Centre de Recherches Mathématiques
CRM Proceedings and Lecture Notes

Aspects of generalized double-bracket flows

Anthony M. Bloch and Arieh Iserles

Abstract. In this paper we analyse generalizations of double-bracket flows to more complex isospectral equations of the form $\boldsymbol{Y}' = [\boldsymbol{Y}, [\boldsymbol{G}(\boldsymbol{Y}-\boldsymbol{N}), \boldsymbol{Y}]]$, $t \ge 0$, over real symmetric matrices. The new flows possess a nontrivial equilibrium structure which we explore. We discuss the stability of these equilibria. We also discuss the relationship of the new flows to the structure of (convex) Horn–Schur polytopes.

1. Introduction

A *double-bracket flow (DBF)* is defined as a matrix differential equation of the form

$$\boldsymbol{Y}' = [[\boldsymbol{N}, \boldsymbol{Y}], \boldsymbol{Y}], \quad t \ge 0, \qquad \boldsymbol{Y}(0) = \boldsymbol{Y}_0 \in \mathrm{Sym}(n), \tag{1.1}$$

where $\boldsymbol{N} \in \mathrm{Sym}(n)$, $\mathrm{Sym}(n)$ is the symmetric space of $n \times n$ real symmetric matrices, while $[\,\cdot\,,\,\cdot\,]$ is the standard matrix commutator. Without loss of generality we may assume that $\boldsymbol{N}$ is a diagonal matrix. Such flows were introduced by Brockett [**Bro91**] and by Chu and Driessel [**CD91**] in connection with control theory and with inverse eigenvalue problems in numerical linear algebra, respectively, and their Lie-algebraic generalization was analysed in [**BBR90**] and [**BBR92**]. Related equations are discussed in [**Bro93**] and [**HM94**].

Double-bracket flows possess a number of interesting features. They are a special case of an *isospectral flow*

$$\boldsymbol{Y}' = [\boldsymbol{B}(t, \boldsymbol{Y}), \boldsymbol{Y}], \quad t \ge 0, \qquad \boldsymbol{Y}(0) = \boldsymbol{Y}_0 \in \mathrm{Sym}(n), \tag{1.2}$$

where $\boldsymbol{B} : \mathbb{R}_+ \times \mathrm{Sym}(n) \to \mathfrak{so}(n)$, the set $\mathfrak{so}(n)$ being the Lie algebra of $n \times n$ real, skew-symmetric matrices. Since $\boldsymbol{Y}(t) = \boldsymbol{Q}(t) Y_0 \boldsymbol{Q}^\top(t)$, $t \ge 0$, where $\boldsymbol{Q}' = \boldsymbol{B}(t, \boldsymbol{Q}\boldsymbol{Y}_0\boldsymbol{Q}^\top)$, $t \ge 0$, $\boldsymbol{Q}(0) = \boldsymbol{I}$, it follows easily that $\boldsymbol{Y}(t)$ for all $t \ge 0$ is similar to $\boldsymbol{Y}_0$. In other words, the eigenvalues of $\boldsymbol{Y}_0$ are integrals of the (1.2). (Hence the name "isospectral flows".) This feature makes isospectral flows of interest in many areas of mathematics including integrable systems and is also crucial in their numerical solution [**CIZ97, Ise02**]. This isospectral property of the double-bracket

---

1991 *Mathematics Subject Classification.* Primary 34C40; Secondary51M20.
Key words and phrases. Double-Bracket Flows, Polytopes.
Supported in part by the National Science Foundation.

flow is highly relevant in the special case when the system reduces to the Toda lattice flow (see [**BBR92**] and references therein).

Another important feature of the DBF (1.1) is that it is a *gradient flow* with respect to the so-called *normal metric* on an appropriate adjoint orbit of a compact Lie group and possesses a global Lyapunov function [**BBR92, Bro91**]. Note that this feature is in general false for general isospectral flows (1.2). Construction of alternative Lyapunov functions for (1.1) was addressed in [**BI03b**].

A key feature of DBF, originally observed by Brockett [**Bro91**] and analysed extensively in [**BI03b**], is its optimality. Roughly speaking, $\hat{\boldsymbol{Y}} = \lim_{t\to\infty} \boldsymbol{Y}(t)$ (which exists because (1.1) is a gradient flow!) minimizes the distance from $\boldsymbol{N}$ along the isospectral orbit with regard to various unitarily-invariant norms, in particular Schatten p-norms for $p \geq 1$.

The theme of the present paper is a generalization of the double-bracket flow to instances when the flow is the gradient of a higher-power trace. This leads a richer equilibrium and stability structure which we proceed to examine.

We also discuss the role of the Schur–Horn polytope in the dynamics of these systems, generalizing the work in [**BBR92**].

The result of Schur is as follows: Let A be a Hermitian matrix with eigenvalues λ_j, $j = 1, \dots, n$. Denote by $\lambda = (\lambda_1, \dots, \lambda_n)$ the vector of eigenvalues of A and by $A^0 = (A_{11}, \dots, A_{nn})$ the diagonal of A. Then $A^0 \in \widehat{(\Sigma_n \lambda)}$. Here the "wide hat" notation denotes the convex hull under the action of the permutation group.

Conversely the result of Horn is: Let $\lambda \in R^n$, with components arranged in non-increasing order. If $A^0 \in \widehat{(\Sigma_n \lambda)}$, there is a Hermitian matrix A with eigenvalues λ whose diagonal is A^0.

