A NEW TRANSVECTANT ALGORITHM FOR NILPOTENT NORMAL FORMS

James Murdock
Department of Mathematics
Iowa State University
Ames, Iowa 50011

and

Jan A. Sanders
Vrije Universiteit Amsterdam
Faculty of Science
Department of Mathematics
Amsterdam, The Netherlands

Running head: NEW TRANSVECTANT ALGORITHM

phone: 515-294-8163
email: jmurdock@iastate.edu
fax: 515-294-5454

ABSTRACT: We offer an algorithm to determine the form of the normal form for a vector field with a nilpotent linear part, when the form of the normal form is known for each Jordan block of the linear part taken separately. The algorithm is based on the notion of transvectant, from classical invariant theory.

1. Introduction

There are well-known procedures for putting a system of differential equations $\dot{x} = Ax + v(x)$ (where v is a formal power series beginning with quadratic terms) into *normal form* with respect to its linear part A. Our concern here is with the *description problem* of normal form theory: given A, to describe the *normal form space* of A, that is, the set of all v such that $Ax + v(x)$ is in normal form. Our main result is a procedure that solves the description problem when A is a nilpotent matrix in Jordan form, provided that the description problem is already solved for each Jordan block of A taken separately. This procedure will be illustrated with several examples that are already known, and one (a 7×7 matrix with three Jordan blocks) that, to our knowledge, has not been handled before. Additional examples will appear in [6]. All of the examples in this paper are calculated by hand, but some of those in [6] are done by machine. The normal form for our 7×7 matrix is so large (i.e. has so many terms) that we will not write them all out, but almost all of the work necessary to find them will be shown, and what remains is entirely mechanical. With examples such as this, our ability to compute normal forms has exceeded our ability to make practical use of them. Further progress will probably depend on the development of machine-based methods to obtain unfoldings, scalings, and bifurcation diagrams, and to select from among the possible bifurcations the ones that are relevant to a specific application.

The normal form space of a matrix A is not unique, but depends on a choice of *normal form style*. When A is semisimple (diagonalizable), the only useful normal form space is the space of all vector fields v that *commute with* A in the sense that the Lie derivative $\mathsf{L}_A v = 0$, where

$$(\mathsf{L}_A v)(x) = v'(x)Ax - Av(x).$$

In this case the normal form space $\ker \mathsf{L}_A$ forms a module (the *module of equivariants*) over the *ring of invariants*, that is, the ring $\ker \mathcal{D}_A$ of scalar formal power series f such that $\mathcal{D}_A f = 0$, where

$$(\mathcal{D}_A f)(x) = f'(x)Ax = (Ax) \cdot \nabla f(x).$$

These formal power series (for both vector and scalar fields) are equivalent, by the Borel-Ritt theorem, to smooth functions modulo flat functions. The invariants are constant along the flow e^{At} of Ax, and the equivariants have flows that commute with the flow of Ax. In the nilpotent case that we consider here, things are not quite so simple.

We adopt the following notations for nilpotent matrices. For each positive integer $k \geq 2$, N_k is the $k \times k$ nilpotent matrix having a single

Jordan block (with the off-diagonal ones above the diagonal):

$$N_k = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & 0 & 1 & \\ & & & 0 & 1 \\ & & & & 0 \end{bmatrix}.$$

The matrices M_k and H_k are the $k \times k$ matrices

$$M_k = \begin{bmatrix} 0 & & & & \\ (k-1) & 0 & & & \\ & 2(k-2) & 0 & & \\ & & & \ddots & \\ & & & & 0 \end{bmatrix}$$

and

$$H_k = \begin{bmatrix} (k-1) & & & & \\ & (k-3) & & & \\ & & (k-5) & & \\ & & & \ddots & \\ & & & & -(k-1) \end{bmatrix}.$$

The diagonal entries of H_k are either all odd or all even, and are symmetrical around zero. More generally, $N_{k,\ell}$ will denote the matrices of the form

$$N_{k,\ell} = \begin{bmatrix} N_k & 0 \\ 0 & N_\ell \end{bmatrix},$$

and this notation is extended in the obvious way to define $M_{k,\ell}$, $H_{k,\ell}$, and matrices with additional subscripts. The commutator brackets of any such matrices M, N, and H (with the same subscripts) satisfy

$$[N, M] = H, \qquad [H, N] = 2N, \qquad [H, M] = -2M.$$

That is, $\{N, M, H\}$ is an *sl(2) triad*, forming a basis for a representation of the Lie algebra sl(2).

There are two major normal form styles for vector fields having nilpotent linear part. The vector field $Nx + v(x)$ is in *inner product normal form* if $v \in \ker \mathsf{L}_{N^*}$, and in *sl(2) normal form* if $v \in \ker \mathsf{L}_M$. Of course $\ker \mathsf{L}_{N^*}$ is a module over the ring $\ker \mathcal{D}_{N^*}$, and $\ker \mathsf{L}_M$ is a module over $\ker \mathcal{D}_M$. The inner product normal form style is more popular than the sl(2) style, both because it is simpler to explain and because the expressions for v in the sl(2) style involve numerical constants ("fudge factors") that make the style seem harder. But the sl(2) style has useful mathematical structure that the inner product style lacks. Therefore we call attention to the following easy observation, which brings the two styles closer together.

Lemma 1. *The system $\dot{x} = Nx + v(x)$ is in inner product normal form if and only if $\dot{x} = M^*x + v(x)$ is in sl(2) normal form.*

Before proving the lemma, we emphasize that it does *not* say that $Nx+v(x)$ and $M^*x+v(x)$ are normal forms *of the same system*; in fact the computational procedures for putting a system into normal form involve different projection maps in the two styles.

Proof. The matrices $\{M^*, N^*, H\}$ satisfy $[M^*, N^*] = H$, $[H, M^*] = 2M^*$, and $[H, N^*] = 2N^*$, so these also form an sl(2) triad. It follows that $M^*x + v(x)$ is in sl(2) normal form if and only if $\mathsf{L}_{N^*}v = 0$. This is the same as the condition for $Nx + v(x)$ to be in inner product normal form. □

From this point on, we focus on the description of ker L_{N^*} as a module over the ring of invariants ker $\mathcal{D}_{N^*}$. In view of Lemma 1, we are describing either the inner product normal form with leading term N, or the sl(2) normal form with leading term M^*. Of course M^* is as good a choice of canonical form for a nilpotent matrix as N, and making this choice (when an sl(2) normal form is desired) has the effect of removing the "fudge factors" from the higher order terms. (The only "fudge factors" now appear in the linear term, and are simpler.) Although our results can be applied to the inner product normal form, the proofs are entirely dependent on sl(2) representation theory and cannot be expressed in the language of the inner product theory alone.

This paper is an outgrowth of methods described in [4]. In that paper, a method was given which "boosts" a description of the invariant ring ker $\mathcal{D}_{N^*}$ to a description of the equivariant module ker L_{N^*}. We have now realized that the same technique, stated differently, allows us to describe the invariant ring ker $\mathcal{D}_{N^*}$ given the invariant rings of the Jordan blocks in N^*. So the natural place to start is with the invariant rings of the Jordan blocks. From there we obtain the invariant ring of N^*, and then (by boosting) the equivariant module for N^*. This is the order that will be followed below. Although we refer to [4] occasionally to avoid repeating some details, this paper is largely independent of [4], and a new reader should begin here and refer to [4] only as needed. For a complete introduction to normal forms using notations consistent with this paper and with [4], see [5]. Another exposition of our results, rather different in style from the present one, will be forthcoming in chapter 12 of [6]. (The paper [1] deals with a related problem, but treats only the Hilbert function for the invariants, which does not completely identify the ring).

A central notion in this paper is the "box product," defined and discussed in section 4. A quick abstract definition is as follows. Recall that for any sl(2) representation space V with triad $\{X, Y, Z\}$, eigenvectors of Z are called *weight vectors*, their eigenvalues are called *weights*, and any weight vector in ker X is the top weight vector of an irreducible subrepresentation. Since ker X is the span of all the top weight vectors, all of V can be obtained from ker X by applying Y repeatedly. So we may consider ker X as expressing the entire representation space in "abridged form." Now let V_k, $k = 1, 2$, be sl(2) representation spaces with sl(2) triads $\{X_k, Y_k, Z_k\}$. Then $V_1 \otimes V_2$ is a representation space with triad $\{X, Y, Z\}$, where $X = X_1 \otimes I + I \otimes X_2$ (and similarly for Y and Z). We now define the *box product* by

$$(\ker X_1) \boxtimes (\ker X_2) = \ker X.$$

The box product is not equal to $(\ker X_1) \otimes (\ker X_2)$. Instead, it is the "abridged form" (in the sense mentioned above) of the full tensor product $V_1 \otimes V_2$.

To begin to put the box product into a computationally useful form, we use the notion of "external transvectant," introduced in [4]. If $a \in \ker X_1$ and $b \in \ker X_2$ are weight vectors with weights w_a and w_b, and i is an integer in the range $0 \le i \le \min(w_a, w_b)$, then $(a, b)^{(i)}$, the *i-th external transvectant* of a and b, is the element of $V_1 \otimes V_2$ defined by

$$(a,b)^{(i)} = \sum_{j=0}^{i} (-1)^j W_{ab}^{ij} (Y_1^j a) \otimes (Y_2^{i-j} b), \tag{1}$$

where

$$W_{ab}^{ij} = \binom{i}{j} \frac{(w_a - j)!}{(w_a - i)!} \cdot \frac{(w_b - i + j)!}{(w_b - i)!}.$$

The external transvectants lie in $(\ker X_1) \boxtimes (\ker X_2)$, and if a and b range over weight bases (bases consisting of weight vectors) for ker X_1 and ker X_2, then their external transvectants range over a basis for $(\ker X_1) \boxtimes (\ker X_2)$. This will be spelled out in more detail, for our specific applications, in section 3 below. These applications cannot be done in the abstract setting, because they depend on the fact that our representation spaces are formal power series rings. (Polynomial rings would also work.) The simplest case of these calculations (because one of the representation spaces is finite-dimensional) occurs in the boosting argument, in section 6. The reader may wish to turn to section 6 after

Lemma 8 in section 4, before studying the harder examples in section 4.

We conclude this section with one more remark that does work in the abstract setting, that will be needed in section 4. With the same notations used above, put $V_k^0 = \ker X_k \subset V_k$ for $k = 1, 2$. Define a vector subspace $W_k^0 \subset V_k^0$ to be *admissible* if it has a basis consisting of weight vectors. In that case, these weight vectors will be the top weight vectors of irreducible representations, and the direct sum of these representations will be a new representation space $W_k = \bigoplus_{j=1}^{\infty} Y^j W_k^0 \subset V_k$ (the sum is actually finite). Notice that W_k is independent of the choice of a weight basis for W_k^0. Now given two admissible subspaces $W_1^0 \subset V_1^0$ and $W_2^0 \subset V_2^0$, everything that we said about box products can be repeated: we can define

$$W_1^0 \boxtimes W_2^0 = \ker\,(X_1|_{W_1} \otimes I + I \otimes X_2|_{W_2}),$$

and this will have for a basis the transvectants $(a, b)^{(i)}$ as a and b range over any (fixed choice of) weight bases for W_1^0 and W_2^0. The box product of subspaces that are not admissible cannot be defined.

