + \sum_{\ell=1}^{m-1} v^P_{J,K_\ell}(z) + v^P_{J,K}(Az) = 0, \quad z \in \mathbb{C}^N.$$

This can be re-written as

$$(A^* - \mu_P I)v^P_{J,K} + \sum_{\ell=1}^{m-1} v^P_{J,K_\ell} = 0 \in (\mathbb{C}^N)^*, \tag{4.10}$$

which proves (4.7). To complete the proof let $P \in \mathcal{P}$, $J \in \mathcal{J}_P$ be fixed. We use induction on $|K| = \sum_{\ell=1}^{m-1} K(\ell)$. The inductive hypothesis is that

$$v^P_{J,K} \in \ker(\mu_P I - A^*)^{(|K|+2-m)} \quad \text{if } m-1 \le |K| \le k.$$

Note first that when $|K| = m-1$, $K(\ell) = 1$ for all ℓ. Hence $K_\ell(\ell) = 0$ for all ℓ and so $(A^* - \mu_P I)v^P_{J,K} = 0$ by (4.10). Since $|K| + 2 - m = 1$ in this case the result is proved when $k = m-1$.

Now suppose $|K| = k+1$. Then for all ℓ either $K_\ell \in \mathcal{K}_{J,P}$ and $|K_\ell| = k$, or $K_\ell(\ell) = 0$ and $v^P_{J,K_\ell} = 0$ by construction. It is now immediate, by the induction hypothesis and (4.10), that $v^P_{J,K} \in \ker(\mu_P I - A^*)^{k+3-m}$, and the result follows since $k+3-m = |K|+2-m$ in this case. □

Remarks. First, note that specifying a particular $P \in \mathcal{P}$ is equivalent to picking a set of $(m-1)$ (not necessarily distinct) eigenvalues of A, and the subsequent choice of J denotes the selection of particular eigenvectors of A corresponding to the eigenvalues already chosen. The system (4.7) is, in fact, a union of uncoupled sub-systems, one for each pair (P, J), each sub-system being parametrized by $K \in \mathcal{K}_{J,P}$. Therefore, it is sufficient, and possibly more convenient, to consider each sub-system separately.

Second, suppose that for a given (P, J) the corresponding sub-system of (4.7) has a non-zero solution. We want to show that there is a solution of the sub-system corresponding to $(\overline{P}, J)$ so that (4.5) holds. If $P \neq \overline{P}$, then (P, J) and $(\overline{P}, J)$ have distinct sub-systems in (4.7), $\mathcal{K}_{J,P} = \mathcal{K}_{J,\overline{P}}$ and it suffices to *define* $v^{\overline{P}}_{J,K}(z)$ to be $\overline{v^P_{J,K}(\overline{z})}$. It is immediate from the construction that this is a non-trivial solution of the sub-system (4.7) for $(\overline{P}, J)$. The case $P = \overline{P}$ occurs if, and only if, $\lambda_{P(\ell)}$ is real for all ℓ, $1 \le \ell \le m-1$, in which case μ_P is also real. Suppose $\{v^P_{J,K} : K \in \mathcal{K}_{J,P}\}$ is a given, non-zero solution of (4.7) for given (P, J). Let

$$w^P_{J,K}(z) = v^P_{J,K}(z) + \overline{v^P_{J,K}(\overline{z})}, \quad z \in \mathbb{C}^N,$$

for all $K \in \mathcal{K}_{J,P}$. Then $\{w^P_{J,K} : K \in \mathcal{K}_{J,P}\}$ is a solution of (4.5) and (4.7). If it is the zero solution, then for all $K \in \mathcal{K}_{J,P}$

$$0 = w^P_{J,K}(x) = 2 \,\mathrm{Real}\, v^P_{J,K}(x), \quad x \in \mathbb{R}^N.$$

If this is so, note that $\{r^P_{J,K} : K \in \mathcal{K}_{J,P}\}$ is also a non-zero solution of (4.7) for given (P, J), where $r^P_{J,K} = \mathrm{i} v^P_{J,K}$. Now define $w^P_{J,K}$ using $r^P_{J,K}$ instead of $v^P_{J,K}$ to obtain a non-zero solution of (4.5) and (4.7).

In all cases, a non-zero solution of (4.7) for given (P, J) leads to a non-zero solution of (4.5) and (4.7). In Theorem 4.6 below it is shown that a solution of (4.5), (4.7) is sufficient, as well as necessary, for the existence of a solution V of (4.4) in R_m. Whether V is non-zero on the diagonal determines whether there is a non-trivial, polynomial invariant W of A in the form $W = \Sigma V$. If however V is a non-trivial, symmetric solution of (4.5) and (4.7), then V must be non-zero on the diagonal, for otherwise differentiating n times gives $V = 0 \in R^\sigma_m$.

