Equivariant bifurcation, quadratic equivariants, and symmetry breaking for the standard representation of $S_k$

Yossi Arjevani$^1$ and Michael Field$^{2,*}$

$^1$ School of Engineering and Computer Science, Hebrew University, Israel
$^2$ Department of Mechanical Engineering, UCSB, Santa Barbara, CA 93106, United States of America

E-mail: yossi.arjevani@gmail.com and mikefield@gmail.com

Received 5 July 2021, revised 11 March 2022
Accepted for publication 28 March 2022
Published 27 May 2022

Abstract
Motivated by questions originating from the study of a class of shallow student-teacher neural networks, methods are developed for the analysis of spurious minima in classes of gradient equivariant dynamics related to neural networks. In the symmetric case, methods depend on the generic equivariant bifurcation theory of irreducible representations of the symmetric group on $k$ symbols, $S_k$; in particular, the standard representation of $S_k$. It is shown that spurious minima (non-global local minima) do not arise from spontaneous symmetry breaking but rather through a complex deformation of the landscape geometry that can be encoded by a generic $S_k$-equivariant bifurcation. We describe minimal models for forced symmetry breaking that give a lower bound on the dynamic complexity involved in the creation of spurious minima when there is no symmetry. Results on generic bifurcation when there are quadratic equivariants are also proved; this work extends and clarifies results of Ihrig & Golubitsky and Chossat, Lauterbach & Melbourne on the instability of solutions when there are quadratic equivariants.

Keywords: equivariant bifurcation, spurious minima, gradient dynamics, symmetric group, forced symmetry breaking, minimal models

Mathematics Subject Classification numbers: 37C80, 37G40, 34C23, 58K70.

(Some figures may appear in colour only in the online journal)
1. Introduction

Using ideas originating in equivariant bifurcation theory, we develop methods that can be used to understand the creation and annihilation of spurious minima (non-global local minima) in shallow neural networks. Specifically, the results apply to student-teacher networks that inherit symmetry from the target model. In the first part of the introduction, there is an overview of the motivation, background and results. This is followed by an outline of contents and description of the mathematical contributions—the focus of the article. At the suggestion of the referees, brief notes are included at the end of the introduction on the relationship of [1–6] to this paper.

1.1. Motivation and background

In general terms, this article developed out of a program to understand why the highly non-convex optimisation landscapes induced by natural distributions allow gradient based methods, such as stochastic gradient descent (SGD), to find good minima efficiently (see [4] for background and sources on neural networks and the student-teacher framework). This article concerns mathematical aspects of these problems, mainly related to bifurcation theory, and no knowledge of neural networks is required for understanding the results or proofs (the specific optimisation problem is briefly discussed in the concluding comments section).

Let $S_k$ denote the permutation group on $k$ symbols. The foundational theory of generic $S_k$-equivariant steady-state bifurcation on the standard (natural) representation of $S_k$ on $H_{k-1} = \{ (x_1, \cdots, x_k) \in \mathbb{R}^k \mid \sum x_i = 0 \}$ was developed by Field and Richardson about thirty years ago. It was shown [16, 17, 19] that generic branching was always along axes of symmetry (‘axial’ in the terminology of [23]) and a complete classification of the (signed indexed) branching patterns was obtained [19, section16]. A feature of generic bifurcation is that all the non-trivial branches of solutions consist of hyperbolic saddles (index $\neq 0, k - 1$). In particular, if the trivial branch of sinks loses stability, no branches of sinks or sources will appear post-bifurcation, other than the trivial branch of sources generated by the change in stability of the trivial solution. This applies if bifurcation occurs on a centre manifold locally $S_k$-equivariantly diffeomorphic to $H_{k-1}$. If all the transverse directions are attracting, then we may see a transition from hyperbolic saddle to hyperbolic sink. More can be said, but first we recall [13, 19] the equations for generic $S_k$-equivariant bifurcation on $H_{k-1}$.

\begin{align}
\mathbf{x}' &= \lambda \mathbf{x} + Q(\mathbf{x}) \quad (k \text{ odd}) \tag{1.1} \\
\mathbf{x}' &= \lambda \mathbf{x} + Q(\mathbf{x}) + T(\mathbf{x}) \quad (k \text{ even}) \tag{1.2}
\end{align}

Here $Q \neq 0$ is a (any) homogeneous quadratic equivariant gradient vector field, and $T$ is the gradient of a homogeneous quartic $S_k$-invariant which is not identically zero on the $S_k$-orbit of one special axis of symmetry (details are given later). Perhaps surprisingly, pre-bifurcation ($\lambda < 0$) all indices of branches are at least $[k/2]$; post bifurcation, all indices are at most $[k/2]$ (equality with $k/2$ occurs iff $k$ is even). A consequence is that the high-dimensional stable manifolds of the saddle branches pre-bifurcation make the attracting trivial solution branch increasingly invisible to trajectories initialised far from the origin as $\lambda \to 0^-$. Similar remarks hold post-bifurcation. If $k$ is odd, there are $2^k - 2$ non-trivial branches of solutions; half of these are backward, and half are forward (two branches are associated to each of the $2^{k-1} - 1$
axes of symmetry). A slightly more complicated formula, depending on the cubic term, can be given when \( k \) is even—see remark 4.2(a).

Although (1.1) and (1.2) have been used as a starting point for physical models (for example, the work of Stewart et al on speciation [23, chapter 2, section 2.7]), from the point of view of local bifurcation and dynamics, the lack of any non-trivial branches of sinks (or sources) limits the use of (1.1) and (1.2) as general or universal models for bifurcation and dynamics. Of course, terms of odd degree may be added, such as \(-c||x||^{2p}x\), where \( c > 0, p \in \mathbb{N} \), to create sinks but these are far from the origin of \( H_{k-1} \) and not part of the bifurcation at \( \lambda = 0, x = 0 \). If \( k \) is small, consideration of secondary bifurcation, mode interactions and unfolding theory can be effective tools for the analysis of specific problems (for example [22, chapter 15, section 4]). For our applications, \( k \) will typically not be small.

The approach in this paper to generic steady-state bifurcation on representations of \( S_k \), in particular the standard representation of \( S_k \), has a different perspective. Thus, we regard the generic \( S_k \)-equivariant steady-state bifurcation, as realised by the equations (1.1) and (1.2), as encoding the solution to a complex problem related to the creation of spurious minima in non-convex optimisation. Roughly speaking, as we increase \( k \) in these problems, we see the formation of spurious minima. These do not arise from bifurcation of the global minima. Careful analysis reveals that, in the symmetric case, the spurious minima—at least those seen in the numerics with the appropriate initialisation scheme (cf Xavier initialisation [46], [4, section 1.2])—arise through a steady-state bifurcation along a copy of the standard representation of \( S_k \) (for example, using a centre manifold reduction). Thus the change in stability, in the symmetric case, occurs through the simultaneous collision at the origin of \( O(2^k) \) hyperbolic saddle points of low index resulting in a branch of spurious minima (directions transverse to the centre manifold are assumed contracting)—in (1.1) and (1.2) this amounts to decreasing \( \lambda \). For example, the type II spurious minima described in [4] appear at \( k \approx 5.58 \). Ignoring for a moment the inconvenient detail that \( k \) is an integer representing the number of neurons, the local mechanism for creation or annihilation of minima via generic bifurcation on the standard representation of \( S_k \) should be clear (vary \( \lambda \) in (1.1) and (1.2)). The argument applies to other irreducible representations of \( S_k \) that have quadratic equivariants (cf [9, 24]; for example, external tensor products of standard representations of the symmetric group). For spurious minima that do not appear in the numerics with Xavier initialisation, bifurcation along the exterior square representation of the standard representation of \( S_k \) may occur. While this representation does not have quadratic equivariants, the mechanism for creation of spurious minima appears similar to that of type II minima.

In practice, rigorous analysis is carried out on fixed point spaces of the action. For a large class of isotropy groups, the associated fixed point spaces have dimension independent of \( k \) and the bifurcation equations, restricted to the fixed point space, depend smoothly on \( k \), now viewed as a real parameter (see [4], where power series in \( 1/\sqrt{k} \) are obtained for families of critical points, and the concluding comments section).

We indicated above that the generic \( S_k \)-equivariant steady-state bifurcation could be viewed as encoding the solution to the problem of the creation of spurious minima. To gain insight into the general problem (no symmetry), it is necessary to describe what happens when we break the symmetry of the model (forced symmetry breaking). We do this by introducing the notion of a minimal model (of forced symmetry breaking). We give this by introducing the notion of a minimal model (of forced symmetry breaking). When \( k \) is odd, this is an explicit local symmetry breaking perturbation of the equation (1.1) to a \( C^2 \)-stable family \( \mathcal{F} = \{ F_\lambda \mid \lambda \in \mathbb{R} \} \) (stable within the space of asymmetric families) which has minimal dynamic complexity. The minimal complexity is described in terms of the minimum number of saddle-node bifurcations and the maximum number of hyperbolic solution curves (defined for \( \lambda \in \mathbb{R} \)) that the family \( \mathcal{F} \) must have. For example, if \( k = 17 \), the minimal model will have exactly 52,666 saddle-node
bifurcations and 12,870 hyperbolic solution curves. The minimal model is indicative of the complex landscape geometry that is involved in the creation of asymmetric spurious minima in non-convex optimisation in neural networks (cf [3, 5]). A similar result holds for \( k \) even—on account of the cubic terms in (1.2), the ‘standard’ model we perturb is defined slightly differently so that no solutions are introduced which are unrelated to the bifurcation. In either case, there is a (small) interval of values of \( k \) for which there are no sinks or sources—a reflection of the previously noted relative ‘invisibility’ of the trivial sink or source near the bifurcation point of the \( S_k \)-equivariant problem.

1.2. Outline of paper and main results

Parts of this article have posed expositional problems on account of missing literature references and foundational definitions. The most important of these issues is the absence in much of the equivariant bifurcation theory reference literature of the definition of a solution branch (for example, [22, 23]). We would argue that this definition should be a key foundational concept in the theory. In [22, 23], the default is that of an axial solution branch—bifurcation along an axis of symmetry (for example, [22, section 2]). The existence of axial solution branches (generically always smooth if the underlying family is smooth) uses Vanderbauwhede’s version of the equivariant branching lemma [11, 48] and, mathematically speaking, the analysis of axial branches (for finite group symmetries) is elementary and depends only on the implicit function theorem (used in the proof of the equivariant branching lemma). However, as has been shown many times [7, 9, 16, 24, 33, 40], generic branches of solutions in equivariant bifurcation theory are typically not axial, even if they are of maximal isotropy, and/or branches of sinks or sources, and/or the family consists of gradient vector fields. Related to this problem of definition is the matter of quadratic equivariants. Ihrig and Golubitsky showed that, under certain conditions, steady-state bifurcation on an absolutely irreducible representation with non-trivial quadratic equivariants was unstable [24, theorem 4.2(B)] (no branches of sinks). Their ‘crucial hypothesis’ (H4) [24, p 20] was that branches were axial. Later Chossat et al showed the result of Ihrig and Golubitsky applied without any restriction on isotropy type [9, theorem 4.2(b)]. Their proof is elementary except for one detail (see below). Unfortunately, their result is not mentioned in [23, section 2.3] and was unknown (or forgotten) by us when we began work on this paper.

The definition of solution branch first appears in [17, section 2] and holds for generic bifurcation—specifically, for an open dense set of one-parameter families (\( C^\infty \)-topology). The proof of genericity is not hard but depends on non-trivial equivariant transversality arguments [13]. The authors of [9] were probably unaware of this definition and instead used an approach based on the curve selection lemma (CSL) [42] which gives the result for real analytic families but not smooth families: the CSL holds for semianalytic families (most generally, sub-analytic families [36]) but does not extend to smooth families. One possible way to extend the result of Chossat et al to smooth families is to use stability and determinacy results for stable families [13–15], but these use non-trivial (and difficult) results of Bierstone on equivariant jet transversality [8]. A simpler and more attractive approach is to use the definition of solution branch (see below).

So as to clarify the foundations, section 2 includes the key definitions of solution branch and (signed, indexed) branching pattern (section 2.4), and a statement of the stability theorem (section 2.6), with brief commentary on the proof.

In section 3, a simple proof is given of the result of Chossat et al on quadratic equivariants that only uses the natural notion of a solution branch, rather than arguments invoking the CSL. The main result of the section is expressed in terms of branching patterns and hyperbolic
branches of solutions rather than unstable branches \[9, 24\]. We state the result below only for gradient vector fields (the result extends to non-gradient quadratic equivariants, subject to an additional condition, and to compact Lie groups, the main setting of \[9\]).

**Theorem.** Let \((V, G)\) be an absolutely irreducible representation of the finite group \(G\) and assume there are non-zero quadratic equivariants, all of which are gradient vector fields. Then for all stable families, every non-trivial branch of solutions is a branch of hyperbolic saddles with index lying in \([1, \dim(V) - 1]\). Generically, therefore, there are no non-trivial branches of sinks or sources.

Also discussed are recent developments in stratification theory giving stability of initial exponents, using the regular arc-wise analytic stratification of Parusinski and Păunescu \[43\], and perturbation theory estimates that apply if vector fields are gradient and analytic. We conclude section 3 with a brief description of open questions about analytic parametrisation of solution branches and notes on the extension to compact Lie groups.

In section 4, we define the notion of a minimal model of forced symmetry breaking, emphasising the case of the standard representation of \(S_k\) on \(H_{k-1}\). After reviewing the classification of the signed indexed branching patterns for the standard representation of \(S_k\) \[19, section 16\], we construct minimal models of forced symmetry breaking for the cases \(k\) odd and even. For the introduction we emphasise the case when \(k\) odd since the description of minimal symmetry breaking models for the stable family (1.1) is relatively simple. If \(k\) is even, the constructions and analysis are more complicated on account of the presence of pitchfork bifurcations (along axes of symmetry with isotropy conjugate to \(S_{k/2} \times S_{k/2}\)); these bifurcations result from the cubic terms in (1.2). The cubic terms also introduce new bifurcations outside of a neighbourhood of the origin. For a global result, we break symmetry in a smooth (not analytic) family which has no extra solutions forced by the presence of cubic terms in (1.2). For \(k\) odd or even, the symmetry breaking is local, supported on arbitrarily small neighbourhood of the bifurcation point, and the family constructed is stable under \(C^3\)-small non-equivariant perturbations. The minimal model will also be \(S_{k-1}\)-equivariant—this is important for the non-elementary part of the proof.

**Definition.** Assume the standard representation \((H_{k-1}, S_k)\) of \(S_k\), and that \(k \geq 3\) is odd. Let \(f\) be a stable family which we assume here to be (1.1). The family \(\hat{f}\) is a minimal symmetry breaking model for \(f\) if

(a) \(f = \hat{f}\) on \(V \times \mathbb{R} \setminus W\), where \(W\) is a compact neighbourhood of \((0, 0) \in H_{k-1} \times \mathbb{R}\).
(b) \(\hat{f}\) has exactly \(\binom{k-1}{k/2}\) crossing curves.
(c) \(\hat{f}\) has exactly \(2^{k-1} - \binom{k-1}{k/2}\) saddle-node bifurcations.

We provide proofs that the notion of minimal symmetry breaking model is well-defined and that given any compact neighbourhood \(W\) of \((0, 0) \in H_{k-1} \times \mathbb{R}\), (a)–(c) can be realised by a \(C^3\)-small perturbation of the model (1.1). Similar results hold for \(k\) even. Although there are many details, most of the proof is elementary with the exception of the argument showing no new solutions are introduced. This uses results from [18], [15, section 4.9] which depend on Bezout’s theorem and the pinning of solutions to the complexification of fixed point spaces. Full statements of the results appear in section 4: theorems 4.4 and 4.30.

Although we have not encountered past work on minimal models, it would be surprising if the phenomenon had not been noticed before.

In the concluding comments, we return to the original motivating problem about the creation and annihilation of spurious minima, indicate how the results of the paper can be used to understand this phenomenon, and discussed related current and proposed developments.
1.3. Companion articles [1–6]

Article [2] identifies spurious minima as examples of symmetry breaking in the student-teacher model and gives an extensive numerical study of the phenomenon in wide range of settings. In [4] several infinite families of critical points of spurious minima are constructed and it is shown that these critical points may be represented by convergent fractional power series (FPS) in $1/\sqrt{k}$ ($k$ is the number of neurons viewed as a real parameter). The FPS result is used in [3], together with results from the representation theory of the symmetric group, to obtain precise results on the Hessian spectrum for several families of spurious minima in shallow neural networks (valid for arbitrarily large $k$). In [5], these results are extended to two-layer ReLU networks where it is shown that to order $O(k^{-\frac{1}{2}})$ the spectra are identical for the global minima and several families of spurious minima. All results to this point assume that the number of inputs $d$ to the network is greater than or equal to the number of neurons $k$. In [6], the over-parametrized case $k > d$ is analysed and it is shown, using FPS methods and representation theory, that the addition of one or two neurons annihilates certain families of spurious minima of types I and II (as defined in [6]). These mechanisms are closely related to the results of this paper (see the concluding comments section) and the results in [6] are robust under symmetry breaking perturbations of the target model. Finally, in [1], it is shown that the symmetry breaking phenomena we describe for networks with ReLU activation are seen also in networks with other classes of activation.

2. Generic equivariant bifurcation

2.1. Preliminaries and notation

Let $\mathbb{N}$ denote the natural numbers—the strictly positive integers—and $\mathbb{Z}$ the set of all integers. Given $k \in \mathbb{N}$, define $k = \{1, \ldots, k\}$ (so that $S_k$ is the symmetric group of permutations of $k$). The symbols $k$, $m$, $n$, $p$, $q$ are reserved for indexing. For example, $\sum_{j=1}^{m} \sum_{j=1}^{n} a_{ij} = \sum_{(i,j) \in n \times m} a_{ij}$, otherwise boldface lower case (resp. uppercase) is used to denote vectors (resp. matrices).

We use $I_V$ (or $I$) to denote the identity map of the vector space $V$, and $I_k$ (or $I$) for the identity map of $\mathbb{R}^k$.

Some familiarity with the definitions and results of steady-state equivariant bifurcation theory is assumed as well as generalities on $G$ spaces (isotropy group, fixed point space, etc). We refer to [15] for more details. The books [10, 21–23] provide an introduction to aspects of equivariant bifurcation theory and its applications but the methods and focus are somewhat different from what is required here.

2.2. Representations

Our exposition is directed towards our applications and establishing conventions (see [3, section 3], [4, section 3] for more detailed introductions to the theory and references).

Let $V$ be a finite dimensional real vector space with inner product $(\cdot, \cdot)$ and norm $\|\|$. Let $O(V)$ denote the orthogonal group of $V$. If $\dim(V) = m$, we often identify $V$ with Euclidean space $\mathbb{R}^m$ and $O(V)$ with $O(m)$ (group of orthogonal $m \times m$ matrices).

Given a finite$^4$ group $G$ acting orthogonally on $V$, we refer to $(V, G)$ as an orthogonal representation of $G$ on $V$. Usually, we say $(V, G)$ is a representation of $G$ and assume orthogonality.

$^4$Everything in sections 2 and 3 can be formulated for compact Lie groups—but the prerequisites and technical details are harder; see [15, chapter 10] and section 3.9.
The representation is trivial if each element of \( G \) acts as the identity on \( V \), and is irreducible if there are no proper \( G \)-invariant (linear) subspaces of \( V \). If \((V, G), (W, G)\) are representations of \( G \), a linear map \( A : V \to W \) is a \( G \)-map if \( A(gv) = gA(v) \), for all \( g \in G \). The representations \((V, G), (W, G)\) are isomorphic if there is a \( G \)-map \( A : V \to W \) which is a linear isomorphism (necessarily \( A^{-1} : W \to V \) will be a \( G \)-map). Suppose that \((V, G), (W, G)\) are irreducible. If the representations are not isomorphic, every \( G \)-map \( A : V \to W \) is zero. Indeed, since the orthogonal complement of image(\( A \)) is a \( G \)-invariant subspace of \( W \), and \((V, G)\) is irreducible, \( A \) is either onto or the zero map. If \( A \) is onto, apply the same argument to kernel(\( A \)) \( \subset V \) to deduce that \( A \) is 1:1 and so a linear isomorphism. If \( V = W \) the space \( E_G(V) \) of self-\( G \)-maps of \( V \) is a real division algebra containing \( \{ aI_V \mid a \in \mathbb{R} \} \). By a famous theorem of Frobenius, \( E_G(V) \) is isomorphic to either \( \mathbb{R}, \mathbb{C} \) or \( \mathbb{H} \) (the quaternions). Our focus will always be on the case when \( E_G(V) \approx \mathbb{R} \).

**Remark 2.1.** A \( G \)-map is \( G \)-equivariant but we prefer the term \( G \)-map when dealing with representations and linear maps.

**Definition 2.2.** An irreducible representation \((V, G)\) is of real type or absolutely irreducible if all \( G \)-maps of \( V \) are real multiples of \( I_V \).

**Remark 2.3.**

(a) In representation theory, the term real representation is most commonly used but may be confusing here since all vector spaces are real. We follow the conventions in the bifurcation literature and use only the term absolutely irreducible.

(b) To avoid uninteresting special cases, an absolutely irreducible representation is always assumed non-trivial.

**Example 2.4 (Representations of \( S_k \)).** Every nontrivial irreducible representation of \( S_k \) is absolutely irreducible \([20, 25]\). In particular, the standard representation of \( S_k \), \( k \geq 2 \), on the hyperplane \( H_{k-1} = \{ x \in \mathbb{R}^k \mid \sum_{i \in k} x_i = 0 \} \subset \mathbb{R}^k \) is absolutely irreducible. Here \( S_k \) acts by permuting coordinates: \( \sigma(x_1, \cdots, x_k) = (x_{\sigma^{-1}(1)}, \cdots, x_{\sigma^{-1}(k)}) \), \( \sigma \in S_k \). Let \( s_k \) denote the isomorphism class of \((H_{k-1}, S_k)\) and \( t \) denote the isomorphism class of the trivial representation \((\mathbb{R}, S_k)\), omitting the subscript \( k \). The isomorphism class of \((\mathbb{R}^k, S_k)\) is \( s_k + t \), since \( \mathbb{R}^k \) is the orthogonal direct sum of \( H_{k-1} \) and \( H_{k-1}^\perp \), on which \( S_k \) acts trivially. We sometimes abuse notation and identify \( s_k \) (resp. \( t \)) with the standard (resp. trivial) representation of \( S_k \).

### 2.3. Spaces of invariant and equivariant polynomial maps

If \((V, G)\) is a representation of the finite (or compact Lie) group \( G \), let \( P_G(V) \) denote the \( \mathbb{R} \)-algebra of invariant polynomial maps \( p : V \to \mathbb{R} \) and \( P_G(V, V) \) denote the \( P_G(V) \)-module of \( G \)-equivariant polynomial maps \( P : V \to V \). Recall that \( P_G(V) \) is finitely generated—the Hilbert–Weyl theorem—and that \( P_G(V, V) \) is finitely generated as a \( P_G(V) \)-module. In either case, a minimal set of homogeneous polynomials can be chosen for the generating set. It follows from Schwarz’ theorem on smooth invariants that these results extend to \( C^\infty \) (smooth) invariants and equivariants with the same sets of polynomial generators (see \([15, 6.6]\) for more details and references).

**Example 2.5 (The symmetric polynomials).** Take the standard representation \((H_{k-1}, S_k)\) of \( S_k \). By classical invariant theory, a set of generators for \( P_{S_k}(H_{k-1}) \) is given by the symmetric polynomials \( \rho_j(x) = \frac{1}{j!} \sum_{i \in k} x_i^j \), \( 2 \leq j \leq k \), \( x \in H_{k-1} \subset \mathbb{R}^k \) (\( \rho_1 | H_{k-1} \equiv 0 \) by irreducibility). Every \( p \in P_{S_k}(H_{k-1}) \) can be written uniquely as a polynomial in \( \rho_2, \cdots, \rho_k \) and so \( \{ \rho_2, \cdots, \rho_k \} \) is a basis for \( P_{S_k}(H_{k-1}) \) (the basis result holds for all finite reflection groups \([16]\)). A minimal
generating set for the $S_k$ equivariants is given by $\text{grad}(\rho_2), \cdots, \text{grad}(\rho_k)$; this is also a basis. If $x = (x_1, \cdots, x_k) \in H_{k-1}$, then
\[
\text{grad}(\rho_j)(x) = (x_j^{j-1}, \cdots, x_k^{j-1}) - \left( \frac{1}{k} \sum_{i=k}^{j-1} x_i^{j-1} \right) (1, \cdots, 1).
\]
That is, we take the gradient of $\rho_j : \mathbb{R}^k \to \mathbb{R}$, restrict to $H_{k-1}$, and project orthogonally onto $H_{k-1}$.