2. Describing invariant rings by Stanley decompositions

The most effective way of describing the invariant ring associated with a nilpotent matrix N is by a device from commutative algebra called a *Stanley decomposition*, introduced for this purpose in [2]. In this section we define Stanley decomposition and state the Stanley decompositions for N_2, N_3, and N_4. These will be used later to obtain Stanley decompositions for the invariants of $N_{2,3}$ and other nilpotent matrices with more than one Jordan block. Derivations of the results in this section may be found in section 4.7 of [5] and section 4 of [4].

We write $\mathbb{R}[[x_1, \ldots, x_n]]$ for the ring of (scalar) formal power series in $x_1, \ldots, x_n$. A *subalgebra* $\mathcal{R}$ of $\mathbb{R}[[x_1, \ldots, x_n]]$ is a subset that is both a subring and a vector subspace. The subalgebra is *graded* if

$$\mathcal{R} = \bigoplus_{d=0}^{\infty} \mathcal{R}_d,$$

where $\mathcal{R}_d$ is the vector subspace of $\mathcal{R}$ consisting of elements of degree d. (The infinite direct sum should technically be called a direct product, since an element of $\mathcal{R}$ can be a sum of infinitely many nonzero terms. But a direct product is usually regarded as in “infinite tuple” rather than a sum.) To define a *Stanley decomposition* of a graded subalgebra $\mathcal{R} \subset \mathbb{R}[[x_1, \ldots, x_n]]$, we begin with the definition of a Stanley term. A *Stanley term* is an expression of the form $\mathbb{R}[[f_1, \ldots, f_k]]\varphi$, where the elements $f_1, \ldots, f_k, \varphi$ are homogeneous polynomials and $f_1, \ldots, f_k$ (not including φ) are required to be algebraically independent. The Stanley term $\mathbb{R}[[f_1, \ldots, f_k]]\varphi$ denotes the set of all expressions of the form $F(f_1, \ldots, f_k)\varphi$, where F is a formal power series in k variables. When $\varphi = 1$, φ is omitted, and the Stanley term is a subalgebra, otherwise it is only a subspace. A *Stanley decomposition* is a finite direct sum of Stanley terms. (The integer k, and the entries $f_1, \ldots, f_k$, will in general be different in different terms of the decomposition. The case in which all terms have the same $f_1, \ldots, f_k$, and only the φ differ, is known as a *Hironaka decomposition.*) The algebraic independence and direct sum conditions in the definition of a Stanley decomposition imply that each element of the subalgebra has a unique expression in the form dictated by the Stanley decomposition. (The Stanley decomposition itself is, however, not unique. For instance, the formal power series ring $\mathbb{R}[[x]]$ in one variable has Stanley decompositions $\mathbb{R}[[x]]$, $\mathbb{R} \oplus \mathbb{R}[[x]]x$, and $\mathbb{R} \oplus \mathbb{R}x \oplus \mathbb{R}[[x]]x^2$, among others.)

Throughout this paper we are concerned with *doubly graded Stanley decompositions* graded by degree and weight. If N is any $n \times n$ nilpotent matrix in upper Jordan form with sl(2) triad $\{N, M, H\}$ as defined above, we write

$$\mathcal{X} = \mathcal{D}_{N^*}, \quad \mathcal{Y} = \mathcal{D}_{M^*}, \quad \mathcal{Z} = \mathcal{D}_H.$$

For instance,

$$\mathcal{X}_2 = \mathcal{D}_{N_2^*} = (N_2^* x) \cdot \nabla = x_1 \frac{\partial}{\partial x_2}.$$

The linear operators $\{\mathcal{X}, \mathcal{Y}, \mathcal{Z}\}$ themselves form an sl(2) triad operating on $\mathbb{R}[[x_1, \ldots, x_n]]$. Eigenfunctions of $\mathcal{Z}$ are called *weight functions*, and the eigenvalues are called *weights*. (The weight of f is denoted w_f.) Because H is diagonal, all monomials in $(x_1, \ldots, x_n)$ are weight functions. Since each formal power series is a sum of monomials having a degree and a weight, $\mathbb{R}[[x_1, \ldots, x_n]]$ is doubly graded by degree and weight. A polynomial f is called *doubly homogeneous of type* (d, w) if every monomial in f has degree d and weight w. Weights are integers, and unlike degrees can be negative, but invariants cannot have negative weights. (This is because an invariant is the top weight vector of an irreducible subrepresentation of sl(2).) A vector subspace V (which may also be a subalgebra $\mathcal{R}$) of ker $\mathcal{X}$ is *doubly graded* if

$$V = \bigoplus_{d=0}^{\infty} \bigoplus_{w=0}^{\infty} V_{dw},$$

where V_{dw} is the vector subspace of V consisting of doubly homogeneous polynomials of degree d and weight w.

A *(doubly graded) Stanley decomposition* of a doubly graded subalgebra $\mathcal{R}$ of ker $\mathcal{X}$ is an expression of $\mathcal{R}$ as a direct sum of vector subspaces of the form $\mathbb{R}[[f_1, \dots, f_k]]\varphi$, where $f_1, \dots, f_k, \varphi$ are doubly homogeneous polynomials (which, being invariants, have nonnegative weights) and $f_1, \dots, f_k$ are algebraically independent. From here on, all Stanley decompositions we consider are of this kind, and we omit the words "doubly graded."

A *standard monomial* associated with a Stanley decomposition is an expression of the form $f_1^{m_1} \cdots f_k^{m_k}\varphi$, where $\mathbb{R}[[f_1, \dots, f_k]]\varphi$ is a term in the Stanley decomposition. Notice that "monomial" here means a monomial in the basic invariants, which are polynomials in the original variables $x_1, \dots, x_n$. The term "standard monomial" comes from Gröbner basis theory, which is used to prove the existence of Stanley decompositions (see [7], section 4 of [4], and appendix A5 of [5]). Given a Stanley decomposition of ker $\mathcal{X}$, its standard monomials of a given degree (or of a given type) form a basis for the (finite-dimensional) vector space of invariants of that degree (or type).

Next we give Stanley decompositions for rings of invariants associated with N_2, N_3, and N_4.

The ring of invariants of N_2 in $R[x_1, x_2]$ is ker $\mathcal{X}_2$. This ring clearly contains

$$\alpha = x_1,$$

which is of type (1,1), and in fact every element of ker $\mathcal{X}_2$ can be written uniquely as a formal power series $f(x_1)$ in x_1 alone. We express this by the Stanley decomposition

$$\ker \mathcal{X}_2 = \mathbb{R}[[\alpha]].$$

The invariants of N_3 in $\mathbb{R}[[x_1, x_2, x_3]]$ are described by the Stanley decomposition

$$\ker \mathcal{X}_3 = \mathbb{R}[[\alpha, \beta]]$$

with

$$\alpha = x_1, \qquad \beta = x_2^2 - 2x_1x_3.$$

Here α is of type (1,2) and β is of type (2,0). Notice that although α has the same form as for N_2, it has a different weight.

The invariants of N_4 in $\mathbb{R}[[x_1, x_2, x_3, x_4]]$ are described by the Stanley decomposition

$$\ker \mathcal{X}_4 = \mathbb{R}[[\alpha, \beta, \delta]] \oplus \mathbb{R}[[\alpha, \beta, \delta]]\gamma$$

where

			degree	weight
α	$=$	x_1	1	3
β	$=$	$x_2^2 - 2x_1x_3$	2	2
γ	$=$	$x_2^3 - 3x_1x_2x_3 + 3x_1^2x_4$	3	3

and

$$\delta = 9x_1^2x_4^2 - 3x_2^2x_3^2 - 18x_1x_2x_3x_4 + 8x_1x_3^3 + 6x_2^3x_4$$

which is of type (4,0). The meaning of this Stanley decomposition is that every element of ker $\mathcal{X}_4$ can be written uniquely in the form

$$f(\alpha, \beta, \delta) + g(\alpha, \beta, \delta)\gamma,$$

where f and g are formal power series. Thus α, β, and γ may occur to any power, but γ can only occur to the first power. The reason for this is that

$$\gamma^2 = \beta^3 + \alpha^2\delta,$$

so any appearances of γ^2 can be replaced by expressions in the other basic invariants. This illustrates how Stanley decompositions enforce uniqueness in the expression of invariants. A standard monomial for this decomposition of ker $\mathcal{X}_4$ is any monomial $\alpha^i\beta^j\gamma^k\delta^\ell$ with $k = 0$ or 1.

3. Invariants of matrices with multiple Jordan blocks

Consider a system with nilpotent linear part

$$N = \begin{bmatrix} \widehat{N} & 0 \\ 0 & \widetilde{N} \end{bmatrix},$$

where $\widehat{N}$ and $\widetilde{N}$ are nilpotent matrices of sizes $\widehat{n} \times \widehat{n}$ and $\widetilde{n} \times \widetilde{n}$ respectively ($\widehat{n} + \widetilde{n} = n$), in (upper) Jordan form, and each may consist of one or more Jordan blocks. Let $\{\mathcal{X}, \mathcal{Y}, \mathcal{Z}\}$, $\{\widehat{\mathcal{X}}, \widehat{\mathcal{Y}}, \widehat{\mathcal{Z}}\}$, and $\{\widetilde{\mathcal{X}}, \widetilde{\mathcal{Y}}, \widetilde{\mathcal{Z}}\}$ be the associated triads of operators. Notice that the first triad acts on $\mathbb{R}[[x_1, \dots, x_n]]$, the second on $\mathbb{R}[[x_1, \dots, x_{\widehat{n}}]]$, and the third on $\mathbb{R}[[x_{\widehat{n}+1}, \dots, x_n]]$.

Suppose that $f = f(x_1, \dots, x_{\widehat{n}}) \in \ker \widehat{\mathcal{X}}$ and $g = g(x_{\widehat{n}+1}, \dots, x_n) \in \ker \widetilde{\mathcal{X}}$ are weight invariants of weights w_f and w_g, and i is an integer in the range $0 \le i \le \min(w_f, w_g)$. Then we define *external transvectant* of f and g of order i to be the polynomial $(f, g)^{(i)} \in \mathbb{R}[[x_1, \dots, x_n]]$ given by

$$(f, g)^{(i)} = \sum_{j=0}^{i} (-1)^j W_{f,g}^{i,j} (\widehat{\mathcal{Y}}^j f)(\widetilde{\mathcal{Y}}^{i-j} g), \tag{2}$$

where

$$W^{i,j}_{f,g} = \binom{i}{j}\frac{(w_f - j)!}{(w_f - i)!} \cdot \frac{(w_g - i + j)!}{(w_g - i)!}.$$

(The proof of the next lemma explains why we omit the $\otimes$ occurring in the abstract definition (1) in the introduction.) We say that a transvectant $(f,g)^{(i)}$ is *well-defined* if i is in the proper range for f and g. Notice that the zeroth transvectant is always well-defined and reduces to the product: $(f,g)^{(0)} = fg$. Given Stanley decompositions for $\ker \widehat{\mathfrak{X}}$ and $\ker \widetilde{\mathfrak{X}}$, the following theorem provides a basis for $\ker \mathfrak{X}$ in each degree. This is a first step toward obtaining a Stanley decomposition for $\ker \mathfrak{X}$.

