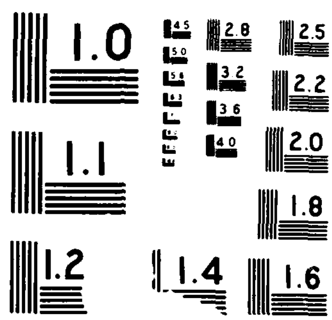
AD-A192 353 GENERALIZED JORDAN CHAINS AND TWO BIFURCATION THEOREMS 1/1
OF KRASNOSELSKII (U) PITTSBURGH UNIV PA INST FOR
COMPUTATIONAL MATHEMATICS AND APP P J RABIER
UNCLASSIFIED 02 JAN 88 ICMA-88-115 AFOSR-TR-88-0246 F/G 12/2 NL


1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.2
1.4
1.6

AD-A192 353

INSTITUTE FOR COMPUTATIONAL MATHEMATICS AND APPLICATIONS

Technical Report ICMA-88-115 January 1988

GENERALIZED JORDAN CHAINS AND TWO BIFURCATION THEOREMS OF KRASNOSELSKII

by

Patrick J. Rabier

Institute of Computational Mathematics and Applications

Department of Mathematics and Statistics

University of Pittsburgh

(2)

Technical Report ICMA-88-115 January 1988

GENERALIZED JORDAN CHAINS AND

TWO BIFURCATION THEOREMS OF KRASNOSELSKII

by

Patrick J. Rabier

Institute of Computational Mathematics and Applications
Department of Mathematics and Statistics
University of Pittsburgh
Pittsburgh, PA 15260

January, 1988

This work was in part supported under Air Force contract AFOSR-84-0131.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION: Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY: UNCLASSIFIED	3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release, distribution unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S): AFOSR-TR- 88-0246
6a. NAME OF PERFORMING ORGANIZATION: University of Pittsburgh; 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION: AFOSR
6c. ADDRESS (City, State and ZIP Code): Pittsburgh, PA 15260	7b. ADDRESS (City, State and ZIP Code): BLDG #410 Bolling AFB, DC 20332-6448
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: AFOSR; 8b. OFFICE SYMBOL (If applicable): NM	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: AFOSR- 84-0131
8c. ADDRESS (City, State and ZIP Code): BLDG #410 Bolling AFB, DC 20332-6448	10. SOURCE OF FUNDING NOS.: PROGRAM ELEMENT NO. 61192F; PROJECT NO. 2304; TASK NO. A3; WORK UNIT NO

11. TITLE (Include Security Classification): Generalized Jordan Chains and Two Bifurcation Theorems of Krasnoselskii

12. PERSONAL AUTHOR(S): Rabier, Patrick J.

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
Technical Report	FROM ______ TO ______	88/02/01	54

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB. GR.	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Given two Banach spaces X and Y over $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ and a parametrized family $A(\mu) \in \mathcal{L}(X, Y)$ with $\mu \in \mathbb{K}$, partial and algebraic multiplicities of any value $\mu_0 \in \mathbb{K}$ such that $A(\mu_0)$ if Fredholm with index zero are defined by the means of generalized Jordan chains. These notions are developed in close connection with bifurcation problems and we show that partial and algebraic multiplicities are not affected by Lyapunov-Schmidt reduction. Properties of invariance under equivalence are also established. These general results are used to give a proof of Magnus' generalization of the classical bifurcation theorem by Krasnoselskii through a somewhat more natural approach than his.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: UNCLASSIFIED/UNLIMITED XX SAME AS RPT ☐ DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION: UNCLASSIFFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL: Arje Nachman	22b. TELEPHONE NUMBER (Include Area Code): 202-767- 5025; 22c. OFFICE SYMBOL

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE.

UNCLASSIFIED

GENERALIZED JORDAN CHAINS AND

TWO BIFURCATION THEOREMS OF KRASNOSELSKII

by

Patrick J. Rabier

Department of Mathematics and Statistics

University of Pittsburgh

Pittsburgh, PA 15260

ABSTRACT

Given two Banach spaces X and Y over $K = R$ or C and a parametrized family $A(\mu) \in \mathcal{L}(X, Y)$ with $\mu \in K$, partial and algebraic multiplicities of any value $\mu_0 \in K$ such that $A(\mu_0)$ if Fredholm with index zero are defined by the means of generalized Jordan chains. These notions are developed in close connection with bifurcation problems and we show that partial and algebraic multiplicities are not affected by Lyapunov-Schmidt reduction. Properties of invariance under equivalence are also established. These general results are used to give a proof of Magnus' generalization of the classical bifurcation theorem by Krasnoselskii through a somewhat more natural approach than his. But the convincing evidence of the usefulness of the notions developed here has to be found in a new and wide extension of the Böhme-Marino-Rabinowitz theorem on bifurcation for gradient operators, the ancestor of which is also due to Krasnoselskii.

1. INTRODUCTION

In 1976, Magnus [13] introduced a generalized notion of algebraic multiplicity for values μ_0 of the scalar parameter μ at which a given parametrized family $A(\mu)$ of linear operators acting between Banach spaces becomes singular. He used this concept to prove various local and global bifurcation theorems generalizing the well known result by Krasnoselskii [11] about bifurcation at a characteristic value with odd algebraic multiplicity.

Magnus' definition of algebraic multiplicity requires using a sequence of projections, as is briefly reviewed later on in this section. Here, we shall also develop a generalized notion of algebraic multiplicity which will eventually be shown to coincide with Magnus' one, but our approach is rather different. Instead of a sequence of projections, our definition involves generalized Jordan chains. We have found many advantages in doing so. First, it is our feeling that the resulting definition is both simpler and more intuitive. More important is the fact that generalized Jordan chains permit the use of the so called root functions which often allow one to replace tedious combinatoric or other arguments by an elementary proof. Most significant of this property is our proof that Lyapunov-Schmidt reduction does not affect the algebraic multiplicity. This was not established by Magnus and, indeed, seems difficult to prove in his approach. Using this result, it becomes possible to parallel one of the classical proofs of Krasnoselskii's theorem to obtain Magnus' generalization of it in a somewhat more natural way than his.

But the decisive argument in favor of generalized Jordan chains is that algebraic multiplicity is derived from the more refined notion of partial multiplicities. It is to be noted that Magnus, too, introduces partial multiplicities which, however, differ from ours. Incidentally, he makes no use of them except as a sometimes convenient terminology. In sharp contrast, the partial multiplicities considered here play a crucial role to establish generalizations of the theorem by Krasnoselskii [11] on bifurcation for gradient operators, subsequently improved by Böhme [2], Marino [14] and Rabinowitz [17]. Our contribution differs in that the versions we give are not limited to nonlinear eigenvalue problems and deal with general equations of the form $F(\mu, x) = 0$, far beyond the currently available results. Aside partial multiplicities, the related proofs involve analytic perturbation theory for linear operators and a preliminary bifurcation theorem in the finite dimensional case based on Conley index and showing that, for gradient operators, bifurcation is guaranteed by a change of Morse index of the linearization.

Other by-products of our analysis are a third definition for the algebraic multiplicity, a fourth one being that of Ize [9], known to be identical with the definition by Magnus. Invariance of partial and algebraic multiplicities under equivalence is also established and used to formulate corresponding bifurcation theorems when bifurcation is studied from a known branch of solutions instead of the trivial branch.

Our approach to multiplicities through generalized Jordan chains closely follows Gohberg, Lancaster and Rodman [7] who consider the finite dimensional and polynomial case. Because neither of these assumptions is typical in applications to bifurcation problems, variants of the proof presented in [7] need to be given, especially those involving determinants. Despite that the infinite dimensional framework has reportedly been investigated in the complex and analytic case by the Russian school (also responsible for the introduction of generalized Jordan chains in the mid forties) we have not found available references for the real and nonanalytic case.

We will be considering a parametrized family $A(\mu) \in \mathcal{L}(X, Y)$ with X and Y Banach spaces over $K = R$ or C and μ varying in some open and connected neighborhood of 0 in K. However, in our exposition, we shall limit ourselves to $K = R$ since all the definitions and results that make sense can be extended to $K = C$ without modification (except for shorter proofs at times). Also, the parametrized family $A(\mu)$ will always be "smooth", which means "of class $\mathcal{C}^\infty$. Whenever a result is proved for smooth families, it can be generalized to families with only some finite regularity, but the smoothness assumption is convenient for expository purposes.

Before we introduce generalized Jordan chains, let us briefly recall how (algebraic) multiplicity is defined by Magnus in [13]. Suppose that $A(0)$ is Fredholm with index zero. If $A(0) \in \mathrm{Isom}(X, Y)$, the algebraic multiplicity of $\mu = 0$ in $A(\mu)$ is defined to be 0. Otherwise, set $A_{(0)}(\mu) = A(\mu)$ and, given an arbitrary projection π_0 with range the null-space $\mathrm{Ker}\, A(0)$, set, for $\mu \neq 0$

$$A_{(1)}(\mu) = \frac{1}{\mu} A_{(0)}(\mu)\pi_0 + A_{(0)}(\mu)(I - \pi_0)$$

and, for $\mu = 0$

$$A_{(1)}(0) = A'_{(0)}(0)\pi_0 + A_{(0)}(0)(I - \pi_0) .$$

If $A_{(1)}(0) \in \mathrm{Isom}(X, Y)$, the multiplicity of $\mu = 0$ in $A(\mu)$ is defined to be $\dim \mathrm{Ker}\, A(0)$. As $A_{(1)}(0)$ is Fredholm with index zero in any case, and if $A_{(1)}(0) \in \mathrm{Isom}(X, Y)$, one may define a new family $A_{(2)}(\mu)$ by repeating the same procedure (note that $A_{(1)}(\mu)$ is smooth). In other words, choosing a projection π_1 with range the null-space $\mathrm{Ker}\, A_{(1)}(0)$, set, for $\mu \neq 0$

$$A_{(2)}(\mu) = \frac{1}{\mu} A_{(1)}(\mu)\pi_1 + A_{(1)}(\mu)(I - \pi_0) ,$$

and, for $\mu = 0$

$$A_{(2)}(0) = A'_{(1)}(0)\pi_1 + A_{(1)}(0)(I - \pi_1) .$$

If $A_{(2)}(0) \in \mathrm{Isom}(X, Y)$, the multiplicity of $\mu = 0$ in $A(\mu)$ is defined to be $\dim \mathrm{Ker}\, A_{(1)}(0) + \dim \mathrm{Ker}\, A(0)$. Otherwise, $A_{(2)}(0)$ is Fredholm with index zero and $A_{(2)}(\mu)$ is smooth, so that a new family $A_{(3)}(\mu)$ is defined. Assuming that this process stops at rank κ (i.e. $A_{(\kappa)}(0) \in \mathrm{Isom}\,(X, Y)$), the multiplicity of $\mu = 0$ in $A(\mu)$ is defined by

$$\sum_{j=0}^{\kappa-1} \dim \mathrm{Ker}\, A_{(j)}(0) .$$

Multiplicity of μ_0 in $A(\mu)$ is defined to be the multiplicity of $\mu = 0$ in $A(\mu + \mu_0)$.

As mentioned before, our approach will be different. But, in any case, let us make it clear right now that there is an actual need for a <u>definition</u> of algebraic multiplicity. In particular, it cannot be overemphasized that the generalized null space of $A(0)$ has <u>nothing</u> <u>to</u> <u>do</u> with the algebraic multiplicity of $\mu = 0$ in $A(\mu)$, except when $A(\mu) = A_0 - \mu I$. This is because algebraic multiplicity must relate to the whole family $A(\mu)$ and not merely to the operator $A(0)$. If $X = Y = R^m$, the order of the root $\mu = 0$ in $\det A(\mu)$ provides a useful definition (totally unrelated to the generalized null-space of $A(0)$) and the problem becomes one of suitably generalizing this notion when X and Y are arbitrary Banach spaces.

2. GENERALIZED JORDAN CHAINS; ROOT FUNCTIONS

Let then X and Y be real Banach spaces and $A(\mu)$ a smooth mapping of the real parameter μ with values in the space $\mathcal{L}(X, Y)$. The function $A(\mu)$ need only be defined for μ on a neighborhood of the origin. For $j \geq 0$, we shall set

$$A_j = (1/j!)A^{(j)}(0) . \qquad (2.1)$$

A family $(e_0, \ldots, e_\ell)$ of $\ell+1$ vectors of X such that $e_0 \neq 0$ and

$$\sum_{i=0}^{j} A_i e_{j-i} = 0 , \quad 0 \leq j \leq \ell, \qquad (2.2)$$

will be called a <u>generalized Jordan chain</u> of $A(\mu)$. For $j = 0$, note that $A_0 = A(0)$. From (2.2), one has $A_0 e_0 = 0$: as $e_0 \neq 0$, existence of generalized Jordan chains requires Ker $A_0 \neq \{0\}$. It is readily seen from the definition that e_ℓ exists if and only if

$$\sum_{i=1}^{\ell} A_i e_{\ell-1} \in \text{Range } A_0 .$$

From this observation, a natural definition for a maximal chain follows as being one that cannot be continued. In this paper, we shall exclusively be concerned with the case when Ker A_0 is <u>finite dimensional</u> and when no generalized Jordan chain can be continued indefinitely. Actually, the stronger assumption that the length of all maximal chains is uniformly bounded from above by a positive integer will be made. But in the case of interest when $A_0 = A(0)$ is Fredholm with index zero, we shall later see that the two requirements are the same. Of course, the length of a chain must be understood as the number of its elements.