As we show in the sequel, the structure of the Schur–Horn polytope plays an important role in understanding the equilibria and stability of the generalized double-bracket flows.

This work will be further elaborated and extended in [**BI03a**].

2. Generalized double-bracket equations

2.1. The main equations. Let $\boldsymbol{Y}_0, \boldsymbol{N} \in \mathrm{Sym}(n)$ and suppose that $g : \mathrm{Sym}(n) \to \mathbb{R}$ is a smooth function. In particular, we are interested in the functions

$$g_m(\boldsymbol{X}) = \mathrm{tr}\, \boldsymbol{X}^m, \qquad m \in \mathbb{Z}_+.$$

We denote by $\boldsymbol{G} : \mathrm{Sym}(n) \to \mathrm{Sym}(n)$ the Jacobian of g. In particular,

$$[\boldsymbol{G}_m(\boldsymbol{X})]_{k,l} = \frac{\partial g_m(X)}{\partial \boldsymbol{X}_{k,l}}, \quad 1 \leq k, l \leq n.$$

It is an easy matter to prove that

$$\boldsymbol{G}_m(X) = m\boldsymbol{X}^{m-1}, \qquad m \in \mathbb{Z}_+. \tag{2.1}$$

The *generalized double-bracket equation (GDBE)* is

$$\boldsymbol{Y}' = [\boldsymbol{Y}, [\boldsymbol{G}(\boldsymbol{Y} - \boldsymbol{N}), \boldsymbol{Y}]], \quad t \geq 0, \qquad \boldsymbol{Y}(0) = \boldsymbol{Y}_0. \tag{2.2}$$

More specifically, in the present paper we focus on $\boldsymbol{G} = \boldsymbol{G}_m$, in which case we drop the multiplicative factor m (which merely corresponds to reparametrisation) and consider

$$\boldsymbol{Y}' = [\boldsymbol{Y}, [(\boldsymbol{Y} - \boldsymbol{N})^{m-1}, \boldsymbol{Y}]], \quad t \geq 0, \qquad \boldsymbol{Y}(0) = \boldsymbol{Y}_0. \tag{2.3}$$

This, clearly, makes sense and is nontrivial only for $m \geq 2$.

We are interested in the stability and optimality of the equilibria of this generalized system. For $m \geq 3$, as we shall discuss below, not all equilibria are diagonal. The diagonal equilibria are of great interest however and we begin by considering these.

Given

$$\boldsymbol{Y}' = [\boldsymbol{Y}, [\boldsymbol{Y}, (\boldsymbol{Y} - \boldsymbol{N})^m]]$$

the linearization at $\boldsymbol{Y} = \boldsymbol{D} = \operatorname{diag} \boldsymbol{d}$ is

$$\begin{aligned}(\delta \boldsymbol{Y})' &= [\boldsymbol{D}, [\delta \boldsymbol{Y}, (\boldsymbol{D} - \boldsymbol{N})^m]] \\ &\quad + [\boldsymbol{D}, [\boldsymbol{D}, \delta \boldsymbol{Y} (\boldsymbol{D} - \boldsymbol{N})^{m-1} + (\boldsymbol{D} - \boldsymbol{N}) \delta \boldsymbol{Y} (\boldsymbol{D} - \boldsymbol{N})^{m-2} + \cdots]].\end{aligned}$$

Writing this in components with $\eta_j = d_j - \lambda_j$ we have

$$(\delta y_{ij})' = \delta y_{ij} (\eta_j^m - \eta_i^m)(d_i - d_j) + (\delta y_{ij} \eta_j^{m-1} + \eta_i \delta y_{ij} \eta_j^{m-2} + \cdots)(d_i - d_j)^2$$

or

$$(\delta y_{ij})' = \delta y_{ij} \left[(\eta_j^m - \eta_i^m)(d_i - d_j) + \frac{\eta_j^m - \eta_i^m}{\eta_j - \eta_i} (d_i - d_j)^2 \right].$$

Hence for stability we require

$$(d_i - d_j)(\eta_j^m - \eta_i^m) \left(1 + \frac{d_i - d_j}{\eta_j - \eta_i} \right) < 0$$

or

$$(d_i - d_j)(\lambda_i - \lambda_j) \frac{\eta_j^m - \eta_i^m}{\eta_j - \eta_i} < 0, \qquad 1 \leq i, j \leq n, \quad i \neq j.$$

3. The dynamics of GDBE

3.1. The case $m = 3$. The stability conditions are

$$(d_k - d_l)(\lambda_k - \lambda_l)(\lambda_k + \lambda_l - d_l - d_k) > 0, \qquad 1 \leq k, l \leq n. \tag{3.1}$$

In other words, given distinct real $\lambda_1, \ldots, \lambda_n, d_1, \ldots, d_n$, we seek a permutation of $d_1, d_2, \ldots, d_n$ so that (3.1) holds.

The combinatorial problem (3.1) has a solution for $n = 3$, as can be demonstrated in a direct manner, by considering a welter of special cases. Such direct demonstration is, unfortunately, fairly complicated and it is highly unlikely that it can be extended even to $n = 4$. It is more natural and promising to define the function $h(t) = \operatorname{tr} (\boldsymbol{Y} - \boldsymbol{N})^3$ and hope that it is strictly monotone for all non-diagonal $\boldsymbol{Y}$. Numerical computations confirm this but, so far, analysis is *very* messy. The beauty of this approach is that, once we prove it, we can claim that the above combinatorial problem has solution for every n, a pleasing, roundabout, homotopy argument for the solution of a discrete problem.