Theorem 2. *Each well-defined transvectant* $(f,g)^{(i)}$ *of* $f \in \ker \widehat{\mathfrak{X}}$ *and* $g \in \ker \widetilde{\mathfrak{X}}$ *belongs to* $\ker \mathfrak{X}$. *If f and g are doubly homogeneous polynomials of types (d_f, w_f) and (d_g, w_g) respectively, $(f,g)^{(i)}$ is a doubly homogeneous polynomial of type $(d_f + d_g, w_f + w_g - 2i)$. Suppose that Stanley decompositions for* $\ker \widehat{\mathfrak{X}}$ *and* $\ker \widetilde{\mathfrak{X}}$ *are given. Then a basis for the (finite-dimensional) subspace* $(\ker \mathfrak{X})_d$ *of homogeneous polynomials in* $\ker \mathfrak{X}$ *with degree d is given by the set of all well-defined transvectants $(f,g)^{(i)}$ where f is a standard monomial of the Stanley decomposition for* $\ker \widehat{\mathfrak{X}}$, *$g$ is a standard monomial of the Stanley decomposition for* $\ker \widetilde{\mathfrak{X}}$, *and $d_f + d_g = d$.*

Proof. The proof of this theorem is given in section 6 of [4] and will not be repeated here in full. We will briefly outline the ideas used in the proof, and make a small correction to Lemma 4 of [4]. Let $\mathbb{R}[[x_1, \ldots, x_{\widehat{n}}]]_{\widehat{d}}$ denote the subspace of $\mathbb{R}[[x_1, \ldots, x_{\widehat{n}}]]$ consisting of elements that are homogeneous of degree $\widehat{d}$, with similar notations for the other rings. Then we may view $\mathbb{R}[[x_1, \ldots, x_n]]_d$ as the direct sum of tensor products

$$\mathbb{R}[[x_1, \ldots, x_n]]_d = \bigoplus_{\widehat{d} + \widetilde{d} = d} \mathbb{R}[[x_1, \ldots, x_{\widehat{n}}]]_{\widehat{d}} \otimes \mathbb{R}[[x_{\widehat{n}+1}, \ldots, x_n]]_{\widetilde{d}}.$$

The tensor product may be replaced with the ordinary product of polynomials, because there is no overlap between the variables appearing in the polynomials in the two spaces being tensored. (This nonoverlap condition implies that the ordinary product satisfies the algebraic requirements for a tensor product map.) Furthermore, the sl(2) representation on $\mathbb{R}[[x_1, \ldots, x_n]]_d$ given by $\{\mathfrak{X}, \mathfrak{Y}, \mathfrak{Z}\}$ is the direct sum of the tensor products of the sl(2) representations on the other spaces. (Recall that the tensor product of the two Lie algebra representations $\{\widehat{\mathfrak{X}}, \widehat{\mathfrak{Y}}, \widehat{\mathfrak{Z}}\}$ and $\{\widetilde{\mathfrak{X}}, \widetilde{\mathfrak{Y}}, \widetilde{\mathfrak{Z}}\}$ is defined to be $\{\widehat{\mathfrak{X}} \otimes I + I \otimes \widetilde{\mathfrak{X}}, \widehat{\mathfrak{Y}} \otimes I + I \otimes \widetilde{\mathfrak{Y}}, \widehat{\mathfrak{Z}} \otimes I + I \otimes \widetilde{\mathfrak{Z}}\}$.

In this case it reduces to $\{\widehat{\mathcal{X}} + \widetilde{\mathcal{X}}, \widehat{\mathcal{Y}} + \widetilde{\mathcal{Y}}, \widehat{\mathcal{Z}} + \widetilde{\mathcal{Z}}\}$.) It follows that the irreducible subrepresentations in $\mathbb{R}[[x_1, \dots, x_n]]_{\widehat{d}+\widetilde{d}}$ are specified by the Clebsch-Gordan theorem, and the top weight vectors (chain tops) of these subrepresentations (chains) are given by the transvectants. There are two small errors in Lemma 4 of [4]; s should be the minimum weight, not the minimum length, of the two chains, and the transvectant is undefined, not zero, when $i > s$. □

The bases given by Theorem 2 are sufficient to determine ker $\mathcal{X}$ one degree at a time, but to find all of ker $\mathcal{X}$ in this way would require finding infinitely many transvectants. A Stanley decomposition for ker $\mathcal{X}$ must be based on a finite number of basic invariants. To construct such a decomposition, we must first find an alternative basis for each (ker $\mathcal{X})_d$ that uses only a finite number of transvectants overall. (We do not count zeroth transvectants, which are simply products. A Stanley decomposition can produce an infinite number of products). Such alternative bases can be found by the following replacement theorem.

Theorem 3. *Any transvectant $(f, g)^{(i)}$ in the basis given by Theorem 2 can be replaced by a product $(f_1, g_1)^{(i_1)} \cdots (f_j, g_j)^{(i_j)}$ of transvectants, provided that $f_1 \cdots f_j = f$, $g_1 \cdots g_j = g$, and $i_1 + \cdots + i_j = i$.*

Since a zeroth transvectant is a product, the replacements given by this theorem are best viewed as products of standard monomials in ker $\widehat{\mathcal{X}}$ and ker $\widetilde{\mathcal{X}}$ and transvectants (of order greater than zero) of such monomials, subject to the conditions that the *stripped form* of the product equals fg and the *total transvectant order* equals i. (The "stripped form" of such a product is the form obtained by erasing the transvectant signs. Thus the stripped form of $(f_1, g_1)^{(i_1)}(f_2, g_2)^{(i_2)}$ is $f_1g_1f_2g_2$.)

Proof. The main task is to show that the replacements proposed in the theorem are linearly independent. We do this by contradiction. We suppose that they are linearly dependent, that is, there exists a nontrivial linear combination of replacements that is equal to zero. We show that when $\widehat{\mathcal{X}}$ is applied enough times to this linear combination, the result is a nontrivial linear combination of terms that are already known to be linearly independent. Since this is impossible, the original supposition (of linear dependence) is impossible. The details depend on the representation theory of sl(2). The proof is modelled on that of Lemma 2 in [4], and the reader may wish to study this easier theorem for motivation.

First, observe that

$$\widehat{\mathcal{X}}^i (f, g)^{(i)} = cfg$$

for some nonzero constant c. In fact $\widehat{\mathcal{X}}^i$ annihilates all terms of $(f,g)^{(i)}$ except the term that is a constant times $(\widehat{\mathcal{Y}}^i f)g$, and multiplies this term by a strictly positive number (positive because it is a sum of weights of invariants). In this calculation we use the following facts: that $f \in \ker \widehat{\mathcal{X}}$; that g depends only on $(x_{\widehat{n}+1}, \ldots, x_n)$ and hence is also annihilated by $\widehat{\mathcal{X}}$; and that $\widehat{\mathcal{Y}}^i f \neq 0$ because the weight of f is at least i (or else the transvectant would not be defined). These remarks can be extended to products of two or more transvectants. For instance,

$$\widehat{\mathcal{X}}^{i+j}(f,g)^{(i)}(h,k)^{(j)} = cfhgk$$

for some nonzero c because only the "dominant" term containing $(\widehat{\mathcal{Y}}^i f)(\widehat{\mathcal{Y}}^j h)gk$ survives. In the following argument we use the word "replacement" to mean "a product of the form proposed in the theorem as a replacement for a transvectant."

Next, observe that no two basis elements from Theorem 2 that have the same stripped form can have the same weight. (If the stripped form is fg, the basis elements will be $(f,g)^{(i)}$ for various i, and these all have different weights.) Now any replacement for one of these basis elements will have the same stripped form and the same weight. Therefore in any full set of replacements, no two elements with the same stripped form will have the same weight.

Now suppose there exists a nontrivial linear combination of replacements that is equal to zero. Then (since the weight subspaces are independent) there will exist a nontrivial linear combination of replacements of some fixed weight that equals zero. Let r denote the highest total transvectant order occurring in these replacements. Apply $\widehat{\mathcal{X}}^r$ to the linear relation. This will annihilate all terms with total transvectant order less than r, but at least one term will survive. Since we began with replacements of equal weight, and no two replacements of the same weight can have the same stripped form, we have a nontrivial linear combination of distinct terms, each of which is a product of two standard monomials, one from $\ker \widehat{\mathcal{X}}$ and one for $\ker \widetilde{\mathcal{X}}$. These terms must be linearly independent. Therefore the supposition at the beginning of this paragraph is false. Therefore any set of replacements is linearly independent.

Thus the map sending each basis element to its replacement is a one-to-one correspondence of linearly independent vectors that preserves degree and weight, and therefore restricts to a one-to-one map of a basis for each "type subspace" $(\ker \mathcal{X})_{dw}$ of $\ker \mathcal{X}$ to a linearly independent set in the same type subspace having the same cardinality, which must then be another basis for the type subspace. Since $\ker \mathcal{X}$ is the (infinite) direct sum of its type subspaces, the one-to-one correspondence holds for the entire space. □

Of course we could equally well use $\widetilde{\mathfrak{X}}$ in place of $\widehat{\mathfrak{X}}$ to prove the theorem. This would only change the dominant term that survives.

The following corollary of the Replacement Theorem will play a crucial role in our calculations.

Corollary 4. *If $w_h = w_k = r$, so that $(h,k)^{(r)}$ has weight zero, then whenever $(fh, gk)^{(i+r)}$ is well-defined, it may be replaced by $(f,g)^{(i)}(h,k)^{(r)}$.*

Proof. Clearly $(fh, gk)^{(i+r)}$ and $(f,g)^{(i)}(h,k)^{(r)}$ have the same stripped form and total transvectant order. It is only necessary to observe that $(f,g)^{(i)}$ is well-defined. But $w_{fh} = w_f + w_h = w_f + r \geq i + r$, so $w_f \geq i$, and similarly $w_g \geq i$. □

The Replacement Theorem by itself is sufficient for doing some simple computations of Stanley decompositions. We illustrate this with the examples $N_{2,2}$, $N_{2,3}$, and $N_{2,2,2}$. In the next section we develop a more powerful technique.

Knowing that ker $\mathfrak{X}_2 = \mathbb{R}[[\alpha]]$ where $\alpha = x_1$, we can calculate ker $\mathfrak{X}_{2,2}$ as follows. Let $\alpha = x_1$, $\beta = x_3$, $\gamma = (a,b)^{(1)}$.

Theorem 5. ker $\mathfrak{X}_{2,2} = \mathbb{R}[[\alpha,\beta,\gamma]]$.

Proof. According to Theorem 2, a basis for ker $\mathfrak{X}_{2,2}$ is given by $(\alpha^n, \beta^m)^{(i)}$ for $i = 0, \ldots, \min\{n,m\}$. By Theorem 3, the map $(\alpha^m, \beta^n)^{(i)} \mapsto \alpha^{m-i}\beta^{n-i}\gamma^i$ gives a replacement for each basis element. The span of these replacements is the indicated Stanley decomposition. □

For $N_{2,3}$ we have $\widehat{N} = N_2$ and $\widetilde{N} = N_3$, with $\widehat{\mathfrak{X}} = x_1\partial/\partial x_2$ and $\widetilde{\mathfrak{X}} = x_3\partial/\partial x_4 + x_4\beta/\partial x_5$. Then, from the results above for $\mathfrak{X}_2$ and $\mathfrak{X}_3$ expressed in the proper variables, ker $\widehat{\mathfrak{X}} = \mathbb{R}[[\alpha]]$ with $\alpha = x_1$, and ker $\widetilde{\mathfrak{X}} = \mathbb{R}[[\beta,\gamma]]$ with $\beta = x_3$ and $\gamma = x_4^2 - 2x_3x_4$. The types of α, β, γ are $(1,1), (1,2), (2,0)$ respectively. We can now compute ker $\mathfrak{X}_{2,3}$ as follows.