2.4. Bifurcation and families of equivariant vector fields

We often omit the prefix ‘$G$’ from $G$-equivariant (or $G$-invariant) maps if no ambiguity results.

Let $(V, G)$ be absolutely irreducible. A family (strictly, one-parameter family) of equivariant vector fields on $V$ is a smooth ($C^\infty$) equivariant map $f : V \times \mathbb{R} \to V$ (the $G$-action on $V \times \mathbb{R}$ is the product of the action on $V$ with the trivial action on $\mathbb{R}$). For $\lambda \in \mathbb{R}$, define the equivariant vector field $f_\lambda : V \to V$ by $f_\lambda(v) = f(v, \lambda)$, $v \in V$. We denote the $V$-derivative of $f_\lambda$ at $v \in V$ by $Df_\lambda : V \to V$ and the derivative of $f$ at $(v, \lambda)$ by $Df(v, \lambda) : V \times \mathbb{R} \to V$. Both $Df_\lambda : V \to L(V, V)$ and $Df : V \times \mathbb{R} \to L(V \times \mathbb{R}, V)$ are $G$-equivariant. For example, $Df_{\lambda, \sigma}g = gDf_{\lambda, \sigma}$, $g \in G$, $v \in V$ (see lemma 3.12).

**Remark 2.6** Maps and families are assumed $C^\infty$. Differentiability requirements can be relaxed though this can be non-trivial [14]. For our main application to the standard representation of $S_k$, $C^1$ suffices.

By equivariance, $f_\lambda(0) = 0$ for all $\lambda \in \mathbb{R}$ (cf remark 2.3(b)). We refer to the curve $\{(0, \lambda) \mid \lambda \in \mathbb{R}\} \subset V \times \mathbb{R}$ as the trivial solution of $f = 0$. Since $(V, G)$ is absolutely irreducible, $Df_{\lambda, 0} = \sigma(\lambda)I_V$, where $\sigma : \mathbb{R} \to \mathbb{R}$ is $C^\infty$. The equilibrium $0$ of $f_\lambda$ is hyperbolic iff $\sigma(\lambda) \neq 0$. If $\sigma(\lambda) > 0$ (resp. $< 0$), $0$ is a source (resp. sink) of $f_\lambda$. If $\sigma(\lambda) \neq 0$, it follows from the implicit function theorem that there is an open neighbourhood $U$ of $(0, \lambda) \in V \times \mathbb{R}$ such that the only solutions to $f = 0$ in $U$ are trivial solutions.

We are interested in bifurcation of the trivial solution. Roughly speaking, this means the existence of a continuous curve $(x(t), \lambda(t))$, defined for $t \in [0, \delta]$, such that $(x(0), \lambda(0))$ is a trivial solution (so $x(0) = 0$) and, for $t > 0$, $(x(t), \lambda(t))$ a non-trivial solution. It follows by the previous paragraph that a necessary condition for bifurcation is $\sigma(\lambda(0)) = 0$.

Without loss of generality, assume that $\sigma(0) = 0$, and refer to the point $(0, 0) \in V \times \mathbb{R}$ as a bifurcation point (since $Df_{0, 0}$ is singular). We make the generic assumption on $f$ that $\sigma'(0) \neq 0$. After a smooth reparametrization, we may assume $\sigma(\lambda) = \lambda$ for $\lambda$ near zero. Since our interest is in bifurcation at $\lambda = 0$, it is no loss of generality to assume
\[
f(v, \lambda) = \lambda v + F(v, \lambda), \quad \text{for all } (v, \lambda) \in V \times \mathbb{R},
\]
where $F$ is equivariant, $C^\infty$, and $DF_{0, 0} = 0$, $\lambda \in \mathbb{R}$. Thus bifurcation of the trivial solution can only occur at $\lambda = 0$.

Let $V(V, G)$, $V$ denote the space of all families $f$ satisfying (2.3). Families $f, f' \in V$ are $C'$-close on a compact $K \subset V \times \mathbb{R}$ if the derivatives of $f$ and $g$ of order at most $r$ are close on $K$. If we define the semi-norm $\|f\|_{K,r}$ on $V$ by
\[
\|f\|_{K,r} = \max_{0 \leq t \leq 1} \sup_{v \in K} \|D^rf(v, \lambda)\|, \quad f \in V
\]
then the set of semi-norms $\|\|_{K,r}$, where $K$ runs over all compact subsets of $V \times \mathbb{R}$, defines the (weak) $C'$-topology on $V$. Write $V'(V, G) = V'$ for $V$ equipped with the $C'$ topology.
(1 ≤ r ≤ ∞). Since the results we need are local, the Whitney $C^r$-topology is not required. Later, we use the semi-norm $\|f\|_{C^r} = \max_{0 \leq s \leq r} \sup_{(x,\lambda) \in K} \|D^s f(x,\lambda)\|$ which uses no $\lambda$-derivatives.

2.5. Branches of solutions

We need to review the core notions of solution branch and branching pattern. A brief overview may be found in [17]; more detail is in the original papers [18, 19] (details for when $G$ is a compact Lie group, and equilibria are replaced by relative equilibria, are in [15, chapter 10]). The problem is to characterise the zero set of $f$, near the bifurcation point $(0,0) \in V \times \mathbb{R}$, for generic equivariant families $f \in V$. More formally, we want a description of the germ of $f^{-1}(0)$ at the bifurcation point $(0,0)$ that is (implicitly) framed in terms of conditions on the partial derivatives of $F(x,0)$, where $F$ is defined by (2.3).

Definition 2.7 (cf [15, section 4.2]). A solution branch of (2.3) consists of a $C^1$-embedding $\gamma = (x,\lambda) : [0,\delta] \to V \times \mathbb{R}$ satisfying

(a) $\gamma(0) = (0,0)$.
(b) $f(\gamma(s)) = f_{\lambda_0}(x(s)) = 0$, for all $s \in [0,\delta]$.

If we can choose $\delta > 0$ so that

(a) $x \not\equiv 0$ on $(0,\delta]$, the branch is non-trivial.
(b) $Df_{\lambda_0}(x(s))$ is non-singular for $s \in (0,\delta]$, the branch is non-singular.
(c) $\lambda(s) > 0$ (resp. $\lambda(s) < 0$) for $s \in (0,\delta]$, the branch is forward (resp. backward).
(d) $x(s)$ is a hyperbolic zero for $s \in (0,\delta]$, the branch is hyperbolic (necessarily non-singular).

Recall that the index of a hyperbolic equilibrium $x$ of $X$ is the number of eigenvalues of $DX_x$ with strictly negative real part (counting multiplicities) and is denoted by $\text{index}(X, x)$.

The family (2.3) has two trivial solution branches $\tau_{\pm}$ defined by

$$\tau_{\pm}(s) = (0, \pm s), \quad (s \in \mathbb{R}),$$

and $\tau_+$ (resp. $\tau_-$) is a hyperbolic forward (resp. backward) branch of index zero (resp. dim($V$)).

Solution branches $\gamma, \rho$ are equivalent if (roughly) the germs of the images of $\gamma$ and $\rho$ at $(0,0)$ are equal. More precisely, if there is a $C^1$ diffeomorphism $\alpha : [0,\varepsilon_1] \to [0,\varepsilon_2]$, mapping $0$ to $0$, such that $\gamma \circ \alpha = \rho$ on $[0,\varepsilon_1]$. We denote the equivalence class of $\gamma$ by $[\gamma]$ and let $\Sigma(f)$ (resp. $\Sigma^*(f)$) denote the set of all equivalence classes of solution (resp. non-trivial solution) branches for $f$. Clearly, $\Sigma(f)$ and $\Sigma^*(f)$ are $G$-sets and $[\tau_{\pm}] \in \Sigma(f)$ are the fixed points of the $G$-action on $\Sigma(f)$.

Lemma 2.8. Let $\gamma = (x,\lambda) : [0,\delta] \to V \times \mathbb{R}$ be a non-trivial solution branch for $f \in V$.

(a) The direction of branching $d(\gamma) = x'(0)/\|x'(0)\| \in V$ is well-defined, non-zero and independent of the parametrisation.
(b) If $\gamma$ is a non-singular branch, $\gamma$ is either forward or backward.
(c) If $\gamma$ is hyperbolic, then $\text{index}(f_{\alpha(0)}, x(s))$ is constant on $(0,\delta]$.

Proof. See [15, section 4.2], [18] for the elementary proof for (a), note that if $x'(0) = 0$, then $\|x(s)\| = o(s)$ and so, using (2.3), $|\lambda(s)| = o(s)$, contradicting the $C^1$-embedding requirement on $\gamma$. □

Definition 2.9. Let $f \in V$ and suppose that $\Sigma(f)$ is finite and consists of hyperbolic solution branches. The signed indexed branching pattern of $f$ is the triple $(\Sigma^*(f), \text{sgn}, \text{index})$ where
(a) \( \text{sgn} : \Sigma^*(f) \to \{-1, +1\} \) and \( \text{sgn}(\gamma) = +1 \) (resp. \(-1\)) if \( \gamma \) is a forward (resp. backward) solution branch (sgn is the sign function).

(b) \( \text{index} : \Sigma^*(f) \to \{0, \ldots, \dim(V)\} \) and \( \text{index}(\gamma) \) is the index of \( Df_{\gamma(t)}, s \neq 0 \).

**Remark 2.10.** The sign and index functions are \( G \)-invariant.

**Definition 2.11.** If \((\Sigma^*(f_i), \text{sgn}_i, \text{index}_i), i \in 2\), are signed indexed branching patterns, they are isomorphic if there is a \( G \)-equivariant bijection \( \beta : \Sigma^*(f_1) \to \Sigma^*(f_2) \) such that \( \text{sgn}_1 = \text{sgn}_2 \circ \beta \) and \( \text{index}_1 = \text{index}_2 \circ \beta \).

**Remark 2.12.** Since the general theory develops from definitions 2.7 and 2.9, it is essential to prove that generic bifurcation can be expressed in terms of solution branches and branching patterns. In particular, solution branches are defined in terms of \( C^1 \)-embeddings (not \( C^0 \) or \( C^\infty \)), and a branching pattern is a finite union of solution branches. The proof requires ideas from the geometry and stratification of semialgebraic sets and equivariant transversality.

### 2.6. Stable and weakly stable families

**Definition 2.13.** A family \( f \in V \) is **stable** if

(a) \( \Sigma(f) \) is finite and consists of hyperbolic solution branches (necessarily, either forward or backward).

(b) For some \( r \geq 1 \), there is a neighbourhood \( U \) of \( f \in V^r \) such that if \((f_t)_{t \in [0,1]} \) is a continuous curve in \( U \) with \( f_0 = f \), then

1. There exists \( \delta > 0 \) such that for all \( [\gamma] \in \Sigma(f) \), there is a continuous family \((\gamma_t)_{t \in [0,1]} \) of \( C^1 \)-maps \([0, \delta] \to V \times \mathbb{R} \) such that each \( \gamma_t \) is a branch of hyperbolic zeros of \( f_t \) and \([\gamma_0] = [\gamma] \).

2. \( \Sigma^*(f) \) and \( \Sigma^*(f_t) \) are isomorphic for all \( t \in [0,1] \).

Denote the set of stable families by \( S(V, G) = \mathcal{S} \).

**Remark 2.14.**

(a) It follows from 2(b) of the definition that \( \Sigma(h) \) is isomorphic as a \( G \)-set to \( \Sigma(f) \) for all \( h \) in the path connected component of \( U \) containing \( f \). Similarly, using 2(a), the signed index branching patterns for \( \Sigma^*(h) \) and \( \Sigma^*(f) \) are isomorphic.

(b) If \( f \) is stable and \([\gamma], [\eta] \in \Sigma^*(f), [\gamma] \neq [\eta] \), then we cannot exclude the possibility that \( d([\gamma]) = d([\eta]) \) but this does not happen if the stability is determined by quadratic or cubic terms (for example, if quadratic terms are of relatively hyperbolic type [15, section 4.6.4]).

We also need the concept of weak stability [15, section 4.2.1].

**Definition 2.15.** A family \( f \in V \) is **weakly stable** if

(a) \( \Sigma(f) \) is finite.

(b) For some \( r \geq 1 \), there is a \( U \) of \( f \in V^r \) such that if \((f_t)_{t \in [0,1]} \) is a continuous curve in \( U \) with \( f_0 = f \), then

1. There exists \( \delta > 0 \) such that for every \([\gamma] \in \Sigma(f)\), there is a continuous family \((\gamma_t)_{t \in [0,1]} \) of \( C^1 \)-maps \([0, \delta] \to V \times \mathbb{R} \) such that each \( \gamma_t \) is a solution branch of \( f_t \) and \([\gamma_0] = [\gamma] \).
2. \( \Sigma(f) \) and \( \Sigma(f_t) \) are isomorphic as \( G \)-sets, \( t \in [0, 1] \).

Denote the set of weakly stable families by \( K(V, G) \).

**Example 2.16.** Let \( S_2 \approx \mathbb{Z}_2 \) act as multiplication by \( \pm 1 \) on \( \mathbb{R} \) (isomorphic to the standard representation of \( S_2 \)) and \( g : \mathbb{R} \to \mathbb{R} \) be any smooth map. Define the smooth \( S_2 \)-equivariant family \( f_t(x) = \lambda x + ax^3 + bx^5g(x^2) \) on \( \mathbb{R} \), where \( a, b \) are real constants. If \( a \neq 0 \), the family is stable and models a non-degenerate pitchfork bifurcation. If \( a = 0 \), the family is weakly stable but not stable. Weak stability is obvious, indeed we can take \( \delta = 2.7 \). The stability theorem

Nonlinearity 35 (2022) 2809

Y Arjevani and M Field

2. The stability theorem

Given \( d \in \mathbb{N} \), let \( P^d_G(V, V) \) denote the space of \( G \)-equivariant homogeneous polynomials from \( V \) to \( V \) of degree \( d \). For \( d > 1 \), define \( \tilde{P}^d_G(V, V) = \bigoplus_{i=0}^d P^i_G(V, V) \)—equivariant polynomial maps of degree at most \( d \) with zero linear part. If \( f \in V \), let \( J^d(f) = j^d f(0) \in \tilde{P}^d_G(V, V) \) denote the \( d \)-jet of \( f_0 \) at 0 (Taylor polynomial of degree \( d \) of \( f_0 \) at 0).

**Theorem 2.17** ([13]). (Assumptions and notation as above.) There exist a minimal \( \delta = \delta(V, G) \in \mathbb{N} \) and an open and dense semi-algebraic subset \( \mathcal{P}^\delta = \mathcal{P} \) of \( \tilde{P}^1_G(V, V) \) such that if \( f \in \mathcal{P} \) and \( J^d(f) \in \mathcal{P}^\delta \), then \( f \in \mathcal{S} \). In particular, \( \mathcal{S} \) contains a \( C^\infty \)-open and dense subset of \( \mathcal{V} \).

Similar results hold for weak stability with a value \( \delta_w \leq \delta(V, G) \).

**Proof.** The result uses equivariant jet transversality [8]. Details may be found in [15, chapter 7] or [13]; some brief notes are at the end of this section. In section 3.4, there is an outline proof for weak stability. This result addresses the points raised in remark 2.12 and does not use equivariant jet transversality.

**Remark 2.18.**

(a) The result extends to absolutely irreducible representations of compact Lie groups [15, section 7.6] and irreducible representations of complex type [15, chapter 10]—see also section 3.9.

(b) There is an upper bound for \( \delta(V, G) \). If \( \{p_1 = \|\cdot\|^2, \ldots, p_m\} \) is a minimal set of homogeneous generators for the \( \mathbb{R} \)-algebra \( \mathcal{P}^1_G(V) \) and \( \{F_1 = I_V, \ldots, F_m\} \) is a minimal set of homogeneous generators for the \( \mathcal{P}^1_G(V) \)-module \( \mathcal{P}^1_G(V, V) \), then \( \delta(V, G) \leq \max_i \deg(p_i) + \max_j \deg(F_j) \) (see section 2.3 for notational conventions and definitions). Often \( \delta \) can be chosen much smaller. If \( (V, G) = g_k \), then \( \max_i \deg(p_i) + \max_j \deg(F_j) = 2k - 1 \), \( k \geq 2 \), but we can take \( \delta = 2 \), if \( k \) is odd, and \( \delta = 3 \) if \( k \) is even. For weak stability, the corresponding minimal \( \delta_w \) satisfies \( \delta_w \leq \max_j \deg(F_j) \) in general. For \( (V, G) = g_k \), \( \delta_w = 2 \), if \( k \geq 3 \). If \( k = 2 \), \( \delta_w = 1 \) (cf example 2.16).

(c) Theorem 2.17 does not imply \( f \) is stable only if \( J^d(f) \in \mathcal{P} \). The resolution of this point is subtle as it depends on the specific stratification used in equivariant transversality (see section 3.4).

(d) Increasing \( \delta \) will not change the space of stable maps given by the theorem. See the section 2.8 for this point.
We have a very useful corollary of the stability theorem.

**Corollary 2.19.** Let \( f \in \mathcal{V} \) and suppose \( f(\mathbf{v}, \lambda) = \lambda \mathbf{v} + F(\mathbf{v}, \lambda) \). Define \( \hat{f}(\mathbf{v}, \lambda) = \lambda \mathbf{v} + F'(\mathbf{v}, 0) \). Then \( J^j(f) \in \mathcal{P} \) iff \( J^j(\hat{f}) \in \mathcal{P} \). If either condition holds, both families are stable and have isomorphic signed indexed branching patterns.

**Remark 2.20.** The corollary allows us to work with polynomial families and use methods based on the CSL [42]. In this way we obtain real analytic parametrisations of solution branches for stable polynomial families \( \lambda \mathbf{x} + P(\mathbf{x}), P \in \mathcal{P} \). Results obtained using this approach may extend to stable smooth families and allow for the sharp analytic estimates; for example on eigenvalues (see section 3).

Define the subspace \( \mathcal{V}_0(V, G) = \mathcal{V}_0 \) of \( \mathcal{V} \) by requiring that \( f \in \mathcal{V}_0 \) iff \( f(\mathbf{v}, \lambda) = \lambda \mathbf{v} + F(\mathbf{v}) \). Give \( \mathcal{V}_0 \) the \( C^\infty \)-topology defined by the semi-norms \( k_\delta \) (section 2.4). Let \( \mathcal{S}_0(V, G) = \mathcal{S}_0 = \mathcal{S} \cap \mathcal{V}_0 \) denote the set of stable families in \( \mathcal{V}_0 \).

### 2.8. Determinacy

Following the statement of theorem 2.17, equivariant bifurcation problems on \( (V, G) \) are \( \delta = \delta(V, G) \)-determined. This notion of determinacy is quite different from that used in [22].

We conclude with brief details about the constructions used to prove the stability theorem and determinacy, avoiding the technicalities of equivariant jet transversality.

Let \( p_1, \ldots, p_\ell \) be a minimal set of homogeneous generators for the \( \mathbb{R} \)-algebra \( P_G(V) \) and \( F_1 = I_V, F_2, \ldots, F_m \) be a minimal set of homogeneous generators for the \( P_G(V) \)-module \( P_G(V, V) \) (section 2.3). Let \( f \in \mathcal{V} \). By Schwarz’ theorem on smooth invariants,

\[
f(\mathbf{v}, \lambda) = \lambda \mathbf{v} + \sum_{i \in \mathbf{m}} g_i(p_1(\mathbf{v}), \ldots, p_\ell(\mathbf{v}), \lambda) F_i(\mathbf{v}),
\]

where the \( g_i : \mathbb{R}^\ell \times \mathbb{R} \to \mathbb{R} \) are \( C^\infty \) functions and \( g_1(0, \lambda) \equiv 0 \) since \( F_1 = I_V \). Setting \( t_j = g_j(0, 0), j \geq 2, \{ t_j | j \geq 2 \} \) is uniquely determined by our choice of generating set \( \{ F_i | i \in \mathbf{m} \} \) [15, section 6.6.2].

**Remark 2.21.** The family \( f \) is weakly stable provided that \( (t_2, \ldots, t_m) \) avoids a codimension 1 semi-algebraic subset of \( \mathbb{R}^{m-1} \) (the branches may not be hyperbolic but have a direction of branching (lemma 2.8(b)) and deform continuously under perturbation of \( f \)). This result [15, theorem 7.1.1] plays an important role in our applications where there are non-trivial quadratic equivarianats.

Given the coefficient functions \( g_i \), we may write \( f(\mathbf{v}, \lambda) \) uniquely as

\[
f(\mathbf{v}, \lambda) = \lambda \mathbf{v} + \sum_{j=2}^m t_j F_j(\mathbf{v}) + \sum_{i, j \in \ell, i \neq m} t_{ij} p_i(\mathbf{v}) F_j(\mathbf{v}) + H(\mathbf{v}, \lambda),
\]

where \( t_j, t_{ij} \) are smooth functions of \( \lambda \) and \( H(\mathbf{v}, \lambda) \) consists of higher order terms in the invariants. The conditions for stability that come from equivariant jet transversality depend only on the \( m - 1 + m \ell \) real numbers \( t_j(0), t_{ij}(0) \). Viewed in this way, once we have found the minimum \( \delta(V, G) \) (that depends on which of the \( t_j(0), t_{ij}(0) \) do not affect the stability of the family), \( \mathcal{P}^d \) is determined for all \( d \geq \delta(V, G) \) with \( \mathcal{P}^d \) projecting naturally onto \( \mathcal{P}^{d-1} \) for \( d > \delta(V, G) \).
3. Quadratic equivariants

Definition 3.1. If \((V, G)\) is an absolutely irreducible representation, then \((V, G)\) has quadratic equivariants if \(\dim(P_G^G(V, V)) \geq 1\).

Example 3.2.

(a) The standard representation \((H_{k-1}, S_k)\) of \(S_k\) has quadratic equivariants for \(k \geq 3\) and \(P_{S_k}^G(H_{k-1}, H_{k-1})\) has basis \(\text{grad}(\rho_3)\), where \(\rho_3(x) = \sum_{i \leq k} x_i^3\), \(x \in H_{k-1}\) (example 2.5).

(b) Let \(\tilde{M}(k, n)\) (resp. \(M^*(k, n)\)) denote the space of real \(k \times n\)-matrices such that all rows and columns sum to zero (resp. the sum of all matrix entries is zero). Obviously \(\tilde{M}(k, n) \subset M^*(k, n)\), if \(k, n \geq 2\). The external tensor product representation \(S_k \otimes S_n\) of \(S_k \times S_n \subset S_{k,n}\) on \(H_{k-1} \otimes H_{n-1} \cong \tilde{M}(k, n)\) is absolutely irreducible and has quadratic invariants iff \(k, d \geq 3\).

In order to show this, observe that the cubic \(\tilde{M}(k, n)\) is one-dimensional for all \(k, d \geq 3\) with basis given by \(C(x) = \sum_{i,j} x_i^3\). For example, if \(k = d = 3\), \(\dim(\tilde{M}(3, 3)) = 4\) and

\[
C(x_1, x_2, x_3, x_4) = \sum_{1 \leq i < j < k \leq 4} x_i x_j x_k + x_1 x_4 (x_1 + x_4) + x_2 x_3 (x_2 + x_3),
\]

where if \(X = [x_{ij}] \in \tilde{M}(3, 3)\), \(x_1 = x_{11}, x_2 = x_{12}, x_3 = x_{21}, x_4 = x_{22}\), and the remaining entries are determined by the row and column sum zero condition. The general formula is an easy induction.