By the ascent of an operator A is meant $\inf\{n \in \mathbb{N} \cup \{0\} : \ker(A^n) = \ker(A^{n+1})\}$. (Since $A^0 = I$, the ascent of A is 0 if, and only if, A is injective.) If μ is an eigenvalue of A then we will refer to the ascent of $(\mu I - A)$ as the ascent of μ. By classical theory μ has the same ascent as an eigenvalue of A and of A^*, and the ascent α of the eigenvalues μ and $\overline{\mu}$ of a real transformation A are equal. Also, for all $\mu \in \mathbb{C}$,

$$\ker(\mu I - A)^\alpha = \bigcup_{k \in \mathbb{N}} \ker(\mu I - A)^k = \mathcal{N}(\mu I - A),$$

$$\mathrm{range}\,(\mu I - A)^\alpha = \bigcap_{k \in \mathbb{N}} \mathrm{range}\,(\mu I - A)^k = \mathcal{R}(\mu I - A),$$

$$\mathbb{C}^N = \mathcal{N}(\mu I - A) \oplus \mathcal{R}(\mu I - A), \quad \mathcal{N}(\mu I - A) \subset \mathcal{R}(\lambda I - A) \text{ if } \lambda \neq \mu,$$

and

$$\mathcal{N}(\mu I - A) \cap \mathcal{N}(\lambda I - A) = \{0\} \quad \text{if } \lambda \neq \mu.$$

Recall that A is a real transformation and a basis of root vectors $\{e^p_{j,k} \colon 1 \le p \le n, 1 \le j \le n(p), 1 \le k \le m(j,p)\}$, which is closed under conjugation, has been chosen for $\mathbb{C}^N$. Let $\hat{\mathcal{P}}$ denote the set of $\hat{P} : \{1, \ldots, m\} \to \{1, \ldots, n\}$, let $\hat{\mathcal{J}}_{\hat{P}}$ denote the set of $\hat{J} : \{1, \ldots, m\} \to \{1, \ldots, n(\hat{P})\}$ and $\hat{\mathcal{K}}_{\hat{J}, \hat{P}}$ the set of $\hat{K} : \{1, \ldots, m\} \to \{1, \ldots, m(\hat{J}, \hat{P})\}$. Then, by Theorem 3.2, $\{V^{\hat{P}}_{\hat{J},\hat{K}} \colon \hat{P} \in \hat{\mathcal{P}}, \hat{J} \in \hat{\mathcal{J}}_{\hat{P}}, \hat{K} \in \hat{\mathcal{K}}_{\hat{J},\hat{P}}\}$ is a basis for C_m where

$$V^{\hat{P}}_{\hat{J},\hat{K}}(z_1, \ldots, z_m) = \prod_{\ell=1}^{m} \left(e^{\hat{P}(\ell)}_{\hat{J}(\ell), \hat{K}(\ell)}\right)^* (z_\ell). \tag{4.11}$$

With respect to this basis an element $V \in C_m$ has coefficient $\alpha^{\hat{P}}_{\hat{J},\hat{K}}$ defined by

$$\begin{aligned}\alpha^{\hat{P}}_{\hat{J},\hat{K}} &= V\left(e^{\hat{P}(1)}_{\hat{J}(1),\hat{K}(1)}, \ldots, e^{\hat{P}(m-1)}_{\hat{J}(m-1),\hat{K}(m-1)}, e^{\hat{P}(m)}_{\hat{J}(m),\hat{K}(m)}\right) \\ &= v^P_{J,K}\left(e^{\hat{P}(m)}_{\hat{J}(m),\hat{K}(m)}\right),\end{aligned} \tag{4.12}$$

where here and later P, J, K denote $\hat{P}, \hat{J}$ and $\hat{K}$, respectively, restricted to $\{1, \ldots, m-1\}$. Therefore, if $\{V^p_{j,k} \colon p \in \mathcal{P}, j \in \mathcal{J}_p, K \in \mathcal{K}_{J,P}\}$ is given, a function $V \in C_m$ is uniquely determined in terms of the basis (4.11) by the coefficients (4.12).

Now, by Theorem 4.1, $v^P_{J,K} \in \mathcal{N}(\mu_P I - A^*)$, the generalised eigenspace of μ_P as an eigenvalue of A^*, and

$$e^{\hat{P}(m)}_{\hat{J}(m),\hat{K}(m)} \in \mathcal{N}(\lambda_{\hat{P}(m)} I - A) \subset \mathcal{R}(\lambda I - A)$$

for any $\lambda \in \mathbb{C} \backslash \{\lambda_{\hat{P}(m)}\}$, where $\mathcal{R}(\lambda I - A)$ denotes the generalised range of $(\lambda I - A)$. In particular, if $\mu_P \neq \lambda_{\hat{P}(m)}$ then

$$e^{\hat{P}(m)}_{\hat{J}(m),\hat{K}(m)} \in \mathcal{R}(\mu_P I - A).$$

Since

$$v^P_{J,K} \in \mathcal{N}(\mu_P I - A^*) \quad \text{and} \quad \mu_P = -\sum_{j=1}^{m-1} \lambda_{\hat{P}(\ell)}$$

it is immediate that

$$\alpha^{\hat{P}}_{\hat{J},\hat{K}} = 0 \quad \text{for all } \hat{P} \text{ with } \sum_{\ell=1}^{m} \lambda_{\hat{P}(\ell)} \neq 0. \tag{4.13}$$