In the hypothesis that the length of all maximal chains is uniformly bounded from above, we can define "canonical sets" of generalized Jordan chains according to the following process: given $e_0 \neq 0$ in Ker A_0, call $\kappa(e_0)$ the maximal length of all the generalized Jordan chains originating at e_0. Such an integer $\kappa(e_0)$ is defined without ambiguity and $\kappa(e_0)$ is uniformly bounded as e_0 runs over Ker $A_0 - \{0\}$. As $\kappa(e_0)$ takes integral values, it follows that an element $e_{0,1} \in$ Ker $A_0 - \{0\}$ exists for which $\kappa(e_{0,1}) = \kappa_1$ is maximal. We thus obtain a generalized Jordan chain

$$(e_{0,1}, \dots, e_{\kappa_1-1,1})$$

with maximal length among all possible generalized Jordan chains. In a second step, we select $e_{0,2}$ by requiring that $\kappa(e_{0,2}) = \kappa_2$ be maximal among all values $\kappa(e_0)$ for $e_0 \in$ Ker A_0 not collinear with $e_{0,1}$. This yields a generalized Jordan chain

$$(e_{0,2}, \dots, e_{\kappa_2-1,2}) .$$

Of course, $\kappa_2 \leq \kappa_1$ follows from the definitions. More generally, having chosen j generalized Jordan chains originating at $e_{0,1}, \dots, e_{0,j}$ respectively, we select $e_{0,j+1}$ by requiring that $\kappa(e_{j+1}) = \kappa_{j+1}$ be maximal among all values $\kappa(e_0)$ for $e_0 \in$ Ker A_0 not in the span of $\{e_{0,1}, \dots, e_{0,j}\}$. Hence, there is a generalized Jordan chain

$$(e_{0,j+1}, \dots, e_{\kappa_{j+1}-1,j+1}) ,$$

with $\kappa_{j+1} \leq \kappa_j$. Clearly, if Ker A_0 is finite-dimensional, the process can be repeated until $n = \dim$ Ker A_0 elements $e_{0,1}, \dots, e_{0,n}$ have been selected. The set of generalized Jordan chains

$$\{(e_{0,j}, \dots, e_{\kappa_j-1,j}) , 1 \leq j \leq n\}$$

with $\kappa_1 \geq \kappa_2 \geq \cdots \geq \kappa_n \geq 1$ is called a <u>canonical</u> <u>set</u> of generalized Jordan chains.

Remarks: 1) If $X = Y$ and $A(\mu) = \mu I - L$, $L \in \mathcal{L}(X)$, a generalized Jordan [illegible] of $A(\mu)$ is nothing but a Jordan chain of L corresponding to the [illegible] $\mu = 0$. 2) Unlike Jordan chains, generalized ones need not be made of [illegible] independent elements. 3) Our definitions (in particular that of a [illegible] set) duplicate those in [7] when $X = Y = R^m$ and $A(\mu)$ is a polynomial. [illegible] Canonical sets of generalized Jordan chains are not uniquely defined. [illegible] ever, we shall see below that the integers $\kappa_1, \dots, \kappa_n$ are independent [illegible] canonical set. □

From the definitions, it is clear that κ_1 is the same for every canonical set of generalized Jordan chains. Set

$$K_1 = \{e_0 \in \text{Ker } A_0 - \{0\} \ , \ \kappa(e_0) = \kappa_1\} \ ,$$

and let E_1 denote the subspace of Ker A_0 generated by K_1, namely the smallest subspace of Ker A_0 containing K_1. It follows that dim E_1 equals the maximum number of linearly independent elements in K_1. Denote by ℓ the dimension of E_1. Let $e_{0,1}, \dots, e_{0,n}$ be the first elements of a canonical set of generalized Jordan chains. From the definition, it is immediate that $e_{0,1}, \dots, e_{0,\ell}$ are linearly independent elements of K_1, so that $\kappa_1 = \cdots = \kappa_\ell$ and

$$E_1 = \text{span } \{e_{0,1}, \dots, e_{0,\ell}\} \ . \tag{2.3}$$

The set K_1 and hence both the space E_1 and its dimension ℓ are independent of any particular choice for a canonical set of generalized Jordan chains, and the above thus shows that $\kappa_1, \dots, \kappa_\ell$ are also independent of the canonical set. Independence is then proved if $\ell = n$. If $\ell < n$, $\kappa_{\ell+1}$ is (from (2.3) and the definition) the maximum of $\kappa(e_0)$ as e_0 runs over $(\text{Ker } A_0) - E_1$. Then, just as κ_1, $\kappa_{\ell+1}$ is independent of the canonical set. Consider

$$K_2 = \{e_0 \in \text{Ker } A_0 - \{0\} \ , \ \kappa(e_0) \geq \kappa_{\ell+1}\} \ ,$$

so that $\kappa(e_0) = \kappa_1$ or $\kappa_{\ell+1}$ for $e_0 \in K_2$. If E_2 denotes the space generated by K_2, one has $E_2 \supset E_1$ so that $\dim E_2 = \ell+k$ $(k \geq 1)$. Clearly, $\ell + k$ coincides with the maximum number of linearly independent elements in K_2 and $e_{0,1}, \ldots, e_{0,\ell+k}$ are linearly independent elements of K_2, hence

$$E_2 = \operatorname{span}\{e_{0,1}, \ldots, e_{0,\ell+k}\} ,$$

while $\kappa_{\ell+1} = \cdots = \kappa_{\ell+k}$. As E_1 and E_2 as well as their respective dimensions ℓ and $\ell+k$ are independent of the canonical set of generalized Jordan chain, so is $k = \ell + k - \ell$. It follows that $\kappa_{\ell+1}, \ldots, \kappa_{\ell+k}$ are independent of the canonical set. Repeating this procedure as many times as necessary yields independence of $\kappa_1, \ldots, \kappa_n$ regarding the choice of the canonical set of generalized Jordan chains. The integers $\kappa_1, \ldots, \kappa_n$ will be called the <u>partial multiplicities</u> of $\mu = 0$ in $A(\mu)$ and the number

$$\gamma = \kappa_1 + \cdots + \kappa_n , \tag{2.4}$$

its <u>algebraic multiplicity</u>.

We now introduce the notion of <u>root function</u> (root polynomial in [7]). Given any $e_0 \neq 0$ in $\operatorname{Ker} A_0$, we call root function corresponding to e_0 a smooth function $e(\mu)$ with values in X such that $e(0) = e_0$ and define the <u>order</u> of $e(\mu)$ to be that of the zero $\mu = 0$ in $A(\mu)e(\mu)$. Suppose then that $e(\mu)$ has order $\ell + 1 \geq 1$. Setting

$$e_j = \frac{1}{j!} e^{(j)}(0) , \ 0 \leq j \leq \ell ,$$

and equating to zero the coefficient of μ^j in the Taylor expansion of $A(\mu)e(\mu)$ (obtained from those of $A(\mu)$ and $e(\mu)$) one finds that $(e_0, \ldots, e_\ell)$ is a generalized Jordan chain of $A(\mu)$. In particular, the order of a root function never exceeds the largest partial multiplicity κ_1. Conversely, given a generalized Jordan chain $(e_0, \ldots, e_\ell)$ of $A(\mu)$, then

$$e(\mu) = \sum_{j=0}^{\ell} \mu^j e_j + \mu^{\ell+1}\delta(\mu) ,$$

is a root function of order $\geq \ell + 1$ for every smooth function $\delta(\mu)$. Root functions are essential to make generalized Jordan chains a powerful tool for theoretical investigation. They even appear to be virtually indispensable in some of the proofs. A first example of their usefulness can be found in the proof of the following important result.

Proposition 2.1: Let $B(\mu)$ and $C(\mu)$ be smooth mappings with values in $\mathcal{L}(Y)$ and $\mathcal{L}(X)$ respectively and suppose that $B(0)$ and $C(0)$ are linear isomorphisms. Then, $(\tilde{e}_0, \ldots, \tilde{e}_\ell)$ is a generalized Jordan chain of $B(\mu)A(\mu)C(\mu)$ if and only if

$$e_j = \sum_{i=0}^{j} \frac{1}{i!} C^{(i)}(0)\tilde{e}_{j-i} , \quad 0 \leq j \leq \ell ,$$

is a generalized Jordan chain of $A(\mu)$.

When $X = Y = R^m$ and $A(\mu)$ is a polynomial, the proof of this statement can be found in [7, Proposition 1.11, p. 29]. In the context of this paper, it has also been given in [16, Proposition 2.2]. Although [16] deals with the case $\dim \operatorname{Ker} A_0 = 1$, the proof is equally valid in general and hence will not be given again. Proposition 2.1 can actually be established without the help of root functions but becomes a rather cumbersome exercise in combinatorics.

The following corollary is immediate.

Corollary 2.1: Let $B(\mu)$ and $C(\mu)$ be as in Proposition 2.1. Then, $\mu = 0$ has the same partial and algebraic multiplicities in $A(\mu)$ and in $B(\mu)A(\mu)C(\mu)$.

Remark 5: More generally, in both Proposition 2.1 and Corollary 2.1 the mapping $B(\cdot)$ (resp. $C(\cdot)$) may take values in $\mathcal{L}(Y, \tilde{Y})$ (resp. $\mathcal{L}(\tilde{X}, X)$) with Y and $\tilde{Y}$

(resp. X and $\tilde{X}$) isomorphic Banach spaces. □

Another characterization of the algebraic multiplicity.

For every integer $p \geq 1$, define $\mathcal{A}_p \in \mathcal{L}(X^p, Y^p)$ by

$$\mathcal{A}_p = \begin{bmatrix} A_0 & & & \\ A_1 & A_0 & & \bigcirc \\ & \ddots & \ddots & \\ A_{p-1} & & A_1 & A_0 \end{bmatrix} \tag{2.5}$$

We shall repeatedly use the trivial observation that for $p \geq 2$, $\mathcal{A}_{p-1}$ identifies with the $(p-1)\times(p-1)$ upper left corner as well as with the $(p-1)\times(p-1)$ lower right corner of $\mathcal{A}_p$.

Suppose now that $\kappa_1 < \infty$ is the maximal length of the generalized Jordan chains of $A(\mu)$ and let $p > \kappa_1$. Then, if $(e_0, \ldots, e_{p-1}) \in \operatorname{Ker}\mathcal{A}_p$, one must have $e_0 = 0$ for otherwise it is immediate that $(e_0, \ldots, e_{p-1})$ is a generalized Jordan chain of $A(\mu)$ with length $p > \kappa_1$, a contradiction. Thus, for $p > \kappa_1$, an element of $\operatorname{Ker}\mathcal{A}_p$ has the form $(0, e_1, \ldots, e_{p-1})$. It is then obvious that $(e_1, \ldots, e_{p-1}) \in \operatorname{Ker}\mathcal{A}_{p-1}$. The converse is just as trivial since the mapping

$$(e_0, \ldots, e_{p-2}) \in X^{p-1} \to (0, e_0, \ldots, e_{p-2}) \in X^p$$

induces a canonical injection of $\operatorname{Ker}\mathcal{A}_{p-1}$ into $\operatorname{Ker}\mathcal{A}_p$. Summing up, we have that, to within canonical identifications, the sequence $\operatorname{Ker}\mathcal{A}_p$ stablizes for $p \geq \kappa_1$.

Conversely, suppose that there is a smallest integer $k \geq 1$ such that $\operatorname{Ker}\mathcal{A}_{k+1} = \operatorname{Ker}\mathcal{A}_k$. Then $k = \kappa_1$. Indeed, let $(e_0, \ldots, e_{p-1})$ be a generalized

Jordan chain of $A(\mu)$ with length $p > k$. Clearly, $(e_0, \ldots, e_{p-1}) \in \operatorname{Ker}\mathcal{A}_p$. In particular, $(e_0, \ldots, e_k) \in \operatorname{Ker}\mathcal{A}_{k+1}$. But, since $\operatorname{Ker}\mathcal{A}_{k+1} = \operatorname{Ker}\mathcal{A}_k$, one must have $e_0 = 0$, contradicting the fact that $(e_0, \ldots, e_{p-1})$ is a generalized Jordan chain. This shows that $\kappa_1 \leq k$. To prove $\kappa_1 \geq k$, recall that $\operatorname{Ker}\mathcal{A}_{k-1} \neq \operatorname{Ker}\mathcal{A}_k$ by definition of k (assuming $k \geq 2$, but the problem is obvious if $k = 1$). Hence, there is $(e_0, \ldots, e_{k-1}) \in \operatorname{Ker}\mathcal{A}_k$ with $e_0 \neq 0$, which means that $(e_0, \ldots, e_{k-1})$ is a generalized Jordan chain of $A(\mu)$ with length k so that $k \leq \kappa_1$.

As a result of $\dim \operatorname{Ker} A_0 < \infty$, one finds that $\operatorname{Ker}\mathcal{A}_p$ is finite dimensional for every $p \geq 1$. It turns out that the algebraic multiplicity $\gamma = \kappa_1 + \cdots + \kappa_n$ is nothing but

$$\gamma = \dim \operatorname{Ker}\mathcal{A}_{\kappa_1} = \dim \operatorname{Ker}\mathcal{A}_p \ , \ \forall p \geq \kappa_1 \ . \tag{2.6}$$

The proof of this statement is postponed until we can make use of a local Smith form (cf. Section 4). Relation (2.6) shows that calculation of the algebraic multiplicity amounts to calculation of a null-space, which may have some importance in the applications.