A weaker, yet interesting, alternative course of action is to show that $h'(t) < 0$ for an open, nonempty set of initial values: at the very least, this will imply the existence of an attractive fixed point, i.e. of a solution to the set of inequalities (3.1).

Let $\boldsymbol{\lambda} \in \mathbb{R}^n$ be given and assume that its components are distinct. We say that a vector $\boldsymbol{v} \in \mathbb{R}^n$ is *consistent with* $\boldsymbol{\lambda}$ if

$$(\lambda_k - \lambda_l)(v_k - v_l)(v_k + v_l - \lambda_k - \lambda_l) > 0, \qquad 1 \leq k, l \leq n, \ k \neq l. \tag{3.2}$$

We denote the set of all such vectors by $\mathscr{C}_{\boldsymbol{\lambda}}$.

LEMMA 1. *The set $\mathscr{C}_{\boldsymbol{\lambda}}$ is nonempty.*

PROOF. We commence by assuming that $\lambda_k > 0$, $1 \le k \le n$. Let $\rho > 0$ and set $v_k = \rho\lambda_k$, $1 \le k \le n$. Then

$$(\lambda_k - \lambda_l)(v_k - v_l)(v_k + v_l - \lambda_k - \lambda_l) = \rho(\rho-1)(\lambda_k-\lambda_l)^2(\lambda_k+\lambda_l) > 0$$

and (3.2) is satisfied.

We next observe that, for any $\alpha \in \mathbb{R}$,

$$\boldsymbol{v} \in \mathscr{C}_{\boldsymbol{\lambda}} \qquad \Leftrightarrow \qquad \boldsymbol{v} + \alpha\mathbf{1} \in \mathscr{C}_{\boldsymbol{\lambda}+\alpha\mathbf{1}}.$$

We can thus always choose α so that $\lambda_k + \alpha > 0$, $1 \le k \le n$ and reduce everything to the framework of the previous paragraph. □

The last lemma is less useful than it sounds, because we are not interested in arbitrary members of $\mathscr{C}_{\boldsymbol{\lambda}}$. Given $\boldsymbol{Y}_0 \in \mathrm{Sym}(n)$ and $\boldsymbol{N} = \mathrm{diag}\,\boldsymbol{\lambda}$, we interpret $\boldsymbol{v}$ as $\mathrm{diag}\,\boldsymbol{Y}_0$, therefore we wish $\boldsymbol{v}$ to live in the Horn–Schur polytope of $\boldsymbol{d} = \sigma(\boldsymbol{Y}_0)$: according to the Horn lemma, in that case, for every $\boldsymbol{v}$ in the polytope we can always find $\boldsymbol{Y}_0 \in \mathrm{Sym}(n)$ with $\sigma(\boldsymbol{Y}_0) = \boldsymbol{\lambda}$ and $\mathrm{diag}\,\boldsymbol{Y}_0 = \boldsymbol{v}$.

3.2. Stability sectors in the Horn–Schur polytope. What is going on in the Horn–Schur polytope? We start from the case $m = 3$, $n = 3$ and note that, since every point in the polytope is normal to $\mathbf{1}$, it lives in $\mathbb{R}^2$. Thus, plotting it, we can project it on a plane: Given $\boldsymbol{x} \in \mathbb{R}^3$ such that (as is the case in the HSP) $\boldsymbol{x}^\top\mathbf{1} = c$, where $c = \boldsymbol{d}^\top\mathbf{1}$, we plot the coordinates

$$\mathbb{R}^2 \ni \boldsymbol{z} = \begin{bmatrix} \frac12(1+\sqrt{3}) & \frac12(-1+\sqrt{3}) & 0 \\ \frac12(-1+\sqrt{3}) & \frac12(1+\sqrt{3}) & 0 \end{bmatrix} (\boldsymbol{x} - \tfrac13 c\mathbf{1}).$$

Our first collection of plots takes two vectors, $\boldsymbol{d}, \boldsymbol{\lambda} \in \mathbb{R}^3$ and colours differently the set of all points in the HS polytope according to the following scheme: Take a point $\boldsymbol{y}$ therein and check for which permutation $\boldsymbol{\eta} = \boldsymbol{\pi}(\boldsymbol{\lambda})$ of $\boldsymbol{\lambda}$ it is true that

$$(\eta_k - \eta_l)(y_k - y_l)(y_k + y_l - \eta_k - \eta_k) \ge 0, \qquad k, l = 1, 2, 3.$$

We use the following colour code, consistent with RGB colouring:

$$\begin{array}{ll} \boldsymbol{\pi} = [1,2,3]: & \text{red,} \\ \boldsymbol{\pi} = [1,3,2]: & \text{yellow,} \\ \boldsymbol{\pi} = [2,3,1]: & \text{green,} \\ \boldsymbol{\pi} = [3,2,1]: & \text{cyan,} \\ \boldsymbol{\pi} = [3,1,2]: & \text{blue,} \\ \boldsymbol{\pi} = [2,1,3]: & \text{magenta.} \end{array}$$

In monochrome, "colours" correspond to different shades of gray, ordered from the lightest to the darker as

$$\text{yellow} \prec \text{cyan} \prec \text{green} \prec \text{magenta} \prec \text{red} \prec \text{blue}.$$

This scheme of permutations is consistent with 'travelling' along the faces of the HS polytope.