Theorem 4.2. *A real, linear transformation has a non-zero, homogeneous, polynomial invariant of degree m if, and only if, there exist m eigenvalues, $\alpha_1, \ldots, \alpha_m$, of A with*

$$\sum_{\ell=1}^{m} \alpha_\ell = 0.$$

Proof. If no set of m eigenvalues of A adds up to zero, then $v^P_{J,K}$ is zero for all P, J, K, by Theorem 4.1(d). Therefore if V satisfies (4.4) then $V \equiv 0$, by (4.13). Hence A has no non-zero, polynomial invariant of degree m, by the remark in italics preceding expression (3.15).

Conversely, suppose $\alpha_1, \dots, \alpha_m$ are eigenvalues of A which add up to zero, and let $\beta_1, \dots, \beta_{m'}$ be the distinct elements of $\{\alpha_1, \dots, \alpha_m\}$. If β_i is not real, let $g_i^* \in \ker(\beta_i I - A^*)$. From the definition of $\overline{g_i^*}$ in Section 3, it follows that $\overline{g_i^*} \in \ker(\overline{\beta_i} I - A^*)$. If β_i is real, let $w_i^* \in \ker(\beta_i I - A^*)$ and let $g_i^* = w_i^* + \overline{w_i^*}$. Then $g_i^* \in \ker(\beta_i I - A)$ and $\overline{g_i^*} = g_i^*$ when β_i is real. Moreover, $\{g_i^* :\ 1 \le i \le m'\}$ is a linearly independent set in $(\mathbb{C}^N)^*$ and hence there exists $\{g_i :\ 1 \le i \le m'\} \subset \mathbb{C}^N$ with $g_i^*(g_j) = \delta_{ij}$.

Now let

$$f_\ell^* = g_i^* \quad \text{if}\, \alpha_\ell = \beta_i,\ 1 \le i \le m$$

and define $V \in C_m$ by

$$V(z_1, \dots, z_m) = \prod_{\ell=1}^{m} f_\ell(z_\ell) + \prod_{\ell=1}^{m} \overline{f_\ell^*}(z_\ell) = \prod_{\ell=1}^{m} f_\ell^*(z_\ell) + \overline{\prod_{\ell=1}^{m} f_\ell^*(\overline{z_\ell})}.$$

It is immediate, from Theorem 3.2(b), that $V \in R_m$. Moreover, for $z_1, \dots, z_m \in \mathbb{C}^N$,

$$\sum_{\ell=1}^{m} f_1^*(z_1) \dots f_\ell^*(Az_\ell) \dots f_m^*(z_m) = \sum_{\ell=1}^{m} f_1^*(z_1) \dots ((A^* f_\ell^*)(z_\ell)) \dots f_m^*(z_m)$$
$$= \left(\sum_{\ell=1}^{m} \alpha_\ell\right) \prod_{\ell=1}^{m} f_1^*(z_1) \dots f_m^*(z_m) = 0.$$

Hence V satisfies (3.14). Now let

$$z = \sum_{\ell=1}^{m'} g_\ell \in \mathbb{C}^N.$$

Then $V(z, z, \dots, z) = 2$. Now let

$$W(x) = V(x, \dots, x), \quad x \in \mathbb{R}^N.$$

Then $W \not\equiv 0$, by the closing remark of Section 3, and W is a homogeneous, polynomial invariant of degree m of A. □

Remark. In many cases when A has m eigenvalues which sum to zero there are at least two distinct (i.e. linearly independent in the space of real-valued functions on $\mathbb{R}^N$) polynomial invariants of A. This follows from Theorem 4.7, which is the converse of Theorem 4.1. But there are exceptions. For example, when there is only one set of m eigenvalues which sums to zero, each element of which has geometric multiplicity one and all but one of which is semi-simple, then there is only one polynomial invariant of A.