3.1. Generic branching when there are quadratic equivariants

We refer to section 1.2 for background on the results of Ihrig and Golubitsky [24], and Chossat et al [9], on the instability of branching when there are quadratic equivariants. In this section we reprove theorem 4.2(b) [9], using only the definition of solution branch, as well as prove a stronger version that uses theorem 2.17.

We show that if the quadratic equivariants satisfy ‘property \((G)\)’, then the stable signed indexed branching patterns consist of branches of hyperbolic saddles (no non-trivial branches of sinks or sources). Property \((G)\) always holds if the quadratic equivariants are gradient vector fields; indeed, the ‘\(G\)’ is short for ‘gradient like’.

In what follows, \(S(V)\) will denote the unit sphere of \(V\).

3.2. Property \((G)\)

Given \(Q \in P_G^G(V, V), Q \neq 0\), define \(Z(Q) = \{u \in S(V) \mid Q(u) = 0\}\). Since \(Q \neq 0\), \(Z(Q)\) is a proper closed subset of \(S(V)\). Let \(Q_0\) be the set of \(Q \in P_G^G(V, V)\) such that for all \(u \in Z(Q), DQ_u\) has an eigenvalue with non-zero real part.

Definition 3.3 (cf [24, theorem 4.2(B)]). If the absolutely irreducible representation \((V, G)\) has quadratic equivariants, then \((V, G)\) satisfies Property \((G)\) if \(Q_0\) is an open and dense subset of \(P_G^G(V, V)\).
Example 3.4.
(a) If every quadratic equivariant is a gradient vector field, then \((V, G)\) satisfies Property (G) with \(Q_0 = P_G^{(2)}(V, V) \setminus \{0\} \) [24, remarks 4.3(g)]. This is obvious since if \(Q = \text{grad} (C) \neq 0\), for some \(C \in P_G^{(2)}(V)\), then \(DQ_u\) is a symmetric matrix for all \(u \in \mathcal{S}(V)\) and so, since \(Q \neq 0\), \(DQ_u\) has at least one non-zero real eigenvalue.

(b) If the only \(Q \in P_G^{(2)}(V, V)\) for which \(Z(Q) \neq \emptyset\) is \(Q = 0\), then \((V, G)\) satisfies Property (G).

(c) Absolutely irreducible representations may have quadratic equivariants which are not gradient. For example, the group \(G = \text{Aff}(\mathbb{R}^3)\) of affine linear transformations of the field with five-elements is isomorphic to the subgroup of \(S_5\) generated by \(t = (12345)\) and \(s = (2453)\) [15, section 5.4] and so acts on \(\mathbb{R}^4\) by restriction of the standard action of \(S_5\) on \(\mathbb{R}^4\). Since \(G\) is a doubly transitive subgroup of \(S_5\), the representation \((G, \mathbb{R}^4)\) is absolutely irreducible [15, lemma 4.10.1]. After some work, it may be shown that on \(G\) acts on \(\mathbb{C}^2 \cong \mathbb{R}^4\) by

\[
\begin{align*}
\gamma(z_1, z_2) &= (\omega z_1, \omega^2 z_2), \\
\sigma(z_1, z_2) &= (z_2, z_1),
\end{align*}
\]

where \(\omega = \exp(2\pi i/5)\) [15, section 5.4.1]. We have \(\dim(P_G^{(2)}(V, V)) = 2\) and \(P_G^{(2)}(V, V)\) has \(\mathbb{R}\)-basis, \(Q_1 = (z_1^2, z_2^2), Q_2 = (z_1 z_2, z_1 z_2)\). It is easy to verify that \(Q_1 + \beta Q_2\) is gradient iff \(\beta = 2\alpha\). Note that if \(Q = Q_1 - Q_2\), then the cubic invariant \(\langle Q(z_1, z_2), (z_1, z_2) \rangle\) is identically zero.

Computing we find that if \(u_1, u_2 \neq 0\), then \(DQ_{u_1}\) has two non-zero real and a complex conjugate imaginary pair of eigenvalues. If \(Q = \alpha Q_1, \alpha \in \mathbb{R}, \alpha \neq 0\), then \(Z(Q) = (\{(z_1, z_2) | z_1 z_2 = 0\})\). Direct computation verifies that \(DQ_{(z_1, z_2)}\) has the pair \(\pm \alpha ||(z_1, z_2)||\) of non-zero real eigenvalues. If we set \(Q^* = Q_1 \pm Q_2\), then \(Z(Q) \neq \emptyset\) iff \(Q \in \mathbb{R}Q^* \cup \mathbb{R}Q_2\). Direct computation verifies that if \(Q \in \mathbb{R}Q^*, Q \neq 0\), then \(DQ_u\) has eigenvalues with non-zero real part for \(u \in Z(Q)\). If \(Q = Q^*\), then \(u \in Z(Q)\) iff \(u\) lies on an axis of symmetry (the group orbit of \(\mathbb{R}(1, 0, 1, 0)\)); if \(Q = Q^*\), then \(u \in Z(Q)\) iff \(u\) lies on the group orbit of \(\mathbb{R}(1, 0, 0, -1)\). Hence \((\mathbb{R}^4, \text{Aff}(\mathbb{R}^3))\) satisfies Property (G) and \(Q_0 = P_G^{(2)}(V, V) \setminus \{0\}\).

Remark 3.5. In the last example, there is a non-zero \(Q \in P_G^{(2)}(V, V)\) satisfying \(\langle Q(v), v \rangle = 0\), for all \(v \in \mathbb{R} = \mathbb{R}^4\). This non-gradient behaviour suggests there may well exist absolutely irreducible representations \((V, G)\) for which Property (G) fails and the eigenvalues of \(DQ_u\) along \(Ru\) are all either zero or pure imaginary. A natural place to look is the works of Lauterbach and Matthews [34], Lauterbach [33], and Lauterbach and Schwenker [35] on low dimensional families of absolutely irreducible representations with no odd dimensional fixed point spaces. However, these families do not have quadratic equivariants and the question appears open.

3.3. Statement of the main theorems

Theorem 3.6. If \((V, G)\) is an absolutely irreducible representation of the finite group \(G\) satisfying Property (G), then for all \(f \in \mathcal{S}\), index : \(\Sigma^f(f) \to [1, \dim(V) - 1]\). In particular, every non-trivial branch is a branch of hyperbolic saddles and so there are no non-trivial branches of sinks or sources.

Remark 3.7. (a) The result applies to all stable families, not just the open and dense set of stable families given by theorem 2.17.

(b) If \(H \subset G\) and the quadratic invariants vanish identically on the fixed point space \(V^H\), then a backward branch lying in \(V^H\) will not be a branch of maximal index. This is part (a) of
stratification—and define

\[ [\gamma] \in \Sigma^r(f). \]

The direction of branching \( d(\gamma) = \mathbf{u} \in \mathbb{S}(V) \). Setting \( Q = J^2(f) \), \([\gamma] \) will be a branch of hyperbolic saddles if either (a) \( Q(\mathbf{u}) \neq 0 \) or (b) \( Q(\mathbf{u}) = 0 \) and \( DQ_u \) has an eigenvalue with non-zero real part. The failure of Property \((\mathcal{G})\) only concerns solution branches which are tangent to \( \mathbb{R}\mathbf{u}, \mathbf{u} \in \mathbb{Z}(J^2(f)) \).

**Definition 3.8.** Suppose \((V, G)\) has quadratic equivariants. Let \( f \in S \) and set \( \hat{J}(f) = Q \). Suppose \([\gamma] \in \Sigma^r(f)\) has direction of branching \( \mathbf{u} \in \mathbb{S}(V) \). If \( \mathbf{u} \notin \mathbb{Z}(Q) \), \([\gamma] \) is a branch of type \( S \), otherwise \([\gamma] \) is a branch of type \( C \).

### 3.4. Weak stability and equivariant transversality

Assume that \((V, G)\) is an absolutely irreducible representation of the finite group \( G \) (no assumption yet about quadratic equivariants). We review the use of equivariant transversality and Whitney regular stratifications in the proof of weak stability (for full details, see [15, chapters 6 and 7]).

**Equivariant transversality.** Fix a minimal homogeneous basis \( \mathcal{F} = \{F_1 = I_V, F_2, \ldots, F_m\} \) of the \( PG(V)\)-module \( PG(V, V) \). Set \( d_i = \deg(F_i) \) and index the polynomials \( F_i \) so that \( 1 = d_1 < d_2 < \cdots < d_m \).

Define \( \vartheta : V \times \mathbb{R}^m \to V \) by \( \vartheta(\mathbf{x}, \mathbf{t}) = \sum_{i\in\mathbb{N}} t_i F_i(\mathbf{x}) \) and set \( \vartheta^{-1}(0) = \Lambda \):

\[
\Lambda = \left\{ (\mathbf{x}, \mathbf{t}) \in V \times \mathbb{R}^m \left| \sum_{i\in\mathbb{N}} t_i F_i(\mathbf{x}) = 0 \right. \right\} \subset V \times \mathbb{R}^m.
\]

Clearly, \( \mathbb{R}^m, V \subset \Lambda \) (where \( \mathbb{R}^m \overset{\text{def}}{=} \{0\} \times \mathbb{R}^m, V \overset{\text{def}}{=} V \times \{0\} \)). Since \( G \) is finite, \( \dim(\Lambda) = m \) and \( \dim(V) \leq m \) [15, remark 6.9.3].

Let \( f \in \mathcal{V}_0 \) (arguments are similar if \( f \in \mathcal{V} \)). There exist \( C^\infty \) invariant functions \( g_i : V \to \mathbb{R} \) such that

\[
f(\mathbf{x}, \lambda) = \lambda \mathbf{x} + \sum_{i\in\mathbb{N}} g_i(\mathbf{x}) F_i(\mathbf{x}), \quad (\mathbf{x}, \lambda) \in V \times \mathbb{R},
\]

where \( g_i(\mathbf{0}) = 0 \) (see [15, section 6.6] for the details which use the Malgrange division theorem [37]). The maps \( g_1, \ldots, g_m \) are generally not uniquely determined by \( f \) but the values \( g_i(\mathbf{0}), 2 \leq i \leq m \), are uniquely determined [15, lemma 6.6.2] and depend linearly and continuously on \( f \) (cf [39]).

Define the smooth equivariant embedding \( \Gamma_f : V \times \mathbb{R} \to V \times \mathbb{R}^m \) by

\[
\Gamma_f(\mathbf{x}, \lambda) = (\mathbf{x}, (\lambda + g_1(\mathbf{x}), g_2(\mathbf{x}), \ldots, g_m(\mathbf{x}))), \quad (\mathbf{x}, \lambda) \in V \times \mathbb{R}.
\]

The tangent space to \( \Gamma_f(V \times \mathbb{R}) \) at \( \Gamma_f(\mathbf{0}, 0) \in V \times \mathbb{R}^m \) is \( V \times \mathbb{R} \), where \( \mathbf{e} = (1, 0, \ldots, 0) \in \mathbb{R}^m \).

The family \( f \) factorises through \( V \times \mathbb{R}^m \) as \( f = \vartheta \circ \Gamma_f \), and \( \Gamma_f(\mathbf{0}, 0) \in \mathbb{R}^{m-1} \overset{\text{def}}{=} \{(0, \mathbf{t}) \in \mathbb{R}^m : t_1 = 0\} \). Define \( \gamma_f : \mathbb{R} \to \mathbb{R}^m \) by \( \gamma_f(\lambda) = (\lambda, g_2(\mathbf{0}), \ldots, g_m(\mathbf{0})) \). Clearly \( \gamma_f = \Gamma_f|\{0\} \times \mathbb{R} \).

We recall some results on Whitney regular stratifications (see [15, sections 3.9 and 6.8] for basic definitions and results and [47] for a recent review which includes an extensive bibliography). Fix a Whitney regular semialgebraic stratification \( \mathcal{S} \) of \( \Lambda \)—for example, the canonical stratification [38]—and define
\[ \mathcal{K}_{0,0} = \{ f \in \mathcal{V}_0 \mid \Gamma_f \cap \mathcal{S} \text{ at } (0,0) \}. \]

It follows by the isotopy theorem for equivariant transversality that \( \mathcal{K}_{0,0} \) consists of weakly stable families [15, theorem 7.7.1].

It is easily shown [15, theorem 6.10.1] that \( \mathcal{S} \) induces a stratification \( \mathcal{S}^* \) of \( \mathbb{R}^{m-1} \)—the strata of \( \mathcal{S}^* \) consist of the strata of \( \mathcal{S} \) which are subsets of \( \mathbb{R}^{m-1} \) intersected, if necessary\(^5\), with \( \mathbb{R}^{m-1} \). We have \( \Gamma_f \cap \mathcal{S} \text{ at } (0,0) \) iff \( \gamma_f \cap \mathcal{S}^* \text{ at } \lambda = 0 \).

**Remark 3.9** The condition for \( f \) to lie in \( \mathcal{K}_{0,0} \) depends only on the values of \( g_2, \cdots, g_m \) at \( x = 0 \). The same arguments apply if \( f \in \mathcal{V} \) and this allows us to assume the coefficients \( g_i \) do not depend on \( \lambda \).

**C\(^1\) parametrisation of solution branches.** We need to examine the stratifications \( \mathcal{S} \) of \( \Lambda \) and \( \mathcal{S}^* \) of \( \mathbb{R}^{m-1} \). If \( S \in \mathcal{S} \) is a connected stratum of dimension \( p \), then \( \partial S \) will be a union of connected strata of dimension less than or equal to \( p - 1 \) (this uses Whitney regularity). Let \( \mathcal{S}_0 \) be the union of all connected strata \( N \) of \( \mathcal{S} \) which are of dimension \( m \) and for which \( \partial N \) has at least one connected stratum \( M \in \mathcal{S}^* \) of dimension \( m - 1 \) (so defining an open subset of \( \mathbb{R}^{m-1} \)). Let \( \mathcal{S}_0^* \subset \mathcal{S}^* \) denote the set of all connected \( m - 1 \)-dimensional strata which are boundary components of some \( N \in \mathcal{S}_0 \). Let \( M \in \mathcal{S}_0^*, N \in \mathcal{S}_0 \) with \( M \subset \partial N \). If \( f \in \mathcal{K}_{0,0} \) and \( \gamma_f(0) \in M \), then \( \Gamma_f \cap \mathcal{S} \text{ at } (0,0) \) and, by Whitney regularity, \( \Gamma_f \) has a non-trivial transversal intersection with \( N \) at \( (0, \gamma_f(0)) \) and therefore \( \Gamma_f^{-1}(N \cup M) \) contains a one-dimensional Whitney regular stratified set \( C \) with \( (0,0) \in \partial C \). Using Whitney regularity, \( C \) is a \( C^1 \) submanifold of \( V \times \mathbb{R} \) with boundary point \( (0,0) \). Hence \( \Gamma_f^{-1}(N \cup M) \) contains a non-trivial solution branch in \( \Sigma(f) \), with \( C^1 \)-parametrisation as defined in definition 2.7. Alternatively, we may invoke Pawłucki’s theorem [44] which implies that \( N \cup M \) is a \( C^1 \)-submanifold of \( V \times \mathbb{R}^m \) and so the intersection is a one-dimensional \( C^1 \) submanifold by the transversality theorem\(^6\). More generally, by openness of transversality, we may choose a closed neighbourhood \( D \) of \( (0,0) \in V \times \mathbb{R} \) such that \( \Gamma_f/D \cap \mathcal{S} \) and \( (\Gamma_f/D)^{-1}(\Lambda) \) gives the branching pattern \( \Sigma(f) \). In particular, if \( S \in \mathcal{S} \), then \( \Gamma_f(D) \cap S \neq \emptyset \) only if \( S \in \mathcal{S}_0 \cup \mathcal{S}_0^* \).

**Remark 3.10.**

(a) The argument given above proves that solution branches and the finiteness of the branching pattern are generic in equivariant bifurcation theory.

(b) Pawłucki’s theorem applies to Whitney regular stratifications of subanalytic sets—it is not true for general Whitney regular stratifications with smooth strata [44].

### 3.5. Analytic parametrisation of solution branches

The \( C^1 \)-parametrisation of branches given by equivariant transversality is precisely what is needed for the proof of theorem 4.2(b) [9], see lemma 3.16 below. In this section, we give conditions for analytic parametrisation of solution branches for polynomial and analytic families that make use of the CSL [42]. We only give the details for polynomial maps (using [42]). The results extend easily to real analytic families using the CSL for semianalytic sets [36, II, section 3, III, section 8] (see [30, section 9] for historical notes on the CSL and its significance in singularity theory).

---

\(^5\) See [15, remark 7.1.1] if \( \exists \) an open \( X \subset \mathbb{R}^{m-1} \) with \( \Sigma(f) = \emptyset \) when \( \gamma_f(0) \in X \). At this time, no examples where this happens are known.

\(^6\) Thom’s transversality isotopy theorem only gives \( C^0 \)-local trivialisation of \( N \cup M \) since vector fields on \( M \cup N \) are only \( C^0 \).
Assume that \( n \geq \delta_0(V, G) \) and let \( V^n \) (resp. \( V_0^n \)) be the subset of \( V \) (resp. \( V_0 \)) consisting of families \( f \in V \) (resp. \( f \in V_0 \)) such that \( f \) is polynomial in \((x, \lambda)\) of degree at most \( n \). Let \( V^n \) and \( V_0^n \) be the corresponding spaces of real analytic families. Details below are only given for \( V^n \); results and methods are the same for \( V_0^n \) and \( V_0 \). We have

\[
V_0^n = \{ f \in V_0 | f(x, \lambda) = \lambda x + P(x), \quad P \in \hat{P}_G^0(V, V) \}.
\]

It follows from theorem 2.17 that if \( n \geq \delta_0(V, G) \) then \( V^n_0 \) contains an open and dense subset \( S^n_0 = V^n \cap S \) of stable families (standard vector space topology on \( \hat{P}_G^0(V, V) \subset P_G^0(V, V) \)).

If \( f \in V^n_0 \), then \( f(x, \lambda) = \lambda x + \sum_{i \in \mathbb{N}} g_i(x) P_i(x) \), where the coefficient functions \( g_i \) are polynomial invariants, \( g_i(0) = 0 \), and the values \( g_0(0), \ldots, g_n(0) \) are uniquely determined by \( f \). As above, we define the real algebraic proper embedding \( \Gamma_f : V \times \mathbb{R} \to V \times \mathbb{R}_n \) by

\[
\Gamma_f(x, \lambda) = (x, (\lambda + g_1(x), g_2(x), \ldots, g_n(x))), \quad (x, \lambda) \in V \times \mathbb{R}.
\]

Given a Whitney regular semialgebraic stratification \( \mathcal{S} \) of \( \Lambda \), and \( n \geq \delta_0(V, G) \), define the space \( K_{n,0} \subset V^n_0 \) of weakly stable real polynomial families by

\[
K_{n,0} = \{ f \in V_0^n | \Gamma_f \cap \mathcal{S} = \emptyset, \mathcal{S} \text{ at } (0, 0) \}.
\]

Repeating the arguments given in the previous section, it follows that if \( f \in K_{n,0} \), then we can choose a neighbourhood \( D \) of \((0, 0) \in V \times \mathbb{R}\) such that \((\Gamma_f \cap D)^{-1}(\Lambda) \) is a finite union of connected one-dimensional real algebraic subsets of \( V \times \mathbb{R} \) with common boundary \((0, 0) \). Hence, by the CSL, each non-trivial \( [\gamma] = [(x, \lambda)] \in \Sigma(f) \) may be parametrised as a real analytic branch

\[
x(s) = s^q \nu + \sum_{j \geq q} s^j \nu_j, \quad \lambda(s) = as^d, \quad s \geq 0,
\]

where \( q \geq p \geq 1, \nu \in V \) is non-zero and we assume is \( p \) is minimal. If the branch has isotropy \( H \), then \( \nu_j \in V^H \), all \( j \geq p \). If \( a \neq 0 \), we can always reparametrize so that \( a = \pm 1 \) and so, since \( p \) is minimal, the parametrisation of \( x(s) \) is then unique. Note that \( a \neq 0 \) iff the branch is either forward or backwards. Hence, for families in \( S^n_0 \), the parametrisation with minimal exponent is unique with \( a \in \{ \pm 1 \} \).

**Remark 3.11.**

(a) In our setting, the regular arc-wise analytic stratification \( \mathcal{P} \) of Parusiński and Paunescu [43, section 1.2] is a more natural choice of stratification than the canonical stratification since it gives a Whitney regular stratification of \( \Lambda \) into semialgebraic strata which satisfy a strong trivialisation property which fails in general for the canonical stratification. Thus \( \mathcal{P} \) gives regular arc-wise analytic trivialisations of strata pairs of dimension \( m \) and \( m - 1 \) and this implies that initial exponents of real analytic parametrisations are locally constant on \( K^n \) (this follows from propositions 1.6 and 7.4 op. cit.).

(b) If \((V, G)\) has no quadratic equivariants, \( q = 2p \) for stable families. However, if there are quadratic equivariants and \( D^2 f_0(\nu_p) = 0 \), all that can be claimed in general seems to be \( q > p \) (for axial branches, obviously \( p = 1, q \in \{1, 2\} \)). See also section 3.8.

If \( p = 1 \), then \( \gamma \) is a real analytic embedding. If \( p > 1 \), set \( t = s^q \) so that

\[
x(t) = t \nu_p + t \left( \sum_{j \geq 1} t^j \nu_p + j \right), \quad \lambda(t) = at^d, \quad t \geq 0
\]
Lemma 3.12. First, an elementary lemma about the derivative of an equivariant map. To obtain a $C^{1+\frac{1}{2}}$ FPS parametrisation of $\gamma$ which is a $C^1$-embedding. Hence we obtain a solution curve with $x'(0) = v_p \neq 0$. Note that if $q > p$, then $\lambda(0) = 0$ and so the branch is tangent to $V \times \{0\}$ at $t = 0$. As indicated in remark 3.11(c), $p$ may not depend continuously on $f$ unless the stratification satisfies additional conditions going beyond Whitney regularity. If we use the regular arc-wise analytic stratification $\mathcal{Q}$ of $\Sigma$, then $p$ is locally constant and it may be shown (using [43, section 7]) that $\nu_p$, and so the direction of branching $d(\gamma)$, depend continuously on $f \in K^n_{0,0}$. However, nothing is said about the exponent $q$.

3.6. Proof of theorem 3.6: branches of type $S$

Assume that $(V, G)$ has quadratic equivariants. We start with some preliminary results before proving a version of theorem 3.6 that applies to branches $\gamma(s)$ that are not tangent to $V$ at $s = 0$.

First, an elementary lemma about the derivative of an equivariant map.

Lemma 3.12. If $f: V \to V$ is equivariant and $C^1$, then $Df: V \to L(V, V)$ is $G$-equivariant:

$$g^{-1}Df_{gx}g = Df_{x}, \quad g \in G, \quad x \in V.$$ 

Proof. Differentiate $f(gx) = g(f(x))$ using the chain rule to get $Df_{gx}g = gDf_{x}$, for all $x \in V, \ g \in G$.

Given $f \in V_0$, define $Tr(f): V \times \mathbb{R} \to \mathbb{R}$ by

$$Tr(f)(x, \lambda) = \text{trace}(Df_{\lambda x}), \quad (x, \lambda) \in V \times \mathbb{R},$$

where $Df_{\lambda x}$ denotes the derivative of $f_\lambda$ at $x$.

Lemma 3.13. $Tr(f)$ is $G$-invariant.

Proof. The invariance of $Tr(f)$ is immediate from lemma 3.12.

Remark 3.14. A similar result holds for the symmetric polynomials (classical definition) in the eigenvalues of $Df_{\lambda x}$—the traces of $\lambda^r Df_{\lambda x}$, $2 \leq r \leq \dim(V)$.

Lemma 3.15 ([24, lemma 4.4]). (Notation and assumptions as above.) Let $Q \in P^2_G(V, V)$. If $F_\lambda(x) = \lambda x + Q(x)$, then

(a) $Tr(F)(x, \lambda) = \dim(V) \lambda$, $(x, \lambda) \in V \times \mathbb{R}$.
(b) $Tr(Q) \equiv 0$.

Proof. Since $DQ_x$ is linear in $x$ and there are no non-zero linear invariants, $Tr(DF_\lambda)$ is independent of $x$ and (a) follows since $DF_\lambda(0) = \lambda I_V$. Hence $Tr(Q) \equiv 0$, proving (b).