Theorem 6. ker $\mathfrak{X}_{2,3} = \mathbb{R}[[\alpha,\beta,\gamma,(\alpha^2,\beta)^{(2)}]] \oplus \mathbb{R}[[\alpha,\beta,\gamma,(\alpha^2,\beta)^{(2)}]](\alpha,\beta)^{(1)}$.

Proof. The basis elements are of the form $(\alpha^\ell, \beta^m\gamma^n)^{(r)}$ with $\ell \geq r$ and $2m \geq r$. We divide these into two classes, $r = 2j$ (even) and $r = 2j+1$ (odd), noting that in the former case $m \geq j$ and in the latter case $m > j$. For $r = 2j$ the basis elements are $(\alpha^{2j+p}, \beta^{j+q}\gamma^n)^{(2j)}$. For $j = 0$ we get $\mathbb{R}[[\alpha,\beta,\gamma]]$. For $j \geq 1$ we replace these first by $\alpha^p\beta^q\gamma^n(\alpha^{2j},\beta^j)^{(2j)}$ and then by $\alpha^p\beta^q\gamma^n\varepsilon^j$, where $\varepsilon = (\alpha^2,\beta)^{(2)}$, and get $\mathbb{R}[[\alpha,\beta,\gamma,\varepsilon]]\varepsilon$; the two rings computed so far sum to $\mathbb{R}[[\alpha,\beta,\gamma,\varepsilon]]$. For

$r = 2j+1$ the basis elements are $(\alpha^{2j+1+p}, \beta^{j+1+q}\gamma^n)^{2j+1})$, which may be replaced first by $\alpha^p\beta^q\gamma^n(\alpha^{2j+1}, \beta^{j+1})^{(2j+1)}$ and then by $\alpha^p\beta^q\gamma^n\varepsilon^j\delta$, where $\delta = (\alpha, \beta)^{(1)}$. This gives $\mathbb{R}[[\alpha, \beta, \gamma, \varepsilon]]\delta$. □

To interpret the result of Theorem 6, observe that $\widehat{\mathcal{Y}} = x_2\partial/\partial x_1$ and $\widetilde{\mathcal{Y}} = 2x_4\partial/\partial x_3 + 2x_5\beta/\partial x_4$. Then, from the definition of transvectant, we find that

$$\begin{aligned}(\alpha, \beta)^{(1)} &= 2x_1x_4 - 2x_2x_3\\ (\alpha^2, \beta)^{(2)} &= 8x_1^2x_5 - 8x_1x_2x_4 + 4x_2^2x_3.\end{aligned}$$

Thus the theorem states that every invariant for $N_{2,3}$ can be written as

$$\begin{aligned}&f(x_1, x_3, x_4^2 - 2x_3x_4, 8x_1^2x_5 - 8x_1x_2x_4 + 4x_2^2x_3)\\ &\quad + g(x_1, x_3, x_4^2 - 2x_3x_4, 8x_1^2x_5 - 8x_1x_2x_4 + 4x_2^2x_3)(2x_1x_4 - 2x_2x_3).\end{aligned}$$

In the sequel we shall omit this kind of calculation.

Next we compute $\ker \mathcal{X}_{2,2,2}$, using the notation of [3]. That is, we work in $R[x_1, y_1, x_2, y_2, x_3, y_3]$ with $\alpha_i = x_i$, of type $(1,1)$, and $\beta_{1j} = (\alpha_i, \beta_j)^{(1)}$ for $i < j$, of type $(2,0)$. From the calculation of $\ker \mathcal{X}_{2,2}$ above, we have $\ker \mathcal{X}_{2,2} = \mathcal{R}[[\alpha_1, \alpha_2]]$ where $\mathcal{R} = \mathbb{R}[[\beta_{12}]]$. This is convenient because β_{12} has weight 0, and therefore can be factored out of any transvectant in which it appears, using Theorem 3.

Theorem 7. $\ker \mathcal{X}_{2,2,2} = \mathbb{R}[[\alpha_1, \alpha_2, \alpha_3, \beta_{12}, \beta_{13}]] \oplus \mathbb{R}[[\alpha_2, \alpha_3, \beta_{12}, \beta_{13}, \beta_{23}]]\beta_{23}$.

Proof. The basis elements (ignoring β_{12}) are $(\alpha_1^{m_1}\alpha_2^{m_2}, \alpha_3^{m_3})^{(r)}$ with $m_1 + m_2 \geq r$ and $m_3 \geq r$. We take the following cases.

If $r = 0$ we get $\mathcal{R}[[\alpha_1, \alpha_2, \alpha_3]]$, with $\mathcal{R} = \mathbb{R}[[\beta_{12}]]$ as above.

If $r \geq 1$ and $m_1 \geq r$: Write $m_1 = r + s$ and $m_3 = r + t$ and replace first by $\alpha_1^s\alpha_2^{m_2}\alpha_3^t(\alpha_1^r, \alpha_2^r)^{(r)}$ and then by $\alpha_1^s\alpha_2^{m_1}\alpha_3^t\beta_{13}^r$. We get $\mathcal{R}[[\alpha_1, \alpha_2, \alpha_3, \beta_{13}]]\beta_{13}$.

If $r \geq 1$ with $m_1 < r$: Write $m_2 = (r - m_1) + u$ and $m_3 = r + t$. Replace first by $\alpha_2^u\alpha_3^t(\alpha_1^{m_1}\alpha_2^{r-m_1}, \alpha_3^r)^{(r)}$ and then by $\alpha_2^u\alpha_3^t\beta_{13}^{m_1}\beta_{23}^{r-m_1}$. Remembering $r - m_1 > 0$, we get $\mathcal{R}[[\alpha_2, \alpha_3, \beta_{13}, \beta_{23}]]\beta_{23}$.

Now we observe that the sum of the first two cases is

$$\mathcal{R}[[\alpha_1, \alpha_2, \alpha_3]] \oplus \mathcal{R}[[\alpha_1, \alpha_2, \alpha_3, \beta_{13}]]\beta_13 = \mathcal{R}[[\alpha_1, \alpha_2, \alpha_3, \beta_{13}]].$$

Adding the third case and using the definition of $\mathcal{R}$ gives the result stated in the theorem. Notice that α_1 is missing from the second term and β_{23} from the first. This result agrees with [3]. □

4. Box products of Stanley decompositions

The examples in the last section were worked using some elementary number theory to classify the terms. This requires a certain amount of thought in each case. A more mechanical (and therefore programmable) way of classifying terms is by using "expanded" Stanley decompositions. We will illustrate this technique with an easy example, redoing ker $\mathcal{X}_{2,3}$. Then we provide general definitions and a proof that the technique is capable of computing all possible examples. The Stanley decompositions produced by this method are long, but can be simplified by combining terms to undo some of the expansion performed at the beginning of the process.

Using the notation of Theorem 6, we first expand the Stanley decomposition of ker $\mathcal{X}_2$ as

$$\ker \mathcal{X}_2 = \mathbb{R}[[\alpha]] = \mathbb{R} \oplus \mathbb{R}\alpha \oplus \mathbb{R}[[\alpha]]\alpha^2. \tag{3}$$

The three summands are constants, terms with exactly one factor of α, and terms with at least two factors. Similarly we expand ker $\mathcal{X}_3$ as

$$\ker \mathcal{X}_3 = \mathbb{R}[[\beta,\gamma]] = \mathbb{R}[[\gamma]] \oplus \mathbb{R}[[\beta,\gamma]]\beta, \tag{4}$$

classifying the terms into those having no factor of β and those with at least one factor. The reason for these expansions will appear in a moment, but notice that we stop at α^2 and β because these terms have equal weight (in this case 2). It is also important here that γ has weight zero. We never expand on terms with weight zero. Finally, notice that although we have expanded the originally given Stanley decompositions, the standard monomials of the decomposition have not changed.

We want to consider all well-defined transvectants $(f, g)^{(i)}$ with $f \in$ ker $\mathcal{X}_2$ and $g \in$ ker $\mathcal{X}_3$, and provide most of them with replacements. We consider cases for f and g according to the expanded Stanley decompositions.

1. If $f \in \mathbb{R}$ and $g \in \mathbb{R}[[\gamma]]$, no transvectants beyond the zeroth, which is just the product fg, are possible; but f is just a number, so the space obtained in this way is $\mathbb{R}[[\gamma]]$.
2. If $f \in \mathbb{R}$ and $g \in \mathbb{R}[[\beta,\gamma]]\beta$, no transvectants beyond zero can be formed and we obtain $\mathbb{R}[[\beta,\gamma]]\beta$.
3. If $f \in \mathbb{R}\alpha$ and $\gamma \in \mathbb{R}[[\gamma]]$, no transvectants beyond the zeroth can be formed because $w_\gamma = 0$, and we get $\mathbb{R}[[\gamma]]\alpha$.
4. If $f \in \mathbb{R}\alpha$ and $g \in \mathbb{R}[[\beta,\gamma]]\beta$, then the zeroth and first transvectants can be formed, because $w_\alpha = 1$.
   a. The zeroth transvectants give $\mathbb{R}[[\beta,\gamma]](\alpha,\beta)^{(0)} = \mathbb{R}[[\beta, g]]\alpha\beta$.

   b. The first transvectants give $\mathbb{R}[[\beta,\gamma]](\alpha,\beta)^{(1)}$.
5. If $f \in \mathbb{R}\alpha^2$ and $g \in \mathbb{R}[[\gamma]]$, no transvectants beyond the zeroth can be formed, giving $\mathbb{R}[[\gamma]]\alpha^2$.
6. Finally, if $f \in \mathbb{R}[[\alpha]]\alpha^2$ and $g \in \mathbb{R}[[\beta,\gamma]]\beta$, we may write $f = h\alpha^2$ and $g = k\beta$, with $h \in \mathbb{R}[[\alpha]]$ and $k \in \mathbb{R}[[\beta,\gamma]]$. Noting that $w_\alpha = 1$ and $w_\beta = 2$, we divide the possible transvectants into order zero, order one, and orders ≥ 2.
   a. The zeroth transvectants give $\mathbb{R}[[\alpha,\beta,\gamma]]\alpha^2\beta$.
   b. For the first transvectants, it follows from Theorem 3 that $(f,g)^{(1)} = (h\alpha^2, k\beta)^{(1)}$ can be replaced by $hk\alpha(\alpha,\beta)^1$. The space spanned by these replacements is $\mathbb{R}[[\alpha,\beta,\gamma]]\alpha(\alpha,\beta)^{(1)}$.
   c. Since α^2 and β have equal weight 2, it follows from Corollary 4 that $(f,g)^{(i)}$ with $i \geq 2$ can be replaced by $(h,k)^{(i-2)}(\alpha^2,\beta)^{(2)}$ in all cases. Here $i-2 \geq 0$, and the transvectants $(h,k)^{(i-2)}$ span the entire space ker $\mathcal{X}_{2,3}$. Although it may appear that we are going in circles, since ker $\mathcal{X}_{2,3}$ is just what we are trying to find, the resulting terms span the space (ker $\mathcal{X}_{2,3})(\alpha^2,\beta)^{(2)}$.