Theorem 4.3. *Suppose W is an active, polynomial invariant of A of degree $m \geq 2$. If λ_1 is an eigenvalue of A there exist $m-1$ eigenvalues of A, not necessarily distinct, such that*

$$\sum_{\ell=1}^{m} \lambda_\ell = 0.$$

Proof. Let $W = \Sigma V, V \in R_m^\alpha$. Let the basis $\{e_{j,k}^P\}$ be chosen as in (4.1). Suppose that λ_1 is an eigenvalue of A with eigenvector e. Now, by the hypothesis that W is active, $\boldsymbol{H}_{k-1}$ holds and therefore $v_{J,K}^P(e) \neq 0$ for some $P \in \mathcal{P}, J \in \mathcal{J}_P, K \in \mathcal{K}_{J,P}$. Suppose $\lambda_1 \neq \mu_P$. Then

$$v_{J,K}^P \in \mathcal{N}(\mu_P I - A^*) \subset \mathcal{R}(\lambda I - A^*) = (\mathcal{N}(\lambda I - A))^\perp.$$

Hence $v_{J,K}^P(e) = 0$ which is a contradiction. Hence $\mu_P = \lambda_1$ which proves the result. □

For given (J, P)

$$\underline{K}(\ell) \leq K(\ell) \leq \overline{K}(\ell), \quad 1 \leq \ell \leq m-1, \quad K \in \mathcal{K}_{J,P},$$

where $\underline{K}, \overline{K} \in \mathcal{K}_{J,P}$ are functions of (J, P) defined by

$$\underline{K}(\ell) = 1, \qquad \overline{K}(\ell) = m(J(\ell), P(\ell)), \quad 1 \leq \ell \leq m-1.$$

Theorem 4.1 says that $(\mu_P I - A^*)^{\alpha_P} v_{J,K}^P = 0$, where $\alpha_P > 0$ is the ascent of μ_P as an eigenvalue of A. This leads to the following theorem.

Theorem 4.4. *Let $P \in \mathcal{P}, J \in \mathcal{J}_P$ and suppose that*

$$(m-1) + \alpha_P \leq |\overline{K}|.$$

Then $v_{J,\underline{K}}^P = 0$.

Proof. Let $K^\alpha \in \mathcal{K}_{J,P}$ be such that $|K^\alpha| = (m-1) + \alpha_P$. Such K^α exists by hypothesis. Then by (4.7)

$$(\mu_P I - A^*) v_{J,K^\alpha}^P = \sum_{\ell=1}^{m-1} v_{J,K_\ell^\alpha}^P,$$

where either $K_\ell^\alpha(\ell) = 0$ or $|K_\ell^\alpha| = |K^\alpha| - 1$. Hence, by induction,

$$(\mu_P I - A^*)^{\alpha_P} v_{J,K^\alpha}^P = r\, v_{J,\underline{K}}^P,$$

where r is some positive integer. But $v_{J,\underline{K}}^P \in \ker(\mu_P I - A^*)$, by Theorem 4.1(b), whence $v_{J,\underline{K}}^P \in \ker(\mu_P I - A^*)^{\alpha_P} \cap \text{range}(\mu_P I - A^*)^{\alpha_P} = \{0\}$, by definition of α_P. This completes the proof. □

It is clear from the proof of the preceding theorem that Eq. (4.7) forces many of the $v^P_{J,K}$ to be zero, but it is difficult to give a more systematic statement of a result in that direction. The significance of (4.5) and (4.7) is that they give a necessary condition for (4.4). In Theorem 4.7 we will observe that this is also sufficient.

Theorem 4.5. *A linear transformation A has a non-degenerate, homogeneous, polynomial invariant W of even degree $m \geq 4$ if, and only if, it is diagonalisable and all its eigenvalues are imaginary.*

Proof. Suppose that the eigenvalues of A, counted according to multiplicity, are $\pm i\alpha_1, \ldots, \pm i\alpha_k$, and possibly 0. Say $e_\ell = a_\ell + ib_\ell$ is an eigenvalue of $i\alpha_\ell$, $1 \leq \ell \leq k$, and if necessary $Af_j = 0$, $f_j \in \mathbb{R}^N$, $j = 2k+1, \ldots, N$. Then $\{a_\ell, b_\ell, f_j, 1 \leq \ell \leq k, 2k+1 \leq j \leq N\}$ is a basis for $\mathbb{R}^N$ and we can choose an inner-product $\langle\, ,\rangle$ relative to which it is orthonormal. Since $Aa_\ell = -\alpha_\ell b_\ell$ and $Ab_\ell = \alpha_\ell a_\ell$ for all ℓ, $1 \leq \ell \leq k$, there results that $\langle Ax, x\rangle = 0$, $x \in \mathbb{R}^N$. Now for even $m \geq 4$, let $W(x) = \langle x, x\rangle^{m/2}$. Therefore, for $y \in \mathbb{R}^N$,

$$\langle \nabla W(x), y\rangle = m\langle x, x\rangle^{(m-2)/2}\langle x, y\rangle,$$

whence

$$\nabla W(x) \neq 0, \ \ x \in \mathbb{R}^N\backslash\{0\} \quad \text{and} \quad \langle \nabla W(x), Ax\rangle = 0, \ \ x \in \mathbb{R}^N.$$

Therefore W is a non-degenerate invariant for A which, by its definition, is clearly a homogeneous polynomial of even degree $m \geq 4$.