<u>Remark</u> <u>6</u>: If $A_0 = A(0) \in \operatorname{Isom}(X, Y)$, it is consistent with the definitions and results of this section to define the partial and algebraic multiplicities of $\mu = 0$ in $A(\mu)$ to be 0. Also, for any $\mu_0 \neq 0$, the partial and algebraic multiplicities of μ_0 in $A(\mu)$ are defined as those of $\mu = 0$ in $A(\mu + \mu_0)$. □

3. GENERALIZED JORDAN CHAINS AND LYAPUNOV-SCHMIDT REDUCTION

With X and Y being real Banach spaces as in the previous section, let us consider a $\mathcal{C}^1$ mapping $F(= F(\mu, x))$ from $\mathbb{R} \times X$ to Y, locally defined near the origin and satisfying

$$F(\mu, 0) = 0 . \tag{3.1}$$

It is the aim of this paper to discuss conditions ensuring bifurcation of solutions to $F = 0$ from the trivial branch $x = 0$ and near the origin (0, 0). Assuming that $D_xF(0, 0)$ is a Fredholm operator with index zero, it is a standard procedure to make the problem into a finite dimensional one through the so called Lyapunov-Schmidt reduction. We shall denote by X_0 and Y_0 the null-space and range of $D_xF(0, 0)$ respectively, and make the choice of (topological) complements X_1 of X_0 and Y_1 of Y_0. For $x \in X$, we shall set

$$x = \epsilon + x_1 ,$$

according to $X = X_0 \oplus X_1$, and call Q_0 and Q_1 the (continuous) projections onto Y_0 and Y_1 respectively. Writing the equation $F(\mu, x) = 0$ as the system

$$\begin{aligned} Q_0F(\mu, \epsilon + x_1) = 0 , \\ Q_1F(\mu, \epsilon + x_1) = 0 , \end{aligned} \tag{3.2}$$

it follows from $Q_0D_xF(0, 0)|_{X_1}$ being an isomorphism of X_1 to Y_0 that the first equation is solved through the Implicit function theorem in the form

$$x_1 = \phi(\mu, \epsilon) ,$$

where ϕ is of class $\mathcal{C}^1$ on a neighborhood of the origin in $\mathbb{R} \times X_0$ with values in X_1 and verifies

$$\phi(\mu, 0) = 0 . \tag{3.3}$$

For future use, note that differentiating the identity $Q_0F(\mu, \epsilon + \phi(\mu, \epsilon)) = 0$

w.r.t ϵ and setting $\epsilon = 0$ results in

$$Q_0 D_x F(\mu, 0)(I_{X_0} + D_\epsilon \phi(\mu, 0)) = 0 . \tag{3.4}$$

In particular, this yields

$$D_\epsilon \phi(0, 0) = 0 . \tag{3.5}$$

Substituting $x_1 = \phi(\mu, \epsilon)$ in the second equation (3.2) leads to the <u>reduced equation</u>

$$f(\mu, \epsilon) = Q_1 F(\mu, \epsilon + \phi(\mu, \epsilon)) = 0 , \tag{3.6}$$

equivalent to $F(\mu, x) = 0$ near the origin. As usual, the mapping f will be referred to as the <u>reduced mapping</u> of F.

Whenever it is smooth, the mapping

$$A(\mu) = D_x F(\mu, 0) \in \mathcal{L}(X, Y) \tag{3.7}$$

satisfies the general conditions of Section 2. In particular, $A_0 = A(0) = D_x F(0, 0)$ has a finite dimensional null-space and all the notions previously developed for $A(\mu)$ make sense. Now, setting

$$a(\mu) = D_\epsilon f(\mu, 0) \in \mathcal{L}(X_0, Y_1) , \tag{3.8}$$

the same comment applies to $a(\mu)$. Indeed, although $D_\epsilon f$ is merely continuous, $D_\epsilon f(\mu, 0)$ is smooth. To see this, note that from (3.4) and (3.7)

$$D_\epsilon \phi(\mu, 0) = -[Q_0 A(\mu)|_{X_1}]^{-1}(Q_0 A(\mu)|_{X_0}) ,$$

is smooth since $A(\mu)$ is smooth and hence, upon differentiating (3.6)

$$a(\mu) = Q_1 A(\mu)(I_{X_0} + D_\epsilon \phi(\mu, 0)) , \tag{3.9}$$

is smooth, too. Theorem 3.1 below shows that Lyapunov-Schmidt reduction preserves all the properties of generalized Jordan chains.

<u>Theorem 3.1</u>: <u>The length of the generalized Jordan chains of</u> $A(\mu)$ <u>is (uniformly) bounded if and only if the length of the generalized Jordan chains of</u> $a(\mu)$ <u>is (uniformly) bounded. Moreover, in this case, the partial multiplicities of</u> $\mu = 0$ <u>are the same in</u> $A(\mu)$ <u>and</u> $a(\mu)$.

<u>Proof</u>: From the definitions, it suffices to show that, given any generalized Jordan chain of $A(\mu)$ (resp. $a(\mu)$), say $(e_0, \dots, e_\ell)$, one can find a generalized Jordan chain of $a(\mu)$ (resp. $A(\mu)$), say $(e_0, \tilde{e}_1, \dots, \tilde{e}_\ell)$ with the same length and starting with the same element e_0.

First, consider a generalized Jordan chain $(e_0, \dots, e_\ell)$ of $A(\mu)$ and the root function

$$e(\mu) = \sum_{j=0}^{\ell} \mu^j e_j \in X \ . \tag{3.10}$$

Then, $A(\mu)e(\mu)$ vanishes together with its first ℓ derivatives at $\mu = 0$, hence

$$A(\mu)e(\mu) = \mu^{\ell+1}\alpha(\mu) \ , \tag{3.11}$$

for some smooth Y-valued mapping $\alpha(\mu)$. Let ψ: $\mathbb{R}\times X_0\times X_1 \to X$ be defined by

$$\psi(\mu, \epsilon, x_1) = (I_{X_0} + D_\epsilon\phi(\mu, 0))\epsilon + x_1 - e(\mu) \ .$$

From (3.5) and (3.10), $\psi(0, e_0, 0) = 0$. On the other hand, for $(h_0, h_1) \in X_0\times X_1 \approx X$, one finds

$$D_{(\epsilon,x_1)}\psi(0, e_0, 0) \cdot (h_0, h_1) = h_0 + h_1 \ .$$

Hence $D_{(\epsilon,x_1)}\psi(0, e_0, 0) \in \mathrm{Isom}(X_0\times X_1, X)$ and the solutions to $\psi = 0$ near $(0, e_0, 0)$ are given by a curve $(\epsilon = e_0(\mu), x_1 = e_1(\mu))$ with smooth $e_0(\cdot)$ and $e_1(\cdot)$. From the definition of ψ

$$e(\mu) = (I_{X_0} + D_\epsilon\phi(\mu, 0))e_0(\mu) + e_1(\mu) \ .$$

Explicit formulas for $e_0(\mu)$ and $e_1(\mu)$ are easily obtained from the observation that ψ above is linear in (ϵ, x_1) (so that the use of the Implicit function theorem is rather artificial here). But these formulas are not useful in what follows. With the above expression for $e(\mu)$, one may rewrite relation (3.11) as

$$A(\mu)(I_{X_0} + D_\epsilon\phi(\mu, 0))e_0(\mu) + A(\mu)e_1(\mu) = \mu^{\ell+1}\alpha(\mu) \quad . \tag{3.12}$$

Recall that $Q_0 + Q_1 = I_Y$ and $Q_0A(\mu)(I_{X_0} + D_\epsilon\phi(\mu, 0)) = 0$ (cf.(3.4) and (3.7)). Thus, (3.12) becomes

$$Q_0A(\mu)e_1(\mu) + Q_1A(\mu)(I_{X_0} + D_\epsilon\phi(\mu, 0))e_0(\mu) + Q_1A(\mu)e_1(\mu) = \mu^{\ell+1}\alpha(\mu) \quad .$$

Writing $\alpha(\mu) = Q_0\alpha(\mu) + Q_1\alpha(\mu)$ and equating components yields

$$Q_0A(\mu)e_1(\mu) = \mu^{\ell+1}Q_0\alpha(\mu) \tag{3.13}$$

$$Q_1A(\mu)(I_{X_0} + D_\epsilon\phi(\mu, 0))e_0(\mu) + Q_1A(\mu)e_1(\mu) = \mu^{\ell+1}Q_1\alpha(\mu) \quad . \tag{3.14}$$

As $Q_0A(0)|_{X_1} \in \text{Isom}\ (X_1, Y_0)$, one has $Q_0A(\mu)|_{X_1} \in \text{Isom}\ (X_1, Y_0)$ for $|\mu|$ small enough and hence, from (3.13)

$$e_1(\mu) = \mu^{\ell+1}\beta(\mu) \quad , \tag{3.15}$$

with

$$\beta(\mu) = [Q_0A(\mu)|_{X_1}]^{-1}Q_0\alpha(\mu) \quad .$$

Substituting (3.15) into (3.14), we get

$$Q_1A(\mu)(I_{X_0} + D_\epsilon\phi(\mu, 0))e_0(\mu) = \mu^{\ell+1}(Q_1\alpha(\mu) - Q_1A(\mu)\beta(\mu)) \quad . \tag{3.16}$$

Due to (3.9) and since $e_0(0) = e_0 \neq 0$, this means that $e_0(\mu)$ is a root function of $a(\mu)$ with order at least $\ell+1$, i.e. $(e_0, \tilde{e}_1, \ldots, \tilde{e}_\ell)$ with $\tilde{e}_j = (1/j!)e_0^{(j)}(0)$, $1 \leq j \leq \ell$, is a generalized Jordan chain of $a(\mu)$.

Conversely, let $(e_0, \ldots, e_\ell)$ be a generalized Jordan chain of $a(\mu)$ and let $e(\mu)$ denote the root function

$$e(\mu) = \sum_{j=0}^{\ell} \mu^j e_j \in X_0 .$$

Using (3.9), this means

$$Q_1 A(\mu)(I_{X_0} + D_\epsilon \phi(\mu, 0))e(\mu) = \mu^{\ell+1} \alpha_1(\mu)$$

for some smooth Y_1-valued function $\alpha_1(\mu)$. As $Q_0 A(\mu)(I_{X_0} + D_\epsilon \phi(\mu, 0)) = 0$ (cf.(3.4) and (3.7)), this also reads

$$A(\mu)(I_{X_0} + D_\epsilon \phi(\mu, 0))e(\mu) = \mu^{\ell+1} \alpha_1(\mu) .$$

From $(I_{X_0} + D_\epsilon \phi(0, 0))e(0) = e(0) = e_0$ (cf.(3.5)), we infer that $\tilde{e}(\mu) = (I_{X_0} + D_\epsilon \phi(\mu, 0))e(\mu)$ is a root function of $A(\mu)$ with order at least $\ell+1$, i.e $(e_0, \tilde{e}_1, \ldots, \tilde{e}_\ell)$ with $\tilde{e}_j = (1/j!)e^{(j)}(0)$, $1 \le j \le \ell$, is a generalized Jordan chain of $A(\mu)$, and we are done. □

Theorem 3.1 provides a good motivation for further study of generalized Jordan chains in the finite dimensional case, which will be taken up in the next section. As a by-product of the results presented there and Theorem 3.1, we shall see that our definition of algebraic multiplicity coincides with that given by Magnus in [13] (see Section 1). Indeed, let $A(\mu) \in \mathcal{L}(X, Y)$ be a smooth parametrized family and set $F(\mu, x) = A(\mu)x$, which agrees with $A(\mu) = D_x F(\mu, 0)$. Now, define $C(\mu) \in \mathcal{L}(X)$ by (setting $x = \epsilon + x_1$ again)

$$C(\mu)x = (I_{X_0} + D_\epsilon \phi(\mu, 0))\epsilon + x_1 .$$

As a result of $Q_0 A(\mu)(I_{X_0} + D_\epsilon \phi(\mu, 0)) = 0$, the operator $A(\mu)C(\mu)$ has the block decomposition

$$A(\mu)C(\mu) = \begin{bmatrix} a(\mu) & Q_1A(\mu)|_{X_1} \\ & \\ 0 & Q_0A(\mu)|_{X_1} \end{bmatrix},$$

upon identifying Y with $Y_1 \times Y_0$, and where $a(\mu)$ is as in (3.9). Next, define $B(\mu) \in \mathcal{L}(Y)$ by

$$B(\mu) = \begin{bmatrix} I_{Y_1} & -Q_1A(\mu)|_{X_1}[Q_0A(\mu)|_{X_1}]^{-1} \\ & \\ 0 & I_{Y_0} \end{bmatrix},$$

so that

$$B(\mu)A(\mu)C(\mu) = \begin{bmatrix} a(\mu) & 0 \\ 0 & Q_0A(\mu)|_{X_1} \end{bmatrix}. \tag{3.17}$$

As $B(0) = I_Y$ and $C(0) = I_X$, it follows that $\mu = 0$ has the same algebraic multiplicity in $A(\mu)$ and $B(\mu)A(\mu)C(\mu)$ (Corollary 2.1) and also in $B(\mu)A(\mu)C(\mu)$ and $a(\mu)$ (Theorem 2.1). As $a(\mu) \in \mathcal{L}(X_0, Y_1)$ and X_0 and Y_1 have the same dimension, it will follow from Section 4 and after obvious identifications that the algebraic multiplicity of $\mu = 0$ in $a(\mu)$ equals the order of the zero $\mu = 0$ in det $a(\mu)$. If now multiplicity is understood in the sense of Magnus, $\mu = 0$ has the same algebraic multiplicity in $A(\mu)$ and $B(\mu)A(\mu)C(\mu)$ ([13, Theorem 2.4]) and also in $B(\mu)A(\mu)C(\mu)$ and $a(\mu)$ ([13, Theorem 2.7] and relation (3.17) above). Finally, in [13, Theorem 2.6] it is shown that the algebraic multiplicity of $\mu = 0$ in $a(\mu)$ coincides with the order of the zero $\mu = 0$ in det $a(\mu)$ and the assertion follows.