Lemma 3.16 (Proof of theorem 2.2(b) [9]). If $f \in K(V, G)$ and $\gamma(s) = (x(s), \lambda(s))$ is a type $S$ solution branch of $x' = f_\lambda(x)$, then $\gamma$ is unstable.

Proof. Since $f \in K(V, G)$, we may require that $\gamma$ is a $C^1$-embedding and $\gamma(s) = (sv + o(s), \pm s)$, where $v/\|v\| = d(\gamma)$. Set $F(f) = Q$. Without loss of generality, suppose the branch is forward: $\lambda(s) = s$. We have $f_\lambda(x) = \lambda x + Q(x) + O(|\lambda||x|^2) + O(||x||^3)$ (it is not assumed that $f \in V_0$). Substituting for $(x, \lambda)$, dividing by $s^2$ and setting $s = 0$ gives $Q(v) = -v$. By Euler’s theorem, $DQ_{x(s)}(x(s)) = 2Q(x(s))$ and so, substituting for $x(s)$, we see that $DQ_x$ has eigenvalue $\mu_x(s) = -2s$ and so $D(sI_V + Q)_x$ has the eigenvalue $-s$. Hence, by lemma 3.15(a), $D(sI_V + Q)_x$ has an eigenvalue with strictly positive real part $\alpha(s) = ax, \ a > 0$. The terms we have omitted from $Df_{\lambda x \theta}$ are all $o(s)$ and so, by the continuous dependence of eigenvalues on the coefficients of the characteristic equation, there exists $\varepsilon > 0$ such that for $s \in (0, \varepsilon)$, there is an eigenvalue with strictly positive real part. Hence the branch is unstable.
Remark 3.17. The proof is similar to that in [9] except no use is made of the CSL which does not apply here unless it is assumed that \( f \) is polynomial (or real analytic) in \((x, \lambda)\) [36] and use is made of theorem 2.17 (to extend the result to families in \( K(V, G) \cap S(V, G) = S(V, G) \)). Modulo the use of equivariant transversality in obtaining solution branches, the proof is simple but gives no quantitative information on the interval \((0, \varepsilon]\) for which there are eigenvalues of opposite sign. As is shown below, it is possible to obtain estimates on eigenvalues and \( \varepsilon \) using the CSL if we assume families are analytic or polynomial.

Suppose that \( f \in S^\omega_\ell \) (similar results hold for families in \( S^\omega \), including polynomial families). Let \( \gamma = (x, \lambda) : [0, \delta) \to V \times \mathbb{R} \) be a non-trivial branch of solutions for \( f = 0 \). By the CSL, we may write

\[
x(t) = \sum_{j=p}^{\infty} v_j t^j
\]

(3.4)

\[
\lambda(t) = \text{sgn}([\gamma]) t^p,
\]

(3.5)

where \( v_j \in V, j \geq p, v_p \neq 0, p, q > 0 \) and \( p \) is minimal. The power series for \( x \) is unique granted the minimality of \( p \) and the expression for \( \lambda(t) \). In what follows, we often assume the branch is forward, so that \( \lambda(t) = t^p \) (the arguments we give apply equally to the case \(-t^p\)).

Proposition 3.18 (Notation and assumptions as above). Let \( f \in S^\omega_\ell \) and \( \gamma = (x, \lambda) \) be a solution branch with unique analytic parametrisation (3.4) and (3.5). Suppose that

\[
J^2(f) = Q \in P^2_G(V, V), \quad \text{and} \quad Q(v_p) \neq 0.
\]

(3.6)

([\gamma] \in \Sigma^\ast(f) is of type S). Then \( q = p \) and \( \gamma \) is a branch of hyperbolic saddles: index \((\gamma) \in [1, \dim(V) - 1]\). If (3.6) holds for all \([\gamma] \in \Sigma^\ast(f)\), then every non-trivial branch \( \gamma \) of solutions of \( f \) is a branch of hyperbolic saddles with index \((\gamma) \in [1, \dim(V) - 1]\).

In particular, if \( \mu \) is an eigenvalue of \( DQ_{v_p} \), then \( Df_{\lambda(0,x(0))} \) will have an eigenvalue \( \tilde{\mu}(t) \) where

(a) If \( \mu \neq 2 \), or \( \mu = 2 \) and the associated generalised eigenspace is not \( \mathbb{R} v_p \), then

\[
\tilde{\mu}(t) = t^p \left[ \text{sgn}([\gamma]) + \mu + O \left(t^1\right) \right].
\]

If \( \mu \neq 2 \) has multiplicity \( m \), then \( Df_{\lambda(0,x(0))} \) will have \( m \) eigenvalues of this type (counting multiplicities).

(b) If \( \mu = 2 \), with associated eigenvector \( v_p \), then

\[
\tilde{\mu}(t) = t^p \left[ -\text{sgn}([\gamma]) + O \left(t^1\right) \right].
\]

In either case, we may take \( \ell = 1 \) if \( f \) is a family of gradient vector fields or if \( \mu \) is a simple eigenvalue.

The result continues to hold if \( f \in S^\omega \) or \( S^n \), \( n \geq 3 \).

Remark 3.19. The main step in the proof is to show that \( Df_{\lambda(0,x(0))} \) has an eigenvalue \( \mu(t) = -\lambda(t)(1 + O(t^1)) \), and associated eigenvector \( e(t) = v_p + O(t^1) \), for some \( \ell \in \mathbb{N} \). The estimate holds with \( \ell = 1 \) if the eigenvalue \(-1 \) of \( I_V + DQ_{v_p} \) is simple (using an implicit function theorem argument, this only requires that \( f \) is \( C^\infty \)) or if \( f \) is an analytic family of gradient vector fields (and so \( Df_\lambda \) is symmetric). This is a well-known result in perturbation theory [45], [26, chapter 2]. In either case, eigenvalues and eigenvectors depend analytically on \( t \).
In general, we have to allow for the eigenvalue \(-1\) of \(I + DQ_{\nu P}f\) to be multiple and/or for \(DF_x\) to be asymmetric. Here we rely on Puiseaux’s theorem to obtain FPS expansions for the eigenvalues (and eigenvectors) of \(DF_{\lambda(\lambda,0)}\), viewed as a perturbation of \(t^n(I + DQ_{\nu P}f)\). That is, the characteristic equation of \(t^nDF_{\lambda(\lambda,0)}\) is a polynomial with coefficients depending analytically on \(t\) and Puiseaux’s theorem is used to parametrise the eigenvalues ([27, chapter 5] or [50]). We can assume analyticity of coefficients since \(f \in S^k\) and the CSL gives an analytic parametrisation of the solution branch. If we only assume \(f \in S\), then the terms \(O(t^k)\) in the estimates are replaced by \(o(t)\) (as in the proof of lemma 3.16—sometimes this can be improved by approximation of \(f\) by a Taylor polynomial in \(S_0\)). We allow for more than one solution branch with the same direction of branching \(\nu_p/\|\nu_p\|\): analyticity implies these branches will be distinct branches of equilibria for sufficiently small non-zero values of the parameter. Similarly, if \(2\) is not a simple eigenvalue of \(DQ_{\nu P}f\), the eigenvector \(\nu_p\) might split into several eigenvectors when we add in the higher order terms. However, these eigenvectors will be close to \(\nu_p\) and the associated sum of (generalised) eigenspaces will be close to the original generalised eigenspace of \(I + DQ_{\nu P}f\). The exponent \(1/\ell\) may be small.

**Proof of proposition 3.18.** Substituting \(x(t) = \sum_{p=1}^{\infty} x_p t^p\), \(\lambda(t) = t^q\) in \(f(x, \lambda) = 0\) and equating lowest powers of \(t\) we find that \(q = p\) and \(Q(\nu_p) = -\nu_p\). By Euler’s theorem, 
\[
DQ_{\nu P}f(x(t)) = 2Q(x(t)).
\]
Substituting for \(x(t)\) we find that \(DQ_{\nu P}f\) has the eigenvalue \(\tilde{\mu}_p(t) = -2t^p = -2\lambda(t)\) and associated eigenvector \(e_p = \nu_p\). If we write \(f(x) = \lambda x + Q(x) + H(x)\), where \(H(x) = O(||x||^3)\), then \(DF_{\lambda(\lambda,0)} = \lambda I + DQ_{\nu P}f + t^pA(t)\), where \(A(t)\) is an analytic family of linear maps. Since \(DQ_{\nu P}f = DQ_{\nu P}f + O(t^{p+1})\), it follows by perturbation theory (see the discussion above), that \(DF_{\lambda(\lambda,0)}\) has eigenvalue \(\mu_p(t) = -t^p + O(t^{p+1/2}) = -\lambda(t)(1 + O(t^{1/2}))\) and associated eigenvector \(e_p(t) = \nu_p + O(t^{1/2})\) (we may take \(\ell = 1\) if \(\mu_p(t)\) is simple or \(Q\) is gradient). Hence for sufficiently small \(t > 0\), \(DF_{\lambda(\lambda,0)}\) has an eigenvalue with strictly negative real part. Setting \(F(x, \lambda) = \lambda x + Q(x)\), it follows from lemma 3.15 and the standard remainder estimate for Taylor’s theorem that
\[
\text{Tr}(f(x(t), \lambda(t)) = \text{Tr}(F(x(t), \lambda(t)) + O(t^{2p})) = |V| \lambda(t) + O(t^{2p}).
\]
Since we have shown that \(DF_{\lambda(\lambda,0)}\) has an eigenvalue \(-\lambda(t) (1 + O(t^{1/2}))\), it follows that for sufficiently small \(t > 0\), \(DF_{\lambda(\lambda,0)}\) has an eigenvalue with strictly positive real part. The estimates on the remaining eigenvalues of \(DF_{\lambda(\lambda,0)}\) follow using perturbation theory.

Finally, if \(f \in V^\omega\), then \(f^\omega\) is stable if \(f^\omega(f) \in S^\omega\), \(\delta = \delta(V, G)\), and then \(\Sigma^\omega(f)\) is isomorphic to \(\Sigma^\omega(f^\omega(f))\).

**Corollary 3.20.** Theorem 3.6 is true if there is an open dense set \(\Xi \subseteq P_{\nu P}^G(V, V)\) such that \(Z(Q) = \emptyset\) if \(Q \in \Xi\).

**Proof.** We may assume that if \(f \in S\), then \(f^\omega(f) \in \Xi\) (this is an open and dense condition). By proposition 3.18 (or lemma 3.16), if \(f \in S^k\), then all solution branches are branches of hyperbolic saddles with index(\(\gamma\)) \(\in [1, \dim(V) - 1]\). If \(f \in V\), then \(f\) is stable if \(f^\omega(f) \in S_0^\omega\), and then \(\Sigma^\omega(f)\) is isomorphic to \(\Sigma^\omega(f^\omega(f))\).

**Example 3.21.** The standard representation \(s_k\) satisfies the conditions of corollary 3.20 if \(k\) is odd. Of course, the result is straightforward to prove directly (see [19, section 16]).
3.7. **Completion of the proof of theorem 3.6**

It remains to consider branches of type C and Property (G). The next lemma is the final step needed for the proof of theorem 3.6.

**Lemma 3.22.** Suppose that Property (G) holds. Let \( f \in \mathcal{S} \) and set \( f^\delta(x, \lambda) = \lambda x + f^\delta(f) \in \mathcal{S}_0^\delta \). Let \( \gamma = (x, \lambda) \) be a solution branch of \( f^\delta \) with unique analytic parametrisation (3.4) and (3.5). Assume that \( \gamma \in \Sigma^*(f^\delta) \) is of type C:

(a) \( f^\delta(f) = \xi \in \mathcal{P}^2_G(V, V) \).

(b) \( \xi \neq 0, f(\xi) = 0 \).

Then \( q > p \) and \( \gamma \) is a branch of hyperbolic saddles.

**Proof.** Substituting the series for \( \gamma(t) \) in \( f^\delta(x) \), it follows that \( q > p \) since \( \xi = 0 \). By Property (G) and lemma 3.15, \( DQ_{\lambda_p} \) has at least one eigenvalue with strictly positive real part and one eigenvalue with strictly negative real part. We have

\[
Df^\delta(\lambda(t), x(t)) = \lambda I + DQ_{\lambda_0} + O(t^{p+1}) = t^p(DQ_{\lambda_p} + O(t)),
\]

and so it follows, as in the proof of lemma 3.16, that for sufficiently small \( t > 0 \), \( Df^\delta(\lambda(t), x(t)) \) has eigenvalues with strictly positive and negative real parts. \( \square \)

**Remark 3.23.**

(a) Little use is made of the analytic parametrisation. In particular, nothing is said about \( q \) except that \( q > p \).

(b) Although the proof of proposition 3.22 is easier than the proof of proposition 3.18, it is unsatisfying as we do not address the ‘radial’ eigenvalue associated to the direction \( \xi \).

If we assume \( f^\delta(\lambda(t), x(t)) \equiv 0 \), then it is straightforward to show that \( q = 2p \) and the radial eigenvalue is \(-2\lambda(t) + O(t^{p+1})\) (this is related to [9, theorem 4.2(a)]). However, if \( f^\delta(\lambda(t), x(t)) \neq 0, q = p + 1 \), and \( p > 1 \), then there is the possibility that \( \xi = 0 \) is a non-zero multiple of \( \xi \), and in this case the radial eigenvalue will not be determined by the cubic terms though it is still dominated by the non-radial eigenvalues of \( DQ_{\lambda_p} \). Of course, all this is easy for axial branches (for example, the branches with isotropy conjugate to \( S_L \times S_L \) for \( g_{2L} \)).

**Proof of theorem 3.6.** If \( f \in \mathcal{S} \), then \( f^\delta \) and \( f \) have isomorphic signed indexed branching patterns by theorem 2.17. By proposition 3.18 (or remark 3.17) and lemma 3.22, all non-trivial branches in \( \Sigma^*(f^\delta) \) are branches of hyperbolic saddles. \( \square \)

3.8. **Notes on the analytic parametrisation for type C branches**

Suppose \( f \in \mathcal{S}_0^\delta \) and \( \gamma = (x, \lambda) : [0, \delta] \to V \times \mathbb{R} \) is a non-trivial branch of solutions of \( f = 0 \). Write

\[
x(t) = \sum_{i=p}^{\infty} v_i t^i, \quad \lambda(t) = \pm t^p
\]

where \( v_i \in V, i \geq p, v_p \neq 0, p, q > 0 \) and \( p \) is minimal. Assume the branch is forward, so that \( \lambda(t) = t^p \). The power series for \( x \) is unique granted the minimality of \( p \) and the expression for \( \lambda(t) \). Suppose that \( Q(\xi) = 0 \) so that \( \gamma \) is a type C branch and \( q > p \). Since \( Q \) is a homogeneous quadratic polynomial, there is a unique symmetric bilinear form \( A : V^2 \to V \) satisfying \( Q(x) = A(x, x), x \in V \), defined by \( A(x, y) = \frac{1}{2}(Q(x + y) - Q(x) - Q(y)) \), \( x, y \in V \).
Substituting $x(t)$, $\lambda(t)$ in $\lambda x + Q(x) = F(x, \lambda)$ we find that if $p < q < 2p$ and $p > 1$, then
\[
F(x(t), \lambda(t)) = t^{p+q}v_p + \sum_{\ell=1}^{p-1} t^{2p+\ell} K_t(v_p, \ldots, v_{p+\ell}) + O(t^{3p}),
\] (3.8)
where $K_t(v_p, \ldots, v_{p+\ell}) = \sum_{i+j=\ell} A(v_{p+i}, v_{p+j})$, and
\[
k_t(v_p, \ldots, v_{p+\ell}) = 0, \quad \ell < q - p,
\] (3.9)
\[
k_{q-p}(v_p, \ldots, v_q) = -v_p,
\] (3.10)
\[
k_t(v_p, \ldots, v_{p+\ell}) = 0, \quad q - p < \ell < p.
\] (3.11)

We give one computation to illustrate some of the issues. Suppose $1 < p < q = p + 1 < 2p$, $\dim(V^{Gr_p}) > 1$ (the branch is not axial), and the branch is forward. Set $G(t) = t^{-q}DF_{x(t)}$. Substituting the series for $x(t)$ given by (3.7) in $F$ and equating lowest order coefficients we find that
\[
2A(v_{p+1}, v_p) + v_p = 0.
\] (3.12)

Since $v_p \neq 0$, $A(v_{p+1}, v_p) \in V^{Gr_p}$ is non-zero. Computing $G(t)$, we find
\[
G(t) = tI_V + DQ_{v_p} + tDQ_{v_{p+1}} + O(t^2).
\]

Hence, by (3.12),
\[
G(t)(v_p) = tv_p + 2tA(v_{p+1}, v_p) + O(t^2) = O(t^2)
\]
\[
G(t)(v_{p+1}) = -v_p + t(v_{p+1} + Q(v_{p+1})) + O(t^2).
\]

Hence $\lambda = 0$ is a non-simple eigenvalue of $G(0)$ and without further information on $Q$ it is possible that $G(0)\big|_{V^{Gr_p}}$ has pure imaginary eigenvalues. If $p = 1$, a cubic term $C(x)$ will contribute an $O(t)$-term to $G(t)(v_{p+1})$ and $G(t)(v_p)$ now has a term $tC(v_p)$. Even if Property $(G)$ holds, it seems difficult to determine an analytic form for the radial direction as the low order contribution given by $DQ_{v_p}(v_{p+1})$ may dominate those coming from higher order terms. Many questions remain.

### 3.9. Extensions to compact Lie groups and spaces of gradient vector fields

The extension of theorem 2.17 to absolutely irreducible representations of compact Lie groups is straightforward but requires consideration of relative equilibria, even if the low degree equivariants are all gradient vector fields; this, in turn, requires generalisation of the definition of signed indexed branching pattern (see [15, 10.1]). The reduced spectrum of the linearisation along a relative equilibrium is well defined as a real number [12, proposition J], [15, section 8.5.2]. It follows easily that the trace argument used for the proof of theorem 3.6, extends to compact Lie groups. Historically, bifurcation from relative equilibria was first systematically studied by Krupa [28] who proved the ‘tangent and normal form’ theorem for equivariant vector fields. Subsequent developments include the work of Wulff et al., allowing for bifurcation from relative periodic orbits (for example, [31, 32, 49]). There are also methods based on reduction to the orbit space; for example, the work of Koenig [29]. However, it is not clear that the genericity of stability theorem can be more easily proved using an orbit space approach. There is the question of proving theorem 2.7 if we restrict to the space of equivariant gradient vector fields. This space is not a $C^\infty_G(V)$-module generated by a finite set.
of polynomial gradient equivariants. However, if \( p_1, \ldots, p_i \) are a minimal set of homogeneous polynomial generators for \( P_G(V) \), then by Schwarz’s theorem on smooth invariants, every gradient vector field \( K \) on \( V \) may be written \( K(v) = \text{grad}(H \circ P)(v) \), where \( H : \mathbb{R}^f \to \mathbb{R} \) is \( C^\infty \) and \( P = (p_1, \ldots, p_i) : V \to \mathbb{R}^f \) is the orbit map. Along similar lines to those used in section 2, one can frame weak stability in terms of transversality to a Whitney regular stratification and stability in terms of equivariant jet transversality and prove the associated genericity theorems. Alternatively, arguments can surely be developed using reduction to the orbit space; the work of Koenig op. cit. is one starting point. Finally, though the optimisation problem motivating our paper is framed in terms of gradient dynamics, it is not necessary to use a gradient formulation of the stability theorem since the equivariants for the standard representation \((H_{k-1}, S_k)\) of degree at \( \leq 3 \) are all gradient and the problems are either two- or three-determined. Of course, if the quadratic equivariants vanished identically, our results would not apply; this is not the case for the motivating problem.

4. Minimal models of forced symmetry breaking of generic bifurcation on \( s_k \)

Suppose given a generic steady-state bifurcation defined on an absolutely irreducible representation \((V, G)\). For example, the pitchfork bifurcation on \((\mathbb{R}, \mathbb{Z}_2)\) or the \( S_1 \)-equivariant bifurcation on the standard irreducible representation \( s_k \). In order to understand forced symmetry breaking perturbations, it is natural to ask if there is a way to embed the bifurcation in a non-equivariant multiparameter family of vector fields which typically exhibit only generic bifurcation (that is, saddle-node bifurcations). More formally, given a smooth one-parameter family \( f_\lambda \) of \( G \)-equivariant vector fields defined on an open neighbourhood \( U \) of \((0,0) \in V \times \mathbb{R} \), an unfolding of the family consists of a smooth map \( K : (U \times \mathbb{R}) \times \mathbb{R}^m \to V; ((x, \lambda), \eta) \mapsto K_\eta(x, \lambda) \) such that \( K_0(x, \lambda) = f(x, \lambda) \), \((x, \lambda) \in U \). Roughly speaking, the unfolding is universal if for every sufficiently small smooth perturbation \( f \) of \( f \), there exists \( \eta \in \mathbb{R}^m \) (close to 0) such that \( K_\eta \) is equivalent in some sense to \( f \) (see [22, chapter 2, p 51] for ‘strong equivalence’ and more details on the singularity approach to unfoldings which we do not follow here). The aim is to be able to describe all small perturbations of \( f \) in terms of the extended family \( K \). Unfortunately, it is not realistic to look for universal unfoldings of generic \( S_k \)-equivariant bifurcation if \( k \gg 3 \) as the number of parameters \( m \) required grows rapidly with \( k \) (the case \( k \) even is especially awkward).

Our approach will be to show that there are ‘minimal models’ for symmetry-breaking perturbations of generic bifurcation on \( s_k \) and to emphasise deformation to a minimal model rather than an explicit construction of a universal unfolding. The minimal model can be viewed as an unfolding but we do not emphasise that aspect. Although we restrict to \( s_k \), the methods we describe can likely be extended to \( s_k \otimes s_\ell, k, \ell \geq 3 \), and possibly to families of equivariant gradient vector fields on general absolutely irreducible representations admitting quadratic equivariants.

The qualifier minimal is intended to suggest a minimal level of complexity in the dynamics of the symmetry breaking perturbation given by a minimal model. Our motivation lies in describing mechanisms for creating or annihilating local minima in gradient systems without creating new local minima in the process. The question arises from problems in non-convex optimisation in neural networks and the occurrence of spurious minima. The spurious minima do not come from bifurcation of the global minima (or spontaneous symmetry breaking) but are created locally through changes in the geometry of the optimisation landscape. We refer to the concluding comments section for recent results related to shallow neural networks.
4.1. Generic bifurcation for $s_k$

Generic steady-state equivariant bifurcation of the trivial solution for families of $S_k$-equivariant vector fields on $\rho_3$ is well understood [19, section 16], [16]. If $k$ is odd, equivariant bifurcation is two-determined, and for stable families all branches are of type $S$. If $k$ is even, equivariant bifurcation is three-determined and for stable families all branches are of type $S_k$ with isotropy conjugate to $S_k \times S_1$; these are of type $C$. All branches of solutions arising from generic bifurcation of the trivial solution are axial [16] and the set $\{ \Sigma^*(f) \mid f \in S(H_{k-1}, S_k) \}$ is known [19] (see below).

We recall definitions and results from [15, 19] needed later. Recall (example 2.5) that the space $P_{2s}^2(H_{k-1}, H_{k-1})$ of homogeneous $S_k$-equivariant quadratic polynomials is one-dimensional with basis given by the gradient of $\rho_3(x) = \sum_{i=1}^{2s} x_i^2 : H_{k-1} \to \mathbb{R}$. Set $Q = \text{grad}(\rho_3)$. We have (example 2.5)

$$Q(x) = \begin{pmatrix} x_1^2 - \frac{1}{k} \sum_{i=1}^{k} x_i^2, \cdots, x_k^2 - \frac{1}{k} \sum_{i=1}^{k} x_i^2 \end{pmatrix}, \quad x = (x_1, \cdots, x_k) \in H_{k-1}.$$

The phase vector field $\mathcal{P}_Q$ of $Q$ is defined on the unit sphere $S^{k-2} \subset H_{k-1}$ by

$$\mathcal{P}_Q(u) = Q(u) - \langle Q(u), u \rangle u, \quad u \in S^{k-2},$$

and $\mathcal{P}_Q = \text{grad}(\rho_3)_1[S^{k-2}]$. Denote the zero set of $\mathcal{P}_Q$ by $Z(\mathcal{P}_Q)$ and note that $u \in Z(\mathcal{P}_Q)$ iff $-u \in Z(\mathcal{P}_Q)$ iff $Q(u) \in \mathbb{R}u$.