For the converse, suppose that A has an eigenvalue with real part non-zero. Let β denote the eigenvalue of A whose real part has largest absolute value and let $f \in \mathbb{C}^N$ denote a corresponding eigenvector of A. Then $\bar{f}$ is an eigenvector of A with eigenvalue $\bar{\beta}$. (We do not exclude the possibility that β is real and $\bar{f} = f$.) If λ is any eigenvalue of A and $1 \leq k \leq m-1$,

$$\text{real}(k\beta + (m-1-k)\bar{\beta} + \lambda) = (m-1)\,\text{real}\,\beta + \text{ real}\,\lambda \neq 0$$

since $m > 2$, because of the choice of β. If W is a polynomial invariant of A of degree $m \geq 4$ let $W = \Sigma V$ where $V \in R^\sigma_m$. Therefore, if $\xi \in \mathbb{C}^N$ it follows from Theorem 4.1(d) that

$$V(f, \ldots, f, \bar{f}, \ldots, \bar{f}, \xi) = 0,$$

where f appears k times and $\bar{f}$ appears $m-1-k$ times. It is now easy to infer from the multi-linearity of V that

$$V(x, x, \ldots, x, \xi) = V(y, y, \ldots, y, \xi) = 0$$

if $f = x + iy$, for $\xi \in \mathbb{C}^N$. Hence

$$V(x, \ldots, x, z) = V(y, \ldots, y, z) = 0, \quad z \in \mathbb{R}^N,$$

whence $\nabla W(x) = \nabla W(y) = 0$, since $W = \Sigma V$.

This proves that if W is a non-degenerate, polynomial invariant of A of degree $m > 2$ then all the eigenvalues of A are purely imaginary. Now we must prove that they are semi-simple. Let $\mathrm{i}\gamma$ be an eigenvalue of A of largest ascent, and let $g = u + \mathrm{i}v$ be a corresponding eigenvector. Suppose the ascent of $\mathrm{i}\gamma$ is $\alpha \geq 2$. Let $P \in \mathcal{P}$ and $J \in \mathcal{J}_P$ be chosen so that for k, $1 \leq k \leq m-1$

$$\lambda_{P(\ell)} = \mathrm{i}\gamma \quad \text{and} \quad e^{P(\ell)}_{J(\ell),1} = f, \quad 1 \leq \ell \leq k,$$
$$\lambda_{P(\ell)} = -\mathrm{i}\gamma \quad \text{and} \quad e^{P(\ell)}_{J(\ell),1} = \bar{f}, \quad k+1 \leq \ell \leq m-1.$$

Note that since the ascent of the eigenvalues $\mathrm{i}\gamma$ and $-\mathrm{i}\gamma$ are equal,

$$|\overline{K}| = \alpha(m-1).$$

Moreover, α is the largest ascent of any eigenvalue of A and hence either $\mu_P = (2k - m + 1)\gamma\mathrm{i}$ is not an eigenvalue of A or it has ascent $\alpha_P \leq \alpha$. Since $m \geq 3$ and $\alpha \geq 2$

$$|\overline{K}| = \alpha(m-1) = \alpha + \alpha(m-2) \geq \alpha_P + 2(m-2) \geq \alpha_P + m - 1.$$

Therefore, by Theorem 4.3, for $1 \leq k \leq m-1$,

$$V(f, \ldots, f, \bar{f}, \ldots, \bar{f}, z) = 0, \quad z \in \mathbb{C}^N,$$

where f and $\bar{f}$ appear k and $m - 1 - k$ times, respectively. The multi-linearity of V now gives

$$V(u, \ldots, u, z) = V(v, v, \ldots, v, z) = 0, \quad z \in \mathbb{R}^N.$$

Therefore

$$\nabla W(u) = \nabla W(v) = 0,$$

since $W = \Sigma V$, and this contradicts the non-degeneracy of W.

This completes the proof. □

Remarks. The proof that when W is non-degenerate all the eigenvalues of A are purely imaginary generalises somewhat to yield a weaker result under weaker hypotheses: if W satisfies $\boldsymbol{H}_k$ for some $k < \frac{1}{2}m$ and is a polynomial invariant of A of degree m, then all the eigenvalues of A are imaginary.

The next theorem has, as a special case, the result that if a non-zero, imaginary eigenvalue with largest absolute value of a non-singular transformation A is simple, then A does not have a pair of strictly independent first integrals of any degree. Note that the hypotheses of parts (e), (f) below are not mutually exclusive.

Theorem 4.6. *Suppose that A is a real, linear transformation on $\mathbb{R}^N$.*

(a) *If A has a non-degenerate, polynomial invariant of any degree $m \geq 3$ then all its eigenvalues are imaginary and semi-simple.*

(b) *If A has a non-degenerate, polynomial invariant of odd degree $m \geq 3$, then A is singular.*
(c) *Suppose A is non-singular and has a non-degenerate, polynomial invariant. Then N is even.*
(d) *Each component of a strictly independent pair of homogeneous polynomials is non-degenerate and both have odd, or even, degrees.*
(e) *If 0 is a simple eigenvalue of $A \neq 0$, then A does not have a strictly independent pair of polynomial invariants of odd degrees $m \geq 3$.*
(f) *If $\pm i\gamma$, $\gamma \in \mathbb{R}$, are the eigenvalues of $A \neq 0$ of largest absolute value and are simple, then A does not have a strictly independent set of polynomial invariants of even degrees.*

Proof

(a) An examination of the second half of the proof of Theorem 4.5 yields the required result if $m \geq 3$ is arbitrary.