<u>Remarks</u>: 1) Although Magnus' algebraic multiplicity coincides with ours, his partial multiplicities are different. The partial multiplicities introduced here agree with the definition in Gohberg et al. [7]. 2) Coincidence of the algebraic multiplicities makes complementary properties established by Magnus available (and indeed easier to prove in his approach). For instance, given two families $A(\mu) \in \mathcal{L}(X, Y)$ and $B(\mu) \in \mathcal{L}(Y, Z)$ with Z another real Banach space and A(0) and B(0) Fredholm with index zero, the algebraic multiplicity of $\mu = 0$ in $A(\mu)B(\mu)$ is the sum of those in $A(\mu)$ and $B(\mu)$ ([13, Theorem 2.4]). Another interesting result is that given a smooth function $\sigma(\tilde{\mu})$ with $\sigma(0) = 0$, the algebraic multiplicity of $\tilde{\mu} = 0$ in $A(\sigma(\tilde{\mu}))$ is k times the multiplicity of $\mu = 0$ in $A(\mu)$ where k is the order of the zero $\tilde{\mu} = 0$ in $\sigma(\tilde{\mu})$ ([13, Theorem 2.9]). An obvious corollary is that the multiplicity of $\mu = 0$ in $A(\mu)$ cannot be made odd through any smooth change of scale in μ (not necessarily a diffeomorphism) if it is not odd in the first place, but there are other applications. 3) Another approach is taken by Ize [9], who defines the algebraic multiplicity of $A(\mu) = D_x F(\mu, 0)$ via the reduced mapping, i.e. in $a(\mu) = Q_1 A(\mu)(I_{X_0} + D_\epsilon \phi(\mu, 0))$ (cf. (3.9)), to be the order of the zero $\mu = 0$ in $\det a(\mu)$. He next proves some independence of his definition regarding Lyapunov-Schmidt reduction. More precisely, he shows that the parity of the algebraic multiplicity is independent of the Lyapunov-Schmidt reduction. Theorem 3.1 is much stronger since it asserts that the partial multiplicities, hence the algebraic multiplicity and a fortiori its parity, are independent of the Lyapunov Schmidt reduction. □

4. LOCAL SMITH FORM FOR THE FINITE-DIMENSIONAL CASE.

In this section, we continue the analysis of Section 2 when X and Y are finite dimensional spaces with the same dimension. Fixing bases, we can then assume $X = Y = R^m$ and that $A(\mu)$ is identified with its matrix in the canonical basis of R^m. With $B(\mu)$ and $C(\mu)$ independent of μ in Corollary 2.1, it is immediate that the partial multiplicities of $\mu = 0$ in $A(\mu)$ are independent of these identifications. On the other hand, the vocabulary and methods of matrix theory are especially convenient for the purposes of this section. Our first aim here is to show that, provided that the length of no generalized Jordan chains of $A(\mu)$ is infinite, then $A(\mu)$ admits a "local Smith form". This means that for $|\mu|$ small enough $A(\mu)$ can be written as

$$A(\mu) = M(\mu)D(\mu)N(\mu) \ , \tag{4.1}$$

where $M(\mu)$ and $N(\mu)$ are $m\times m$ matrices (operators) with constant nonzero determinants and $D(\mu)$ is diagonal of the form

$$D(\mu) = \operatorname{diag}\,(\mu^{\kappa_1} d_1(\mu), \ \dots \ , \mu^{\kappa_n} d_n(\mu), \ d_{n+1}(\mu), \ \dots \ , d_m(\mu)) \ , \tag{4.2}$$

where $\kappa_1 \geq \cdots \geq \kappa_n \geq 1$ are the partial multiplicities of $\mu = 0$ in $A(\mu)$ and $d_i(0) \neq 0$, $1 \leq i \leq m$. Among other things, we shall see that assuming that the length of all the generalized Jordan chains of $A(\mu)$ is uniformly bounded is equivalent to the seemingly weaker requirement that no chain can be continued indefinitely. Other results of some practical importance will also be derived.

When $A(\mu)$ is a polynomial, existence of a local Smith form is easily deduced from an available "global" Smith form (cf. [7]). If $A(\mu)$ is not a polynomial, there is a no global Smith form but obtaining a local one is actually technically simpler than in the polynomial case (although the basic idea is the same). In what follows, we suppose that no generalized Jordan chain of $A(\mu)$ can be continued indefinitely, and argue by induction on the dimension m of the space R^m. If $m = 1$, $A(\mu)$ is a real-valued function $a(\mu)$ with $a(0) = 0$. It is readily checked that a generalized Jordan chain $(e_0, \ldots, e_\ell)$ exists if and only if $a^{(j)}(0) = 0$, $0 \leq j \leq \ell$. Impossibility of continuing a chain indefinitely thus amounts to saying that $a^{(k)}(0) \neq 0$ for some index $k \geq 1$. If so, one may write $a(\mu) = \mu^k d(\mu)$ with $d(0) \neq 0$ and the result follows with $M(\mu) = 1$, $D(\mu) = \mu^k d(\mu)$ and $N(\mu) = 1$. Let then $m \geq 2$ and suppose that $(m-1)\times(m-1)$ matrix-valued functions $M_0(\mu)$, $D_0(\mu)$ and $N_0(\mu)$ with $M_0(\mu)$ and $N_0(\mu)$ having constant nonzero determinants and $D_0(\mu) = \mathrm{diag}(\mu^{k_1} d_1(\mu), \ldots, \mu^{k_{m-1}} d_{m-1}(\mu))$, $k_i \geq 0$, $d_i(0) \neq 0$, $1 \leq i \leq m-1$, can be found so that

$$A_0(\mu) = M_0(\mu) D_0(\mu) N_0(\mu) \quad , \tag{4.3}$$

whenever $A_0(\mu)$ is a smooth parametrized family of $(m-1)\times(m-1)$ matrices, none of whose generalized Jordan chains can be continued indefinitely. To prove that the same is true with $A(\mu) = (a_{ij}(\mu))$ being $m\times m$, let us first observe that one among the coefficients $a_{ij}(\mu)$ does not vanish to infinite order at the origi . Otherwise, $A^{(j)}(0) = 0$ for every $j \geq 0$, so that every root function has infinite order, a contradiction. Multiplying on the left and on the right by appropriate permutation matrices, one may then assume that $a_{mm}(\mu)$ has finite order k at the origin, say

$$a_{mm}(\mu) = \mu^k \, b_{mm}(\mu) \quad ,$$

with $k \geq 0$ an integer and $b_{mm}(0) \neq 0$. Since any element $a_{ij}(\mu)$ can be put in the last row and last column through this procedure, we may as well assume that the order of $a_{mm}(\mu)$ is the smallest possible among all $a_{ij}(\mu)$'s, namely that

$$a_{ij}(\mu) = \mu^k \, b_{ij}(\mu) \ , \ 1 \le i, \ j \le m \ ,$$

for some smooth function $b_{ij}(\mu)$. Therefore, multiplying on the right by the matrix (with determinant one)

$$\begin{bmatrix} 1 & & & \bigcirc \\ 0 & \ddots & & \\ & & \ddots & \\ -\dfrac{b_{m1}}{b_{mm}} & -\dfrac{b_{m2}}{b_{mm}} & \cdots \quad \cdot & 1 \end{bmatrix}$$

the matrix $A(\mu)$ is transformed into one which has zero elements on its last row, except for $a_{mm}(\mu) = \mu^k b_{mm}(\mu)$ which is unchanged. For simplicity of notation, we now assume that $A(\mu)$ had this particular structure in the first place. Then, multiplying on the left by the matrix (with determinant one)

$$\begin{bmatrix} 1 & 0 & \cdots & 0 & -\dfrac{b_{1m}}{b_{mm}} \\ & \ddots & & & -\dfrac{b_{2m}}{b_{mm}} \\ & & \ddots & & \vdots \\ \bigcirc & & & \ddots & 1 \end{bmatrix}$$

does not affect the last row but transforms $A(\mu)$ into a matrix having zero elements in its last column, except for $a_{mm}(\mu) = \mu^k \, b_{mm}(\mu)$ which is unchanged. In summary, these operations reduce $A(\mu)$ to the case when

$$A(\mu) = \begin{bmatrix} A_0(\mu) & 0 \\ 0 & \mu^k \, b_{mm}(\mu) \end{bmatrix} , \tag{4.4}$$

with $A_0(\mu)$ a smooth family of $(m-1)\times(m-1)$ matrices. No generalized Jordan chain of $A_0(\mu)$ can be extended indefinitely since the block diagonal decomposition (4.4) of $A(\mu)$ shows that the same would be true of at least one generalized Jordan chain of $A(\mu)$, a contradiction. If $A(\mu)$ has been put into the form (4.4) after multiplication by smooth parametrized families of invertible matrices as above, the same argument works via Proposition 2.1. Using the decomposition (4.3) for $A_0(\mu)$, one finds

$$A(\mu) = \begin{pmatrix} M_0(\mu) & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} D_0(\mu) & 0 \\ 0 & \mu^k b_{mm}(\mu) \end{pmatrix}\begin{pmatrix} N_0(\mu) & 0 \\ 0 & 1 \end{pmatrix}.$$

This yields the desired decomposition for $A(\mu)$. To be complete, one must also observe that, upon multiplying by appropriate permutation matrices, it is possible to assume that the diagonal elements of $D(\mu)$ are arranged in decreasing order, say $D(\mu) = \mathrm{diag}\,(\mu^{k_i} d_i(\mu))$ with $k_1 \geq \cdots \geq k_m$. As Ker $A(0)$ and Ker $D(0)$ have the same dimension n, one has

$$k_1 \geq \cdots \geq k_n \geq 1\ ,$$

$$k_{n+1} = \cdots = k_m = 0\ .$$

From the diagonal structure of $D(\mu)$, it is easily checked that no generalized Jordan chain of $D(\mu)$ has length $\geq k_1$, hence the length of the generalized Jordan chains of $D(\mu)$ is uniformly bounded. The partial multiplicities of $\mu = 0$ in $D(\mu)$ are thus well defined. Using root functions, it is straightforward to show that they coincide with $k_1, \ldots, k_n$. From Proposition 2.1, $A(\mu)$ may replace $D(\mu)$ in this statement. In the notation of Section 2, this means

$$\kappa_i = k_i\ ,\ 1 \leq i \leq n\ .$$

Also, det $M(\mu)$ = det $N(\mu) = 1$ in our construction, so that det $A(\mu)$ = det $D(\mu)$. As det $D(\mu) = \mu^\gamma d_1(\mu)\cdots d_m(\mu)$ with $d_i(0) \neq 0$, $0 \leq i \leq m$ and

$$\gamma = \kappa_1 + \cdots + \kappa_n ,$$

we infer that det $A(\mu)$ vanishes up to order γ at the origin. Conversely, assuming that det $A(\mu)$ vanishes up to finite order at the origin, it is easily seen that $A(\mu)$ possesses a local Smith form through arguments similar to those used above. But we have just seen that existence of a local Smith form implies existence of partial and algebraic multiplicities, the latter coinciding with the order of the zero $\mu = 0$ in det $A(\mu)$. In particular, this shows that assuming that no generalized Jordan chain of $A(\mu)$ can be continued indefinitely amounts to assuming that $\mu = 0$ has finite order as a zero of det $A(\mu)$. If so, the length of the generalized Jordan chains of $A(\mu)$ is uniformly bounded.

<u>Application</u>: <u>Proof of the multiplicity formula</u> (<u>2</u>.<u>6</u>).

Existence of a local Smith form allows for a simple proof of the multiplicity formula (2.6). To begin with, in a notation similar to (2.6), denote by $\mathcal{B}_p \in \mathcal{L}(Y^p)$ and $\mathcal{C}_p \in \mathcal{L}(X^p)$ the operators associated with smooth families $B(\mu) \in \mathcal{L}(Y)$ and $C(\mu) \in \mathcal{L}(X)$ like the operator $\mathcal{A}_p$ is associated with $A(\mu)$. Clearly, $\mathcal{B}_p$ and $\mathcal{C}_p$ are isomorphisms whenever $B(0)$ an $C(0)$ are isomorphisms. In this case

$$\dim \operatorname{Ker} \mathcal{A}_p = \dim \operatorname{Ker} \mathcal{B}_p \mathcal{A}_p \mathcal{C}_p .$$

Recall that $B(\mu)$ and $C(\mu)$ can be chosen so that, in the notation of Section 3, $B(\mu)A(\mu)C(\mu)$ has the block-diagonal decomposition (cf. (3.17))

$$\begin{bmatrix} a(\mu) & 0 \\ 0 & Q_0 A(\mu)|_{X_1} \end{bmatrix} \tag{4.5}$$

with $a(\mu)$ given by $a(\mu) = Q_1 A(\mu)(I_{X_0} + D_\epsilon \phi(\mu, 0))$ (cf. (3.9)) where ϕ is defined through $F(\mu, x) = A(\mu)x$ and condition (3.4). It follows that formula (2.6) can be proved with $A(\mu)$ of the form (4.5). Set

$$b(\mu) = Q_0 A(\mu)|_{X_1} \in \mathcal{L}(X_1, Y_0) ,$$

and, for $j \geq 0$

$$a_j = (1/j!)a^{(j)}(0) \in \mathcal{L}(X_0, Y_1) ,$$

$$b_j = (1/j!)b^{(j)}(0) \in \mathcal{L}(X_1, Y_0) .$$

With this notation, the operator A_j in (2.1) has the block decomposition

$$A_j = \begin{bmatrix} a_j & 0 \\ 0 & b_j \end{bmatrix} .$$

Thus

$$\mathcal{A}_p = \begin{bmatrix} \begin{bmatrix} a_0 & 0 \\ 0 & b_0 \end{bmatrix} & & & \bigcirc \\ \begin{bmatrix} a_1 & 0 \\ 0 & b_1 \end{bmatrix} & \ddots & & \\ \vdots & & \ddots & \\ \begin{bmatrix} a_{p-1} & 0 \\ 0 & b_{p-1} \end{bmatrix} & & & \begin{bmatrix} a_0 & 0 \\ 0 & b_0 \end{bmatrix} \end{bmatrix} .$$

As $b_0 \in \mathrm{Isom}\,(X_1, Y_0)$, it is immediate that $\dim \mathrm{Ker}\,\mathcal{A}_p = \dim \mathrm{Ker}\,\mathfrak{a}_p$ where

$$\mathfrak{a}_p = \begin{bmatrix} a_0 & & & & \bigcirc \\ a_1 & \ddots & & & \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ a_{p-1} & & & a_1 & a_0 \end{bmatrix} \in \mathcal{L}(X_0^p, Y_1^p) \ .$$

From Theorem 3.1, the algebraic multiplicity γ is equally computed from $A(\mu)$ or $a(\mu)$, and $\mathfrak{a}_p$ plays the same role as $\mathcal{A}_p$ upon replacing $A(\mu)$ by $a(\mu)$. Therefore, the problem comes down to proving formula (2.6) in the finite dimensional case. From the existence of a local Smith form, it is possible to duplicate the proof given in [7, Proposition 1.15, p. 35] for the case when $a(\mu)$ is a polynomial.