Summing up the subspaces we have calculated gives

$$
\begin{aligned}
(5)\qquad \ker \mathcal{X}_{2,3} = {} & \mathbb{R}[[\gamma]] \oplus \mathbb{R}[[\beta,\gamma]]\beta \oplus \mathbb{R}[[\gamma]]\alpha \oplus \mathbb{R}[[\beta,\gamma]]\alpha\beta \\
& \oplus \mathbb{R}[[\beta,\gamma]](\alpha,\beta)^{(1)} \oplus \mathbb{R}[[\gamma]]\alpha^2 \oplus \mathbb{R}[[\ a,\beta,\gamma]]\alpha^2\beta \\
& \oplus \mathbb{R}[[\alpha,\beta,\gamma]]\alpha(\alpha,\beta)^{(1)} \oplus (\ker \mathcal{X}_{2,3})(\alpha^2,\beta)^{(2)}.
\end{aligned}
$$

This is almost a Stanley decomposition, except for the last term. But it is naturally set up for an iteration. Writing $\mathcal{R} = \ker \mathcal{X}_{2,3}$, letting $\mathcal{S}$ denote the sum of all the terms in (5) except the last, and temporarily putting $\varepsilon = (\alpha^2,\beta)^{(2)}$ (as in the proof of Theorem 6), we have

$$\mathcal{R} = \mathcal{S} \oplus \mathcal{R}\varepsilon = \mathcal{S} \oplus \mathcal{S}\varepsilon \oplus \mathcal{R}\varepsilon^2 = \cdots = \mathcal{S}[[\varepsilon]].$$

That is, the zero-weight element ε should be entered into all of the square brackets in the expression for $\mathcal{S}$, and we will have the complete Stanley decomposition for $\mathcal{R}$. Therefore

$$
\begin{aligned}
\ker \mathcal{X}_{2,3} = {} & \mathbb{R}[[\gamma,\varepsilon]] \oplus \mathbb{R}[[\beta,\gamma,\varepsilon]]\beta \oplus \mathbb{R}[[\gamma,\varepsilon]]\alpha \oplus \mathbb{R}[[\beta,\gamma,\varepsilon]]\alpha\beta \\
& \oplus \mathbb{R}[[\beta,\gamma,\varepsilon]](\alpha,\beta)^{(1)} \oplus \mathbb{R}[[\gamma,\varepsilon]]\alpha^2 \\
& \oplus \mathbb{R}[[\alpha,\beta,\gamma,\varepsilon]]\alpha^2\beta \oplus \mathbb{R}[[\alpha,\beta,\gamma,\varepsilon]]\alpha(\alpha,\ b)^{(1)}.
\end{aligned}
$$

As mentioned above, this comes out longer at first than the result of Theorem 6, but the terms can be grouped and summed to give the same result in the end.

Now we formalize this process and prove that it always works. Notice that if we were to "multiply" the Stanley decompositions (3) and (4) with a multiplication that distributes over direct sum, this distributive law would correspond exactly to the classification of cases that we have used in the example. So we define the *box product* of the two spaces of invariants, $(\ker \widehat{\mathfrak{X}}) \boxtimes (\ker \widetilde{\mathfrak{X}})$, to be the space spanned (over $\mathbb{R}$, allowing infinite sums) by the well-defined transvectants $(f, g)^{(i)}$ as f ranges over the standard monomials of some Stanley decomposition for $\ker \widehat{\mathfrak{X}}$ and g ranges over the standard monomials of some Stanley decomposition for $\ker \widetilde{\mathfrak{X}}$. It follows at once from Theorem 2 that the result does not depend on the Stanley decompositions that are used, and that

$$(\ker \widehat{\mathfrak{X}}) \boxtimes (\ker \widetilde{\mathfrak{X}}) = \ker \mathfrak{X}.$$

(This equation is the same as the abstract definition of box product given at the end of section 1.) Furthermore, it follows from Theorem 3 that the box product is also spanned by any set of replacements for the well-defined transvectants $(f, g)^{(i)}$.

In order to obtain a distributive law for the box product (over direct sums), we must extend the definition of the box product to certain subspaces of $\ker \widehat{\mathfrak{X}}$ and $\ker \widetilde{\mathfrak{X}}$. Suppose that $\mathbb{R}[[f_1, \dots, f_k]]\varphi \subset \ker \widehat{\mathfrak{X}}$ and $\mathbb{R}[[g_1, \dots, g_\ell]]\psi \subset \ker \widetilde{\mathfrak{X}}$ are Stanley terms selected from given Stanley decompositions for $\ker \widehat{\mathfrak{X}}$ and $\ker \widetilde{\mathfrak{X}}$. We want to define $(\mathbb{R}[[f_1, \dots, f_k]]\varphi) \boxtimes (\mathbb{R}[[g_1, \dots, g_\ell]]\psi)$ as the space spanned by the well-defined transvectants $(f, g)^{(i)}$, where f is a standard monomial in $\mathbb{R}[[f_1, \dots, f_k]]\varphi$ and g is a standard monomial in $\mathbb{R}[[g_1, \dots, g_\ell]]\psi$. At first sight it appears that the box product of two such subspaces may depend on the Stanley decomposition used, because the same subspace may be spanned by a different set of standard monomials. But we have already dealt with this question in the last paragraph of section 1: standard monomials are weight vectors, so the subspaces of $\ker \widehat{\mathfrak{X}}$ and $\ker \widetilde{\mathfrak{X}}$ that we are dealing with are admissible (in the sense of section 1), and therefore the box product is well defined.

The next lemma is now trivial, but essential to the method.

Lemma 8. *Box distributes over direct sums of admissible subspaces: If $\widehat{V} \subset \ker \widehat{\mathfrak{X}}$, $\widetilde{V}_1 \subset \ker \widetilde{\mathfrak{X}}$, and $\widetilde{V}_2 \subset \ker \widetilde{\mathfrak{X}}$ are admissible subspaces, with $\widetilde{V}_1 \cap \widetilde{V}_2 = \{0\}$, then $\widetilde{V}_1 \oplus \widetilde{V}_2$ is admissible and*

$$\widehat{V} \boxtimes (\widetilde{V}_1 \oplus \widetilde{V}_2) = (\widehat{V} \boxtimes \widetilde{V}_1) \oplus (\widehat{V} \boxtimes \widetilde{V}_2),$$

and similarly for $(\widehat{V}_1 \oplus \widehat{V}_2) \boxtimes \widetilde{V}$.

Proof. The standard monomials of the Stanley decomposition for $\ker \widetilde{\mathfrak{X}}$ that belong to $\widetilde{V}_1 \oplus \widetilde{V}_1$ are partitioned into those in $\widetilde{V}_1$ and those in $\widetilde{V}_2$, since the subspaces are admissible. □

One further issue must be settled before proceeding: Is it legitimate to use replacements for the transvectants in box products of Stanley terms? The answer is a qualified yes. When the transvectants in a basis for the box product are replaced, the span of the replacements may not be exactly the same as the "true" box product, but (according to Theorem 3) the new space will remain linearly independent of the other subspaces in the direct sum, and will serve as a valid replacement for the box product. In the sequel, when we compute box products we actually compute replacements for box products in this sense. In fact, we are able to show that many of the replacements that we use do not modify the box product at all, but we omit the proof because the result will not be used. (This applies to replacements that contain only a single transvectant of order greater than zero. For replacements that contain a product of transvectants, we do not have a clear answer, and the box product space is probably modified.)

The main theorem is that the box product of $\ker \widehat{\mathcal{X}}$ and $\ker \widetilde{\mathcal{X}}$ is computable from given Stanley decompositions of these spaces, and the result is a Stanley decomposition of $\ker \mathcal{X}$. The proof also serves as a description of the computation procedure. Certain choices are required at various points in the proof, and the efficiency of the calculation may depend on the way these choices are made. In the proof, we write $=$ between subspaces that are clearly equal, and $\cong$ between subspaces that serve as replacements for each other in the direct sums (even when we know that these are actually equal).

Theorem 9. *A Stanley decomposition of* $\ker \mathcal{X} = \ker \widehat{\mathcal{X}} \boxtimes \ker \widetilde{\mathcal{X}}$ *is computable in a finite number of steps given Stanley decompositions of* $\ker \widehat{\mathcal{X}}$ *and* $\ker \widetilde{\mathcal{X}}$*.*

Proof. The given Stanley decompositions define standard monomials in $\ker \widehat{\mathcal{X}}$ and $\ker \widetilde{\mathcal{X}}$. During the course of the proof we will preform various expansions of these given decompositions, but the expansions will not change the standard monomials. Therefore the notion of "admissible subspace" does not change as we proceed. Each Stanley term in each Stanley decomposition will be an admissible subspace. By Lemma 8, we can compute $\ker \mathcal{X}$ if we can compute any box product of the form $\mathbb{R}[[f_1, \ldots, f_k]]\varphi \boxtimes \mathbb{R}[[g_1, \ldots, g_\ell]]\psi$, where each factor is a Stanley term from the given decompositions of $\ker \widehat{\mathcal{X}}$ and $\ker \widetilde{\mathcal{X}}$. It turns out that to do so, we must be able to compute any box product of this form in which the factors are admissible.

Let p be the number of elements of weight > 0 in $f_1, \ldots, f_k$, and q the number of such elements in $g_1, \ldots, g_\ell$. We proceed by double induction on p and q. We first construct the box product explicitly in the case $p = q = 0$. Next we handle the general case $p = 0$ by induction on q, by reducing calculations of the form $(0, q)$ to calculations of the form $(0, q-1)$ plus one calculation that is handled explicitly. Since cases (p, q) and (q, p) are symmetric, we will also have handled $(0, p)$. Finally we handle the general case (p, q) by reduction to calculations of the forms $(p-1, q)$, $(p, q-1)$, and $(p-1, q-1)$, and some terms that are handled explicitly. There is a special trick involved in this last reduction that involves a formal iteration argument (as in the example above).

Suppose $p = q = 0$. Then the box product is spanned by transvectants of the form $(f_1^{m_1} \cdots f_k^{m_k}\varphi, g_1^{n_1} \cdots g_\ell^{n_\ell}\psi)^{(i)}$, which is well-defined if and only if $0 \le i \le r$, where $r = \min(w_\varphi, w_\psi$. (The f and g factors add no weight, and cannot support any higher transvectants.) By Theorem 3 each transvectant may be replaced by $f_1^{m_1} \cdots f_k^{m_k}\gamma_1^{n_1} \cdots g_\ell^{n_\ell}(\varphi, \psi)^{(i)}$, which remains well-defined. Therefore

$$\mathbb{R}[[f_1, \dots, f_k]]\varphi \boxtimes \mathbb{R}[[g_1, \dots, g_\ell]]\psi \cong \bigoplus_{i=0}^{r} \mathbb{R}[[f_1, \dots, f_k, g_1, \dots, g_\ell]](\varphi, \psi)^{(i)}.$$

Now we make the induction hypothesis that all cases with $p = 0$ are computable up through the case $q - 1$, and we discuss case q. Choose one of the q elements of $g_1, \dots, g_\ell$ having positive weight; we assume the chosen element is g_1. Then we may expand

$$\mathbb{R}[[g_1, \dots, g_\ell]]\psi = \left(\bigoplus_{\nu=0}^{t-1} \mathbb{R}[[g_2, \dots, g_\ell]] g_1^\nu \psi \right) \oplus \mathbb{R}[[g_1, \dots, g_\ell]] g_1^t \psi,$$

where t is the smallest integer such that $w_{g_1^t\psi} > w_\varphi$. This decomposition corresponds to classifying monomials according to the power of g_1 that occurs, with all powers greater than or equal to t assigned to the last term. (Note that g_1 is missing from the square brackets except in the last term.) Now take the box product of $\mathbb{R}[[f_1, \dots, f_k]]\varphi$ times this expression, and distribute the product according to Lemma 8. All of the terms except the last are computable by the induction hypothesis. We claim the last term is computable by the formula

$$\mathbb{R}[[f_1, \dots, f_k]]\varphi \boxtimes \mathbb{R}[[g_1, \dots, g_\ell]] g_1^t \psi \cong \bigoplus_{i=0}^{w_\varphi} \mathbb{R}[[f_1, \dots, f_k, g_1, \dots, g_\ell]](\varphi, g_1^t\psi)^{(i)}.$$

This is because w_φ is an absolute limit to the order of transvectants in this box product that will be well-defined, and any such transvectant $(f_1^{m_1} \cdots f_k^{m_k}\varphi, \gamma_1^{n_1} \cdots g_\ell^{n_\ell} g_1^t \psi)^{(i)}$ can be replaced by
$f_1^{m_1} \cdots f_k^{m_k}\gamma_1^{n_1} \cdots g_\ell^{n_\ell}(g_1^t\varphi, \psi)^{(i)}$.