(b) Let $\pm \mathrm{i}\gamma \neq 0$ be the eigenvalues of A of largest absolute value and suppose that $Af = \mathrm{i}\gamma f$, where $f = u + \mathrm{i}v$. (This is possible by part (a).) Then, for any k, $1 \leq k \leq m-1$,

$$\mathrm{i}k\gamma + \mathrm{i}(m-1-k)(-\mathrm{i}\gamma) = \mathrm{i}(2k-m+1)\gamma \notin -\sigma(A)$$

if $0 \notin \sigma(A)$, since $2k - m + 1$ is even for all k and $\mathrm{i}\gamma$ is the eigenvalue of largest absolute value. As in the proof of Theorem 4.5, it follows that if $N = \Sigma V$, $V \in R_m^\sigma$, then

$$V(u, u, \ldots, u, z) = V(v, v, \ldots, v, z) = 0, \quad z \in \mathbb{C}^N.$$

Hence $\nabla W(a) = \nabla W(v) = 0$, which contradicts the non-degeneracy of W.

(c) Suppose A is non-singular and W is a non-degenerate, polynomial invariant for A. Then

$$\pm\lambda Ax + (1-\lambda)\nabla W(x) \neq 0, \quad x \in \mathbb{R}^N, \ \|x\| = 1, \ \lambda \in [0, 1], \tag{4.14}$$

because $\langle \nabla W(x), Ax \rangle = 0, x \in \mathbb{R}^N$. Hence (4.14) defines an admissible homotopy for Brouwer degree on the unit ball Ω. Hence

$$\deg(\Omega, A, 0) = \deg(\Omega, \nabla W, 0) = \deg(\Omega, -A, 0).$$

Since $\deg(\Omega, A, 0) = \operatorname{sign}(\operatorname{Det} A)$, this implies that N is even.

(d) Suppose W_1, W_2 is a strictly independent pair of polynomials. Then

$$\lambda\nabla W_1(x) + (1-\lambda)\nabla W_2(x) \neq 0, \quad x \in \mathbb{R}^N, \ \|x\| = 1, \ \lambda \in [0, 1]. \tag{4.15}$$

When $\lambda = 0, 1$ we find that both W_1 and W_2 are non-degenerate. Also, (4.15) defines an admissible homotopy in the sense of Brouwer degree on the unit ball Ω in $\mathbb{R}^N$ and consequently

$$\deg(\Omega, \nabla W_1, 0) = \deg(\Omega, \nabla W_2, 0).$$

However, $\deg(\Omega, f, 0)$, when it is defined, is odd for an odd function f, and even for an even, homogeneous function f [7, Ch. II, Theorem 4.1 and Ch. IV, Section 2]. Since ∇W_1

is even when W_1 is odd and vice versa for W_2, this proves that W_1 and W_2 both have odd, or even, degrees.

(e) Suppose that $W = \Sigma V, V \in R_m^\sigma$, is any non-degenerate, polynomial invariant of A of odd degree $m \geq 3$ and let 0 be a simple eigenvalue of A with $Au = 0, u \in \mathbb{R}^N \backslash \{0\}$. Since all the eigenvalues of A are imaginary, by part (a), let $\pm \mathrm{i}\gamma$ be the eigenvalues of largest absolute value and suppose $Af = \mathrm{i}\gamma f, f = a + \mathrm{i}b$. Then for any k, $1 \leq k \leq m-1$,

$$k\mathrm{i}\gamma + (m-1-k)(-\mathrm{i}\gamma) \in -\sigma(A),$$

if, and only if, $2k = m-1$, since $2k - m + 1$ is even and $\pm \mathrm{i}\gamma$ are the eigenvalues with largest absolute value. Therefore, if ξ is any generalised eigenvector of A corresponding to a non-zero eigenvalue we find, from (4.13), that

$$V(f, \ldots, f, \bar{f}, \ldots, \bar{f}, \xi) = 0,$$

where f and $\bar{f}$ appear k and $m-1-k$ times, respectively. Hence, by the multi-linearity of V,

$$V(a, \ldots, a, \xi) = V(b, \ldots, b, \xi) = 0 \quad \text{for all such } \xi.$$

Therefore for all x in the real subspace which is invariant under A and complementary to span$\{u\}$,

$$V(a, a, \ldots, a, x) = \langle \nabla W(a), x \rangle = 0.$$

In other words, $\nabla W(a)$ lies in a one-dimensional space determined by A. Since this is true for any polynomial invariant W of A of odd degree, and since $a \neq 0$ is independent of W, the result is proved.