Remark 1: If the algebraic multiplicity is defined, then $A(\mu)$ is invertible for $|\mu| \neq 0$ small enough. This follows from the decomposition (4.5) of $B(\mu)A(\mu)C(\mu)$ since $a(\mu)$ is invertible for $|\mu| \neq 0$ small enough and $Q_0A(0)|_{X_1}$ is invertible. Invertibility of $a(\mu)$ is indeed ensured by existence of a local Smith form. Conversely, if $A(\mu)$ is analytic and $A(\mu)$ is invertible for $|\mu| \neq 0$ small enough (and $A(0)$ is Fredholm with index zero) the converse is true. Indeed, from (4.5) again, $a(\mu)$ is invertible for $|\mu| \neq 0$ and analytic. Hence $\det a(\mu)$ has an isolated zero at $\mu = 0$. Necessarily, this zero has finite order, so that the partial multiplicities of $\mu = 0$ in $a(\mu)$ are defined. The assertion follows from Theorem 3.1 since $a(\mu)$ is obtained from $A(\mu)$ through Lyapunov-Schmidt reduction of $F(\mu, x) = A(\mu)x$. $\square$

5. AN ALTERNATE PROOF OF THE KRASNOSELSKII-MAGNUS THEOREM

Taking $A(\mu) = D_x F(\mu, 0)$ with F of class $\mathcal{C}^1$ as in Section 3, we shall now give a proof of Theorem 2.2 in Magnus [13], a generalization of the famous theorem by Krasnoselskii, based on Lyapunov-Schmidt reduction. Bifurcation for the reduced equation is next proved using a standard argument. In his proof, Magnus also makes use of a reduction to the finite dimensional case, which differs from the usual Lyapunov-Schmidt procedure. The reason for this is that Theorem 3.1 seems to be difficult to prove using Magnus' definition of algebraic multiplicity (reviewed in Section 1). In complementing remarks, we shall also see that the smoothness assumptions can be weakened in a way depending of the <u>partial</u> multiplicities.

Since $A(0) = A_0 = D_x F(0, 0)$ is Fredholm with index zero by hypothesis, let us first observe that algebraic multiplicity of $\mu = 0$ in $A(\mu)$ is defined upon merely assuming that no generalized Jordan chain can be continued indefinitely. Indeed, the same is then true with $a(\mu) = D_\epsilon f(\mu, 0)$ replacing $A(\mu)$, where f denotes any reduced mapping of F (cf. (3.6)), as it follows from arguments in the proof of Theorem 3.1. From the results of Section 4 and Theorem 3.1, the length of the generalized Jordan chains of $A(\mu)$ is uniformly bounded, which suffices to prove the claim. Moreover, the algebraic multiplicity of $\mu = 0$ in $A(\mu)$ equals the order of $\mu = 0$ as a zero of $\det a(\mu)$. With Taylor's formula and since $f(\mu, 0) = 0$, one has

$$f(\mu, \epsilon) = a(\mu)\epsilon + g(\mu, \epsilon) ,$$

with g of class $\mathcal{C}^1$ since f is of class $\mathcal{C}^1$ and $a(\mu)$ is smooth. Obviously, $g(\mu, 0) = 0$ and $D_\epsilon g(\mu, 0) = 0$. If F (hence f, and also g) is of class $\mathcal{C}^2$, bifurcation for $f = 0$ follows from applying [3, Theorem 7.1, p. 201] provided that $\mu = 0$ has odd algebraic multiplicity γ. If F is only $\mathcal{C}^1$ and γ is odd, the same conclusion holds. To see this, it suffices to observe that the same result in [3] remains valid if one makes use of the "strong" version of the Implicit function theorem (see e.g. Lyusternik and Sobolev [12]) in the proof of [3, Lemma 7.2, p. 202]. Accordingly, we have shown

<u>Theorem 5.1</u>: <u>let</u> $F(= F(\mu, x))$ <u>be a</u> $\mathcal{C}^1$ <u>mapping from</u> $R\times X$ <u>to</u> Y <u>with</u> X <u>and</u> Y <u>real Banach spaces, locally defined near the origin and verifying</u> $F(\mu, 0)) = 0$. <u>Suppose that</u> $D_x F(\mu, 0)$ <u>is smooth and that</u> $A_0 = D_x F(0, 0)$ <u>is Fredholm with index zero. Finally, assume that the algebraic multiplicity</u> γ <u>of</u> $\mu = 0$ <u>is defined. Then, bifurcation occurs if</u> γ <u>is odd and the bifurcated solutions contain a continuum</u>.

In a given problem, finding out the algebraic multiplicity γ can be done using Magnus' definition in terms of projections, or our definition in terms of generalized Jordan chain or else from the calculation of an appropriate null-space according to formula (2.6). Whatever the option taken, the success of the procedure depends much on what is actually known about the derivatives $A^{(j)}(0)$. In any case, many corollaries to Theorem 5.1 can be proved formulating ad hoc hypotheses. We shall only mention one particular situation, namely when $\gamma = n(= \dim \operatorname{Ker} A_0)$ and n is odd. Obviously, $\gamma = n$ if and only if $\kappa_1 = \cdots = \kappa_n = 1$. It is straightforward to check that this amounts to assuming

$$A_1^{-1}(\text{Range } A_0) \cap \operatorname{Ker} A_0 = \{0\} , \tag{5.1}$$

and one finds again Theorem A of Westreich [20], who has another proof, not involving algebraic or partial multiplicities. His Theorem B -- on bifurcation from a smooth curve -- is also a particular case of the results of the next section. Note that condition (5.1) generalizes that of Crandall and Rabinowitz in [5].

It is also worth mentioning that smoothness of $D_xF(\mu,0)$ is unnecessarily restrictive in Theorem 5.1. Indeed, assuming $D_xF(\mu, 0)$ to be of class ^{r} with $r \geq \gamma$ is sufficient to obtain a local Smith form as in Section 4 and repeat all other necessary arguments to prove Theorem 5.1. But even this assumption can be weakened to $r \geq \kappa_1$ (of course, $r \geq \gamma$ or $r \geq \kappa_1$ can only be observed a posteriori), although the method of Section 4 to obtain a local Smith form fails. Assuming $r \geq \kappa_1$, reduction of the problem to proving bifurcation for $f = 0$ remain true and the proof of Theorem 5.1 goes through provided that one can show that det $a(\mu)$ changes sign as μ crosses 0. This is where the results of Section 4 help in the smooth case. An alternate procedure is as follows: write

$$a(\mu) = \sum_{j=0}^{\kappa_1} \frac{\mu^j}{j!} a^{(j)}(0) + \mu^{\kappa_1} R(\mu) , \tag{5.2}$$

with $R(\mu)$ continuous and $R(0) = 0$. From the definition of κ_1, $a(\mu)$ and the principal part of its Taylor expansion in (5.2), denoted by $\tilde{a}(\mu)$, have the same generalized Jordan chains. Observe in passing that a chain with length $\leq \kappa_1$ involves derivatives of $a(\mu)$ of order $\leq \kappa_1 - 1$ only, so that it is actually possible to check that no chain has length $\geq \kappa_1$ when $a(\mu)$ is only of class $\mathcal{C}^{\kappa_1}$. From the above, $a(\mu)$ and $\tilde{a}(\mu)$ have the same partial multiplicities, hence the same algebraic multiplicity. As $\tilde{a}(\mu)$ is a polynomial, it possesses a local Smith form after identifying X_0 and Y_1 with R^n. This means that there are smooth parametrized families $\tilde{m}(\mu)$ and $\tilde{n}(\mu)$ of invertible operators with constant nonzero determinants such that

$$\tilde{a}(\mu) = \tilde{m}(\mu)\tilde{d}(\mu)\tilde{n}(\mu)$$

and $\tilde{d}(\mu) = \mathrm{diag}\,(\mu^{\kappa_1}, \ldots , \mu^{\kappa_n})$. From (5.2)

$$\tilde{m}^{-1}(\mu)a(\mu)\tilde{n}^{-1}(\mu) = \tilde{d}(\mu) + \mu^{\kappa_1}\tilde{R}(\mu) , \tag{5.3}$$

with $\tilde{R}(\mu) = \tilde{m}^{-1}(\mu)R(\mu)\tilde{n}^{-1}(\mu)$. In particular, $\tilde{R}(0) = 0$. Now, for $\mu \neq 0$, $\tilde{d}(\mu)$ is invertible and

$$\tilde{d}^{-1}(\mu)\tilde{m}^{-1}(\mu)a(\mu)\tilde{n}^{-1}(\mu) = I + \mu^{\kappa_1}\tilde{d}^{-1}(\mu)\tilde{R}(\mu) .$$

but $\mu^{\kappa_1}\tilde{d}^{-1}(\mu) = \mathrm{diag}\,(1, \mu^{\kappa_1-\kappa_2}, \ldots, \mu^{\kappa_1-\kappa_n})$ has a continuous extension at $\mu = 0$. As $\tilde{R}(0) = 0$, we find that $\det(I + \mu^{\kappa_1}\tilde{d}^{-1}(\mu)\tilde{R}(\mu))$ is positive for $|\mu|$ small enough. The above relation thus shows that $(\det \tilde{d}^{-1}(\mu)) \det a(\mu)$ has the same sign for both $\mu > 0$ and $\mu < 0$ with $|\mu|$ small enough. If $\gamma = \kappa_1 + \cdots + \kappa_n$ is odd, it follows that $\det a(\mu)$ changes sign as μ crosses 0 since this is true of $\det \tilde{d}(\mu)$.

Remark 1: Let $X = Y$ and $A(\mu) = (I - \lambda_0 L) - \mu L$ with $L \in \mathcal{L}(X)$ compact and $1/\lambda_0 \in Sp(L)$. The generalized Jordan chains of $A(\mu)$ are usual Jordan chains of L corresponding to the eigenvalues $1/\lambda_0$ of L, the i^{th} chain of any canonical set generates a cyclic Jordan subspace with dimension κ_i and $\gamma = \kappa_1 + \cdots + \kappa_n$ is the dimension of the generalized null-space of $I - \lambda_0 L$, namely the algebraic multiplicity of $1/\lambda_0$ in the usual sense. In this case, Theorem 5.1 coincides with Krasnoselskii's, apart from extra regularity assumptions. □

Remark 2: If $\dim \mathrm{Ker}\, A_0 = 1$ (i.e. $n = 1$) more can be said about the structure of the bifurcated solutions and another criterion for bifurcation can be found, see [16] for details. □

6. INVARIANCE UNDER CHANGES OF VARIABLES AND APPLICATIONS.

We shall begin with a brief review of results on equivalence between Banach space valued mappings. In the finite dimensional case, this notion lies at the bottom of singularity theory. Its first appearance in bifurcation theory dates back to the paper by Golubitsky and Schaeffer [8] in their study of perturbed bifurcation. In [8], equivalence was used in connection with the theory of unfoldings towards finding normal forms for a given problem but many other important applications exist. One of them will be given here to a generalization of Theorem 5.1 when bifurcation is studied from an arbitrary curve, not necessarily the trivial branch.

Consider four real Banach spaces U, V, $\tilde{U}$ and $\tilde{V}$ and a pair of sufficiently smooth mappings

$$G : U \to V \ , \ \tilde{G} : \tilde{U} \to \tilde{V} \ ,$$

locally defined near the origin. The mappings G and $\tilde{G}$ are said to be equivalent if there are sufficiently smooth mappings

$$\tilde{\tau} : \tilde{V} \to \mathrm{Isom}\ (V, \tilde{V}) \ ,$$

$$\rho : \tilde{U} \to U \ ,$$

both locally defined near the origin and with ρ being an origin preserving local diffeomorphism, such that

$$\tilde{G}(\tilde{u}) = \tilde{\tau}(\tilde{u})G(\rho(\tilde{u})) \ , \tag{6.1}$$

for $\tilde{u}$ in a neighborhood of the origin in $\tilde{U}$. It is well known (and easily checked) that the above notion encompasses the apparently more general notion of equivalence in which

$$\tilde{G}(\tilde{u}) = \tilde{\psi}(\tilde{u},\ G(\rho(\tilde{u}))) \ , \tag{6.2}$$

where $\tilde{\psi} : \tilde{U}\times V \to \tilde{V}$ is a sufficiently smooth mapping with $\tilde{\psi}(\tilde{u}, \cdot)$ being an origin-preserving local diffeomorphism of V to $\tilde{V}$.