The meaning of the experiment is as follows: the red points are arranged consistently with the stable attractor for $N = \mathrm{diag}\,[\lambda_1, \lambda_2, \lambda_3]$, the magenta points consistently with $N = \mathrm{diag}\,[\lambda_2, \lambda_1, \lambda_3]$ and so on. Of course, it does not mean (or, at least, we did not prove it, unlike the case $m = 2$, and it might well be false) that, once you solve the ODE with an initial condition in the red region (and with

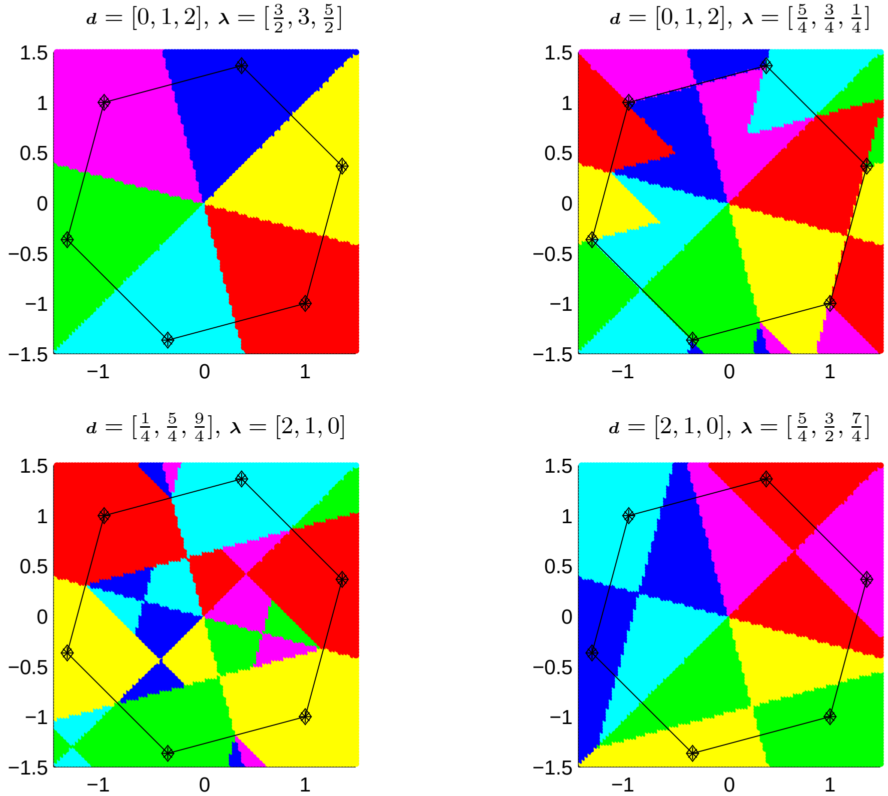

FIGURE 1. Colouring of the HSP.

$N = \text{diag}\,[\lambda_1, \lambda_2, \lambda_3]$) then you *must* tend to the corresponding vertex. All it means is that if a vertex is bordered completely by a red region and we start near enough to it, attractivity of the fixed point will bring us to the vertex in question.

In Fig. 1 we display four such plots, for different choices of $\boldsymbol{d}$ and $\boldsymbol{\lambda}$.

Since the λ_k's are distinct, the colour changes once one of six equalities holds for the vector $\boldsymbol{\eta}$,

$$\eta_k = \eta_l, \qquad \eta_k + \eta_l = \lambda_k + \lambda_l, \qquad 1 \le k < l \le 3.$$

Recall that everything has been rotated in Fig. 1 to a plane, hence we need to express the above equalities as linear equations in the (x, y) plane. Moreover, the one consequence of being in the HS polytope is that $\mathbf{1}^\top \boldsymbol{\eta} = c$ for $c = \mathbf{1}^\top \boldsymbol{d}$.

For example, the condition $\eta_2 = \eta_3$, in tandem with $\eta_1 + \eta_2 + \eta_3 = c$, after rotation, yields

$$x = \tfrac{1}{4}[(3 + \sqrt{3})\eta_1 - (1 + 3\sqrt{3})c], \qquad y = \tfrac{1}{4}[(-3 + \sqrt{3})\eta_1 + (1 - 3\sqrt{3})c].$$