Now we make the induction hypothesis that cases $(p-1,q)$, $(p,q-1)$, and $(p-1,q-1)$ can be handled, and we treat the case (p,q). Choose one of the p functions in $f_1,\dots,f_k$ having positive weight; we assume the chosen element is f_1. Similarly, choose a function of positive weight from $g_1,\dots,g_\ell$, and suppose it is g_1. Let s and t be the smallest integers such that

$$s\cdot w_{f_1}=t\cdot w_{g_1}.$$

Expand

$$\mathbb{R}[[f_1,\dots,f_k]]\varphi=\left(\bigoplus_{\mu=0}^{s-1}\mathbb{R}[[f_2,\dots,f_k]]f_1^{\mu}\varphi\right)\oplus\mathbb{R}[[f_1,\dots,f_k]]f_1^{s}\varphi$$

and

$$\mathbb{R}[[g_1,\dots,g_\ell]]\psi=\left(\bigoplus_{\nu=0}^{t-1}\mathbb{R}[[g_2,\dots,g_\ell]]g_1^{\nu}\psi\right)\oplus\mathbb{R}[[g_1,\dots,g_\ell]]g_1^{t}\psi.$$

Now take the box product of these last two expansions and distribute the product. There are four kinds of terms. Terms that are missing both f_1 and g_1 in square brackets are of type $(p-1,q-1)$. Terms that are missing f_1 in square brackets, but not g_1, are of type $(p-1,q)$, and there are likewise terms of type $(p,q-1)$. All of these can be handled by the induction hypothesis. Finally, there is the term

$$\mathbb{R}[[f_1,\dots,f_k]]f_1^{s}\varphi\boxtimes\mathbb{R}[[g_1,\dots,g_\ell]]g_1^{t}\psi.$$

There is no upper limit to the transvectant order that can occur here, since in general there remain terms of positive weight in the square brackets. However, setting $r=s\cdot w_{f_1}=t\cdot w_{g_1}$, we will show that this box product is $\cong$ to

$$\left(\bigoplus_{i=0}^{r-1}\mathbb{R}[[f_1,\dots,f_k,g_1,\dots,g_k]](f_1^s,g_1^t)^{(i)}\right)$$
$$\oplus\,(\mathbb{R}[[f_1,\dots,f_k]]\varphi\boxtimes\mathbb{R}[[g_1,\dots,g_\ell]]\psi)(f_1^s,g_1^t)^{(r)}.$$

The terms for transvectant orders $\le r-1$ in this expression are obtained in the usual way, by replacing $(f_1^{m_1}\cdots f_k^{m_k}f_1^s\varphi,\gamma_1^{n_1}\cdots g_\ell^{n_\ell}g_1^t\psi)^{(i)}$ by $f_1^{m_1}\cdots f_k^{m_k}\varphi\gamma_1^{n_1}\cdots g_\ell^{n_\ell}\psi(f_1^s,g_1^t)^{(i)}$. The final term is quite different from any others considered until now, since it involves a box product of subspaces as the coefficient of $(f_1^s,g_1^t)^{(r)}$. This term is obtained from Corollary 4 using the fact that $w_{(f_1^s,g_1^t)^{(r)}}=0$: for any $i\ge 0$, we replace $(f_1^{m_1}\cdots f_k^{m_k}\varphi,\gamma_1^{n_1}\cdots g_\ell^{n_\ell}\psi)^{(i+r)}$ by $(f_1^{m_1}\cdots f_k^{m_k}f_1^s\varphi,\gamma_1^{n_1}\cdots g_\ell^{n_\ell}g_1^t\psi)^{(i)}(f_1^s,g_1^t)^{(r)}$.

At this point we have reduced the calculation of $\mathbb{R}[[f_1,\ldots,f_k]]\varphi \boxtimes \mathbb{R}[[g_1,\ldots,g_\ell]]\psi$ in case (p,q) to a number of terms computable by the induction hypothesis or by explicit formula, plus one special term that seems to lead in circles since it involves the very same box product that we are trying to calculate. Thus our result has the form

$$\mathcal{R} = \mathcal{S} \oplus \mathcal{R}\theta,$$

where θ has weight zero. But this implies $\mathcal{R} = \mathcal{S} \oplus (\mathcal{S} \oplus \mathcal{R}\theta)\theta = \mathcal{S} \oplus \mathcal{S}\theta \oplus \mathcal{R}\theta^2$. Continuing in this way we have $\mathcal{R} = \mathcal{S} \oplus \mathcal{S}\theta \oplus \mathcal{S}\theta^2 \oplus \mathcal{S}\theta^3 \oplus \cdots$, which reduces to $\mathcal{R} = \mathcal{S}[[\theta]]$; that is, we add the weight-zero element θ to the ring $\mathcal{S}$, allowing all powers. (Remember that these are formal power series rings, so we allow formal sums to infinity.) This simply means that we erase the "unusual" term $\mathbb{R}[[f_1,\ldots,f_k]]\varphi \boxtimes \mathbb{R}[[g_1,\ldots,g_\ell]]\psi)(f_1^s, g_1^t)^{(r)}$ from our computation, and instead insert $(f_1^s, g_1^t)^{(r)}$ into the square brackets in all the coefficient rings that have already been computed. This does not affect the induction, because the new elements added have weight zero, and the induction is on the numbers p and q of elements of positive weight. □

5. A new example

Since ker $\mathcal{X}_{2,2} = \mathbb{R}[[\alpha,\beta,\gamma]]$ and ker $\mathcal{X}_3 = \mathbb{R}[[\delta,\varepsilon]]$, where the weights of $\alpha,\beta,\gamma,\delta,\varepsilon$ are $1,1,0,2,0$, we have

$$\ker \mathcal{X}_{2,2,3} = \mathbb{R}[[\alpha,\beta,\gamma]] \boxtimes \mathbb{R}[[\delta,\varepsilon]].$$

The following transvectants will appear in the course of the calculation:

$$\begin{aligned}
\zeta &= (\alpha,\delta)^{(1)}\\
\eta &= (\beta,\delta)^{(1)}\\
\theta &= (\alpha^2,\delta)^{(2)}\\
\lambda &= (\alpha\beta,\delta)^{(2)}\\
\mu &= (\beta^2,\delta)^{(2)}.
\end{aligned}$$

Suppressing γ and ε and noticing that these will appear in every square bracket of the box product, we compute

$$\mathbb{R}[[\alpha,\beta]] \boxtimes \mathbb{R}[[\delta]] = (\mathbb{R}[[\beta]] \oplus \mathbb{R}[[\beta]]\alpha \oplus \mathbb{R}[[\alpha,\beta]]a^2) \boxtimes (\mathbb{R} \oplus \ R[[\delta]]\delta).$$

Distributing the box product gives three kinds of terms:

1. Three terms that are immediately computed in final form: $\mathbb{R}[[\beta]] \oplus \mathbb{R}[[\beta]]\alpha \oplus \mathbb{R}[[\alpha,\beta]]\alpha^2$.

2. Two box products that must be computed by further expansion (according to the induction scheme in Theorem 9): $\mathbb{R}[[\beta]]\boxtimes\mathbb{R}[[\delta]]\delta$ and $\mathbb{R}[[\beta]]\alpha\boxtimes\mathbb{R}[[\delta]]\delta$. When these are worked out, they will recycle to themselves but not to the original box product $\mathbb{R}[[\alpha,\beta]]\boxtimes\mathbb{R}[[\delta]]$.
3. One box product, $\mathbb{R}[[\alpha,\beta]]\alpha^2\boxtimes\mathbb{R}[[\delta]]\delta$, that will recycle to $\mathbb{R}[[\alpha,\beta]]\boxtimes\mathbb{R}[[\delta]]$. In fact

$$\mathbb{R}[[\alpha,\beta]]\alpha^2\boxtimes\mathbb{R}[[\delta]]\delta=\mathbb{R}[[\alpha,\beta,\delta]]\alpha^2\delta\oplus\mathbb{R}[[\alpha,\beta,\delta]](\alpha^2,\delta)^{(1)}$$
$$\oplus\,(\mathbb{R}[[\alpha,\beta]]\boxtimes\mathbb{R}[[\delta]])(\alpha^2,\delta)^{(2)}.$$

According to the recycling rule, the last term here will be deleted and $\theta=(\alpha^2,\delta)^{(2)}$, which has weight zero, will be added to all square brackets (along with the other suppressed weight-zero invariants γ and ε).

Now we turn to the calculations in item 2 of the list. The first of these is

$$\mathbb{R}[[\beta]]\boxtimes\mathbb{R}[[\delta]]\delta=(\mathbb{R}\oplus\mathbb{R}\beta\oplus\mathbb{R}[[\beta]]\beta^2)\boxtimes(\mathbb{R}\delta\oplus\mathbb{R}[[\delta]]\delta^2).$$

Notice that δ is playing two roles: the factor $\mathbb{R}[[\delta]]\delta$ would appear in the notation of Theorem 9 as $\mathbb{R}[[g_1]]\psi$. (In working through the details that are not written out here, it was helpful to temporarily set $\psi=\delta$ and write the product as $\mathbb{R}[[\beta]]\boxtimes\mathbb{R}[[\delta]]\psi=(\mathbb{R}\oplus\mathbb{R}\beta\oplus\mathbb{R}[[\beta]]\beta^2)\boxtimes(\mathbb{R}\psi\oplus\mathbb{R}[[\delta]]\delta\psi)$). All of the six terms (after distributing the box product) can be computed explicitly except the last, which recycles to $\mathbb{R}[[\beta]]\boxtimes\mathbb{R}[[\delta]]$. Handling this last term first, we have

$$\mathbb{R}[[\beta]]\beta^2\boxtimes\mathbb{R}[[\delta]]\delta^2=\mathbb{R}[[\beta,\delta]]\beta^2\delta^2\oplus\mathbb{R}[[\beta,\delta]]\beta\delta(\beta,\delta)^{(1)}$$
$$\oplus\,(\mathbb{R}[[\beta]]\boxtimes\mathbb{R}[[\delta]]\delta)(\beta^2,\delta)^{(2)}.$$

The last term will be deleted and $\mu=(\beta^2,\delta)^{(2)}$ will be inserted in all the square brackets resulting from this calculation (but *not* all the brackets in the main calculation). For this reason we do not suppress μ, but state it explicitly. The final result of this calculation, after recombining terms whenever possible, is

$$\mathbb{R}[[\beta]]\boxtimes\mathbb{R}[[\delta]]\delta=\mathbb{R}[[\beta,\delta,\mu]]\delta\oplus\mathbb{R}[[\beta,\delta,\mu]]\eta\oplus\mathbb{R}[[\beta,\mu]]\mu.$$

Notice that μ appears both inside and outside of brackets.