Given $1 \leq p < k$, set $q = k - p$ and define $e_p \in S^{k-2}$ by

$$e_p = \frac{1}{\sqrt{pqk}}(q, \cdots, q, -p, \cdots, -p) = -\sigma e_p, \quad \text{def} \quad \frac{1}{\sqrt{pqk}}(q^p, -p^0), \quad (4.13)$$

where $q^p$ means $q$ is repeated $p$ times. Note that $e_q = -\sigma e_p$, where $\sigma \in S_k$ is given by $\sigma(i) = q + i \mod k$, and $-e_p \in S_k e_p$ iff $k$ is even and $p = q = k/2$.

Write $k = 2\ell + 1$ ($k$ odd) or $k = 2\ell$ ($k$ even). For $p \in \ell$, set $L_p = \mathbb{R} e_p$. The line $L_p$ is an axis of symmetry for $s_k$ and the isotopy of non-zero points on $L_p$ is $S_p \times S_q$. In every case, except when $k$ is even and $p = \ell$, $S_p \times S_q$ is a maximal proper subgroup of $S_k$. The set of axes of symmetry with isotropy conjugate to $S_p \times S_q$ is $A_p = \{ gL_p \mid g \in S_p \times S_q \}$ and $\bigcup_{p \leq \ell} A_p$ is the complete set of $2^{k-1} - 1$ axes of symmetry of $s_k$.

Simple computations [19, section 16] verify $e_p \in Z(\mathcal{P}_Q)$, $p \in k - 1$, and

$$Q(e_p) = \begin{pmatrix} \frac{1}{pqk}(q-p) e_p, \quad p \in k - 1 \end{pmatrix} \quad (4.14)$$

$$\bigcup_{p=1}^{k-1} S_k e_p = Z(\mathcal{P}_Q). \quad (4.15)$$

Moreover, $Z(\mathcal{P}_Q)$ consists of hyperbolic zeros [18, section 4] and for $p \in k - 1$

$$\text{index}(\mathcal{P}_Q; e_p) = k - p - 1, \quad (4.16)$$

$$\text{index}(\mathcal{P}_{sQ}; e_p) = p - 1, \quad (4.17)$$

This is slightly different from [19, section 16] where $e_p$ defined to be $(\frac{1}{p}, -\frac{1}{q})$.
index(\(\mathcal{P}_Q; \varepsilon_p\)) + index(\(\mathcal{P}_Q; -\varepsilon_p\)) = k - 2 = \text{dim}(S^{k-2}).
(4.18)

The statements for \(p = 1\) (resp. \(k - 1\)) follow since points in \(S_kx_1\) (resp. \(S_kx_{k-1}\)) give the absolute maximum (resp. minimum) value of \(C : S^{k-2} \to \mathbb{R}\). The remaining indices can easily be found using an inductive argument or just computed directly. We use the results on \(\mathcal{P}_Q\) to give a complete description of the signed indexed branching patterns of stable families.

Replacing \(x\) by \(-x/a, a \neq 0\), \(x' = \lambda x + aQ(x)\), transforms to

\[x' = \lambda x - Q(x).\]
(4.19)

Analysis of (4.19) gives all the hyperbolic branches except for those with isotropy conjugate to \(S_k \times S_k\), when \(k = 2\ell\).

Representative solution branches \((x^+_p, \lambda) : [0, \infty) \to H_{k-1} \times \mathbb{R}\) of (4.19) are given for \(1 \leq p < k/2\) by

(B) \(x^+_p(s) = s\frac{\sqrt{q-p}}{q-p} \varepsilon_p, \lambda(s) = -s, s \in [0, \infty),\) is a backward branch of hyperbolic saddles of index \(k - p - 1\).

(F) \(x^+_p(s) = s\frac{\sqrt{q-p}}{q-p} \varepsilon_p, \lambda(s) = s, s \in [0, \infty),\) is a forward branch of hyperbolic saddles of index \(p\).

It follows from (B) and (F) that for \(1 \leq p < k/2\), we have

\[
\text{index}(\{x^+_p\}) + \text{index}(\{x^-_p\}) = k - 1.
(4.20)
\]

Excluding the case \(k = 2\ell, p = \ell\), the set of forward and backward solution branches of (4.19) is obtained by taking the \(S_k\)-orbits of each of the representative solution branches. Observe that the radial eigenvalue along the branch is always \(-\lambda\)—since (4.19) is quadratic—and the transverse eigenvalues are given by the eigenvalues of \(\mathcal{P}_{\pm Q}\) at \(\varepsilon_p\), multiplied by \(R = \|x(s)\| \sim s\). If we introduce higher order terms in (4.19), then the solutions and eigenvalues are perturbed by terms of order \(O(s^2)\) and the signed indexed branching patterns are unchanged.

**Lemma 4.1.** If \(\rho > 0\) and \(\delta \geq 4\sqrt{k+1}/\rho\), then every forward solution branch \(\gamma(x) = (x(s), \lambda(s))\) of (4.19) satisfies

\[\gamma([0, \rho]) \subseteq D_0(0) \times [0, \rho] \subseteq H_{k-1} \times \mathbb{R}\]

provided that we exclude branches along axes in \(S_kL_{k/2}\) if \(k\) is even.

A similar statement holds for backward solution branches.

**Proof.** The expressions for solution branches \((x(s), \lambda(s))\) given by (B) and (F) above imply that the slope of the line \(\|s\|, \|x(s)\| \subset \mathbb{R}^2\) is less than \(4\sqrt{k+1}\)—this holds for \(k\) odd or even, provided branches along axes in \(S_k\), are excluded when \(k = 2\ell\).

It remains to look at solution branches when \(k = 2\ell\) and \(p = q = \ell\). Consider the cubic system

\[x' = \lambda x - Q(x) + T(x),\]
(4.21)

where \(T \in P^3(H_{k-1}, H_{k-1})\). For \(k \geq 4, P^3(H_{k-1}, H_{k-1})\) has basis \(R(x) = \|x\|^2x\) and \(\text{grad}(\rho_k)\) (example 2.5). Since \(T : L_\ell \to L_\ell, T(\varepsilon_\ell) = c\varepsilon_\ell\) and the generic case is when \(c \neq 0\). If \(c < 0\) (resp. \(c > 0\)) we have a supercritical (resp. subcritical) pitchfork bifurcation along \(L_\ell\). In what follows we assume \(c < 0\), the analysis for \(c > 0\) is similar.
Suppose first that \( k = 4 \). Minimal models for nonlinearity are given by \((\pm x_i(s), \lambda(s))\), where

\[
x_i(s) = \frac{1}{\sqrt{-c}} s e_i, \quad \lambda(s) = s^2, \quad s \in [0, \infty).
\]

The radial eigenvalue \( \mu_i(s) \) at \( \pm x_i(s) \) is therefore \(-2\lambda(s) = -2s^2\) and so \( \mu_i(s) = O(s^2) = O(R^2) \), where \( R = ||x(s)|| \). On the other hand, the eigenvalues at \( x_i(s) \) in directions transverse to \( L_e \) are given by the eigenvalues of the Hessian of \( P_{eQ} \) at \( e_i \) scaled by \( R \). Hence these eigenvalues are \( O(R) \) and dominate any transverse eigenvalues coming from the cubic term \( S \). Since index(\( P_{eQ}, e_i \)) is \( \ell - 1 \), it follows that index(\( [\pm x_i] \)) = \( \ell \). For \( k \geq 4 \), \( \ell \in [1, k - 2] \), and so \((\pm x_i(s), \lambda(s)) \) is a branch of hyperbolic saddles.

**Remark 4.2.**

(a) If \( k = 2\ell \), then the maximal index \( \ell \) of the forward branches occurs for branches of isotropy type \((S_i \times S_i)\) when \( c < 0 \) (supercritical branching in \( L_e \)). In particular, for \( k \geq 4 \), when the trivial solution loses stability, not only are there no branches of sinks but the maximal index of the new forward solution branches is \( \ell = k/2 \). The minimal index of the backward branches is also \( \ell \). Continuing to assume \( c < 0 \), generic equivariant bifurcation on \( s_k \) results from the simultaneous collision of \( 2^{k-1} - 1 - \left(\frac{2^{k-1}}{\ell}\right) \) branches of saddles of relatively high index \( \geq \ell \), followed by the emergence of \( 2^{k-1} - 1 + \left(\frac{2^{k-1}}{\ell}\right) \) branches of saddles of relatively low index \( \leq \ell \). Similar results hold for \( k \) odd: there are now \( 2^{k-1} - 1 \) forward (resp. backward) non-trivial branches with minimal (resp. maximal) index \([k/2] \). Summarising, for all \( k \geq 3 \), generic equivariant bifurcation on \( s_k \), changing the trivial solution from a sink to a source, results from a collision of non-trivial saddle branches of relatively high index \( \geq [k/2] \), followed by the emergence of non-trivial saddle branches of relatively low index \( \leq [k/2] \).

(b) Lemma 4.1 fails for branches of isotropy type \((S_i \times S_i)\) unless \( \rho > 1/\sqrt{|c|} \).

### 4.2. Minimal models for \( s_k \)

Suppose first that \( k = 2\ell + 1 \), \( \ell \in \mathbb{N} \), and set \( f(x, \lambda) = \lambda x - Q(x) \), where \( Q = \text{grad}(\mu_0) \), \( \rho_1 = \frac{1}{2} \sum_{i=0}^{k} d_i^3 \).

Let \( \rho > 0 \). Choose \( \delta_i = K_i \rho, \ i \in 2 \), where \( K_1 > 0, K_2 = K_3(k) > 0 \), and let \( W \) be the compact neighbourhood of \((0, 0) \in (L_1 \oplus L_1^\perp) \times \mathbb{R} = H_{k-1} \times \mathbb{R} \) defined by

\[
W = W_\rho = ([-\delta_1, \delta_1] \times \overline{D_{S_k}(0)}) \times [-\rho, \rho].
\]

Outside of \( W \), \( f \) has only hyperbolic critical points. Indeed, the only non-hyperbolic zero of \( f \) on \( V \times \mathbb{R} \) occurs when \( \lambda = 0 \) and \( x = 0 \). Our interest is in choosing \( K_i \) so that we can perturb the family \( f \) to obtain a smooth, but not \( S_k \)-equivariant, family \( \tilde{f} \), equal to \( f \) on \( V \times \mathbb{R} \setminus W \), such that

(a) \( \tilde{f} \) is stable under all sufficiently small \( C^2 \) perturbations.

(b) \( \tilde{f} \) has the maximum number of crossing curves \( \gamma = (x, \lambda) : \mathbb{R} \to V \times \mathbb{R} \), where \( \lambda(t) \) takes all values in \( \mathbb{R} \). For every crossing curve, it is required that \( x(t) \) will be a hyperbolic zero of \( \tilde{f}_{\lambda(t)} \), all \( t \in \mathbb{R} \).

(c) \( \tilde{f} \) has the minimum number of saddle node bifurcations (necessarily in \( W \)). All other zeros of \( \tilde{f} \) are hyperbolic.
(d) As $\rho \to 0$, $K_1, K_2 \to 0$ (convergence of $K_2$ is not uniform in $k$).

In section 4.1, we gave the index and branching data for solution curves of $f$ (statements (B) and (F)). We display these in figure 1.

Taking account of the indices of branches, it is clear that crossing curves must be of index $\ell$. All solution branches with index not equal to $\ell$ must have a saddle-node bifurcation. Since the number of branches along $S_k L_\ell$ is $\binom{k}{\ell}$, the number of crossing curves is at most $\left(\frac{2^k}{\ell}\right)$. If there are $m$ crossing curves then the number of saddle-node bifurcations is at least $(2^k - 2m)/2$, since there are $2^k$ solution branches of $f$. If $\lambda < 0$, branches differing by 1 in index can join through a saddle-node bifurcation in the region $W$; similarly for $\lambda > 0$. A straightforward count verifies that $m \leq \left(\frac{2^k}{\ell}\right)$. Hence the number of crossing curves is at most $\left(\frac{2^k}{\ell}\right)$ and the number of saddle-node bifurcations is at least $2^{2\ell} - \left(\frac{2^k}{\ell}\right)$.

**Definition 4.3 (Assumptions and notation as above).** The family $\hat{f}$ is a minimal symmetry breaking model for $f$ if

(a) $f = \hat{f}$ on $V \times \mathbb{R} \setminus W$.
(b) $\hat{f}$ has exactly $\left(\frac{k-1}{k/2}\right)$ crossing curves.
(c) $\hat{f}$ has exactly $2^{k-1} - \left(\frac{k-1}{k/2}\right)$ saddle-node bifurcations.

If $k = 2\ell$ is even, the analysis is slightly different as generically there are pitchfork bifurcations along the axes $S_k L_\ell$ and we need to take account of cubic equivariants. If, for example, we take the family $f(x, \lambda) = \lambda x - Q(x) - c T(x)$, $c > 0$, where $T(x) = ||x||^2 x$, if $x \in L_\ell$, then
Indices for branches, \( k \) even, close to the bifurcation point \((0,0) \in V \times \mathbb{R}\). Branches along \( S_{L_\ell} \) occur only for \( \lambda > 0 \). A crossing curve results from an index \( \ell \) branch along \( L_{\ell-1}, \lambda < 0 \), joining with an index \( \ell \) branch along \( L_{\ell}, \lambda > 0 \).

However, the presence of a non-zero cubic term will create new bifurcations in the family \( f \) outside of a compact \( c \)-dependent neighbourhood of \((0,0) \in V \times \mathbb{R}\). In order to handle this difficulty, we may either modify definition 4.3 so that it is local—everything defined on a \( c \)-dependent neighbourhood of \((0,0)\)—or modify the family \( f \) so that the coefficient \( c \) is a smooth \( S_k \)-invariant which is identically 1 on a thin neighbourhood \( K_1 \) of \( S_k L_\ell \) and 0 outside of a neighbourhood \( K_2 \supset K_1 \) of \( S_k L_\ell \). Here \( K_1, K_2 \) are chosen so that the only bifurcations of the modified family occur at \((0,0) \in V \times \mathbb{R}\). In the latter case the definition of minimal model is formally identical to definition 4.3 (with the modified family replacing \( f \)).

4.3. Minimal symmetry breaking model: \( k \) odd

**Theorem 4.4.** Assume \( k = 2\ell + 1 \) (\( k \) is odd) and let \( f_s(x) = \lambda x - Q(x) \), where \( Q \) is the standard quadratic equivariant on \((H_{k-1},S_k)\). Let \( \eta_0 > 0 \) and define \( \rho = 4\sqrt{\eta_0}, \delta_1 = \rho/2, \delta_2 = \sqrt{4/\rho} \). Following (4.22), define

\[
W = W_\rho = \left([-\delta_1, \delta_1] \times \overline{D_{L_\ell}(0)}\right) \times [-\rho, \rho] \subset (L_1 \oplus L_1^\perp) \times \mathbb{R}.
\]
There exist $C > 0$, depending only on $\eta_0$ and $k$, and $\eta_1 \in (0, \eta_0]$, such that for every $\eta \in (0, \eta_1]$, there is a smooth $S_{k-1}$-equivariant family $\hat{f}_\lambda^\eta$ satisfying

(a) $\hat{f}_\lambda^\eta(x) = f_\lambda(x)$ if $(x, \lambda) \notin W$.
(b) $\|\hat{f}_\lambda^\eta - f\|_{W,2} < C\eta$.
(c) The only bifurcations of the family $\hat{f}_\lambda^\eta$ are saddle-node bifurcations and $\hat{f}_\lambda^\eta$ is stable under $C^2$-small perturbations supported on a compact neighbourhood of $(0, 0) \in H_{k-1} \times \mathbb{R}$ (no assumption of equivariance).
(d) The family $\hat{f}_\lambda^\eta$ has exactly $2^\ell = \left(\frac{2^\ell}{\ell}\right)$ saddle node bifurcations and all of the branches of solutions with index $\neq \ell$ will end or start with a saddle-node bifurcation that occurs in $W$.
(e) There are exactly $\left(\frac{2^\ell}{\ell}\right)$ crossing curves $(x, \lambda) : \mathbb{R} \to H_{k-1} \times \mathbb{R}$ of solutions to $\hat{f}_\lambda^\eta = 0$; each of these curves consists of hyperbolic equilibria of index $\ell$.

**Remark 4.5.**

(a) Suppose $a \in \mathbb{R}$ is non-zero. Recall (4.19) that the change of coordinates $x = -a^{-1}x$ transforms $x' = \lambda x + a\sigma(x) = \lambda y - Q(x)$. Hence theorem 4.4 gives minimal symmetry breaking models for all generic families $f_\lambda(x) = \lambda x + aQ(x), a \neq 0$.
(b) The stability under $C^2$-small perturbations (rather than $C^1$) is required on account of the saddle-node bifurcations which are not necessarily preserved under $C^1$-small perturbations.
(c) The interest of the result lies in small values of $\eta$ and $\rho, \delta_1, \delta_2$. In the proof, $\hat{f}_\lambda^\eta = f_\lambda^\eta$ if $\eta > \eta_1$, and $\hat{f}_\lambda^\eta$ is smooth in $\eta$ on $(0, \eta_1]$. It is straightforward to modify the construction to obtain smooth dependence for $\eta \in (0, \infty)$.

**Sketch of the proof.** We break symmetry from $S_k$ to $S_{k-1}$ using a perturbation parallel to $L_1$. Initially, we assume the perturbation is constant. Next follow a number of lemmas that describe the effect of the perturbation on the dynamics restricted to a sequence of flow-invariant two-planes. Most of these results will hold for $k$ odd or even. The final step is to localise the perturbation to have support in $W$ and show that the localisation process does not introduce (or destroy) solutions, or change stabilities, within $W$. With the exception of one detail, used for localisation, the proof of existence of the minimal symmetry breaking model is elementary.

4.4. **Minimal symmetry breaking model: preliminary results, $k$ odd or even**

Set $V = H_{k-1} \subset \mathbb{R}^2$. For $p \in k - 1$, set $q = k - p$. If $k$ is odd (resp. even), define $\Sigma = \ell + 1$ (resp. $\Sigma = \ell$). Recall from section 4.1 that for $p \in \Sigma$, $L_p$ is the axis of symmetry through $\varepsilon_p$ and

$$L_p = \{(px^p, -px^p) : x \in \mathbb{R}\} = V^{S_p \times S_q}.$$ 

For $2 \leq p \leq \ell + 1$, define

$$L_{p-1}^* = \{(z, y^{p-1}, z^q) : (p-1)y + (q+1)z = 0\} = \sigma L_{p-1},$$

where $\sigma = (1p) \in S_k$.

For $1 < p \leq \Sigma$, define the two-plane $E_p \subset V$ by

$$E_p = \{(x, y^{p-1}, z^q) : x + (p-1)y + qz = 0\} = V^{S_1 \times S_{p-1} \times S_q}.$$
In what follows, $S_{k-1} := S_1 \times S_{k-1} \subset S_k$.

**Remark 4.6.** The representation $(V, S_{k-1})$ is the orthogonal direct sum of $(H_{k-2}, S_{k-1})$ and the trivial representation $(L_1, S_{k-1})$, where we regard $H_{k-2} = \{ x \in H_{k-1} | x_1 = 0 \}$. In particular, the fixed point spaces $V^{S_{k-1} \times S_q}$, $S_{p-1} \times S_q \subset S_{k-1}$, $p + q = k$, are two-dimensional and $V^{S_{p-1} \times S_q} = E_p$, $2 \leq p \leq \ell + 1$.

**Lemma 4.7 (Notation and assumptions as above).**

(a) For $2 \leq p \leq \ell$, $E_p$ contains exactly three axes of symmetry: $L_1, L_p$ and $L^*_p$.  
(b) If $k$ is odd, $E_{l+1}$ contains exactly three axes of symmetry: $L_1, L_\ell$ and $L^*_\ell$.
(c) For the $S_{k-1}$-action on $V$ and $p \leq \Sigma$,

$$E_p \setminus L_1 = \{ x \in V \mid (S_{k-1})_x = S_{p-1} \times S_q \}.$$  
(d) For all $k \geq 4$ we have

$$\theta_{L_1, L_p} = \cos^{-1} \left( \sqrt{\frac{q}{(k-1)p}} \right) \in \left[ \frac{\pi}{3}, \frac{\pi}{2} \right]$$

$$\theta_{L^*_p, L_p} = \cos^{-1} \left( \sqrt{\frac{q(p-1)}{(q+1)p}} \right) \in \left( 0, \cos^{-1}(1/3) \right]$$

$$\theta_{L^*_p, L_1} = \cos^{-1} \left( -\sqrt{\frac{p-1}{(k-1)(q+1)}} \right) \in \left( \pi, \frac{2\pi}{3} \right].$$

(e) If $k = 3$, then $p = 2$, $E_2 = \mathbb{R}^2$ and $L_2, L_1, L^*_1$ are the three axes of symmetry for the standard $S_3$-action on $\mathbb{R}^2$.

(f) For $p \neq p', 2 \leq p, p' \leq \Sigma$, $E_p \cap E_{p'} = L_1$.

**Proof.** All statements are easy to verify and we omit the details.

**Remark 4.8.** Taking the $S_{k-1}$-action on $V$, statement (c) implies that non-zero points on the axes $L_p, L^*_p$ have the same isotropy. Hence, there is the possibility of $S_{k-1}$-equivariant deformations of the original $S_1$-equivariant family on $V$ that allow us to connect zeros on the axes $L_p, L^*_p$ via a saddle-node bifurcation. This observation lies at the core of our construction.

The $S_{k-1}$-symmetry organises the details.

For $p \leq \Sigma$, define the linear map $U_p : \mathbb{R}^2 \to E_p$ by

$$U_p(u, 0) = \frac{1}{\sqrt{(k-1)}} ((k-1)u, -u^{p-1}, -u^q), \quad u \in \mathbb{R}$$

$$U_p(0, v) = \frac{1}{\sqrt{(k-1)(p-1)q}} (0, qv^{p-1}, -(p-1)v^q), \quad v \in \mathbb{R}.$$  

Observe that $\|U_p(1, 0)\| = \|U_p(0, 1)\| = 1$ and $U_p(u, 0) \perp U_p(0, v)$ for all $(u, v) \in \mathbb{R}^2$. Hence $U_p$ maps $\mathbb{R}^2$ isometrically onto $E_p$. Let $\{e_1 = (1, 0), e_2 = (0, 1)\}$ be the standard Euclidean basis of $\mathbb{R}^2$.

**Lemma 4.9 (Assumptions and notation as above).**

(a) $U_p(e_1) = e_1$ and $U_p(Re_1) = L_1$.  
2838
Figure 3. Zeros and dynamics of $x' = \lambda x - Q(x)$ on $E_p$, $2 \leq p \leq \ell$, $\lambda < 0$.

(b) $U_p^{-1}(L_p) = \{(u, m_p u) \mid u \in \mathbb{R}\}$, where $m_p = \sqrt{\frac{k-1}{q}}$ and $\langle U_p^{-1}(e_p), e_2 \rangle > 0$.

(c) $U_p^{-1}(L_p^*) = \{(-u, m_{p-1}^* u) \mid u \in \mathbb{R}\}$, where $m_{p-1}^* = \sqrt{\frac{k}{p-1}}$, $\langle U_p^{-1}(e_{p-1}^*), e_2 \rangle > 0$, and $e_{p-1}^* = (1/p)e_{p-1}$.

In particular, if we take the standard orientation of $\mathbb{R}^2$ and the orientation on $E_p$ defined by $e_1, e_p$, $U_p$ preserves orientation.

**Proof.** (a) is obvious by the definition of $U_p(u, 0)$. For (b), note that since $p \geq 2$, $U_p(u, m_p u) \in L_p$ iff the first two components of $U_p(u, m_p u)$ are equal. This leads to the condition

$$
\frac{k-1}{\sqrt{k(k-1)}}u = -\frac{u}{\sqrt{k(k-1)}} + \frac{q m_p u}{\sqrt{(k-1)(p-1)q}}, \quad u \in \mathbb{R}.
$$

Dividing through by $u$ and simplifying, we obtain $m_p = +\sqrt{\frac{k-1}{q}}$.