Eliminating η_1 from the above equations results in the line

$$y(x) = (\sqrt{3} - 2)[x + 8(1 + \sqrt{3})c].$$

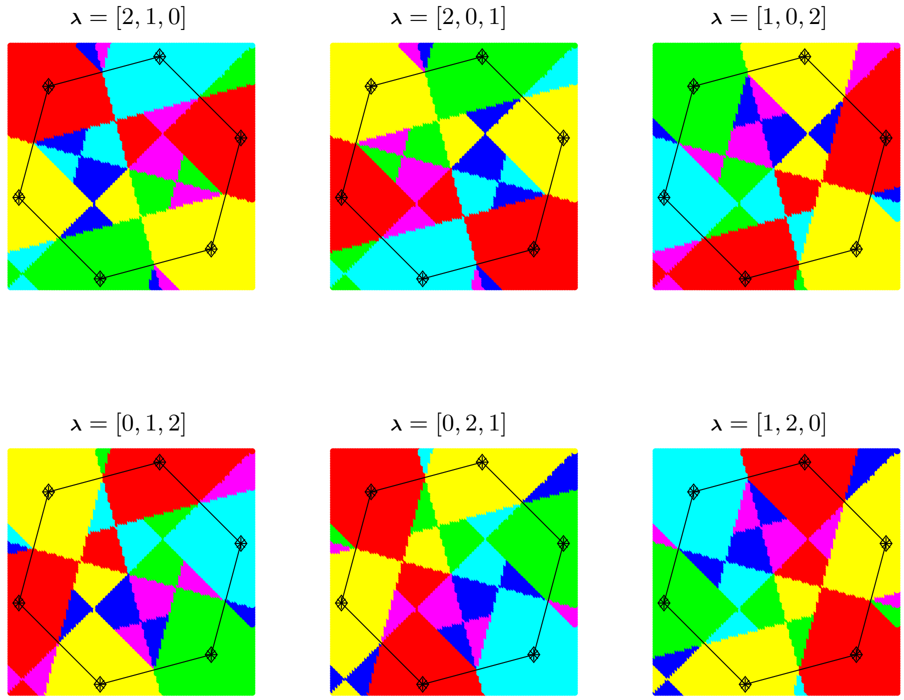

FIGURE 2. $\boldsymbol{d} = [\frac{1}{4}, \frac{5}{4}, \frac{9}{4}]$ and different permutations of $\lambda = [2, 1, 0]$.

In general, we obtain the following lines along which colour might change:

$$\begin{aligned}
\eta_1 = \eta_2 : &\qquad y(x) = x, \\
\eta_1 = \eta_3 : &\qquad y(x) = -(2+\sqrt{3})[x + 4(\sqrt{3}-1)c], \\
\eta_2 = \eta_3 : &\qquad y(x) = (\sqrt{3}-2)[x + 8(1+\sqrt{3})c], \\
\eta_1 + \eta_2 = \lambda_1 + \lambda_2 : &\qquad y(x) = -x - 2\sqrt{3}c + \sqrt{3}(\lambda_1 + \lambda_2), \\
\eta_1 + \eta_3 = \lambda_1 + \lambda_3 : &\qquad y(x) = (2-\sqrt{3})[x - (6+2\sqrt{3})(\lambda_1 + \lambda_3)], \\
\eta_2 + \eta_3 = \lambda_2 + \lambda_3 : &\qquad y(x) = (2+\sqrt{3})[x + (6-2\sqrt{3})(\lambda_2 + \lambda_3)].
\end{aligned}$$

Note that the first three lines are always parallel to a face of the (rotated) HS polytope. Moreover, the *slopes* of all six lines are independent of the choice of $\boldsymbol{d}$ and $\boldsymbol{\lambda}$, all that changes is the affine shift.

In Fig. 2 we display the coloured HS polytope for fixed $\boldsymbol{d}$ and all permutations of $\boldsymbol{\lambda}$. Several symmetries stare us in the face. For example, any permutation of the form $(a, b, c) \mapsto (a, c, b)$ corresponds to red $\mapsto$ magenta, blue $\mapsto$ green and cyan $\mapsto$ yellow. There are other symmetries, which might include both colour swaps and reflections (or rotations?) of the picture. Are they universal or special for this particular choice of $\boldsymbol{d}$ and $\boldsymbol{\lambda}$? We intend to discuss this further in [**BI03a**].

4. General equilibria

As mentioned earlier, it is entirely possible for the matrices $\boldsymbol{Y} \in \mathrm{Sym}(n)$ and $\boldsymbol{N} = \mathrm{diag}\,\boldsymbol{\lambda}$ to be such that $[(\boldsymbol{Y} - \boldsymbol{N})^2, \boldsymbol{Y}] = O$, yet $(\boldsymbol{Y} - \boldsymbol{N})^2$ is *not* diagonal.

The very important consequence is that *we might have fixed points – even attractive fixed points* – inside *the HS polytope!*

4.1. Nontrivial solutions: the solution for $n = 3$. We consider now solutions to the equilibrium condition

$$[\boldsymbol{N}^2, \boldsymbol{Y}] = [\boldsymbol{N}, \boldsymbol{Y}^2] \tag{4.1}$$

The idea is to commence by specifying the *diagonal* of $\boldsymbol{Y} \in \mathrm{Sym}(3)$. Thus, suppose that

$$\boldsymbol{Y} = \begin{bmatrix} \alpha & a & b \\ a & \beta & c \\ b & c & \gamma \end{bmatrix},$$

where α, β, γ are fixed and we seek to determine a, b, c so that (4.1) is satisfied.

The matrix equation (4.1) is equivalent to

$$(\lambda_k + \lambda_l) Y_{k,l} = (\boldsymbol{Y}^2)_{k,l}, \qquad 1 \le k < l \le n \tag{4.2}$$

(in our case $n = 3$) and the latter can be used to find the conditions on a, b, c: they are

$$\begin{aligned} (\lambda_1 + \lambda_2) a &= (\alpha + \beta) a + cb, \\ (\lambda_1 + \lambda_3) b &= (\alpha + \gamma) b + ac, \\ (\lambda_2 + \lambda_3) c &= (\beta + \gamma) c + ab. \end{aligned}$$