The second calculation in item 2 of the list is

$$\mathbb{R}[[\beta]]\alpha\boxtimes\mathbb{R}[[\delta]]\delta=(\mathbb{R}\alpha\oplus\mathbb{R}\alpha\beta\oplus\mathbb{R}[[\beta]]\alpha\beta^2)\boxtimes(\mathbb{R}\delta\oplus\mathbb{R}[[\delta]]\delta^2).$$

All of the distributed terms are computable immediately except the last, which contains a term that recycles to $\mathbb{R}[[\beta]]\alpha\boxtimes\mathbb{R}[[\delta]]d$ and once again brings about the introduction of μ into every bracket (of this subcalculation). A new feature that arises is the need to make choices when faced with a transvectant such as $(\alpha\beta,\delta)^{(1)}$, which could be replaced by either $\alpha(\beta,\delta)^{(1)} = \alpha\eta$ or $\beta(\alpha,\delta)^{(1)} = \beta\zeta$. If we always favor η over ζ in such cases, the final result for this calculation is

$$\begin{aligned}\mathbb{R}[[\beta]]\alpha \boxtimes \mathbb{R}[[\delta]]\delta = \mathbb{R}[[\beta,\delta,\mu]]\alpha\delta \oplus \mathbb{R}[[\delta,\mu]]\zeta \oplus \mathbb{R}[[\beta,\delta,\mu]]\alpha\eta \\ \oplus \mathbb{R}[[\delta,\mu]]\lambda \oplus \mathbb{R}[[\beta,\mu]]\alpha\mu.\end{aligned}$$

To state the final result, let $\mathcal{R} = \mathbb{R}[[\gamma,\varepsilon,\theta]]$. Then

$$\begin{aligned}\ker \mathcal{X}_{2,2,3} = \mathcal{R}[[\alpha,\beta]] & \\ \oplus\, & \mathcal{R}[[\beta,\delta,\mu]]\delta \oplus \mathcal{R}[[\beta,\delta,\mu]]\eta \oplus \mathcal{R}[[\beta,\mu]]\mu \\ \oplus\, & \mathcal{R}[[\beta,\delta,\mu]]\alpha\delta \oplus \mathcal{R}[[\delta,\mu]]\zeta \oplus \mathcal{R}[[\beta,\delta,\mu]]\alpha\eta \\ \oplus\, & \mathcal{R}[[\delta,\mu]]\lambda \oplus \mathcal{R}[[\beta,\mu]]\alpha\mu \\ \oplus\, & \mathcal{R}[[\alpha,\beta,\delta]]\alpha^2\delta \oplus \mathcal{R}[[\alpha,\beta,\delta]]\alpha\zeta.\end{aligned}$$

6. Boosting to equivariants

In this section we describe the procedure for obtaining a Stanley decomposition of the module of equivariants (or normal form space) $\ker \mathsf{X}$ from a Stanley decomposition of the ring of invariants $\ker \mathcal{X}$; here $\mathsf{X} = \mathsf{L}_{N^*}$, just as $\mathcal{X} = \mathcal{D}_{N^*}$. As pointed out in the introduction, this procedure was already completely described in section 5 of [4]. But transvectants were not introduced in that paper until section 6, and the connection between transvectants and section 5 was only briefly explained in section 8. Now that we have recognized the central role of transvectants in this theory, it seems appropriate to restate the "boosting" process from the beginning in the language of transvectants, and to provide examples of calculations in this language.

The starting point is that the module of all formal power series vector fields on $\mathbb{R}^n$ can be viewed as the tensor product $\mathbb{R}[[x_1,\dots,x_n]] \otimes \mathbb{R}^n$, and in fact the tensor product can be identified with the ordinary product (of a scalar field times a constant vector) since (just as in the case of a tensor product of two polynomial spaces with nonoverlapping variables, used in previous sections) the ordinary product satisfies the same algebraic rules as a tensor product. Specifically, every formal power series vector field can be written as

$$f_1(x)e_1 + \cdots + f_n(x)e_n = \begin{bmatrix} f_1(x) \\ \vdots \\ f_n(x)\end{bmatrix},$$

where the e_i are the standard basis vectors of $\mathbb{R}^n$.

Next, the Lie derivative $\mathsf{X} = \mathsf{L}_{N^*}$ can be expressed as the tensor product of $\mathcal{X}$ and $-N^*$, that is, $\mathsf{X} = \mathcal{X} \otimes I + I \otimes (-N^*)$. Under the identification of $\otimes$ with ordinary product, this means $\mathsf{X}(fv) = (\mathcal{X}f)v + f(-N^*v)$, where $f \in \mathbb{R}[[x_1, \ldots, x_n]]$ and $v \in \mathbb{R}^n$ in agreement with the following calculation (in which $v' = 0$ because v is constant):

$$\begin{aligned}\mathsf{X}(fv) &= \mathsf{L}_{N^*}(fv) \\ &= (\mathcal{D}_{N^*}f)v + f(\mathsf{L}_{N^*}v) \\ &= (\mathcal{D}_{N^*}f)v + f(v'N^*x - N^*v) \\ &= (\mathcal{D}_{N^*}f)v + f(-N^*v).\end{aligned}$$

This kind of calculation also shows that the sl(2) representation (on vector fields) with triad $\{\mathsf{X}, \mathsf{Y}, \mathsf{Z}\}$ is the tensor product of the representation (on scalar fields) with triad $\{\mathcal{X}, \mathcal{Y}, \mathcal{Z}\}$ and the representation (on $\mathbb{R}^n$) with triad $\{-N^*, -M^*, -H\}$.

It follows (as in Theorem 2 above) that a basis for the normal form space ker X is given by the well-defined transvectants $(f, v)^{(i)}$ as f ranges over a basis for ker $\mathcal{X} \subset \mathbb{R}[[x_1, \ldots, x_n]]$ and v ranges over a basis for ker $N^* \subset \mathbb{R}^n$. (Of course ker $(-N^*)$ = ker N^*.) The first of these bases is given by the standard monomials of a Stanley decomposition for ker $\mathcal{X}$. The second is given by the standard basis vectors $e_r \in \mathbb{R}^n$ such that r is the index of the bottom row of a Jordan block in N^* (or equivalently, in N). It is useful to note that the weight of such an e_r is one less than the size of the block. The definition (2) of transvectant in this case becomes

$$\begin{aligned}(f, e_r)^{(i)} &= \sum_{j=0}^{i} (-1)^j W^{i,j}_{f,e_r} (\mathcal{Y}^j f)((-M^*)^{i-j} e_4) \\ &= (f, g)^{(i)} = (-1)^i \sum_{j=0}^{i} W^{i,j}_{f,g} (\mathcal{Y}^j f)((M^*)^{i-j} g).\end{aligned}$$

From here on, the computational procedures are the same as those used in previous sections, except that infinite iterations never arise. We illustrate this first by computing a Stanley decomposition for the normal form of vector fields with linear part N_4. This example was treated in section 5 of [4], and thus provides a comparison of the previous method and the new one.

We begin with ker $\mathcal{X}_4 = \mathbb{R}[[\alpha, \beta, \delta]] \oplus \mathbb{R}[[\alpha, \beta, \delta]]\gamma$. Since δ has weight zero, it is convenient to remove it from the calculation by setting $\mathcal{R} =$

$\mathbb{R}[[\delta]]$ and writing

$$\begin{aligned}(6)\qquad \ker \mathcal{X}_4 &= \mathcal{R}[[\alpha,\beta]] \oplus \mathcal{R}[[\alpha,\beta]]\gamma \\ &= \mathcal{R}[[\beta]] \oplus \mathcal{R}[[\alpha,\beta]]\alpha \oplus \mathcal{R}[[\alpha,\beta]]\gamma \\ &= \mathcal{R} \oplus \mathcal{R}\beta \oplus \mathcal{R}[[\beta]]\beta^2 \oplus \mathcal{R}[[\alpha,\beta]]\alpha \oplus \mathcal{R}[[\alpha,\beta]]\gamma.\end{aligned}$$

The only basis element ker N_4^* is e_4, the bottom row of the single Jordan block, and its weight is 3. So ker $N_4^* = \mathbb{R}e_4$. The expansions in (6) are terminated with β^2 and α because these (having weights 4 and 3 respectively) are sufficient to match the weight 3 of e_4, which puts a maximum on the transvectants that can be well-defined. Now we only need to compute the box product

$$\ker \mathsf{X}_4 = (\ker \mathcal{X}_4) \boxtimes (\mathbb{R}e_4).$$

Distributing the box product over the decomposition (6) gives the following cases.

1. If $f \in \mathcal{R}$ the only transvectant that can be formed is $(f, e_4)^{(0)} = fe_4$. So
$$\mathcal{R} \boxtimes \mathbb{R}e_4 \cong \mathcal{R}e_4.$$
2. If $f \in \mathcal{R}\beta$ then $f = g\beta$ with $g \in \mathcal{R}$ (having weight zero). Then $(f, e_4)^{(i)}$ can be formed for $i = 0, 1, 2$, and can be replaced by $g(\beta, e_4)^{(i)}$. Therefore
$$\mathcal{R}\beta \boxtimes \mathbb{R}e_4 \cong \mathcal{R}\beta e_4 \oplus \mathcal{R}(\beta, e_4)^{(1)} \oplus \mathcal{R}(\beta, e_4)^{(2)}.$$
3. If $f \in \mathcal{R}[[\beta]]\beta^2$ then $f = g\beta^2$ with $g \in \mathcal{R}[[\beta]]$ (with unlimited weight). We can form $(f, e_4)^{(i)}$ for $i = 0, 1, 2, 3$, the limit coming from e_4. These can be replaced by $g\beta(\beta, e_4)^{(i)}$ if $i = 0, 1, 2$ and by $g(\beta^2, e_4)^{(3)}$ if $i = 3$. Therefore
$$\begin{aligned}\mathcal{R}[[\beta]]\beta^2 \boxtimes \mathbb{R}e_4 \cong{}& \mathcal{R}[[\beta]]\beta^2 e_4 \oplus \mathcal{R}[[\beta]]\beta(\beta, e_4)^{(1)} \\ &\oplus \mathcal{R}[[\beta]]\beta(\beta, e_4)^{(2)} \oplus \mathcal{R}[[\beta]](\beta^2, e_4)^{(3)}.\end{aligned}$$
4. If $f \in \mathcal{R}[[\alpha,\beta]]\alpha$ then $f = g\alpha$ with $g \in \mathcal{R}[[\alpha,\beta]]$. Transvectants $(f, e_4)^{(i)}$ can be formed with $i = 0, 1, 2, 3$ and can be replaced by $g(\alpha, e_4)^{(i)}$. Therefore
$$\begin{aligned}\mathcal{R}[[\alpha,\beta]]\alpha \boxtimes \mathbb{R}e_4 \cong{}& \mathcal{R}[[\alpha,\beta]]\alpha e_4 \oplus \mathcal{R}[[\alpha,\beta]](\alpha, e_4)^{(1)} \\ &\oplus \mathcal{R}[[\alpha,\beta]](\alpha, e_4)^{(2)} \oplus \mathcal{R}[[\alpha,\beta]](\alpha, e_4)^{(3)}.\end{aligned}$$
5. Since the weights of α and γ are equal, the last calculation is almost the same:
$$\begin{aligned}\mathcal{R}[[\alpha,\beta]]\gamma \boxtimes \mathbb{R}e_4 \cong{}& \mathcal{R}[[\alpha,\beta]]\gamma e_4 \oplus \mathcal{R}[[\alpha,\beta]](\gamma, e_4)^{(1)} \\ &\oplus \mathcal{R}[[\alpha,\beta]](\gamma, e_4)^{(2)} \oplus \mathcal{R}[[\alpha,\beta]](\gamma, e_4)^{(3)}.\end{aligned}$$