Equivalence as in (6.1) preserves local bifurcation properties, for the local zero sets of G and $\tilde{G}$ are deduced from each other through the local diffeomorphisms ρ and ρ^{-1}. A "good" bifurcation theorem should then comply with the requirement that it equally applies with G or any mapping $\tilde{G}$ equivalent to G. To begin with, we will be interested in proving independence of the hypotheses of Theorem 5.1 under changes of variable, i.e. under equivalence (6.1) when $U = \tilde{U} = R\times X$, $= \tilde{V} = Y$ and $G = F$. Clearly, Theorem 5.1 cannot be stated without referring to the trivial branch and hence changes of variable $\rho(\tilde{\mu}, \tilde{x})$ must leave the trivial branch invariant. Any such change of variable has the form

$$\rho(\tilde{\mu}, \tilde{x}) = (\alpha(\tilde{\mu}, \tilde{x}) \ , \ \xi(\tilde{\mu}, \tilde{x})) \ ,$$

with $\xi(\tilde{\mu}, 0) = 0$, and hence $D_{\mu}\alpha(0, 0) \neq 0$, $D_{x}\xi(0, 0) \in \text{Isom}\,(X)$. Setting

$$\tilde{F}(\tilde{\mu}, \tilde{x}) = \tilde{\tau}(\tilde{\mu}, \tilde{x})F(\alpha(\tilde{\mu}, \tilde{x}) \ , \xi(\tilde{\mu}, \tilde{x})) \ , \tag{6.3}$$

an elementary calculation provides (note that $D_{\mu}F(\mu, 0) = 0$ as a result of $F(\mu, 0) = 0$)

$$D_{x}\tilde{F}(\tilde{\mu}, 0) = \tilde{\tau}(\tilde{\mu}, 0)D_{x}F(\alpha(\tilde{\mu}, 0) \ , 0)D_{x}\xi(\tilde{\mu}, 0) \ . \tag{6.4}$$

For notational convenience, we shall set

$$A(\mu) = D_{x}F(\mu, 0) \ , \ \tilde{A}(\tilde{\mu}) = D_{x}\tilde{F}(\tilde{\mu}, 0)$$

and

$$\tilde{B}(\tilde{\mu}) = \tilde{\tau}(\tilde{\mu}, 0) \ , \ C(\tilde{\mu}) = D_{x}\xi(\tilde{\mu}, 0) \ .$$

Letting

$$\sigma(\tilde{\mu}) = \alpha(\tilde{\mu}, 0) \ ,$$

relation (6.4) becomes

$$\tilde{A}(\tilde{\mu}) = \tilde{B}(\tilde{\mu})A(\sigma(\tilde{\mu}))C(\tilde{\mu}).$$

Assuming the mappings $\tilde{\tau}(\cdot, 0)$, $\alpha(\cdot, 0)$ (i.e. σ) and $\xi(\cdot, 0)$ are smooth, the hypotheses of Theorem 5.1 will equally be satisfied with F or $\tilde{F}$ if the algebraic multiplicity of $\mu = 0$ in $A(\mu)$ equals the algebraic multiplicity of $\tilde{\mu} = 0$ in $\tilde{A}(\tilde{\mu})$. Indeed, $\tilde{A}(0)$ is Fredholm with index zero if and only if this is true of $A(0)$ since $\tilde{B}(0)$ and $C(0)$ are isomorphisms. From Corollary 2.1 (see also Remark 5 of Section 2) it suffices to compare the algebraic multiplicity of $\mu = 0$ and $\tilde{\mu} = 0$ in $A(\mu)$ and $A(\sigma(\tilde{\mu}))$ respectively. Equality of these follows from Remark 2 of Section 3 since $\sigma'(0) \neq 0$.

Now, consider a smooth mapping $\mathcal{F} = \mathcal{F}(u)$ defined on a neighborhood of the origin of some Banach space U with values in another Banach space Y. Suppose that $\mathcal{F}(0) = 0$ and $D\mathcal{F}(0)$ is a Fredholm operator with index one such that

$$\dim \operatorname{Ker} D\mathcal{F}(0) = n + 1 \quad (n \geq 1) \ . \tag{6.6}$$

This implies that $D\mathcal{F}(0)$ has rank-deficiency n. Suppose that a smooth curve $u(\mu)$ with $u(0) = 0$ is known to be in the local zero set of $\mathcal{F}$, namely

$$\mathcal{F}(u(\mu)) = 0 \ , \tag{6.7}$$

where $u'(0) = u_0 \neq 0$ (a condition for instance satisfied when $u(\mu)$ is an arclength parametrization). This situation is considered in [16] under the extra assumption $n = 1$, but the entire analysis can be repeated verbatim. We shall therefore pass to the conclusions and refer the interested reader to [16, Section 4] for further detail. Differentiating (6.7), one finds $D\mathcal{F}(0)u_0 = 0$, namely $u_0 \in \operatorname{Ker} D\mathcal{F}(0)$. Choosing X to be an arbitrary complement of $\mathbb{R}u_0$ in U, and for (μ, x) in a neighborhood of the origin in $\mathbb{R}\times X \approx U$, set

$$F(\mu, x) = \mathcal{F}(u(\mu) + x) \ .$$

Clearly, $D_x F(\mu, 0) = D\mathcal{F}(0)|_X$ is Fredholm with index zero while bifurcation from the curve $u(\mu)$ in $\mathcal{F} = 0$ amounts to bifurcation from the trivial branch in $F = 0$. The hypotheses of Theorem 5.1 must then be checked with

$$A(\mu) = D\mathcal{F}(u(\mu))|_X ,$$

and whether or not they are satisfied is independent of the choice of the complement X of the line $\mathbb{R}u_0$ (the tangent to the curve of known solutions at the origin) as well as of the regular parametrization $u(\mu)$. In particular, this allows for an intrinsic definition of the algebraic multiplicity of the singularity of $\mathcal{F}$ at the origin, viewed in a smooth curve of solutions to $\mathcal{F} = 0$. Assumptions of regularity on $\mathcal{F}$ and $u(\cdot)$ can be weakened according to the comments made in Section 5.

7. VARIATIONAL BIFURCATION THEOREMS.

The question that most naturally arises in view of the results obtained in Section 5 is whether the notions previously developed are also appropriate for proving bifurcation theorems for general variational problems. That the answer to this question is positive best illustrates the advantage of our approach. Indeed, no such result is derived in Magnus [13] or Ize [9] and neither the reference book by Chow and Hale [3] nor the recent monograph of Rabinowitz [18] contains any such theorem, except of course for the special class of nonlinear eigenvalue problems.

To begin with, we shall establish a bifurcation theorem for general variational problems in the finite dimensional case. More precisely, let $F = F(\mu, x)$ be a $\mathcal{C}^1$ mapping locally defined near the origin of $R \times R^n$ with values in R^n and of the form

$$F(\mu, x) = \nabla J(\mu, x) , \tag{7.1}$$

where J is some real-valued functional. Here and in what follows, "∇" denotes the gradient operator with respect to the variable x alone. We shall further assume that $F(\mu, 0) \equiv 0$ and set

$$A(\mu) = D_x F(\mu, 0) \in \mathcal{L}(R^n) . \tag{7.2}$$

Of course, $A(\mu)$ is selfadjoint for every μ. To make the question significant of whether nontrivial solutions to $F = 0$ bifurcate from $(0, 0)$, $A(0)$ must be singular. Suppose that $A(\mu)$ is nevertheless invertible for $|\mu| > 0$ small enough: then, the Morse index of $A(\mu)$ (number of positive eigenvalues of $A(\mu)$) is constant for $\mu > 0$ and $\mu < 0$ with $|\mu|$ small enough. Denote these indices by h_+ and h_- respectively. In Theorem 7.1 below, we show that the condition $h_+ \neq h_-$ guarantees bifurcation. This will be obtained as a simple application of C. Conley's theory to the (local) flow defined by the differential equation

$$\frac{dx}{dt} = F(\mu, x) \quad . \tag{7.3$_\mu$}$$

In particular, neither the statement nor the proof of the theorem makes reference to partial or algebraic multiplicities. These will appear in connection with the condition $h_+ \neq h_-$ and used as a substitute for it in the infinite dimensional case when the Morse index is no longer defined and when Conley's theory does not directly apply. Regarding the definition and properties of Conley index, we shall make reference to the easily accessible book by Smoller [19]. Another reference is of course Conley's original monograph [4].

Theorem 7.1: If $h_+ \neq h_-$, nontrivial solutions to $F = 0$ bifurcate from $(0, 0)$.

Proof: Irrespective of the condition $h_+ \neq h_-$, bifurcation obviously occurs if $x = 0$ is not an isolated solution to $F(0, \cdot) = 0$. In the remainder of the proof, we shall then make the non restrictive hypothesis that $x = 0$ is an isolated solution to $F(0, \cdot) = 0$. Then, it follows from [19, Theorem 23.32, p. 503] and the gradient nature of the local flow generated by $(7.3)_0$ that $x = 0$ is an isolated invariant set. Incidentally, note that the result we intend to prove is unaffected by modifying F in such a way that $F(\mu, x)$ is unchanged for, say, $|x| \leq 1$ and $F(\mu, x) \equiv 0$ for $|x| \geq 2$. Doing so allows one to speak of the global flow generated by $(7.3)_\mu$ and makes available the simplified approach taken in [19, Chap. 22].

Let Δ_δ denote the open ball with radius $\delta > 0$ and center 0 in $\mathbf{R}^n$. From [19, Theorem 22.18, p. 468 and Proposition 22.12, p. 464], Δ_δ contains an isolating neighborhood N with $0 \in N$. Here, "isolating neighborhood" is understood in the sense of [19, Definition 22.3, p. 460] (and not merely in the extended sense of [19, Definition 23.4, p. 481]). Then, N remains an isolating neighborhood for the flow generated by $(7.3)_\mu$ for $|\mu|$ small enough, and the Conley index $h(S_\mu)$ of the isolated invariant set $S_\mu \subset N$, is independent of μ. As a result, one has $S_\mu \neq \{0\}$ for either $\mu > 0$ or $\mu < 0$ (or both). Indeed, suppose $\mu > 0$ and $S_\mu = \{0\}$. Then, since $A(\mu)$ is invertible, $h(S_\mu) = h_+$ (see [19, pp. 503-504]). Similarly, if $\mu < 0$ and $S_\mu = \{0\}$, $h(S_\mu) = h_-$. Thus, if $S_\mu = \{0\}$ for both $\mu > 0$ and $\mu < 0$, $h_+ = h_-$, a contradiction. Because of the gradient nature of the flow $(7.3)_\mu$ (1) $S_\mu \neq \{0\}$ must contain a solution to $F(\mu, x) = 0$ different from $x = 0$: assuming that it does not, one finds that S_μ contains a complete orbit. But this orbit must then tend to two different rest points in both time directions (because $F(\mu, \cdot)$ is a gradient), one of which is necessarily nonzero, a contradiction.

In summary, for every $\delta > 0$ and every $\mu > 0$ or $\mu < 0$ (or both) with $|\mu|$ small enough, we have found a pair (μ, x_μ) with $x_\mu \neq 0$, $x_\mu \in S_\mu \subset N \subset \Delta_\delta$ and $F(\mu, x_\mu) = 0$. Since δ is arbitrary, the desired bifurcation property is established. □

Remark 1: The condition $h_+ \neq h_-$ is guaranteed if, for instance, $\det A(\mu)$ changes sign as μ crosses 0. Thus, Theorem 7.1 appears as a generalization of the bifurcation theorem valid under the latter assumption, but only when $F(\mu, \cdot)$ is a gradient. □

When the space R^n is replaced by an infinite-dimensional (real) Hilbert space X, the two problems arise that the Morse index is generally not defined and that, in any case, the method based on Conley index is not applicable. A natural procedure is to seek a reduction to the finite dimensional case in a

(1) Modifying F so that $F(\mu, x) = 0$ for $|x| \geq 2$ can obviously be done without affecting $F(\mu, \cdot)$ being a gradient.

form suitable for the application of Theorem 7.1. Naturally, and this is one of the difficulties, the reduction must comply with the requirement that it does not destroy the gradient nature of the problem. Also, conditions that can be verified directly on the given problem rather than on the reduced one must be given that ensure $h_+ \neq h_-$, in the latter.

We shall retain the same assumptions on F as before, except that now R^m is replaced by the space X. We shall write F in the form

$$F(\mu, x) = A(\mu)x + H(\mu, x) \ . \tag{7.4}$$

Since $A(\mu)$ is selfadjoint, $H(\mu, x)$ is the gradient of the functional

$$\Phi(\mu, x) = J(\mu, x) - \frac{1}{2}(A(\mu)x, x) \ . \tag{7.5}$$

Assuming that J and A are of class $\mathcal{C}^2$, it is obvious that Φ is of class $\mathcal{C}^2$, too, so that H is of class $\mathcal{C}^1$. Also, note that

$$H(\mu, 0) = 0 \ , \ D_x H(\mu, 0) = 0 \ . \tag{7.6}$$

The following lemma is a straightforward but crucial generalization of [3, Theorem 11.1]. Its proof is given for the convenience of the reader. Before, let us recall that a selfadjoint Fredholm operator has necessarily index zero. For this, see Kato [10], or Yosida [21]; the conclusion also follows from Deimling [6, p. 86].