An obvious solution of the above equations is $a = b = c = 0$, which corresponds to a diagonal matrix $\boldsymbol{Y}$ (which, needless to say, obeys (4.1)). Being interested in nontrivial solutions, we exclude this case, whereby

$$a^2(\lambda_1 + \lambda_2 - \alpha - \beta) = b^2(\lambda_1 + \lambda_3 - \alpha - \gamma) = c^2(\lambda_2 + \lambda_3 - \beta - \gamma) = abc.$$

Note that in the present case $a, b, c \neq 0$. We obtain

$$\begin{aligned} a &= (-1)^{k_a} b \sqrt{\frac{\lambda_1 + \lambda_3 - \alpha - \gamma}{\lambda_1 + \lambda_2 - \alpha - \beta}}, \\ c &= (-1)^{k_c} b \sqrt{\frac{\lambda_1 + \lambda_3 - \alpha - \gamma}{\lambda_2 + \lambda_3 - \beta - \gamma}}, \end{aligned}$$

where $k_a, k_b \in \{0, 1\}$. Moreover,

$$abc = (\lambda_1 + \lambda_3 - \alpha - \gamma) b^2 = (-1)^{k_a + k_c} \frac{\lambda_1 + \lambda_3 - \alpha - \gamma}{\sqrt{(\lambda_1 + \lambda_2 - \alpha - \beta)(\lambda_2 + \lambda_3 - \beta - \gamma)}} b^3.$$

Therefore, we deduce that

$$\begin{aligned} a &= (-1)^{k_a} \sqrt{(\lambda_1 + \lambda_3 - \alpha - \gamma)(\lambda_2 + \lambda_3 - \beta - \gamma)}, \\ b &= (-1)^{k_a + k_c} \sqrt{(\lambda_1 + \lambda_2 - \alpha - \beta)(\lambda_2 + \lambda_3 - \beta - \gamma)}, \\ c &= (-1)^{k_c} \sqrt{(\lambda_1 + \lambda_2 - \alpha - \beta)(\lambda_1 + \lambda_3 - \alpha - \gamma)}. \end{aligned} \tag{4.3}$$

This exhausts all possible solutions, except that we must ensure that they are real. This will happen if and only if the three numbers

$$\lambda_1 + \lambda_2 - \alpha - \beta, \lambda_1 + \lambda_3 - \alpha - \gamma, \lambda_2 + \lambda_3 - \beta - \gamma$$

are all of the same sign. Therefore, (α,β,γ) must lie in one of two cones: either

$$\begin{aligned}\alpha+\beta &\geq \lambda_1+\lambda_2,\\ \alpha+\gamma &\geq \lambda_1+\lambda_3,\\ \beta+\gamma &\geq \lambda_2+\lambda_3\end{aligned}$$

or

$$\begin{aligned}\alpha+\beta &\leq \lambda_1+\lambda_2,\\ \alpha+\gamma &\leq \lambda_1+\lambda_3,\\ \beta+\gamma &\leq \lambda_2+\lambda_3.\end{aligned}$$

Letting $\bar\alpha=\alpha-\lambda_1$, $\bar\beta=\beta-\lambda_2$, $\bar\gamma=\gamma-\lambda_3$, the first cone, $\mathcal{C}$, say, becomes

$$\bar\alpha+\bar\beta\geq 0,\quad \bar\alpha+\bar\gamma\geq 0,\quad \bar\beta+\bar\gamma\geq 0,$$

while the second cone is $-\mathcal{C}$. Therefore the condition for the existence of a nontrivial solution is that

$$\operatorname{diag} Y\in \boldsymbol{\lambda}+\mathcal{C}\quad\text{or}\quad \operatorname{diag} Y\in\boldsymbol{\lambda}-\mathcal{C}. \tag{4.4}$$

Of course, the whole point is that $\operatorname{diag} Y\in\mathcal{HS}$, the Horn–Schur polytope. Thus, nontrivial solutions exist if and only if

$$\text{either}\quad \mathcal{HS}\cap\{\boldsymbol{\lambda}+\mathcal{C}\}\neq\emptyset\quad\text{or}\quad \mathcal{HS}\cap\{\boldsymbol{\lambda}-\mathcal{C}\}\neq\emptyset.$$

Recall that all $\boldsymbol{x}\in\mathcal{HS}$ lie on the plane $\mathbf{1}^\top\boldsymbol{x}=\mathbf{1}^\top\boldsymbol{d}$. The intersection of this plane with the union of the two cones is always nonempty, but this is not to say that the intersection of the HS polytope itself with the union of the cones is nonempty. The easiest special case is when $\mathbf{1}^\top\boldsymbol{d}=\mathbf{1}^\top\boldsymbol{\lambda}$, whereby the intersection of the plane with the cones is just the point set $\{\boldsymbol{\lambda}\}$. Therefore, in that case, either $\boldsymbol{\lambda}\in\mathcal{HS}$ and we have a nontrivial solution,[1] or $\boldsymbol{\lambda}\notin\mathcal{HS}$ and all solutions are trivial. For all other values of $\mathbf{1}^\top\boldsymbol{d}$ the set of nontrivial solutions in the HS polytope is itself a polytope. Moreover, every point in this polytope corresponds to four different solutions, with different sign configurations.