(f) Suppose W_1 and W_2 form a strictly independent pair of polynomial invariants of even degree. If both are quadratic then the result is proved in [3], Lemma 1.1. If one of them has higher degree then all the eigenvalues of A are imaginary and semi-simple. We suppose this to be the case henceforth and adopt the hypothesis that $\pm \mathrm{i}\gamma$, the eigenvalues of largest absolute value, are simple.

Suppose that $Af = \mathrm{i}\gamma f$, where $f = u + \mathrm{i}v$. Then the choice of $\mathrm{i}\gamma$ means that

$$k\mathrm{i}\gamma + (m-1-k)(-\mathrm{i}\gamma) \notin -\sigma(A) \backslash \{\pm \mathrm{i}\gamma\}, \quad k \in \mathbb{Z}.$$

Let $W = \Sigma V, V \in R_m^\sigma, m \geq 2$, be any non-degenerate, homogeneous, polynomial invariant of even degree m of A, and let ξ be any generalised eigenvector of A corresponding to any eigenvalue of A other than $\pm \mathrm{i}\gamma$. Then, by (4.13)

$$V(f, \ldots, f, \bar{f}, \ldots, \bar{f}, \xi) = 0,$$

where f and $\bar{f}$ appear k and $m-k-1$ times, respectively. Hence, by the multi-linearity of V,

$$V(u, u, \ldots, u, \xi) = 0$$

for all such ξ. Let $\mathbb{R}^N = E \oplus \text{span}\{u, v\}$ where E is a real, invariant subspace for A. Then since $W = \Sigma V$ and V is symmetric, we have shown that

$$\langle \nabla W(u), e \rangle = 0 \quad \text{for all } e \in E.$$

Also, since $Au = -\mathrm{i}\gamma v, \gamma \neq 0$, and $\langle \nabla W(u), Au \rangle = 0$, we find that

$$\langle \nabla W(u), x \rangle = 0 \quad \text{if } x \in \text{span}\{E, v\}.$$

But span $\{E, v\}$ has real co-dimension 1, and is determined only by the eigenspaces of A. Hence $\nabla W(u)$ lies in a one-dimensional space determined by A. Since u and span$\{v, E\}$ are independent of W, this proves the required result. □

Finally, for completeness, we prove the converse of Theorem 4.1.

Theorem 4.7. *Suppose that* $\{f^P_{J,K} : P \in \mathcal{P}, J \in \mathcal{J}_P, K \in \mathcal{K}_{J,P}\} \subset (\mathbb{C}^N)^*$ *is any solution of* (4.5) *and* (4.7). *Let*

$$V(z_1, \ldots, z_m) = \sum \alpha^{\hat{P}}_{\hat{J},\hat{K}} V^{\hat{P}}_{\hat{J},\hat{K}}(z_1, \ldots, z_m), \tag{4.16}$$

where

$$\alpha^{\hat{P}}_{\hat{J},\hat{K}} = f^P_{J,K}\left(e^{\hat{P}(m)}_{\hat{J}(m),\hat{K}(m)}\right) \quad \textit{and} \quad (P, J, K) = (\hat{P}, \hat{J}, \hat{K})|_{\{1,2,\ldots,m-1\}}. \tag{4.17}$$

Then $V \in R_m$ *and* V *satisfies* (4.4).

Proof. Clearly V defined by (4.13) is an element of C_m. To see that it is in R_m we use Theorem 3.2(c). Now

$$\begin{aligned}
\overline{\alpha^{\hat{P}}_{\hat{J},\hat{K}}} &= \overline{f^P_{J,K}\left(e^{\hat{P}(m)}_{\hat{J}(m),\hat{K}(m)}\right)}, \quad \text{by (4.14)}\\
&= f^{\overline{P}}_{J,K}\left(\overline{e^{\hat{P}(m)}_{\hat{J}(m),\hat{K}(m)}}\right), \quad \text{by (3.5)}\\
&= f^{\overline{P}}_{J,K}\left(e^{\overline{\hat{P}}(m)}_{\hat{J}(m),\hat{K}(m)}\right), \quad \text{by definition of } \overline{\hat{P}}\\
&= \alpha^{\overline{\hat{P}}}_{\hat{J},\hat{K}} \quad \text{since } \overline{P} = \overline{\hat{P}}|_{\{1,\ldots,m-1\}}.
\end{aligned}$$

Since $\mathcal{J}_P = \mathcal{J}_{\overline{P}}, \mathcal{K}_{J,P} = \mathcal{K}_{J,\overline{P}}$ it is immediate that the criterion in Theorem 3.2(c) is satisfied and hence $V \in R_m$.