The first component of $U_p(u, m_p u)$ is strictly positive if $u > 0$. Hence there exists $u > 0$ such that $e_p = U_p(u, m_p u)$ and so $\langle U_p^{-1}(e_p), e_2 \rangle > 0$. The proof of (c) is similar. $\square$

**Lemma 4.10 (Notation and assumptions as above).** Suppose $\lambda < 0$ and denote the zeros of $x' = \lambda x - Q(x)$ lying on $L_p, L_p^*$ by $c_p, c_{p-1}^*$ respectively. For $2 \leq p \leq \ell$, there are connections from $c_p$ to $c^*$ and $c_{p-1}$ (see figure 3). A similar result holds when $\lambda > 0$ with connections now from $c_1$ and $c_{p-1}^*$ to $c_p$.

**Proof.** The fixed point space $E_p$ is invariant by the flow of $f_{\lambda}(x) = \lambda x - Q(x)$. All the zeros of $f_{\lambda}$ occur on axes of symmetry, since the coefficient of $Q$ is non-zero [19]. The index of $c_1$ is $k - 2$ and so the index of $c_1$ for $f_{\lambda}|E_p$ is 1. Observe that if the index of $c_p$ for $f_{\lambda}|E_p$ is
there would have to be an additional zero for the phase vector field in the interior of the region of \( S^1 \subset S(V) \) determined by the wedge defined by \( c_1, c_p \) contradicting the result that all zeros of the phase vector field lie on axes of symmetry. Hence the index of \( c_p \) for \( f_s|E_p \) is 0 and from this it follows easily the there is a connection from \( c_p \) to \( c_1 \). Similar arguments show there must be a connection from \( c_p \) to \( c_{p-1}^* \). The argument when \( \lambda > 0 \) is similar with the zeros \( c_1 \in L_1, c_p \in L_p, c_{p-1} \in L_{p-1}^* \) now lying on the other side of the origin. 

We need to compute the pull-back by \( U_p \) of the family \((\lambda I) - Q|E_p\). For this it suffices to compute the pull back \( Q_p \) of \( Q \) to \( \mathbb{R}^2 \) since \( \lambda I \) pulls back to \( \lambda I_{\mathbb{R}^2} \).

Lemma 4.11 (Assumptions and notation as above). For \( 2 \leq p \leq L \), \( Q_p = (Z_1, Z_2) \) where

\[
Z_1(u, v) = \frac{1}{\sqrt{k(k-1)}} \left( (k-2)u^2 - v^2 \right) \\
Z_2(u, v) = -\frac{2}{\sqrt{k(k-1)}} u v + \frac{q - p + 1}{\sqrt{q(k-1)(p-1)}} u^2.
\]

Proof. Since \( Q \) is a gradient vector field, and \( U_p \) is an isometry, it suffices to find the gradient of the pull back \( C_p(u, v) = c(U_p(u, v)) \) of the cubic \( c : V \rightarrow \mathbb{R} \) defining \( Q \).

Now \( c(x) = \frac{1}{3} \left( \sum_{i} k x_i^3 \right) \), \( x \in V \), and so, taking \( x = U_p(u, v) \), \( (u, v) \in \mathbb{R}^2 \), and an easy computation gives

\[
C_p(u, v) = \frac{1}{3} \left( A^3(k-1)u^3 + (p - 1)(-Au + Bqv)^2 - q(Au + B(p - 1)v)^2 \right),
\]

where \( A = 1/\sqrt{k(k-1)} \) and \( B = 1/\sqrt{(k-1)(p-1)q} \). After some elementary algebra, we find that the components \( Z_1, Z_2 \) of \( Q_p \) are given by (4.23) and (4.24).

Proposition 4.12. In \((u, v)\)-coordinates, the dynamics of \( x' = \lambda x - Q(x) \) on \( E_p \) is given by

\[
\dot{u} = \lambda u - \frac{1}{\sqrt{k(k-1)}} ((k-2)u^2 - v^2) \\
\dot{v} = \lambda v + \frac{2}{\sqrt{k(k-1)}} u v - \frac{q - p + 1}{\sqrt{q(k-1)(p-1)}} u^2.
\]

We have solution curves \( c_1 = (u_1, v_1), c_p = (u_p, v_p), c_{p-1}^* = (u_{p-1}^*, v_{p-1}^*) \) for (4.25) and (4.26) given for \( \lambda \in \mathbb{R} \) by

\[
(u_1, v_1)(\lambda) = \lambda \frac{\sqrt{k(k-1)}}{k-2} e_1 \\
(u_p, v_p)(\lambda) = \lambda \sqrt{k-1} \left( \frac{q}{k-1} \frac{k(p-1)}{q} \right) \\
(u_{p-1}^*, v_{p-1}^*)(\lambda) = \lambda \sqrt{k-1} \left( \frac{p-1}{k-1} \frac{kq}{p-1} \right).
\]

Proof. A straightforward computation using lemma 4.11.
Forced symmetry breaking to $S_{k-1}$.

**Proposition 4.13.** Let $\eta > 0$ and consider the perturbed equations

$$
\dot{u} = \lambda u - \frac{1}{2\sqrt{k(k-1)}}((k-2)u^2 - v^2) - \eta \tag{4.30}
$$

$$
\dot{v} = \lambda v + \frac{2}{\sqrt{k(k-1)}}uv - \frac{q-p+1}{\sqrt{q(k-1)(p-1)}}v^2. \tag{4.31}
$$

For $1 \leq p \leq \ell$, define

$$
\gamma_{k,p} = \begin{cases} 
2\sqrt{\eta} \frac{\sqrt{k-2}}{\sqrt{k(k-1)}}, & p = 1 \\
2\sqrt{\eta} \frac{\sqrt{k-2} q - p + 1}{\sqrt{k(k-1)}} \sqrt{\frac{4q(p-1)}{(q-p+1)^2 k(k-2)}}, & p > 1
\end{cases}. \tag{4.32}
$$

(a) Suppose that $2 \leq p < k/2$. The system (4.30) and (4.31) has saddle-node bifurcations at $\lambda = \pm \gamma_{k,p}$. Specifically, at $\lambda = -\gamma_{k,p}$, the (perturbed) branches $c_p(\lambda, \eta), c_p^{\pm}(\lambda, \eta) \subset E_p$ collide in a saddle-node bifurcation as $\lambda \nearrow -\gamma_{k,p}$ and at $\lambda = +\gamma_{k,p}$, the branches $c_p(\lambda, \eta), c_p^{\pm}(\lambda, \eta)$ are created in a saddle-node bifurcation as $\lambda \nearrow \gamma_{k,p}$. Set $c_p(\pm \gamma_{k,p}, \eta) = \pm b_p \in E_p$.

(b) The system (4.30) and (4.31) has saddle-node bifurcations along $L_1$ at $\lambda = \pm \gamma_{k,1}$. Specifically, at $\lambda = -\gamma_{k,1}$, the (perturbed) branch $c_1(\lambda, \eta)$ collides with the (perturbed) trivial solution branch in a saddle-node bifurcation as $\lambda \nearrow -\gamma_{k,1}$ and at $\lambda = \gamma_{k,1}$, the branch $c_1(\lambda, \eta)$ and the perturbed trivial solution branch are created in a saddle-node bifurcation as $\lambda \nearrow \gamma_{k,1}$. Set $\pm b_1 = c_{\pm}(\pm \gamma_{k,1}, \eta) \in L_1$.

(c) For all $k \geq 3$, and $p \in [1, k/2)$, $|\gamma_{k,p}| \leq 2\sqrt{\eta}$. For fixed $k$, $|\gamma_{k,p}|$ is a strictly monotone decreasing function of $p \in [1, k/2)$ and

$$
|\gamma_{k,p}| \in \left[2\sqrt{\eta} \frac{\sqrt{3}}{\sqrt{k} \sqrt{k(k-1)}}, 2\sqrt{\eta} \frac{\sqrt{k-2}}{\sqrt{k(k-1)}}\right].
$$

(d) $\|b_1\| \leq \sqrt{\eta}$ and for all $p \in [1, k/2)$,

$$
\|b_p\| < 2\sqrt{\frac{\eta k}{3}}
$$

**Remark 4.14.**

(a) If $k$ is even, then $\gamma_{k,\ell} = 0$. Cubic terms are needed to resolve this case.
(b) (a) only applies if $k > 4$; (b) applies for all $k \geq 3$.
(c) (c) implies that as $\lambda \nearrow 0$, there is a sequence of saddle-node bifurcations. The first on $L_1$; the second is an $S_{k-1}$-orbit of the saddle-node bifurcation on $E_2$. The sequence ends with the $S_{k-1}$-orbit of the saddle-node bifurcation on $E_\ell$ (resp. $E_{\ell-1}$) if $k$ is odd (resp. even).
The order of the sequence is reversed when $\lambda$ increases through zero.

Before proving proposition 4.13, we give a lemma that helps simplify and organise the computations.

**Lemma 4.15 (Assumptions of proposition 4.13).** Under the linear coordinate change $u = \frac{\sqrt{\xi} + 1}{4} \bar{u}$, $v = \sqrt{\frac{\xi - 1}{2}} \bar{v}$, (4.30) and (4.31) transforms to

$$\dot{u} = \lambda \bar{u} - \bar{u}^2 + \bar{v}^2 - \bar{\eta},$$

(4.33)

$$\dot{\bar{v}} = \lambda \bar{v} + \frac{2}{(k - 2)} \bar{u} \bar{v} - C_p \bar{v}^2,$$

(4.34)

where

(a) $\bar{\eta} = \frac{k - 2}{\sqrt{(k-1)^3}} \bar{\eta}$ and $C_p = \sqrt{\frac{(q-p+1)}{(q-p+1)-1}} \frac{k}{k-2}$.

(b) For fixed $k$, and $p \in [2, \ell]$, $C_p$ is strictly monotone decreasing in $p$ and

$$C_2 = \frac{k - 3}{k - 2} \sqrt{k},$$

$$C_\ell = \begin{cases} \frac{4\sqrt{k}}{(k+1)(k-1)(k-3)} & \text{if } k \text{ is odd} \\ \frac{2}{(k-2)} & \text{if } k \text{ is even.} \end{cases}$$

**Proof.** A straightforward computation and we omit details. \( \square \)

**Proof of proposition 4.13.** Fix $p \in [2, \ell]$. Following lemma 4.15, transform to the equations (4.33) and (4.34). We look for solutions not lying on $L_1$. That is, solutions with $\bar{v} \neq 0$. It follows from (4.34) that

$$\ddot{u} = \frac{k - 2}{2} (C_p \bar{v} - \lambda).$$

Substituting for $\ddot{u}$ in (4.33), we find that $\bar{v}$ satisfies the equation

$$\bar{v}^2 ((k - 2)^2 C_p^2 - 4) - 2\bar{v} \lambda C_p (k - 2)(k - 1) + \lambda^2 k(k - 2) + 4\bar{\eta} = 0.$$

This equation has a double root iff

$$\lambda^2 C_p^2 (k - 2)^2 - 4 \lambda (4(k - 2)^2 C_p^2 - 4)k(k - 2) + 4((k - 2)^2 C_p^2 - 4)\bar{\eta} = 0.$$

Solving for $\lambda$, and using the expressions for $C_p$ given in lemma 4.15, we find that

$$\lambda = \pm 2\sqrt{\eta} \frac{\sqrt{k - 2}}{\sqrt{k(k - 1)}} \frac{q - p + 1}{(k - 1)} \sqrt{1 - \frac{4q(p - 1)}{(q - p + 1)^2 k(k - 2)}}, \quad p > 1.$$

It is straightforward to verify that these values of $\lambda$ define saddle-node bifurcation points $\pm \gamma_{k,p}$ for the perturbed branches associated to $c_p$, $c_{p-1}$, $\lambda < 0$, and the corresponding pair of branches for $\lambda > 0$. The case $p = 1$ is easy to prove directly—take $\bar{v} = 0$ in (4.33)—but the expression for $\gamma_{k,1}$ follows by taking $p = 1$ in the formula for $\gamma_{k,p}$.

The remaining statements of the proposition follow by straightforward, though lengthy, computation. For the estimate of $\|b_p\|$, we compute the $v$-coordinate of $b_p = (u_p, v_p)$, $p > 1$,
and prove that $|v_p| \leq \sqrt{\frac{2}{k}}$, all $p \in [1, k/2)$. Finally, we show that for $k/2 > p > 1$, $|u_p/v_p|$ is uniformly bounded by 1.

The space $F_\ell = E_{\ell+1}$, $k$ odd. We assume $k = 2\ell + 1$ is odd and set $E_{\ell+1} = F_\ell$ so that

$$F_\ell = \{(x, y', z') \mid x + \ell y + \ell z = 0\}.$$

Setting $U_{\ell+1} = T_\ell$, the isometry $T_\ell : \mathbb{R}^2 \to F_\ell$ is given by

$$T_\ell(u, 0) = \frac{1}{\sqrt{k(k-1)}} (2\ell u, -u^1, -u^\ell), \quad u \in \mathbb{R}$$

$$T_\ell(0, v) = \frac{1}{\ell \sqrt{2\ell}} (0, \ell v^1, -\ell v^\ell), \quad v \in \mathbb{R}.$$

Recall from lemma 4.9 that

(a) $T_\ell(e_1) = e_1$ and $T_\ell([e_1]) = L_1$.
(b) $T_\ell^{-1}(L_1^\ast) = \{(u, -\sqrt{k} u) \mid u \in \mathbb{R}\}$.
(c) $T_\ell^{-1}(L_\ell) = \{(u, \sqrt{k} u) \mid u \in \mathbb{R}\}$.

**Proposition 4.16.** In $(u, v)$-coordinates, dynamics of $x' = \lambda x - Q(x)$ restricted to $F_\ell$ is given by

$$\dot{u} = \lambda u - \frac{1}{\sqrt{k(k-1)}} ((2\ell - 1) u^2 - v^2)$$

$$\dot{v} = \lambda v + \frac{2}{\sqrt{k(k-1)}} uv. \quad (4.35)$$

Denote the zeros of $x' = \lambda x - Q(x)$ lying on $L_1$, $L_1^\ast$ and $L_\ell$ by $c_1(\lambda)$, $c_1^\ast(\lambda)$ and $c_1(\lambda)$ respectively. We have

(a) $c_1(\lambda) = \lambda \left(\frac{\sqrt{k(k-1)}}{2}, 0\right)$.
(b) $c_1^\ast(\lambda) = \lambda \left(-\frac{\sqrt{k+1}}{2}, \frac{\sqrt{k} - 1}{2}\right)$.
(c) $c_1(\lambda) = \lambda \left(-\frac{\sqrt{k-1}}{2}, \frac{\sqrt{k+1}}{2}\right)$.

If $\lambda \neq 0$, then all zeros are of index 1 within $F_\ell$ and there are no connections between $c_1(\lambda)$, $c_1^\ast(\lambda)$ and $c_1(\lambda)$ (see figure 4).

**Proof.** The proof uses the explicit expressions for the zeros together with (4.16)–(4.18) giving the index of zeros of the phase vector field and so stabilities in directions transverse to radial direction.

**Remark 4.17.** When $k = 3$, the dynamics shown in figure 4 is that of the generic $S_3 = D_3$ bifurcation on $\mathbb{R}^2$.

**Proposition 4.18.** Let $\eta > 0$ and consider the perturbed equations on $F_\ell$

$$\dot{u} = \lambda u - \frac{1}{\sqrt{k(k-1)}} ((k - 2) u^2 - v^2) - \eta$$

$$\dot{v} = \lambda v + \frac{2}{\sqrt{k(k-1)}} uv. \quad (4.37)$$

$$\dot{u} = \lambda u - \frac{1}{\sqrt{k(k-1)}} ((k - 2) u^2 - v^2)$$

$$\dot{v} = \lambda v + \frac{2}{\sqrt{k(k-1)}} uv. \quad (4.38)$$
In terms of the parameter \( \eta \), we have a curve \( c_t(\lambda, \eta) \) of zeros such that for \( \lambda < 0 \), \( c_t(\lambda, \eta) \) is close to \( c_t(\lambda) \) and for \( \lambda > 0 \), \( c_t(\lambda, \eta) \) is close to \( c_t^*(\lambda) \). The closest approach of \( c_t(\lambda, \eta) \) to \( L_1 \) occurs when \( \lambda = 0 \), and then \( c_t(0, \eta) = (0, \sqrt{\frac{\eta}{4k}}(k-1)) \). A similar result holds for \( c_t^*(\lambda, \eta) \) with \( c_t^*(0, \eta) = c_t(0, \eta) \).

**Proof.** Straightforward computation of the perturbed zeros not lying on the axis \( L_1 \). \( \square \)

The \( S_{k-1} \)-equivariant family \( f^\eta_t(x) = \lambda x - Q(x) - \eta \varepsilon_1 \). If we regard \( H_{k-1} \) as an \( S_{k-1}(= S_1 \times S_{k-1}) \)-representation then the isotypic decomposition is \( H_{k-1} = L_1 \oplus L_1^\perp \), where \( (L_1, S_{k-1}) \) is the trivial representation and \( (L_1^\perp, S_{k-1}) \) is the standard representation of \( S_{k-1} \). Identify \( L_1 = \mathbb{R} \varepsilon_1 \) with \( \mathbb{R} \) and \( L_1^\perp = \mathbb{R}^{k-2} \). Let \( \pi_1 : \mathbb{R}^{k-1} \to \mathbb{R} \) and \( \pi_2 : \mathbb{R}^{k-1} \to \mathbb{R}^{k-2} \) denote the orthogonal projections onto \( \mathbb{R} \) and \( \mathbb{R}^{k-2} \) respectively. Denote coordinates on \( \mathbb{R} \times \mathbb{R}^{k-2} \) by \((x, y)\). Define the \( H_{k-1} \)-valued bilinear form \( B : H_{k-1}^2 \to H_{k-1} \) by

\[
B(x_1, x_2) = \frac{1}{2} [Q(x_1 + x_2) - Q(x_1) - Q(x_2)], \quad x_1, x_2 \in H_{k-1}.
\]

For all \( x \in H_{k-1} \), \( Q(x) = B(x, x) \) and for all \( x_1, x_2 \in H_{k-1} \), \( g \in S_k \), \( B(gx_1, gx_2) = gB(x_1, x_2) - B \) is \( S_k \)-equivariant.

Define \( A \in L_{S_{k-1}}(\mathbb{R}^{k-2}, \mathbb{R}^{k-2}) \), \( Q_1 \in p_{S_{k-1}}^{(2)}(\mathbb{R}^{k-2}, \mathbb{R}) \), and \( Q_2 \in p_{S_{k-1}}^{(2)}(\mathbb{R}^{k-2}, \mathbb{R}^{k-2}) \) by

\[
A(y) = 2 \pi_2 B(\varepsilon_1, y) = \alpha y, \quad y \in \mathbb{R}^{k-2}
\]
Lemma 4.19 (Notation and assumptions as above). In \((x, y)\) coordinates, the system \(\dot{x} = \lambda x - Q(x) - \eta x_1\) may be written as
\[
\begin{align*}
\dot{x} &= \lambda x - x^2 - Q_1(y) - \eta \\
\dot{y} &= (\lambda + x\alpha)y - Q_2(y),
\end{align*}
\] (4.39) \(4.40\)
where \(\alpha \in \mathbb{R}\) is uniquely determined since \((\mathbb{R}^{k-2}, S_{k-1})\) is absolutely irreducible and so \(A\) is a real multiple of \(I_{\mathbb{R}^{k-2}}\).

**Proof.** If \(x = (x, y)\), then \(Q(x) = B(xe_1 + y, xe_1 + y)\) and the result follows easily by writing \(B(xe_1 + y, xe_1 + y)\) in terms of \(Q(xe_1) = x^2e_1, Q(y)\) and \(2xB(e_1, y) = x\lambda(y) = xy\). To show \(\alpha > 0\), either compute directly or use \((4.31)\) of proposition 4.13.

**Corollary 4.20.** For \(\eta \geq 0\), the only zeros of the \(S_{k-1}\)-equivariant family \(f_3^2(x) = \lambda x - Q(x) - \eta x_1\) are those given by propositions 4.13 and 4.18. All zeros are hyperbolic except the saddle-node bifurcation points listed in proposition 4.13.

**Proof.** It follows from [16, 19] that for each non-zero value of \(\lambda + x\alpha\), (4.40) has exactly \(2^{k-2} - 1\) non-trivial hyperbolic (within \(\mathbb{R}^{k-2}\)) solutions each of which lies on an axis of symmetry for the \(S_{k-1}\)-action and so on the union of the \(S_{k-1}\)-orbits of \(E_p \cap \{x\} \times \mathbb{R}^{k-2} \times \{\lambda\}\), \(2 \leq p \leq \ell + 1\). For these solutions to extend to solutions of (4.39) and (4.40), additional conditions may have to be satisfied (depending on \(x, y, \lambda\)) but no new solutions can be created and so there are no solutions outside \(S_{k-1} \left( \bigcup \cup E_p \right)\). This leaves the question of what happens if \(\lambda = -x\alpha\). Since \(Q_2\) is a quadratic \(S_{k-1}\)-equivariant and \(k = 1\) is even, \(Q_2(y) = 0\) has solutions if \(k > 4\). Substituting in (4.39), we see that if \(\eta \geq 0\) and \(Q_2(y) = 0\), there are no solutions of \(f_3^2\) unless \(\lambda = x = 0\) (crossing solution). In other words, if \(\lambda \neq 0, \lambda, x\) have to be of opposite sign if \(Q_2(y) = 0\) which they are not (see figure 4 for the perturbed solutions \(c_i(\lambda, \eta), c_j(\lambda, \eta)\); for small enough \(\lambda, \eta\), there are no solutions on \(L_1\) on account of the saddle-node bifurcation on \(L_1\).}

**Remark 4.21.**
(a) If \(\eta < 0\), then the argument at the end of proof of corollary 4.20 fails. In this case we expect to find pitchfork bifurcations. This happens already for the case \(k = 3\) and additional symmetry breaking is then required to obtain a minimal model.
(b) The proof of corollary 4.20 implicitly relies on the Bezout’s theorem in that the number of solutions of the homogeneous equation is determined by looking for solutions of the homogeneous equation \(\lambda x - Q(x) = 0\). Introduction of the term \(-\eta\) can destroy solutions through saddle-node bifurcations but the solutions exist over the complexes and are ‘pinned’ to the corresponding complexified fixed point space. See [15, section 4.9] for more details.

Finally, some elementary symmetry and combinatorics needed for the proof of theorem 4.4.

**Lemma 4.22 (Notation and assumptions as above).** Regard \(S_{k-1}\) as the subgroup \(S_1 \times S_{k-1}\) of \(S_k\) and assume \(\ell + 1 \geq p \geq 2\).
(a) For all $\sigma \in S_{k-1}$, $\sigma | L_1 = I_{L_1}$ and so $(L_1, S_{k-1})$ is the trivial representation of $S_{k-1}$. In particular, for all $\sigma, \tau \in S_{k-1}$, $L_1 \subset \sigma E_p \cap \tau E_p$.

(b) If $P_1 = \sigma E_p$, $P_2 = \nu E_p$, then $P_1 = P_2$ iff $\sigma \nu^{-1} \in S_{p-1} \times S_q \subset S_{k-1}$.

(c) There are $\binom{k-1}{p-1}$ distinct planes in the $S_{k-1}$-orbit of $E_p$.

(d) $\sum_{p=0}^{\ell} (-1)^{j} \binom{2r+1}{j} = (-1)^{j} \binom{2r}{j}$. The proof of this is straightforward and is omitted.