4.2. General $n\geq 2$. Let us commence by setting the stage for general n. Denoting $\operatorname{diag}\boldsymbol{Y}=\boldsymbol{x}$, it follows from (4.2) that

$$\begin{aligned}(\lambda_k+\lambda_l)Y_{k,l} &= \sum_{i=1}^{k-1}Y_{i,k}Y_{i,l}+x_kY_{k,l}+\sum_{i=k+1}^{l-1}Y_{k,i}Y_{i,l}+x_lY_{k,l}\\ &\quad+\sum_{i=l+1}^{n}Y_{k,i}Y_{l,i},\qquad 1\leq k<l\leq n.\end{aligned}$$

Therefore, letting

$$\mu_{k,l}=\lambda_k+\lambda_l-x_k-x_l,\qquad 1\leq k<l\leq n,$$

we have

$$\mu_{k,l}Y_{k,l}=\sum_{i=1}^{k-1}Y_{i,k}Y_{i,l}+\sum_{i=k+1}^{l-1}Y_{k,i}Y_{i,l}+\sum_{i=l+1}^{n}Y_{k,i}Y_{l,i},\quad 1\leq k<l\leq n. \tag{4.5}$$

[1]Recall our assumption that the spectra are distinct, hence $\boldsymbol{\lambda}$ cannot be a vertex of the HS polytope.

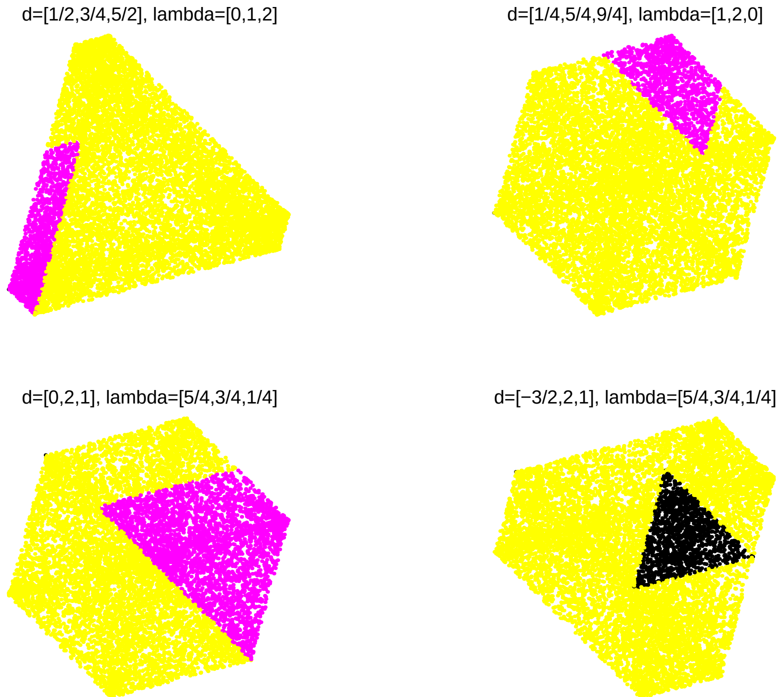

FIGURE 3. Different nontrivial solution regimes. Magenta denotes points in $\boldsymbol{\lambda}+\mathcal{C}$, black points are in $\boldsymbol{\lambda}-\mathcal{C}$ and yellow in neither cone. (In monochrome, "yellow" is the brightest and "black" the darkest.)

Consider now the general case. There is just a single solution for $n=2$, namely a diagonal matrix ($Y_{1,2}=0$). We denote the number of solutions in the $n\times n$ case by ψ_n, therefore

$$\psi_2=1,\quad \psi_3=5,\quad \psi_4=25.$$

(It might be tempting by this stage to conjecture that $\psi_n=5^{n-2}$, but, as will be clear later, is probably false.)

Clearly, for $n\geq 3$ there are lower-dimensional problems hidden in the matrix and they can be obtained simply by taking rows and columns of off-diagonal zeros. Our real interest is in the solutions which cannot be obtained from lower-dimensional cases. We denote the number of such solutions by $\psi_n^\star$: $\psi_2^\star=1$, $\psi_3^\star=4$, $\psi_4^\star=16$. It follows trivially that

$$\psi_n=\sum_{k=0}^{n-3}\binom{n}{k}\psi_{n-k}^\star+1=\sum_{k=3}^{n}\binom{n}{k}\psi_k^\star+1 \tag{4.6}$$

(the "+1" corresponds to the diagonal solution). From now on, we concentrate only on the "truly n-dimensional" solutions.

We say that the symmetric matrix $E=(e_{k,l})_{k,l=1}^n$ is *tri-balanced* if

$$e_{k,l}\in\{0,1\},\quad 1\leq k,l\leq n \qquad \text{and} \qquad e_{i,k}+e_{i,l}=e_{k,l} \bmod 2,\quad 1\leq i,k,l\leq n$$

(note that automatically $e_{k,k}=0$, but we will not use diagonal elements). Assuming that E is tri-balanced, we set

$$Y_{k,l}=(-1)^{e_{k,l}}\sqrt{(\sigma-\lambda_k+x_k)(\sigma-\lambda_l+x_l)},\qquad 1\le k,l\le n,\quad k\ne l, \tag{4.7}$$

where

$$\sigma=\frac{1}{n-2}\sum_{i=1}^{n}(\lambda_i-x_i)=\frac{1}{n-2}\mathrm{tr}\,(\boldsymbol{N}-\boldsymbol{Y}). \tag{4.8}$$