Now to see that (4.4) is satisfied. Since $\mu_P = -\sum_{\ell=1}^{m-1} \lambda_{P(\ell)}$ we find, by (4.7), that

$$\left\{\sum_{\ell=1}^{m-1} \lambda_{P(\ell)} f^P_{J,K} + \sum_{\ell=1}^{m-1} f^P_{J,K_\ell} + f^P_{J,K} \circ A\right\}(z) = 0, \quad z \in \mathbb{C}^N.$$

In particular, if $\hat{P} \in \hat{\mathcal{P}}, \hat{J} \in \hat{\mathcal{J}}_{\hat{P}}, \hat{K} \in \hat{\mathcal{K}}_{\hat{J},\hat{P}}, z = e^{\hat{P}(m)}_{\hat{J}(m),\hat{K}(m)}$ and if $(P, J, K) = (\hat{P}, \hat{J}, \hat{K})|_{\{1,\dots,m-1\}}$, then

$$\left(\sum_{\ell=1}^{m-1} \lambda_{P(\ell)} f^P_{J,K} + \sum_{\ell=1}^{m-1} f^P_{J,K_\ell} + f^P_{J,K} \circ A\right)\left(e^{\hat{P}(m)}_{\hat{J}(m),\hat{K}(m)}\right) = 0. \tag{4.18}$$

However, by definition of V,

$$V\left(e^{\hat{P}(1)}_{\hat{J}(1),\hat{K}(1)}, \dots, e^{\hat{P}(m)}_{\hat{J}(m),\hat{K}(m)}\right) = f^P_{J,K}\left(e^{\hat{P}(m)}_{\hat{J}(m),\hat{K}(m)}\right)$$

and since $Ae^{\hat{P}(\ell)}_{\hat{J}(\ell),\hat{K}(\ell)} = \lambda_{P(\ell)} e^{\hat{P}(\ell)}_{\hat{J}(\ell),\hat{K}(\ell)} + e^{\hat{P}(\ell)}_{\hat{J}(\ell),\hat{K}(\ell)-1}$, (4.15) can be re-written

$$\sum_{\ell=1}^{m} V\left(e^{\hat{P}(1)}_{\hat{J}(1),\hat{K}(1)}, \dots, Ae^{\hat{P}(\ell)}_{\hat{J}(\ell),\hat{K}(\ell)}, \dots, e^{\hat{P}(m)}_{\hat{J}(m),\hat{K}(m)}\right) = 0$$

for all $\hat{P} \in \hat{\mathcal{P}}, \hat{J} \in \hat{\mathcal{J}}_{\hat{P}}, \hat{K} \in \hat{\mathcal{K}}_{\hat{J},\hat{P}}$. But

$$\sum_{\ell=1}^{m} V(z_1, \dots, Az_\ell, \dots, z_m)$$

$$= \sum_{\ell=1}^{m} \sum_{\hat{P}\in\hat{\mathcal{P}},\, \hat{J}\in\hat{\mathcal{J}}_{\hat{P}},\, \hat{K}\in\hat{\mathcal{K}}_{\hat{J},\hat{P}}} \alpha^{\hat{P}}_{\hat{J},\hat{K}} V^{\hat{P}}_{\hat{J},\hat{K}}\left(e^{\hat{P}(1)}_{\hat{J}(1),\hat{K}(1)}, Ae^{\hat{P}(\ell)}_{\hat{J}(\ell),\hat{K}(\ell)}, e^{\hat{P}(m)}_{\hat{J}(m),\hat{K}(m)}\right) = 0.$$

This shows that V satisfies (4.4). This completes the proof. □

Acknowledgements

Professor Dancer wishes to acknowledge the support of a United Kingdom SERC Visiting Fellowship GR J 98158 during the tenure of which this paper was completed.

References

[1] N.G. Chetayev, *The Stability of Motion* (Pergamon Press, Oxford, 1961).
[2] E.N. Dancer and J.F. Toland, A degree theory for orbits of prescribed period of flows with a first integral, Proc. London Math. Soc. 60 (3) (1990) 549–580.
[3] E.N. Dancer and J.F. Toland, Equilibrium states in the degree theory of periodic orbits with a first integral, Proc. London Math. Soc. 63 (1991) 569–594.
[4] E.N. Dancer and J.F. Toland, The index change and global bifurcation for flows with a first integral, Proc. London Math. Soc. 66 (1993) 539–567.
[5] I. Ekeland, *Convexity Methods in Hamiltonian Mechanics* (Springer, Berlin, 1990).
[6] E. Hille and R.S. Phillips, Functional analysis and semi-groups, in: *Colloquium Publications*, Vol. XXXI (American Mathematical Society, Providence, RI, 1957).
[7] M.A. Krasnosel'skii, *Topological Methods in the Theory of Nonlinear Integral Equations* (Pergamon Press, Oxford, 1963).
[8] K.R. Meyer and G.R. Hall, *Introduction to Hamiltonian Dynamical Systems and the N-Body Problem* (Springer, New York, 1992).
[9] V. Yakubovich and V. Starzhinskii, *Linear Differential Equations with Periodic Coefficients*, Vols. I and II (Halstead Press, Wiley, New York, 1975).