**Proof.** (a) is immediate since if $x \in L_1$, the $S_k$ isotropy group of $x$ contains $S_{k-1}$. For (b) it suffices to recall that $E_p \setminus L_1 = V^{S_{p-1}} \times S_q$. Hence the set of distinct planes in the $S_{k-1}$-orbit of $E_p$ has cardinality $\binom{k-1}{p-1}$, proving (c). Finally (d) results from the binomial identity

$$\binom{m}{n} = \sum_{j=0}^{m} (-1)^{j} \binom{m+1}{j} \left( \binom{m}{n-j} + \binom{m-1}{n-1} \right).$$

4.5. Proof of theorem 4.4

We shall assume that $k = 2\ell + 1 \geq 3$ (most of the arguments below are valid for $p < \ell$ if $k$ is even). Fix $\eta > 0$ and consider the $S_{k-1}$-equivariant family

$$f_\lambda^p(x) = \lambda x - Q(x) - \eta \xi_1. \quad (4.41)$$

The first step is to show that (4.41) satisfies (d) and (e) of theorem 4.4. For $p \in [0, \ell + 1]$, set $\chi(p) = (-1)^p \left( \sum_{j=0}^{p} (-1)^{j} \binom{2r+1}{j} \right)$. With the notation of proposition 4.13, $f_\lambda^p$ has exactly $2^{2r-2}$ hyperbolic zeros if $\lambda < -\gamma_{1,k}$. As we increase $\lambda$ there is, by proposition 4.13(b), a saddle-node bifurcation at $(b_1, -\gamma_{1,p})$ in $L_1$ resulting from the collision of the trivial solution branch and the branch $c_1(\lambda, \eta) \subset L_1$ at $\lambda = -\gamma_{1,k}$. If $\ell = 1$ ($k = 3$), we have captured all the single $(\chi(0))$ saddle-node bifurcation that occurs for $\lambda \leq 0$ and $\chi(1) = 2$ branches remain of index 1. If $\ell \geq 2$, we continue to increase $\lambda$, and the remaining $\chi(1)$ branches of index $k - 2$ will collide with $\chi(1)$ branches of index $k - 3$ in $\chi(1)$-saddle-node bifurcations all occurring at $\lambda = -\gamma_{2,k} > -\gamma_{1,k}$. More precisely, by proposition 4.13(a) the curves $c_1(\lambda, \eta), c_2(\lambda, \eta) \subset E_2$ collide in a saddle-node bifurcation at $(b_2, -\gamma_{2,k})$; the other saddle-node bifurcations lie on the $S_{k-1}$-orbit of $(b_2, -\gamma_{2,k})$. There will be $\chi(2)$ remaining branches of index $k - 3$. Proceeding inductively, at the $j$th stage, assuming $p \leq \ell$, there will be $\chi(p)$-branches of index $k - p - 1$ in $\chi(p)$-saddle-node bifurcations all occurring at $\lambda = -\gamma_{p,k}$. The set of saddle-node bifurcations is given by proposition 4.13(a) and is the $S_{k-1}$-orbit of $(b_p, -\gamma_{p,k})$. The process terminates when $p = \ell$. We are then left with $\chi(\ell + 1)$ hyperbolic branches of index $\ell$. These branches connect with hyperbolic branches of index $\ell$, defined for all $\lambda \geq 0$, by proposition 4.18. Specifically, the branches $c_1(\lambda, \eta), c_2(\lambda, \eta) \subset F_\ell$ are connected at $\lambda = 0$ and then we use $S_{k-1}$-equivariance to obtain the remaining $\chi(\ell) - 1$ crossing branches. To complete the process, we now reverse the preceeding steps as we increase $\lambda$ through the sequence of bifurcation points $\gamma_{\ell,k} < \cdots < \gamma_{1,k}$. Application of lemma 4.22(d) then gives statements (c)–(e) of theorem 4.4 (with $\eta_1 = \eta_0 = +\infty$).

It remains to prove that for all $\rho > 0$, a family $f_\lambda^\rho$ can be constructed, using a perturbation supported on a neighbourhood $W_\rho$ of $(0,0) \in V \times \bar{R}$, so as to satisfy all the statements of theorem 4.4.

Fix $\eta_0 > 0$ and set $\rho = 4 \sqrt{\eta_0}, \delta_1 = \rho/2, \delta_2 = \sqrt{\eta_0}$ and $W = W_\rho$, as in the statement of the theorem. It follows from proposition 4.13(c) and (d) that for all $0 \leq \eta \leq \eta_0$, the bifurcation points of $f_\lambda^\rho$ are contained in $W_\rho/2$. That is, for all $p \in [1, \ell], S_{k-1}(\pm b_p, \pm \gamma_{p,k}) \subset W_\rho/2$. 

2846
Choose an even $C^\infty$ function $\varphi : \mathbb{R} \to \mathbb{R}$ such that
\[
\varphi(t) = \begin{cases} 
1, & \text{if } t \in [-1, 1] \\
0, & \text{if } |t| \geq 2 
\end{cases}
\] (4.42)
\[
\varphi(t) \in (0, 1), \quad \text{if } |t| \in (1, 2) 
\] (4.43)
\[
\varphi'(t) \leq 0, \quad \forall t \geq 0. 
\] (4.44)

Define the smooth $S_{k-1}$-equivariant vector field $\hat{\eta}$ on $(L_1 \oplus L_1^\perp) \times \mathbb{R}$ by
\[
\hat{\eta}(x, y, \lambda) = \varphi \left( \frac{2\lambda}{\rho} \right) \varphi \left( \frac{2|x|}{\delta_1} \right) \varphi \left( \frac{2|y|}{\delta_2} \right) \eta e_1, \quad (x, y) \in L_1 \oplus L_1^\perp, \quad \lambda \in \mathbb{R}
\]
Observe that $\hat{\eta}|_{W_{\rho}/2} = \eta e_1$ and $\hat{\eta}|_{(V \times \mathbb{R}) \setminus W_{\rho}} = 0$.

If we replace $\eta e_1$ by $\eta(\lambda) = \varphi \left( \frac{2\lambda}{\rho} \right) \eta e_1$ it is easy to see that $\eta_1$ is supported in $H_{k-1} \times [-\rho, \rho]$ and that conditions (b)-(d) of theorem 4.4 hold with $\eta \in (0, \eta_0]$. Turning to the vector field $\hat{\eta}$, the argument used for corollary 4.20 shows that no new zeros are introduced—first add the $\varphi \left( \frac{2|\lambda|}{\rho} \right)$ multiple, and use $S_{k-1}$-equivariance. Then add $\varphi \left( \frac{2|\lambda|}{\rho} \right)$ and note that for all $\eta \leq \eta_0$, no new zeros are created in $W_{\rho} \setminus W_{\rho}/2$ (using propositions 4.13 and 4.18). However, there is the possibility that stabilities of solutions could be changed in $W_{\rho} \setminus W_{\rho}/2$, $\lambda \in [-\rho, \rho]$, on account of the $x$ and $y$ derivatives of $\varphi$ that occur. However, since these derivatives of $\hat{\eta}$ are supported on a compact set, disjoint from $(0, 0) \in H_{k-1} \times \mathbb{R}$, and all multiplied by $\eta$, we can choose $C > 0$ (which will depend on $\sqrt{\rho}$) and $\eta_1 \in (0, \eta_0]$ so that (a)-(e) of theorem 4.4 are satisfied.

4.6. Theorem 4.4 and the Poincaré–Hopf theorem

Let $\text{ind}_\varphi(X)$ denote the Poincaré–Hopf index of a hyperbolic zero $z$ of the vector field $X$ ($\text{ind}_\varphi(X) = +1$ (resp. $-1$) if the index of $X$ at $z$ is even (resp. odd), see [41, section 6]). Assume $k = 2\ell + 1$ and let $Z(f_\lambda)$ denote the zero set of $f_\lambda = \lambda x - Q(x)$. Since all the zeros of $f_\lambda$ are non-singular for $\lambda \neq 0$, $-I_{k-1}$ is isotopic to $I_{k-1}$, and we may assume all zeros lie inside the sphere $S^{k-2}$ of radius 1 for $|\lambda|$ sufficiently small, it follows that $\sum_{z \in Z(f_\lambda)} \text{ind}_\varphi(f_\lambda)$ is constant on $\mathbb{R}$. Either straightforward direct computation, or theorem 4.4(e), shows that $\sum_{z \in Z(f_\lambda)} \text{ind}_\varphi(f_\lambda) = (-1)^{\ell+1} \left( \frac{2\ell}{\ell} \right)$ and so any perturbation of $f_\lambda$ to a $C^2$-stable family must have at least $\left( \frac{2\ell}{\ell} \right)$ solutions for each $\lambda \in \mathbb{R}$.

4.7. Minimal symmetry breaking model: $k$ even

Assume $k = 2\ell$ is even. Since $Q|L_\ell \equiv 0$, we need to take account of higher order terms. Recall [19, sections 16 and 17] that for $k \geq 4$ a basis for $P_{3k}(H_{k-1}, H_{k-1})$ is given by $\{T_1, T_2\}$ where
\[
T_1(x) = ||x||^2 x, \\
T_2(x) = \text{grad}(\rho_\ell)(x), \quad x \in H_{k-1} \quad \text{(example 2.5)}
\]

Lemma 4.23 (Notation and assumptions as above). For $p \in [1, \ell]$,
\[
T_1(e_p) = e_p, \\
T_2(e_p) = \alpha_p e_p.
\]
where \( \alpha_p = \frac{1}{q} \left( \frac{p}{q} + \frac{p}{q} - 1 \right) \in [\frac{1}{q}, 1] \). \( \alpha_p \) is strictly monotone decreasing on \([1, \ell]\) with minimum value of \( \alpha_1 = \frac{1}{\ell} \) and maximum value of \( \alpha_1 = 1 - \frac{1}{\ell} + O(\frac{1}{\ell^2}) \).

**Proof.** The statement for \( T_1 \) is trivial. The expression for \( T_2 \) is a straightforward computation using (4.13) and the definition of \( T_2 \). □

Every \( T \in \mathcal{P}_3^{(3)}(H_{k-1}, H_{k-1}) \) may be written uniquely as \( T = \alpha_1 T_1 + \alpha_2 T_2 \), \( \alpha = (\alpha_1, \alpha_2) \in \mathbb{R}^2 \). For \( p \in [1, \ell] \), set \( \beta_p = \beta_p(T) = \alpha_1 + \alpha_2 \alpha_p \). Define the open and dense subset \( \mathcal{T}_3 \) of \( \mathcal{P}_3^{(3)}(H_{k-1}, H_{k-1}) \) to consist of all \( T \) for which \( \beta_k(T) \neq 0 \). Since \( T(\varepsilon) = \beta \varepsilon \),

\[ T \in \mathcal{T}_3 \text{ iff } T(\varepsilon) \neq 0. \]

**Proposition 4.24.** Let \( T \in \mathcal{T}_3 \) and

\[ F_\lambda(x) = \lambda x - Q(x) + T(x). \quad (4.45) \]

(a) If \( \beta_{\ell} < 0 \) (resp. \( \beta_{\ell} > 0 \)), then (4.45) has a supercritical (resp. subcritical) pitchfork bifurcation along \( L_p \). The branches are given for \( t \in [0, \infty) \) by

\[ \lambda(t) = -\text{sgn}(\beta_{\ell}) t^2 \]

\[ x_p^\pm(t) = \pm \frac{t}{\sqrt{\text{sgn}(\beta_{\ell})}} \varepsilon_p. \]

(b) If \( p \in [1, \ell - 1] \) and \( \beta_p = 0 \), the branch along \( L_p \) is the same as that for \( x' = \lambda x - Q(x) \) (section 4.1, (B) and (F)) and there are no other non-trivial zeros of (4.45) on \( L_p \).

(c) If \( p \in [1, \ell - 1] \) and \( \beta_p < 0 \), then the forward branch of \( x' = \lambda x - Q(x) \) along \( L_p \) perturbs to \( (x_p^{+}(t), t) = t((\frac{p-2}{4p} + O(t))\varepsilon_p, t) \) and is defined for all \( t \geq 0 \). There is also a branch \( (z_p^{+}(t), t) \) along \( L_p \) for which \( \|z_p^{+}(t)\| > \frac{p-2}{4p} \sqrt{\|z_p^{+}(0)\|} \), all \( t \geq 0 \).

The backward branch of \( x' = \lambda x - Q(x) \) along \( L_p \) is perturbed to \( z_p^{+}(t) = -t((\frac{p-2}{4p} + O(t))\varepsilon_p, t) \) and is defined for \( t \in [0, -\frac{p-2}{4p} \sqrt{\|z_p^{+}(0)\|}] \). At \( t = -\frac{p-2}{4p} \sqrt{\|z_p^{+}(0)\|} \), the branch collides with the branch \( z_p^{+}(t) \) along \( L_p \) in a saddle node bifurcation and neither branch is defined for \( t > \frac{p-2}{4p} \sqrt{\|z_p^{+}(0)\|} \). We have \( \|z_p^{+}(0)\| = \frac{p-2}{2\sqrt{p\|z_p^{+}(0)\|}} \), and for all \( t < -\frac{p-2}{4p} \sqrt{\|z_p^{+}(0)\|} \), \( \|z_p^{+}(t)\| - \|z_p^{+}(0)\| > 0 \). Similar statements hold when \( \beta_p > 0 \).

**Proof.** We omit the straightforward computation. □

**Remark 4.25.** Statement (c) highlights the issues that arise when general cubic invariants are included. It will play no further role.

**Corollary 4.26 (Notation and assumptions as above).** Suppose \( T = -T_1 \) and set \( \lambda_0 = \frac{2}{k(k-1)}, R_0 = \sqrt{\lambda_0} \). The only branches of solutions to (4.45) meeting \( D_{\mathcal{R}_0}(0) \times (-\lambda_0, \lambda_0) \subset H_{k-1} \times \mathbb{R} \) are the \( \mathcal{S}_i \)-orbits of the supercritical branches \( x_p^{\pm} \) and perturbed branches \( x_p^{\pm}(t), t \in [0, \ell] \), given by proposition 4.24. In particular, the branches \( (x_p^{\pm}(t), \pm t) \), \( t \in [0, \lambda_0] \), are contained in \( D_{\mathcal{R}_0}(0) \times (-\lambda_0, \lambda_0) \) for all \( p \in [1, \ell - 1] \). The same result holds with \( T = T_1 \) (subcritical pitchfork).

Choosing \( \lambda_0 > 0 \) smaller if necessary, we may require that the indices of the branches \( x_p^{\pm}(t), p < \ell \), and \( x_p^{\pm}(t) \) are constant on \((0, \lambda_0)\).

**Proof.** The proof follows from proposition 4.24 or by direct computation of the branches. □
Remark 4.27. Unlike what happens when \( k \) is odd, there is no natural family \( f : H_{k-1} \times \mathbb{R} \to H_{k-1} \) for \( S_k \)-equivariant bifurcation if \( k \) is even since the addition of higher order terms typically results in the appearance of additional solution branches which can and do merge with the branches of interest along \( L_p, p < \ell \). However, as indicated by the proposition, the signed indexed branching pattern is uniquely determined by the sign of \( \beta_\ell \). In particular, we may replace \( T \in T_3 \) by \( \text{sgn}(\beta_\ell)T_1 \) without changing the signed indexed branching pattern. What we shall do is modify \( \lambda x - Q(x) + \|x\|^2x \) to define an \( S_k \)-equivariant family that models the bifurcation and has only the branches along \( S_kL_p, p < \ell \), given by the previous model for \( k \) odd, and only super- or subcritical branching along \( S_kL_\ell \).

Let \( Z = Z(P_Q) \subset \mathbb{S}^{k-2} \) be the zero set of \( P_Q \). Let \( \rho \) denote the standard \( O(k-1) \)-invariant metric on \( \mathbb{S}^{k-2} \) and set \( \kappa = \min_{u,v \in Z} \rho(u,v) \). For \( \tau > 0 \), define the \( S_\ell \times S_\ell \)-invariant closed neighbourhood \( B_\tau \) of \( \varepsilon_\ell \) by

\[
B_\tau = \{ u \in \mathbb{S}^{k-2} | \rho(u, \varepsilon_\ell) \leq \tau \}.
\]

Choose \( \tau \ll \kappa \), for example \( \tau = \kappa/100 \), so that the \( S_\ell \)-orbit of \( B_\tau \) is a set of disjoint disc neighbourhoods of the zeros in \( S_\ell \varepsilon_\ell \subset Z(P_Q) \) and \( S_\ell B_\tau \cap Z(P_Q) = S_\ell \varepsilon_\ell \). Choose a smooth \( S_\ell \times S_\ell \)-invariant function \( \psi : B_\tau \to \mathbb{R} \) satisfying

\[
\psi(u) = \begin{cases} 
1, & \rho(u, \varepsilon_\ell) \leq \tau/2 \\
0, & \rho(u, \varepsilon_\ell) \geq 3\tau/4 \\
\in (0, 1), & \rho(u, \varepsilon_\ell) \in (\tau/2, 3\tau/4).
\end{cases}
\]

Extend \( \psi \) \( S_k \)-equivariantly to \( S_k B_\delta \) and then to \( \mathbb{S}^{k-2} \) by taking \( \psi \equiv 0 \) on \( \mathbb{S}^{k-2}\setminus S_k B_\tau \). Thus \( \psi : \mathbb{S}^{k-2} \to [0, 1] \subset \mathbb{R} \) will be a \( C^\infty \) \( S_k \)-equivariant map equal to 1 on \( S_k B_{\tau/2} \) and equal to zero outside \( S_k B_{3\tau/4} \).

Choose \( \varphi \in C^\infty(\mathbb{R}) \) satisfying (4.42)–(4.44). Define the \( C^\infty \) \( S_k \)-equivariant radial vector field \( K \) on \( H_{k-1} \times \mathbb{R} \) by

\[
K(x, \lambda) = [\varphi(2R/R_0)\varphi(2\lambda/\lambda_0) + (1 - \varphi(2R/R_0)\varphi(2\lambda/\lambda_0))\psi(u)]T_1(x),
\]

where \( R_0, \lambda_0 > 0 \) are given by corollary 4.26 and \( x = Ru \) (\( R = \|x\| \), and \( u = x/\|x\|, x \neq 0 \)). Define the family \( F^\pm_{\lambda} \) on \( H_{k-1} \) by

\[
F^\pm_{\lambda}(x) = \lambda x - Q(x) \pm K(x, \lambda).
\]

Proposition 4.28. The family \( F^\pm_{\lambda} \) has exactly \( 2^k - 2 \) non-trivial solution branches. All of these branches are axial and

(a) If \( p \in [1, \ell - 1] \), then the forward (resp. backward) branch \( (x^+_p(t), t) \) (resp. \( (x^-_p(t), -t) \)) along \( L_p \) is defined for all \( t \geq 0 \) and consists of hyperbolic equilibria of index \( p \) (respectively \( k - p - 1 \)). The eigenvalues of \( DF_{\pm\lambda}(x^\pm_p(t)) \) corresponding to eigendirections transverse to \( L_p \) are given by the eigenvalues of the Hessian of \( P_Q \) at \( \pm \varepsilon_p \).

These results are independent of the choice of \( F^+_{\lambda} \) or \( F^-_{\lambda} \).

(b) If \( p = \ell \), then \( F^-_{\lambda} \) (resp. \( F^+_{\lambda} \)) has a supercritical (resp. subcritical) pitchfork bifurcation along \( L_{\ell} \) and the branches are given by \( (x^+_\ell(t), \lambda(t)) = (\pm \varepsilon_{\ell,1}, t^2) \) (resp. \( (x^-_\ell(t), \lambda(t)) = (\pm \varepsilon_{\ell,1}, -t^2) \), \( t \geq 0 \). These branches are hyperbolic of index \( \ell \).

2849
Remark 4.29. Every solution of the family (4.46) lies on a solution branch starting at the bifurcation point that are of the form

Proof. Transform (4.46) to spherical coordinates (\(R, u\)) to obtain

\[
R' = \lambda R - R^2 (Q(u), u) + R^3 H(R, u, \lambda) \\
u' = R P_Q(u),
\]

(4.47)

(4.48)

where the \(C^\infty\) scalar function \(H\) takes values in \([0, 1]\) and is equal to zero iff (a) \(R \geq R_0\) or \(|\lambda| \geq \lambda_0\), and (b) \(u \notin S_1 B_{3,1}\). If \((\tilde{R}, \tilde{u})\) is a zero of (4.47) and (4.48) with \(\tilde{R} \neq 0\), then \(P_Q(\tilde{u}) = 0\). Hence \(\tilde{u}\) (and \(\tilde{x} = \tilde{R} u\)) must lie on an axis of symmetry \([19]\). Substituting in (4.47), and cancelling an \(\tilde{R}\) factor, \(\tilde{R}\) satisfies

\[
\lambda - \tilde{R}(Q(\tilde{u}), \tilde{u}) + \tilde{R}^3 H(\tilde{R}, \tilde{u}, \lambda) = 0.
\]

(4.49)

If \(\tilde{u} \in S_1 L_\ell\), then \(Q(\tilde{u}) = 0\) and so, since \(H\) is equal to one on a neighbourhood of \(S_1 L_\ell \times \mathbb{R}\) in \(H_{k-1} \times \mathbb{R}\), \(\lambda \pm \tilde{R}^2 = 0\), proving statement (b) of the proposition.

It remains to complete the proof of (a). The solutions of \(F^\pm(x, \lambda) = 0\) on \(\{ (x, \lambda) \in H_{k-1} \times \mathbb{R} \mid K(x, \lambda) = 0 \}\) are given by the solutions of \(f(x, \lambda) = \lambda x - Q(x) = 0\)—that is, the solution branches of \(f = 0\) along axes in \(\cup_{p \in [1, \ell-1]} S_{k-1} L_p\). It follows from corollary 4.26 and the definition of \(S\) that the only zeros of (4.47) with \(u = \tilde{u}\) and \(x, \lambda\) lying in the support of \(K\) are those described by corollary 4.26.

Remark 4.29. Every solution of the family (4.46) lies on a solution branch starting at the bifurcation point \((0, 0) \in H_{k-1} \times \mathbb{R}\) and defined for all \(t \geq 0\). The index and direction of branching are constant on each branch and there are no spurious solutions resulting from the presence of higher order polynomial terms.

Just as for the case when \(k\) is odd, we consider symmetry breaking perturbations of \(F^\pm\) that are of the form \(F^\pm_\eta = F^\pm - \eta e_1\), where \(\eta > 0\). The main new feature is the effect of the perturbation on the supercritical branches that occur along axes in \(S_1 L_\ell\).

Theorem 4.30. Assume \(k = 2\ell\) is even and let \(F^\pm(x, \lambda)\) be the model (4.46) as described in proposition 4.28. Given a compact neighbourhood \(W\) of \((0, 0) \in H_{k-1} \times \mathbb{R}\), there exist \(C, \eta_1 > 0\) (depending on \(W\) and \(k\)) such that for \(\eta \in (0, \eta_1]\) there is a smooth \(S_{k-1}\)-equivariant family \(F^\pm_\eta\) satisfying

(a) \(F^\pm_\eta(x) = F^\pm_{\lambda_0}(x)\) if \((x, \lambda) \notin W\).

(b) \(\|F^\pm_\eta - F^\pm\|_{W,3} \leq C\eta\).

(c) The only bifurcations of the family \(F^\pm_\eta\) are saddle-node bifurcations and \(F^\pm_\eta\) is stable under \(C^2\)-small perturbations supported on a compact neighbourhood of \((0, 0) \in H_{k-1} \times \mathbb{R}\).

(d) The family \(F^\pm_\eta\) has exactly \(2^{2\ell-1} - \left(\frac{2\ell-1}{\ell}\right)\) saddle node bifurcations. Specifically, all of the branches of solutions with index \(\ell\) will end or start with a saddle-node bifurcation (connecting possibly to a branch of index \(\ell\)).

(e) There are exactly \(\left(\frac{2\ell-1}{\ell}\right)\) crossing curves \((x, \lambda) : \mathbb{R} \to H_{k-1} \times \mathbb{R}\) of solutions to \(F^\pm_\eta = 0\); each of these curves consists of hyperbolic saddles of index \(\ell\).

Example 4.31. In figure 5, we illustrate theorem 4.30 in the case \(k = 6\) and the supercritical bifurcation is forward. Here exactly half the of the 20 branches generated by the supercritical pitchfork bifurcations along \(S_6 L_3\) connect to 10 index 3 branches to give 10 crossing curves, the remaining 10 branches join 10 index 2 branches through saddle-node bifurcations. In more detail, assume \(\eta > 0\) is sufficiently small. Starting with \(\lambda < 0\), the index 5 trivial solution...
A count verifies there are 10 crossing curves (index 3) and 22 saddle-node bifurcations.