We claim that (4.7) represents a solution of (4.5), thereby of (4.1):

$$\begin{aligned}\sum_{\substack{i=1\\ i\ne k,l}}^{n} Y_{k,i}Y_{l,i} &= \sum_{\substack{i=1\\ i\ne k,l}}^{n}(-1)^{e_{i,k}+e_{i,l}}(\sigma-\lambda_i+x_i)\sqrt{(\sigma-\lambda_k+x_k)(\sigma-\lambda_l+x_l)}\\ &= \sum_{\substack{i=1\\ i\ne k,l}}^{n}(-1)^{e_{i,k}+e_{i,l}-e_{k,l}}(\sigma-\lambda_i+x_i)Y_{k,l}\\ &= \left[\sum_{i=1}^{n}(\sigma-\lambda_i+x_i)-2\sigma+(\lambda_k+\lambda_l)-(x_k+x_l)\right]Y_{k,l}=\mu_{k,l}Y_{k,l}.\end{aligned}$$

for all $k\ne l$. Therefore, for every tri-balanced matrix there corresponds a solution (4.7).

But how many tri-balanced matrices are there? Suppose that $e_{k,1}=e_{1,k}$ is given for all $1\le k\le n$ (of course, $e_{1,1}=0$) and set

$$e_{k,l}=e_{1,k}+e_{1,l},\qquad 1\le k,l\le n.$$

It is easy to see that, with the above definition,

$$e_{k,i}+e_{i,l}=(e_{1,k}+e_{1,i})+(e_{1,l}+e_{1,i})=e_{1,k}+e_{1,l}=e_{k,l} \bmod 2,$$

hence we have a tri-balanced matrix. Moreover, no other choice of $e_{k,l}$ is tri-balanced. This is true because suppose that there exist k,l such that $e_{i,k}\ne e_{1,i}+e_{1,k} \bmod 2$. Then necessarily $e_{i,k}\ne e_{1,i}+e_{1,k}+1 \bmod 2$, hence for every i

$$\begin{aligned} & e_{i,k}+e_{i,l}=e_{1,k}+e_{1,l} \bmod 2\\ \Rightarrow\quad & e_{i,l}=e_{1,l}-(e_{1,i}+e_{1,k}+1)=e_{1,i}+e_{1,l}+1 \bmod 2\\ \overset{i=l}{\Rightarrow}\quad & 2e_{1,l}+1=0 \bmod 2,\end{aligned}$$

which is a contradiction. Therefore, we have precisely 2^{n-1} tri-balanced matrices and $\psi_n^\star\ge 2^{n-1}$.

We deduce that

$$\psi_n\ge 1+\sum_{k=3}^{n}\binom{n}{k}2^{k-1}=\tfrac{1}{2}(3^n+1-2n^2).$$

Stranger number than 5^{n-2}, yet this upper bound coincides with ψ_n for $n=1,2,3$. However, it is presently unclear whether it represents *all* solutions of (4.1) or is just a lower bound.

Again, solutions are valid only if either

$$x_k\ge\lambda_k-\sigma,\qquad 1\le k\le n,$$

or

$$x_k\le\lambda_k-\sigma,\qquad 1\le k\le n.$$

A cone and its negative, again.... Thus, we have a solution (given eigenvalues $\boldsymbol{d}$, rather than the diagonal $\boldsymbol{x}$) if and only if one of these cones intersects the HS polytope.

Clearly, the above analysis falls short of a comprehensive description of the dynamics of generalized double-bracket flows and its connection with the Horn–Schur polytope. Yet, even these partial and often tentative results demonstrate that the underlying problem is fascinating, nontrivial and leads in some highly interesting directions.

References

[BBR90] A. M. Bloch, R. W. Brockett, and T. Ratiu, *A new formulation of the generalized Toda lattice equations and their fixed-point analysis via the moment map*, Bull. AMS **23** (1990), 447–456.

[BBR92] A. M. Bloch, R. W. Brockett, and T. S. Ratiu, *Completely integrable gradient flows*, Comm. Math. Phys. **147** (1992), 57–74.

[BI03a] A. M. Bloch and A. Iserles, *Generalized double bracket equations*, Tech. Report to appear, DAMTP, University of Cambridge, 2003.

[BI03b] ———, *On the optimality of double-bracket flows*, Tech. Report NA 2003/04, DAMTP, University of Cambridge, 2003.

[Bro91] R. W. Brockett, *Dynamical systems that sort lists, diagonalize matrices, and solve linear programming problems*, Linear Algebra Appl. **146** (1991), 79–91.

[Bro93] ———, *Differential geometry and the design of gradient algorithms*, Proc. Symp. Pure Maths **54** (1993), 69–92.

[CD91] Moody T. Chu and Kenneth R. Driessel, *Constructing symmetric nonnegative matrices with prescribed eigenvalues by differential equations*, SIAM J. Math. Anal. **22** (1991), 1372–1387.

[CIZ97] M. P. Calvo, A. Iserles, and A. Zanna, *Numerical solution of isospectral flows*, Maths Comput. **66** (1997), 1461–1486.

[HM94] U. Helmke and J. Moore, *Optimization and dynamical systems*, Springer Verlag, New York, 1994.

[Ise02] Arieh Iserles, *On the discretization of double-bracket flows*, Found. Comp. Maths **2** (2002), 305–329.

Department of Mathematics, University of Michigan, Ann Arbor, MI 48109, USA
E-mail address: abloch@umich.edu

Department of Applied Mathematics and Theoretical Physics, Centre for Mathematical Sciences, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, England
E-mail address: A.Iserles@damtp.cam.ac.uk.