<u>Lemma 7.1</u>: <u>Suppose that</u> $A(0) = A_0$ <u>is a Fredholm operator (with index zero) and</u> Ker A_0 <u>is stable under</u> $A(\mu)$ <u>for</u> μ <u>near the origin</u>. <u>Then, in the notation of Section 3, if Lyapunov-Schmidt reduction is performed with</u> $X_1 = Y_0 =$ Range A_0 <u>and</u> $Y_1 = X_0 =$ Ker A_0, <u>the resulting reduced equation is variational</u>. <u>More precisely</u>, $f(\mu, \epsilon) = 0$ <u>if and only if</u> ϵ <u>is a critical point of the functional</u>

$$\epsilon \in \operatorname{Ker} A_0 \to J(\mu,\ \epsilon + \phi(\mu,\ \epsilon))^{(2)} \qquad (7.7)$$

<u>Proof</u>: Since $A(\mu)$ is selfadjoint and $\operatorname{Ker} A_0$ is stable under $A(\mu)$, so is Range $A_0 = (\operatorname{Ker} A_0)^{\perp}$. Denoting by Q_1 and Q_0 the projections onto $\operatorname{Ker} A_0$ and Range A_0 respectively, according to the orthogonal decomposition

$$X = \operatorname{Ker} A_0 \oplus \text{Range } A_0 \ ,$$

it follows that Q_0 and Q_1 commute with $A(\mu)$. From this observation, it is easily found that the reduced equation is

$$Q_1 A(\mu)\epsilon + Q_1 H(\mu,\ \epsilon + \phi(\mu,\ \epsilon)) = 0 \ , \qquad (7.8)$$

with $\phi(\mu,\ \epsilon)$ characterized by

$$Q_0 A(\mu)\phi(\mu,\ \epsilon) + Q_0 H(\mu,\ \epsilon + \phi(\mu,\ \epsilon)) = 0 \ . \qquad (7.9)$$

On the other hand, as $\nabla J = F$, ϵ is a critical point of the functional (7.7) if and only if

$$(F(\mu,\ \epsilon + \phi(\mu,\ \epsilon))\ ,\ h + D_\epsilon \phi(\mu,\ \epsilon)\cdot h) = 0\ ,\ \forall h \in \operatorname{Ker} A_0\ .$$

But h and $D_\epsilon \phi(\mu,\ \epsilon)\cdot h$ being in $\operatorname{Ker} A_0$ and Range A_0 respectively, this reads

$$(Q_1 F(\mu,\ \epsilon + \phi(\mu,\ \epsilon))\ ,h) = 0\ ,\ \forall h \in \operatorname{Ker} A_0\ ,$$

$$(Q_0 F(\mu,\ \epsilon + \phi(\mu,\ \epsilon)\ ,\ D_\epsilon \phi(\mu,\ \epsilon)\cdot h) = 0\ ,\ \forall h \in \operatorname{Ker} A_0\ .$$

From (7.4) and (7.9), the second equation is automatically satisfied. The above system thus reduces to the first equation alone, which is nothing but $Q_1 F(\mu,\ \epsilon + \phi(\mu,\ \epsilon)) = 0$. From (7.4), this equation coincides with (7.8) and the proof is complete. □

<u>Remark 2</u>: Since $D_x H(\mu,\ 0) = 0$ from (7.4) and $A(\mu)\ \ D_x F(\mu,\ 0)$, differentiating (7.9) yields

(2) Of course, this statement is only local.

$$Q_0A(\mu)D_\epsilon\phi(\mu,\ 0) = 0\ .$$

Hence $D_\epsilon\phi(\mu,0) = 0$ since $Q_0A(\mu) \in \text{Isom (Range } A_0)$. Note however that this result heavily relies on the assumption that Ker A_0 is stable under $A(\mu)$: otherwise, (7.9) is not valid. □

To make Lemma 7.1 available in a general framework, note that the variational character of the equation $F(\mu,\ x) = 0$ is not affected by changing $F(\mu,\ x)$ into

$$\tilde{F}(\mu,\ x) = M^*(\mu)F(\mu,\ M(\mu)x)\ , \tag{7.10}$$

where $M(\mu)$ is any parametrized family of invertible operators. Indeed, $\tilde{F}(\mu,\ x) = \nabla\tilde{J}(\mu,\ x)$ where

$$\tilde{J}(\mu,\ x) = J(\mu,\ M(\mu)x)\ . \tag{7.11}$$

Changing F into $\tilde{F}$ preserves the trivial branch and, in this process, $A(\mu)$ and $H(\mu,\ x)$ are transformed into

$$\tilde{A}(\mu) = M^*(\mu)A(\mu)M(\mu) \tag{7.12}$$

and

$$\tilde{H}(\mu,\ x) = M^*(\mu)H(\mu,M(\mu)x)\ , \tag{7.13}$$

respectively. The question that now arises is to find conditions on $A(\mu)$ ensuring that a parametrized family $M(\mu)$ as above can be found so that $\tilde{A}(\mu)$ as in (7.12) leaves Ker $\tilde{A}(0)$ invariant. As shown in Lemma 7.2 below, a very simple sufficient condition is that $A(\mu)$ be <u>analytic</u>.

<u>Lemma 7.2</u>: <u>Suppose that $A(\mu)$ is an analytic family of selfadjoint operators with $A_0 = A(0)$ Fredholm (with index zero) and $A(\mu)$ invertible for $|\mu| \neq 0$ small enough. Then, there is an orthonormal basis of Ker A_0 and an analytic family $M(\mu)$ of invertible operators of $\mathcal{L}(X)$ such that $\tilde{A}(\mu) = M^*(\mu)A(\mu)M(\mu)$ has</u>

the block diagonal representation

$$\tilde{A}(\mu) = \begin{bmatrix} \tilde{D}(\mu) & 0 \\ 0 & \tilde{B}(\mu) \end{bmatrix} \tag{7.14}$$

with $\tilde{B}(\mu) \in \mathcal{L}(\text{Range } A_0)$ _analytic and invertible and_ $\tilde{D}(\mu) \in \mathcal{L}(\text{Ker } A_0)$ _of the form_

$$\tilde{D}(\mu) = \text{diag } (\sigma_1 \mu^{\kappa_1}, \ldots , \sigma_n \mu^{\kappa_n}) \tag{7.15}$$

in the given basis of Ker A_0, _where_ $\kappa_1, \ldots , \kappa_n$ ($n = \dim \text{Ker } A_0$) _are the partial multiplicities of_ $\mu = 0$ _in_ $A(\mu)$ _and_ $\sigma_j = \pm 1$, $1 \leq j \leq n$.

Note: This statement is different from that of Section 4 guaranteeing existence of a local Smith form.

Proof: Since A_0 is Fredholm and selfadjoint, 0 is an isolated eigenvalue of A_0 with finite multiplicity $n = \dim \text{Ker } A_0$. As $A(\mu)$ is analytic, it follows from Kato [11, pp. 122 and 386] that there is an analytic family $U(\mu) \in \mathcal{L}(X)$ of unitary operators with $U(0) = I_X$ such that

$$\hat{A}(\mu) = U^*(\mu)A(\mu)U(\mu) \tag{7.16}$$

has the representation

$$\hat{A}(\mu) = \begin{bmatrix} \hat{D}(\mu) & 0 \\ 0 & \hat{B}(\mu) \end{bmatrix} ,$$

relative to the decomposition $X = \text{Ker } A_0 \oplus \text{Range } A_0$. In addition, $\hat{D}(\mu)$ has the form

$$\hat{D}(\mu) = \text{diag}(\lambda_1(\mu), \ldots , \lambda_n(\mu)) , \tag{7.17}$$

in some orthonormal basis of Ker A_0, with $\lambda_j(\cdot)$ being analytic, $1 \leq j \leq n$. Although Kato's results are established in the case of a complex Hilbert space, the construction of $U(\mu)$ as a solution to a differential equation can equally be carried out in the real case. Observe that $\hat{D}(0) = 0$ since

$\hat{A}(0) = A_0$ from $U(0) = I_X$. Since $A(\mu)$ is invertible for $|\mu| \neq 0$ small enough, it follow from Remark 1 of Section 4 that the partial and algebraic multiplicities of $\mu = 0$ in $A(\mu)$ are well defined. From Corollary 2.1, they are the same in $A(\mu)$ and $\hat{A}(\mu)$ and, further, in $\hat{A}(\mu)$ and $\hat{D}(\mu)$. To prove the latter statement, note first that $\hat{A}(0)$ and $\hat{D}(0) = 0$ have null-space Ker A_0. Let then $\hat{e}(\mu)$ be a root function of $\hat{A}(\mu)$ of order j. Setting $\hat{e}(\mu) = \hat{e}_0(\mu) + \hat{e}_1(\mu)$ according to $X = \text{Ker } A_0 \oplus \text{Range } A_0$, one has $\hat{e}(0) = \hat{e}_0(0) \neq 0$ and $\hat{e}_0(\mu)$ is obviously a root function of $\hat{D}(\mu)$ whose order is no less than j. Conversely, any root function $\hat{e}_0(\mu)$ of $\hat{D}(\mu)$ is also a root function of $\hat{A}(\mu)$ with the same order since $\hat{A}(\mu)\hat{e}_0(\mu) = \hat{D}(\mu)\hat{e}_0(\mu)$. Equality of the partial multiplicities is an easy consequence of these observations, upon considering root functions corresponding to elements of canonical sets of generalized Jordan chains.

After rearranging the λ_j's in (7.15), which does not affect previous results provided that the vectors of the orthonormal basis of Ker A_0 are rearranged accordingly, one then has

$$\lambda_j(\mu) = \mu^{\kappa_j} d_j(\mu) \ , \ 1 \leq j \leq n \ ,$$

with $d_j(\cdot)$ analytic and $d_j(0) \neq 0$ (see Section 4 where a similar argument is used). Thus, we may write

$$\lambda_j(\mu) = \sigma_j \mu^{\kappa_j} \delta_j^2(\mu) \ ,$$

with $\delta_j(\mu) = |d_j(\mu)|^{\frac{1}{2}}$ and $\sigma_j = \text{sgn } d_j(0)$. From $d_j(0) \neq 0$, $\delta_j^{-1}(\cdot)$ is analytic. Finally, setting

$$S(\mu) = \begin{pmatrix} \Delta(\mu) & 0 \\ 0 & I \end{pmatrix} ,$$

where $\Delta(\mu) = \text{diag } (\delta_1^{-1}(\mu), \ldots , \delta_n^{-1}(\mu))$ in the same basis of Ker A_0, one finds the desired decomposition (7.14) with $\tilde{D}(\mu)$ as in (7.15) by taking

$M(\mu) = S(\mu)U(\mu)$. □

Assume then that $A(\mu) = D_x F(\mu, 0)$ is analytic with $A(\mu)$ invertible for $|\mu| \neq 0$ small enough. Choosing $M(\mu)$ as in Lemma 7.2 and with $\tilde{F}$ given by (7.10), bifurcation in the equation $F = 0$ occur if and only if it occurs in the equation $\tilde{F} = 0$. None of the assumptions on F is affected by so changing F into $\tilde{F}$: in particular, this does not affect the partial multiplicities. Doing so reduces the problem to the case when Ker A(0) is stable under $A(\mu)$ for $|\mu|$ small enough. From Lemma 7.1, bifurcation in $F = 0$ amounts to bifurcation in $f = 0$, where $f(\mu, \epsilon)$ is the gradient of the functional

$$\epsilon \in \operatorname{Ker} A(0) \to J(\mu, \epsilon + \phi(\mu, \epsilon))$$

with ϕ as in Lemma 7.1. Now, recall that not only $\phi(\mu, 0) = 0$ but also $D_\epsilon \phi(\mu, 0) = 0$ (see Remark 2). Using this along with $F(\mu, 0) = 0$ (i.e. $D_x J(\mu, 0) = 0$) a straightforward calculation shows that $f(\mu, 0) = 0$ and

$$D_\epsilon f(\mu, 0) = A(\mu)\epsilon \ , \ \forall \ \epsilon \in \operatorname{Ker} A(0) \ .$$

But, since the above holds upon changing F into $\tilde{F}$, $A(\mu)$ must be understood as being $\tilde{A}(\mu)$ in (7.14). Thus, for $\epsilon \in \operatorname{Ker} A(0)$, $A(\mu)\epsilon = \tilde{D}(\mu)\epsilon$ (after identifying ϵ with its decomposition in the suitable basis of Ker A(0)). From the form of $D(\mu)$ given by (7.15), it is immediate to check whether some change of the Morse index occurs as μ crosses 0. If so, Theorem 7.1 guarantees bifurcation. Note that $\tilde{D}(\mu)$ is invertible for $\mu \neq 0$ as is required by Theorem 7.1.

It is worth expanding a little on the condition ensuring that the Morse index of $\tilde{D}(\mu)$ changes as μ crosses zero. This depends on both the partial multiplicities κ_j and the σ_j's. Clearly, those terms involving an even κ_j do not contribute to any modification of the Morse index. On the contrary, $\sigma_j \mu^{\kappa_j}$ does change sign as μ crosses zero if κ_j is odd, but the Morse index will not change if an equal number of eigenvalues of $\tilde{D}(\mu)$ cross the origin in either direction, namely if $\sum_{\kappa_j \text{ odd}} \sigma_j = 0$. This condition is both necessary and sufficient for the Morse index to be unchanged. The criterion for bifurcation is then that $\sum_{\kappa_j \text{ odd}} \sigma_j \neq 0$ or, equivalently

$$\sum_{j=1}^{n} (1 + (-1)^{\kappa_j})\sigma_j \neq 0 \ . \tag{7.18}$$

Remark 3: A simple examination of the proof of Lemma 7.2 reveals that the $\lambda_j(\mu)$'s in (7.17) are the eigenvalues of $A_0 = A(0)$, and that $\sigma_j \mu^{\kappa_j}$ is the first nonzero term in the Taylor series of $\lambda_j(\mu)$ about the origin, apart from the multiplicative constant $|d_j(0)|$. In other words, σ_j is the sign of the first nonzero derivative of $\lambda_j(\cdot)$ at the origin, and condition (7.18) amounts to saying that the eigenvalues of $A(\mu)$ cross the origin in one direction more than in the other as μ crosses zero. □

For convenience, the family $(\sigma_1, \ldots, \sigma_n)$ obtained in Lemma 7.2 will be called a sign characteristic of $A(\mu)$ (at $\mu = 0$). Sign characteristics are unique to within permutations of $\{1, \ldots, n\}$ compatible with $\kappa_1 \geq \cdots \geq \kappa_n$. Of course, condition (7.18) is independent of the representative for the sign characteristic. We shall summarize the results obtained above in the following theorem.