In large part the proof of theorem 4.30 follows that of theorem 4.4 and only two additional results are needed related to the presence of pitchfork bifurcation along axes lying in the $S_k$-orbit of $\varepsilon_1$. For this we need to quantify dynamics on $E_1 = \{(x, y, z) | x + (\ell - 1)y + \ell z = 0\}$. Recall from section 4.4 that the map $U_\ell : \mathbb{R}^2 \to E_1$ defined by

$$U_\ell(u, 0) = \frac{1}{\sqrt{k(k-1)}} (k-1)u, -u^{\ell-1}, -u^\ell)$$

$$U_\ell(0, v) = \frac{2}{\sqrt{k(k-1)(k-2)}} (0, \ell u^{\ell-1}, -(\ell - 1)u^\ell)$$

is an isometry and $E_1$ contains the axes of symmetry

$$L_1 = \mathbb{R} \varepsilon_1 = \{U_\ell(u, 0) | u \in \mathbb{R}\}$$
Let $f \in C^1$. Proposition 4.33.

A similar result holds for $f^+ \in C^1$.

**Proof.** Statements (a) and (b) follow from proposition 4.12 (the $O(\lambda^3)$ terms come from $-\|x\|^2$; (c) is a standard computation.

**Proposition 4.33.** Let $\eta > 0$. The perturbed equations on $E_t$ are

\begin{align*}
\dot{u} &= \lambda u - \frac{1}{\sqrt{k(k-1)}}((k-2)u^2 - v^2) - (u^2 + v^2)u - \eta \\
\dot{v} &= \lambda v + \frac{2}{\sqrt{k(k-1)}}uv - \frac{2}{\sqrt{k(k-1)(k-2)}}v^2 - (u^2 + v^2)v.
\end{align*}

In terms of the parameter $\eta$, we have a curve $c^+(-\lambda, \eta)$ of non-singular zeros of index $\ell$ such that for $\lambda \ll 0$, $c^-(-\lambda, \eta)$ is close to $c^+(-\lambda, \eta)$ and for $\lambda \gg 0$, $c^+(-\lambda, \eta)$ is close to $c^+(-\lambda, \eta)$. There is also a branch $c^+(\lambda, \eta) \subset E_t \times \mathbb{R}^+$ with a single saddle-node bifurcation near $\lambda = 0.$ The index along the branch changes from $\ell$ to $\ell - 1$, with the index $\ell$ component approximating $c^+(-\lambda, \eta)$ and the index $\ell - 1$ component approximating $c^+(-\lambda, \eta)$ (see figure 6). The $S_{\ell-1}$-equivariance implies analogous results for all the curves lying in the $S_{\ell-1}$-orbit of $E_t$. 

---

**Figure 6.** Zeros, dynamics and bifurcation on $E_t$, $k = 2\ell$. Dynamics for $\lambda < 0$ is shown in blue, that for $\lambda > 0$ in red, and new branches/connections in purple. Note that $\lambda_-$ (resp. $\lambda_+$) denotes the $\eta$-dependent value of $\lambda$ at which the branch of sinks (resp. sources) meets the index $k - 2$ (resp. 1) branch in a saddle-node bifurcation.

$L^*_t = (1, \ell) \mathbb{R} e_{\ell - 1} = \left\{ U_t \left( u, -k \sqrt{\frac{1}{k-2}} u \right) \mid u \in \mathbb{R} \right\}$, where $(1, \ell) \in S_k$.

$L_t = \mathbb{R} e_t = \{ U_t(u, \sqrt{(k-2)}u) \mid u \in \mathbb{R} \}$.

**Lemma 4.32.** Let $f^\lambda(x) = \lambda x - Q(x) - \|x\|^2 x$. The zeros of $f^\lambda|E_t$ are $c_1(\lambda) \in L_t$, $c^+_{t-1}(\lambda) \in L^*_t$ and, for $\lambda > 0$, $c^+_t(\lambda) \in L_t$, where

(a) $c_1(\lambda) = \lambda \sqrt{k(k-1)}U_t(1,0) + O(\lambda^2)$,

(b) $c^+_{t-1}(\lambda) = \frac{1}{k} \sqrt{\frac{1}{k(k-1)}}U_t\left(-2k, k\sqrt{k-2}\right) + O(\lambda^2)$

(c) $c^+_{t}(\lambda) = \pm \sqrt{\frac{\lambda}{(k-1)}}U_t(1, k\sqrt{k-2}) = \pm \sqrt{\frac{1}{\sqrt{k}}} \left( \left( \frac{1}{\sqrt{k}} \right)^t, \left( \frac{1}{\sqrt{k}} \right)^{t+1} \right)$.

A similar result holds for $f^+_\lambda$. 

---

2852
Proof. Similar to that of proposition 4.18 and we omit the details. □

Proof of theorem 4.30. The proof is broadly similar to that of theorem 4.4. First, the saddle-node bifurcations of the perturbed branches of index not equal to $\ell, \ell \pm 1$ are handled along the same lines as in the proof of theorem 4.4. The crossing branches, and saddle-node bifurcation between index $\ell$ and index $\ell - 1$ branches, use proposition 4.33. Finally, for localisation, constants are chosen, as in the proof of theorem 4.4, so that all the perturbation and bifurcation occurs with the assigned neighbourhood $W$. Typically, this will require $\eta_1 > 0$ to be small. In order to show that no new zeros are introduced, we use the known result that $Q$ is of relatively hyperbolic type [18, sections 4 and 10], [19, section 16.2.5]. This implies that, for sufficiently small $\eta > 0$, all the zeros within $W$ are pinned to fixed point spaces (even if no longer real) and by Bezout’s theorem no new zeros are introduced (same argument as in the proof of theorem 4.4). □

Remark 4.34.

(a) The index $\ell$ crossing curves in theorem 4.30 may be compared with unfoldings of the pitchfork bifurcation given in Golubitsky and Schaeffer [21, III, figure 7.2 (1) and (2)]. At first glance, the scenario described above seems to replicate these unfoldings. However, the crossing curves associated to the index $\ell$ branches arise from bifurcation of two-dimensional dynamics in the pencil of two-planes $S_{\ell-1}E_\ell$ through $L_1$, as shown in figure 6 for $E_2$. Consequently, the analysis of the bifurcation cannot be reduced to the study of bifurcation on an axis of symmetry contained in a center manifold.

(b) If $W$ is fixed and $\eta \in \mathbb{R}$ is variable in the minimal model theorems 4.4 and 4.30, then the $(\eta, \lambda)$-dependent family $f^\eta$ may be viewed as an unfolding of $f$ since $\lim_{\eta \to 0} f^\eta = f$ in the $C^\infty$ topology (note the definition of $\eta$ in the proof of theorem 4.4).

5. Concluding comments

The focus in this article has been on the creation of local minima and forced symmetry breaking for the standard representation of $S_k$. Motivation for this work came from an analysis of symmetry properties of a student-teacher shallow neural net used for theoretical investigations in machine learning (we refer to [4] for background and references). In the simplest case, when the number of neurons $k$ equals the number of inputs $d$, the weight space for the student-teacher network is the space $M(k, k)$ of $k \times k$-matrices. If we fix a target weight $V \in M(k, k)$, then the loss is defined by

$$L(W) = \frac{1}{2} \mathbb{E}_{x \sim \mathcal{N}(0, I_d)} \left( \sum_{i \in k} \sigma(w^i x) - \sum_{i \in k} \sigma(v^i x) \right)^2, \quad W \in M(k, k),$$

where $\sigma$ is the ReLU activation function defined by $\sigma(t) = \max\{0, t\}, t \in \mathbb{R}$; $w^i$ denotes the $i$th row of $W \in M(k, k)$; $E$ denotes expectation with respect to the multivariate normal distribution $\mathcal{N}(0, I_d)$ (see [4, section 4] for details, background and an explicit formula for $L(W)$). The key properties of the distribution $\mathcal{N}(0, I_k)$ are orthogonal invariance (this can be weakened) and equivalence to Lebesgue. Note that the minimum value of the parenthesised expression is zero (take $W = V$).

There is a natural action of $\Gamma \overset{\text{def}}{=} S_k^r \times S_k^c$, where $S_k^r$ is defined by permuting rows $S_k^r$-factor) and columns (S_k^c-factor). It is trivial that $L$ is always $S_k^r$-invariant. If we take $V = I_k \in M(k, k)$, then $L$ is $\Gamma$-invariant [2, 4] and the global minimum zero of $L(W)$ is attained.
iff $W \in \Gamma V$ [4, proposition 4.14]. The isotropy group $\Gamma_V$ is equal to $\Delta S_k$ (diagonal subgroup of $\Gamma$). Henceforth, we write $S_p$ rather than $\Delta S_p$, $p \in \mathbb{k}$—only actions of subgroups of $\Delta S_k \approx S_k$ are considered.

In the introduction, we remarked that under gradient descent (or SGD), there is a transition from saddle to local minimum at $k \approx 5.58$ [4, 5]. These minima are strictly positive and are usually referred to as *spurious minima*. They were first seen numerically in this problem for $k \in [6, 20]$ [46]. The associated critical points of the loss function have isotropy conjugate to $S_{k-1} = S_{k-1} \times \{e\}$ and, following [4], we refer to these critical points (or the minima) as being of type II. We indicate next how the results of the article help to understand this transition.

A full analysis of the stability of type II critical points requires the isotypic decomposition of $S_{k-1}$ acting on $M(k,k)$ (the decomposition is independent of $k \geq 5$ and $s_{k-1}$ has multiplicity 5 in the $S_{k-1}$-representation $M(k,k)$ [3, theorem 4]). The fixed point space $F_{k-1,1} \subset M(k,k)$ of $S_{k-1} \times \{e\} \subset S_k$ is of dimension 5 (independent of $k \geq 3$) and contains one critical point $e_5$ of type II. In the standard way, we reduce the analysis to $F_{k-1,1}$. For gradient descent dynamics on $F_{k-1,1}$, $e_5$ is a sink for dynamics restricted to $F_{k-1,1}$, $k \geq 3$—and is easy to find numerically using gradient descent on $F_{k-1,1}$. For dynamics on $M(k,k)$, $e_5$ is a saddle point for $k \leq 5$, and a local minimum if $k \geq 6$ (detectable by gradient descent or SGD not initialised on $F_{k-1,1}$ [46]).

The change in stability of $e_5$ on $M(k,k)$ can be shown by spectral analysis of the Hessian which verifies a change in sign of eigenvalues (using the isotypic decomposition) associated to a bifurcation tangent to a copy of $S_{k-1}$ at $k \approx 5.58$ [3, 5]. The representations $S_4$ and $S_5$ both have two conjugacy classes of axes of symmetry. For generic bifurcation on $S_4$, we expect (generically) $4 + 6d$ branches of hyperbolic saddles of low index to collide with the source at the bifurcation point where $\delta = 1$ (resp. 0) if there are (resp. are not) pitchfork branches $k \leq 5$. Similarly, at $k = 6$, we expect 15 branches of hyperbolic saddles of high index to collide with the sink at the bifurcation point, $k < 6$. The hyperbolic saddles should lie on axes of symmetry. In particular, if $k = 5$, we expect the saddle points to have isotropy conjugate to either $S_4 = S_4 \times \{e\} \subset S_0$ or $S_2 \times S_2 \subset S_4$. If $k = 6$, the saddle points have isotropy conjugate to either $S_4 = S_4 \times \{e\} \subset S_5$ or $S_2 \times S_2 \subset S_5$. The associated fixed point spaces $F_{k-2,1,1}, F_{k-3,2,1}$ have dimensions 10 and 11 respectively and $F_{k-2,1,1} \cap F_{k-3,2,1} = F_{k-1,1}$, for all $k \geq 5$. The natural conjecture is that, within the fixed point spaces, bifurcations along all axes of symmetry occur at the same value $k \approx 5.58$ found by the Hessian analysis and that this a general phenomenon for the creation of spurious minima. For type II critical points, numerical checking of the conjecture shows that bifurcation along axes of symmetry does occur at $k \approx 5.58$.

Although the bifurcation does not occur at an integer value of $k$, analysis of the Hessian and dynamics on fixed point spaces, strongly indicate that at $k = 5$ (resp. $k = 6$), $L$ is close to a generic bifurcation at $e_5$ (resp. $e_6$) along a centre manifold tangent to $S_5$ (resp. $S_6$). If this is so, there are strong implications about the existence of saddle points near $e_5, e_6$. Moreover the minimal unfolding results given in section 4 have implications for forced symmetry breaking—here, of the target $V$ implicit in the definition of the loss function $L$. For example, the possibility of a small range of $k$ values for the perturbed system for which there are no sinks or sources. In general, it is difficult to ‘find’ critical points numerically which are saddles in a fixed point space (gradient descent is not helpful) and this is a significant issue when $k$ is large.

In recent work [6], we show that the standard representation also plays a key role in the annihilation of spurious minima in the *overparametrized* case: $k > d$. For example, if we add two neurons to the hidden layer ($k = d + 2$), then we annihilate type II spurious minima through a similar mechanism to that described above: bifurcation on the standard representation $S_{k-1}$. In particular, no new spurious minima appear. This result holds for all $d, k = d + 2$ and is robust to symmetry perturbation of the target model $V$. 

2854
The creation of spurious minima is not an artifact of the existence of non-smooth points of the loss function. For example, spurious minima can occur when polynomial activation is used in optimisation problems of symmetric tensor decomposition [1].

For large values of $k$, the minimal model we construct implies a high-dimensional local landscape deformation for the creation of spurious minima without introducing additional spurious minima with, for example, lower symmetry. Other constructions of relatively simple minimal models are surely possible—for example, changing the sign of $\eta$ in the perturbation $\eta \varepsilon_1$ used in our construction. This already results in interesting geometry and pitchfork bifurcations in case $k = 3$. Although the generic bifurcation on the standard representation is often viewed as being ‘transcritical’, the reality is that even if $k$ is odd, pitchfork bifurcations play a significant role in understanding dynamics. This is not so surprising, as bifurcations along axes when $k$ is odd are never a simple exchange of stability—whatever the dynamics appear to be by restricting attention to the fixed point space. Rather they are a stability inversion—a consequence of the presence of quadratic equivariants and analysis of the phase vector field (cf (4,20)). We believe the interest of quadratic equivariants lies in this point (rather than just the existence of unstable solution branches).

For reasons of exposition, we have restricted attention to the representation $s_k$. However, our methods likely extend without difficulty to the external products $s_k \boxtimes s_n$, $k, n \geq 3$. Indeed, the space of quadratic gradient equivariants for these representations is one-dimensional and, although we have not checked all the details, every homogeneous quadratic equivariant is likely gradient, as is the case for $s_k$. These irreducible representations may well occur in the mechanisms leading to the annihilation of spurious minima.

Certain spurious minima are associated with bifurcation tangent to the exterior square representation of $S_k$. These minima do not decay to zero as $k \to \infty$ [4, section 8] and are not seen in [46], where Xavier initialisation is used. Thus far, there is no evidence that the exterior square representation plays a role in problems related to annihilation of families of spurious minima [6].

Finally, the phenomena we have described for type II minima with isotropy $S_{k-1}$ (that is, $\Delta S_{k-1} \subset S_k^r \times S_k^r$) also occur for families of spurious minima with isotropy $S_{k-p} \times S_p$, where $p \ll k$ (these minima are referred to as type M in [4]). Modulo terms of order $O(k^{-2})$, these critical points exhibit the same Hessian spectrum [5, theorem 1] suggesting the possibility of underlying self-similar structure in the landscape geometry of $\mathcal{L}$.

Acknowledgments

Many thanks to the referees for their careful reading of the paper and many comments and suggestions. Special thanks to Ian Melbourne, for pointing out his result with Pascal Chossat and Reiner Lauterbach on the instability of branching [9, theorem 4(b)]; to Adam Parusiński and Laurentiu Paunescu for their help with the regular arc-wise analytic stratification; and to David Trotman for his help with Pawlucki’s theorem. Last, but not least, MF would like to express his gratitude to Tzee-Char Kuo for his wisdom and many rewarding mathematical discussions over the years on singularity and stratification theory, for recently telling us about the arc-wise analytic stratification of Parusiński & Paunescu, and for his variation on Steenrod’s comment about the importance of having the correct definition: the definition is wrong, hence the whole theory is wrong [30, sections 8 and 2].
References

[1] Arjevani Y, Bruna J, Field M, Kileel J, Trager M and Williams F 2021 Symmetry breaking in symmetric tensor decomposition (arXiv:2103.06234)

[2] Arjevani Y and Field M Spurious local minima of shallow ReLU networks conform with the symmetry of the target model (arXiv:1912.11939)

[3] Arjevani Y and Field M 2020 Analytic characterization of the Hessian in shallow ReLU models: a tale of symmetry Advances in Neural Information Processing Systems (NeurIPS) (Vancouver, Canada) p 33 (arXiv:2008.01805)

[4] Arjevani Y and Field M 2021 Symmetry & critical points for a model shallow neural network Physica D 427 133014

[5] Arjevani Y and Field M 2021 Analytic study of families of spurious minima in two-layer ReLU neural networks: a tale of symmetry II Advances in Neural Information Processing Systems (NeurIPS) (Sydney, Australia) p 34 (arXiv:2107.10370)

[6] Arjevani Y and Field M Annihilation of spurious minima in two-layer ReLU networks (submitted)

[7] Aronson D G, Golubitsky M and Krupa M 1991 Coupled arrays of Josephson junctions and bifurcation of maps $S_3$ symmetry Nonlinearity 4 861

[8] Bierstone E 1977 Generic equivariant maps Real and Complex Singularities, Oslo 1976: Proc. 9th Nordic Summer School/NAVFSymp. Math. (Sijthoff and Noordhoff, Leyden) pp 127–61

[9] Chossat P, Lauterbach R and Melbourne I 1990 Steady-state bifurcation with $O(3)$ symmetry Arch. Ration. Mech. Anal. 113 313–76

[10] Chossat P and Lauterbach R 2000 Methods in Equivariant Bifurcations and Dynamical Systems (Advanced Series in Nonlinear Dynamics vol 15) (Singapore: World Scientific)

[11] Cicogna G 1981 Symmetry breakdown from bifurcation Lett. Nuovo Cimento 31 600–2

[12] Field M J 1980 Equivariant dynamical systems Trans. Am. Math. Soc. 259 185–205

[13] Field M 1989 Equivariant bifurcation theory and symmetry breaking J. Dyn. Differ. Equ. 1 369–421

[14] Field M J 1996 Symmetry breaking for compact Lie groups Memoir. Am. Math. Soc. 120

[15] Field M J 2007 Dynamics and Symmetry (Advanced Texts in Mathematics vol 3) (London: Imperial College Press)

[16] Field M J and Richardson R W 1989 Symmetry breaking and the maximal isotropy subgroup conjecture for reflection groups Arch. Ration. Mech. Anal. 105 61–94

[17] Field M J and Richardson R W 1990 Symmetry breaking in equivariant bifurcation problems Bull. Am. Math. Soc. 22 79–84

[18] Field M J and Richardson R W 1992 Symmetry-breaking and branching patterns in equivariant bifurcation theory: I Arch. Ration. Mech. Anal. 118 297–348

[19] Field M J and Richardson R W 1992 Symmetry breaking and branching patterns in equivariant bifurcation theory: II Arch. Ration. Mech. Anal. 120 147–90

[20] Fulton W and Harris J 1991 Representation Theory (Graduate Texts in Mathematics vol 129) (Berlin: Springer)

[21] Golubitsky M and Schaeffer D G 1988 Singularities and Groups in Bifurcation Theory (Appl. Math. Sci. Ser. vol 51) (New York: Springer)

[22] Golubitsky M, Stewart I N and Schaeffer D G 1988 Singularities and Groups in Bifurcation Theory (Appl. Math. Sci. Ser. vol 69) (New York: Springer)

[23] Golubitsky M and Stewart I 2002 The Symmetry Perspective: From Equilibrium to Chaos in Phase Space and Physical Space vol 200 (Basel: Birkhäuser)

[24] Ihrig E and Golubitsky M 1984 Pattern selection with $O(3)$ symmetry Physica D 13 1–33

[25] James G D 1978 The Representation Theory of the Symmetric Groups (Springer Lecture Notes in Mathematics vol 682) (Berlin: Springer)

[26] Kato T 1976 Perturbation Theory for Linear Operators (Grundlehren vol 132) (Berlin: Springer)

[27] Knopp K 1996 Theory of Functions, Parts I and II (New York: Dover)

[28] Krupa M 1990 Bifurcations of relative equilibria SIAM J. Math. Anal. 21 1453–86

[29] Koenig M 1997 Linearization of vector fields on the orbit space of a compact Lie group Math. Proc. Camb. Phil. Soc. 121 401–24

[30] Kuo T C 2017 An old man’s mathematical stories Proc. JARCS Australian-Japanese Real and Complex Singularities Workshop

2856
[31] Lamb J S W and Melbourne I 2007 Normal form theory for relative equilibria and relative periodic solutions Trans. Am. Math. Soc. 359 4537–57
[32] Lamb J S W, Melbourne I and Wulff C 2003 Bifurcation from periodic solutions with spatiotemporal symmetry, including resonances and mode interactions J. Differ. Equ. 191 377–407
[33] Lauterbach R 2015 Equivariant bifurcation and absolute irreducibility in \(\mathbb{R}^3\): a contribution to ize conjecture and related bifurcations J. Dyn. Differ. Equ. 27 841–61
[34] Lauterbach R and Matthews P 2010 Do absolutely irreducible group actions have odd dimensional fixed point spaces? (arXiv:1011.3986)
[35] Lauterbach R and Schwenker S N 2017 Equivariant bifurcations in four-dimensional fixed point spaces Dyn. Syst. 32 117–47
[36] Łojasiewicz S 1995 On Semi-Analytic and Subanalytic Geometry vol 34 (Warsaw: Banach Center Publications) pp 89–104
[37] Malgrange B 1966 Ideals of Differentiable Functions (London: Oxford University Press)
[38] Mather J N 1973 Stratifications and mappings Proc. Dynamical Systems Conf. (Salvador, Brazil) ed M Peixoto (New York: Academic) pp 195–232
[39] Mather J N 1977 Differentiable invariants Topology 16 145–55
[40] Melbourne I 1994 Maximal isotropy subgroups for absolutely irreducible representations of compact Lie groups Nonlinearity 7 1385–93
[41] Milnor J 1997 Topology from the Differentiable Viewpoint (Princeton, NJ: Princeton University Press)
[42] Milnor J 1968 Singular Points of Complex Hypersurfaces (Annals of Mathematics Studies vol 61) (Princeton, NJ: Princeton University Press)
[43] Parusiński A and Paunescu L 2017 Arc-wise analytic stratification, Whitney fibering conjecture and Zariski equisingularity Adv. Math. 309 254–305
[44] Pawłucki W 1985 Quasi-regular boundary and Stokes formula for a sub-analytic leaf Seminar on Deformations (Springer Lectures Notes in Mathematics vol 1165) (Łódź-Warsaw 1981–1983) pp 235–52
[45] Rellich F 1969 Perturbation Theory of Eigenvalue Problems (Lecture Notes Reprinted by New York University) (New York: Gordon and Breach)
[46] Safran I and Shamir O 2018 Spurious local minima are common in two-layer ReLU neural networks Proc. 35th Int. Conf. Machine Learning vol 80 pp 4433–41 (for data sets, see https://github.com/ItaySafran/OneLayerGDconvergence)
[47] Trotman D 2020 Stratification theory Handbook of Geometry and Topology of Singularities I ed J L ed Cisneros-Milina, D T Lê and S José (Berlin: Springer)
[48] Vanderbauwhede A 1982 Local Bifurcation and Symmetry (Research Notes in Mathematics vol 75) (Boston, MA: Pitman)
[49] Wulff C, Lamb J S W and Melbourne I 2001 Bifurcation from relative periodic solutions Ergod. Theory Dynam. Syst. 21 605–35
[50] Walker R J 1978 Algebraic Curves (Berlin: Springer)