Before adding the terms from these different cases, we observe the following collapses that take place:

$$\begin{aligned}
&\mathcal{R}e_4 \oplus \mathcal{R}\beta e_4 \oplus \mathcal{R}[[\beta]]\beta^2 e_4 \oplus \mathcal{R}[[\alpha,\beta]]\alpha e_4 = \mathcal{R}[[\alpha,\beta]]e_4\\
&\mathcal{R}(\beta,e_4)^{(1)} \oplus \mathcal{R}[[\beta]]\beta(\beta,e_4)^{(1)} = \mathcal{R}[[\beta]](\beta,e_4)^{(1)}\\
&\mathcal{R}(\beta,\varepsilon_4)^{(2)} \oplus \mathcal{R}[[\beta]]\beta(\beta,e_4)^{(2)} = \mathcal{R}[[\beta]](\beta,e_4)^{(2)}.
\end{aligned}$$

Therefore we finally have

(7)

$$\begin{aligned}
\ker \mathsf{X}_4 = {} & \mathcal{R}[[\alpha,\beta]]e_4 \oplus \mathcal{R}[[\beta]](\beta,e_4)^{(1)} \oplus \mathcal{R}[[\beta]](\beta,e_4)^{(2)} \oplus \mathcal{R}[[\beta]](\beta^2,e_4)^{(3)}\\
& \oplus \mathcal{R}[[\alpha,\beta]](\alpha,e_4)^{(1)} \oplus \mathcal{R}[[\alpha,\beta]](\alpha,e_4)^{(2)} \oplus \mathcal{R}[[\alpha,\beta]](\alpha,e_4)^{(3)}\\
& \oplus \mathcal{R}[[\alpha,\beta]]\gamma e_4 \oplus \mathcal{R}[[\alpha,\beta]](\gamma,e_4)^{(1)} \oplus \mathcal{R}[[\alpha,\beta]](\gamma,e_4)^{(2)} \oplus \mathcal{R}[[\alpha,\beta]](\gamma,e_4)^{(3)}.
\end{aligned}$$

To complete the calculation, it is necessary to compute the transvectants that appear in the Stanley decomposition. These are all of the form $(f,e_4)^{(i)}$ for $i = 0,1,2,3$. These can be computed once and for all, and then the individual f that are needed (namely α, β, and β^2, expressed in terms of $x_1,\dots,x_4$) can be substituted in. Of course

$$(f,e_4)^{(0)} = fe_4 = \begin{bmatrix}0\\0\\0\\f\end{bmatrix}.$$

For $i = 1$ we have, from the definition of transvectant,

$$\begin{aligned}
(f,e_4)^{(1)} &= w_f f(-M^*)e_4 - w_{e_4}(\mathcal{Y}f)e_4\\
&= w_f f\begin{bmatrix}0\\0\\-3\\0\end{bmatrix} - \begin{bmatrix}0\\0\\0\\3\mathcal{Y}f\end{bmatrix}\\
&= -3\begin{bmatrix}0\\0\\w_f f\\\mathcal{Y}f\end{bmatrix}\\
&= -3\begin{bmatrix}0\\0\\\mathcal{X}g\\g\end{bmatrix} \qquad \text{with} \qquad g = \mathcal{Y}f.
\end{aligned}$$

Similar calculations show that

$$(f, e_4)^{(2)} = 6 \begin{bmatrix} 0 \\ \mathcal{X}^2 g \\ \mathcal{X} g \\ g \end{bmatrix} \qquad \text{with} \qquad g = \mathcal{Y}^2 f$$

$$(f, e_4)^{(3)} = -6 \begin{bmatrix} \mathcal{X}^3 g \\ \mathcal{X}^2 g \\ \mathcal{X} g \\ g \end{bmatrix} \qquad \text{with} \qquad g = \mathcal{Y}^3 f.$$

The nonzero constant factors 3, 6, and -6 may be ignored, because we are only concerned with computing basis elements. The forms using g (in place of f) are convenient because they avoid confusing constant factors (including w_f). But in practice, one does not want to apply $\mathcal{Y}$ several times to f to find g, and then undo this by applying $\mathcal{X}$ several times to g. Instead, the rule that $\mathcal{X}(\mathcal{Y}^i f) = i(w_f + 1 - i)\mathcal{Y}^{i-1} f$ allows the constant factors to be restored.

For comparison with [4] and section 4.7 of [5], it is helpful to notice that $\mathcal{Y}^i f \in \ker \mathcal{X}^{i+1}$. The procedure in [4] and [5] calls for putting elements such as g from a Stanley decomposition of ker $\mathcal{X}^3$, filtered by ker $\mathcal{X} \subset$ ker $\mathcal{X}^2 \subset$ ker $\mathcal{X}^3 \subset$ ker $\mathcal{X}^4$, into the into the bottom position of a vector to form

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ g \end{bmatrix}.$$

These vectors form a basis for the simplified normal form, and the inner product normal form is then "reconstructed" by passing to

$$\begin{bmatrix} \mathcal{X}^3 g \\ \mathcal{X}^2 g \\ \mathcal{X} g \\ g \end{bmatrix}.$$

Some of the top entries may be zero, depending on the position of g in the filtration. This is clearly the same result, up to a constant factor, as we have obtained here. Also note that our Stanley decomposition has 11 terms, just as the one in equation (40) of [4] or (4.7.30) of [5]. One small difference is that the earlier method allows expressions in g such as $\mathcal{Y}\beta\mathcal{Y}^2\beta$, which do not exactly fit the form coming from transvectants; the current method would use $\mathcal{Y}^3\beta^2$.

Finally, we turn to the example of $N_{2,2,3}$. As mentioned in the introduction, the normal form space for this example is quite large, and we only compute a few terms. The basis elements for ker $N^*_{2,2,3}$ are e_2, e_4, and e_7. Therefore we need to compute the box product of the invariant ring ker $\mathcal{X}_{2,2,3}$, computed in section 5, with $\mathbb{R}e_2 \oplus \mathbb{R}e_4 \oplus \mathbb{R}e_7$. We consider the part coming from the particular term $\mathcal{R}[[\delta, \mu]]\eta$ in ker $\mathcal{X}_{2,2,3}$. We have

$$\begin{aligned}
&\mathcal{R}[[\delta,\mu]]\eta \boxtimes (\mathbb{R}e_2 \oplus \mathbb{R}e_4 \oplus \mathbb{R}e_7) = \\
&\quad (\mathcal{R}[[\delta,\mu]]\eta \boxtimes \mathbb{R}e_2) \oplus (\mathcal{R}[[\delta,\mu]]\eta \boxtimes \mathbb{R}e_4) \oplus (\mathcal{R}[[\delta,\mu]]\eta \boxtimes \mathbb{R}e_7) = \\
&\quad (\mathcal{R}[[\delta,\mu]]\eta \boxtimes \mathbb{R}e_2) \oplus (\mathcal{R}[[\delta,\mu]]\eta \boxtimes \mathbb{R}e_4) \oplus (\mathcal{R}[[\mu]]\eta \boxtimes \mathbb{R}e_7) \oplus (\mathcal{R}[[\delta,\mu]]\delta\eta \boxtimes \mathbb{R}e_7) = \\
&\quad \mathcal{R}[[\delta,\mu]](\eta,e_2)^{(0)} \oplus \mathcal{R}[[\delta,\mu]](\eta,e_2)^{(1)} \oplus \\
&\quad \mathcal{R}[[\delta,\mu]](\eta,e_4)^{(0)} \oplus \mathcal{R}[[\delta,\mu]](\eta,e_4)^{(1)} \oplus \\
&\quad \mathcal{R}[[\mu]](\eta,e_7)^{(0)} \oplus \mathcal{R}[[\mu]](\eta,e_7)^{(1)} \oplus \\
&\quad \mathcal{R}[[\delta,\mu]](\delta\eta,e_7)^{(0)} \oplus \mathcal{R}[[\delta,\mu]](\delta\eta,e_7)^{(1)} \oplus \mathcal{R}[[\delta,\mu]](\delta\eta,e_7)^{(2)}.
\end{aligned}$$

Notice that $\mathcal{R}[[\delta, \mu]]\eta$ does not need to be expanded when the box product with $\mathbb{R}e_2$ or $\mathbb{R}e_4$ is taken, because e_2 and e_4 have weight 1, which equals the weight of η. But it does need to be expanded (in δ) when the box product with $\mathbb{R}e_7$ is taken, because e_7 has weight 2 and can support transvectants up to order 2. Bringing out one δ forms $\delta\eta$, which has weight 3 and can support the required transvectants. In the last line we can now replace $(\delta\eta, e_7)^{(i)}$ with $\delta(\eta, e_7)^{(i)}$ for $i = 0$ and 1, but not for $i = 2$. After these replacements, the terms for $i = 0$ and 1 can be recombined with terms in the previous line:

$$\mathcal{R}[[\mu]](\eta,e_7)^{(i)} \oplus \mathcal{R}[[\delta,\mu]]\delta(\eta,e_7)^{(i)} = \mathcal{R}[[\delta,\mu]](\eta,e_7)^{(i)}.$$

References

[1] Bram Broer. On the generating functions associated to a system of binary forms. *Indag. Math. (N.S.)*, 1(1):15–25, 1990.

[2] R. Cushman and J. A. Sanders. A survey of invariant theory applied to normal forms of vectorfields with nilpotent linear part. In Dennis Stanton, editor, *Invariant Theory and Tableaux*, pages 82–106. Springer, New York, 1990.

[3] David Malonza. Normal forms for coupled Takens-Bogdanov systems. *Journal of Nonlinear Mathematical Physics*, 11:376–398, 2004.

[4] James Murdock. On the structure of nilpotent normal form modules. *Journal of Differential Equations*, 180:198–237, 2002.

[5] James Murdock. *Normal Forms and Unfoldings for Local Dynamical Systems.* Springer, New York, 2003.

[6] Jan Sanders, Ferdinand Verhulst, and James Murdock. *Averaging Methods in Nonlinear Dynamical Systems.* Springer, New York, 2007.

[7] B. Sturmfels and N. White. Computing combinatorial decompositions of rings. *Combinatorica*, 11:275–293, 1991.