Theorem 7.2: let $J : \mathbb{R}\times X \to \mathbb{R}$ be a $\mathcal{C}^2$ functional and set $F(\mu, x) = \nabla J(\mu, x)$. Suppose that $F(\mu, 0) = 0$ and $A(\mu) = D_x F(\mu, 0)$ is analytic and invertible for $\mu \neq 0$ near 0, with $A(0)$ Fredholm (with index zero). Finally, suppose that the condition

$$\sum_{j=1}^{n} (1 + (-1)^{\kappa_j})\sigma_j \neq 0 \ ,$$

holds, where $\kappa_1 \geq \cdots \geq \kappa_n$ *are the partial multiplicities of* $\mu = 0$ *in* $A(\mu)$ ($n = \dim \operatorname{Ker} A(0) \geq 1$) *and* $(\sigma_1, \ldots, \sigma_n)$ *is any sign characteristic of* $A(\mu)$ *at* $\mu = 0$. *Then, bifurcation in* $F = 0$ *occurs at* $(0, 0)$.

In practice, Theorem 7.2 has the inconvenience of requiring $A(\mu)$ to be analytic. Even if this should not be a severe restriction in many applications, it would be desirable to have a result that, at least, does not require checking analyticity of $A(\mu)$. It turns out that any problem of the form considered here involving an operator $A(\mu) = D_x F(\mu, 0)$ which is smooth - or even smooth enough - with $A(0)$ being Fredholm and such that the partial multiplicities $\kappa_1 \geq \cdots \geq \kappa_n$ ae well defined, is equivalent to a problem in which $A(\mu)$ is a polynomial (hence analytic). This result, that shows how Theorem 7.2 can be applied in the nonanalytic case as well, is now established.

Exposition 7.1: *Suppose that* $A(\mu)$ *is a smooth (or smooth enough) family of selfadjoint operators with* $A_0 = A(0)$ *Fredholm (with index zero) and set* $\dim \operatorname{Ker} A_0 = n \geq 1$. *Suppose that the partial multiplicities* $\kappa_1 \geq \cdots \geq \kappa_2$ *of* $\mu = 0$ *in* $A(\mu)$ *are well defined. For every integer* $k \geq \kappa_1$, *set*

$$A_k(\mu) = \sum_{j=0}^{k} \frac{\mu^j}{j!} A^{(j)}(0) .$$

Then, there is a smooth (or smooth enough) family $T_k(\mu) \in \mathcal{L}(X)$ *of operators invertible for* $|\mu|$ *small enough such that*

$$A_k(\mu) = T_k^*(\mu) A(\mu) T_k(\mu) .$$

Proof: To begin with, observe that the partial multiplicities $\kappa_1, \ldots, \kappa_n$ are the same for $A(\mu)$ and $A_k(\mu)$. This follows from finiteness of κ_1 involving only the derivatives $A^{(j)}(0)$, $0 \leq j \leq \kappa_1$, which, also, are all that is needed to characterize all possible generalized Jordan chains. Write

$$A(\mu) = A_k(\mu) + \mu^{k+1}R(\mu) ,$$

with $R(\mu)$ smooth. Recall that $A_k(\mu)$ is invertible for $|\mu| > 0$ small enough (cf. Remark 1 of Section 4). Hence, for $|\mu| > 0$ small enough

$$A(\mu) = A_k(\mu)(I + C_k(\mu)) , \tag{7.19}$$

where

$$C_k(\mu) = \mu^{k+1}A_k^{-1}(\mu)R(\mu) .$$

Since $A_k(\mu)$ is selfadjoint and $A_0 = A_k(0)$, one may use Lemma 7.2 with $A_k(\mu)$ and write

$$A_k(\mu) = M_k^*(\mu)\tilde{A}_k(\mu)M_k(\mu) ,$$

with $M_k(\mu)$ being analytic in μ and invertible for $|\mu|$ small enough (including $\mu = 0$) and with

$$\tilde{A}_k(\mu) = \begin{bmatrix} \tilde{D}_k(\mu) & 0 \\ 0 & \tilde{B}_k(\mu) \end{bmatrix} ,$$

where

$$\tilde{D}_k(\mu) = \mathrm{diag}(\sigma_1\mu^{\kappa_1} , \ldots , \sigma_n\mu^{\kappa_n}) ,$$

and $\tilde{B}_k(\mu) \in \mathcal{L}(\text{Range } A_0)$ (recall that $A_0 = A_k(0)$ as well) is analytic in μ and invertible for $|\mu|$ small enough. These properties immediately show that $\mu^{k+1}A_k^{-1}(\mu)$, and hence $C_k(\mu)$ above, is smooth and vanishes at the origin since $k \geq \kappa_1$.

Since both $A_k(\mu)$ and $A(\mu)$ are selfadjoint, it follows from (7.19) that

$$A(\mu) = A_k(\mu)(I + C_k(\mu)) = (I + C_k^*(\mu))A_k(\mu) ,$$

and hence

$$A_k(\mu)C_k(\mu) = C_k^*(\mu)A_k(\mu) .$$

Therefore, more generally, for every integer $p \geq 0$

$$A_k(\mu)(C_k(\mu))^p = (C_k^*(\mu))^p A_k(\mu) \quad . \tag{7.20}$$

As $C_k(0) = 0$, the operator

$$(I + C_k(\mu))^{\frac{1}{2}} \in \mathcal{L}(X)$$

is well defined for $|\mu|$ small enough through the series

$$(I + C_k(\mu))^{\frac{1}{2}} = \sum_{p=0}^{\infty} \alpha_p (C_k(\mu))^p \quad , \tag{7.21}$$

where the real coefficients α_p are as in the Taylor series of $(1 + x)^{\frac{1}{2}}$ at $x = 0$.

Combining (7.20) and (7.21), one finds

$$A_k(\mu)(I + C_k(\mu))^{\frac{1}{2}} = [(I + C_k(\mu))^{\frac{1}{2}}]^* A_k(\mu) \quad .$$

This shows that (7.19) may be rewritten as

$$A(\mu) = [(I + C_k(\mu))^{\frac{1}{2}}]^* A_k(\mu)(I + C_k(\mu))^{\frac{1}{2}} \quad .$$

The desired result follows by taking

$$T_k(\mu) = (I + C_k(\mu))^{-\frac{1}{2}} \quad .$$

Smoothness of $T_k(\mu)$ is ensured by smoothness of $(I + C_k(\mu))^{\frac{1}{2}}$, the latter resulting from smoothness of $C_k(\mu)$ and normal convergence of the series in (7.21) (for $|\mu|$ small enough). □

The way Proposition 7.1 can be combined with Theorem 7.2 to prove bifurcation in the equation $F = 0$ when $F(\mu, 0) \equiv 0$ and $A(\mu) = D_x F(\mu, 0)$ verifies the assumptions of Proposition 7.2 is clear: with $A_k(\mu)$ and $T_k(\mu)$ as above, set

$$F_k(\mu, x) = T_k^*(\mu) F(\mu, T_k(\mu)x) \quad .$$

then, $F_k(\mu, 0) = 0$ and $D_x F_k(\mu, 0) = A_k(\mu)$ is analytic. Moreover, $F_k(\mu, x)$ is the gradient of the functional

$$J_k(\mu, x) = J(\mu, T_k(\mu)x) .$$

As $T_k(\mu)$ is invertible for $|\mu|$ small enough, bifurcation in $F = 0$ occurs if and only if it occurs in $F_k = 0$, and bifurcation in $F_k = 0$ may be examined through Theorem 7.2.

Remark 4: If $X = R^n$, the condition $\sum_{j=0}^{n} (1 - (-1)^{\kappa_j})\sigma_j \neq 0$ where $(\sigma_1, \ldots, \sigma_n)$ is a sign characteristic of $A_k(\mu)$ at $\mu = 0$ is equivalent to saying that the Morse index of $A_k(\mu)$ changes as μ crosses zero. This follows from the argument that the $\sigma_j \mu^{\kappa_j}$'s represent (essentially) the dominant term of the eigenvalues $\lambda_{j,k}(\mu)$ of $A_k(\mu)$ (see Remark 3). From Proposition 7.1, it is easily inferred that, in addition, this condition is equivalent to saying that the Morse index of $A(\mu)$ itself changes as μ crosses zero. Indeed, $A(\mu)$ and $A_k(\mu)$ are congruent through the invertible operator $T_k(\mu)$, which can only happen if $A(\mu)$ an $A_k(\mu)$ have the same number of positive and negative eigenvalues (as is well known) and hence the same Morse index. Thus, the combination of Theorem 7.2 and Proposition 7.1 yields again a particular case of Theorem 7.1 (in which existence of partial multiplicities is not required). Also, the above arguments easily yield that the condition $\sum_{j=0}^{n} (1 - (-1)^{\kappa_j})\sigma_j \neq 0$ is independent not only of the sign characteristic of $A_k(\mu)$ at $\mu = 0$ but also of the Taylor polynomial $A_k(\mu)$ provided that $k \geq \kappa_1$. Although it is natural to conjecture that the same result is true if X is an infinite dimensional Hilbert space, we have not found a completely general proof of this as yet. □

Many bifurcation theorems can be derived from the combination of Theorem 7.2 and Proposition 7.1 by formulating assumptions that guarantee that the partial multiplicities are well defined and that condition (7.18) holds. For instance, assuming that the partial multiplicities are well defined and that the algebraic multiplicity $\gamma = \kappa_1 + \cdots + \kappa_n$ is odd yields bifurcation. Indeed, γ is odd if and only if an odd number of κ_j's are odd and, if so, condition (7.18) is necessarily satisfied. This result is nothing but Theorem 5.1 for a variational problem. Another (new) bifurcation theorem whose hypotheses should be especially easy to check in practice and that can be obtained according to this procedure is as follows.

Theorem 7.3: Let $J : R\times X \to R$ be a $\mathcal{C}^2$ functional and set $F(\mu, x) = \nabla J(\mu, x)$. Suppose that $F(\mu, 0) = 0$ and $A(\mu) = D_x F(\mu, 0)$ is smooth (or, more generally, $\mathcal{C}^1$; e.g. if J is $\mathcal{C}^3$). Suppose that A(0) is Fredholm (with index zero) and that $A'(0)$ is positive definite on Ker $A(0) \neq \{0\}$. Then, bifurcation in $F = 0$ occurs at (0, 0).

Note: Changing F into $-F$ shows that the result is valid upon replacing "positive definite" by "negative definite". For another generalization, see Remark 5 later.

Proof: First, we shall show that no generalized Jordan chain of $A(\mu)$ has length >1. This ensures that $\kappa_j = 1$, $1 \leq j \leq n$ ($n = \dim \operatorname{Ker} A(0)$). Otherwise, one can find $e_0 \in \operatorname{Ker} A(0) - \{0\}$ and $e_1 \in X$ such that

$$A(0)e_1 + A'(0)e_0 = 0 .$$

Hence

$$(A(0)e_1, e_0) + (A'(0)e_0, e_0) = 0 .$$

But $(A(0)e_1, e_0) = (e_1, A(0)e_0) = 0$ from $e_0 \in \operatorname{Ker} A(0)$. Thus, one finds $(A'(0)e_0, e_0) = 0$, in contradiction with $e_0 \neq 0$. Replacing $A(\mu)$ by, say,

$$A_1(\mu) = A(0) + \mu A'(0)$$

does not affect the hypotheses on $A(\mu)$, nor the partial multiplicities. As explained before, Proposition 7.1 allows one to reduce the problem to the case when $A(\mu)$ is analytic. The next step consists in proving that the sign characteristic of $A(\mu)$ at $\mu = 0$ is $(1, 1, \ldots, 1)$. If so, condition (7.19) is trivially fulfilled and bifurcation follows from Theorem 7.2.

In the assumption that $A(\mu)$ is analytic, let $(\lambda_j(\mu))$ denote the family of eigenvalues of $A(\mu)$ and consider an orthonormal family $(x_j(\mu))$ of associated eigenvectors. The λ_j's are analytic in μ and it is a standard result that the x_j's can be taken analytic, too. Differentiating the identity

$$A(\mu)x_j(\mu) = \lambda_j(\mu)x_j(\mu)$$

at $\mu = 0$ yields

$$A'(0)x_j(0) + A(0)x_j'(0) = \lambda_j'(0)x_j(0) + \lambda_j(0)x_j'(0) .$$

From $|x_j(\mu)| = 1$, one finds

$$(x_j'(0), x_j(0)) = 0 .$$

Hence

$$\lambda_j'(0) = (A'(0)x_j(0), x_j(0)) + (A(0)x_j'(0), x_j(0)) .$$

But $(A(0)x_j'(0), x_j(0)) = (x_j'(0), A(0)x_j(0)) = 0$ from $x_j(0) \in \operatorname{Ker} A(0)$ and we are left with

$$\lambda_j'(0) = (A'(0)x_j(0), x_j(0)) > 0 .$$

Since σ_j is the sign of the first nonzero derivative of λ_j at the origin (cf. Remark 3), one has $(\sigma_1, \ldots, \sigma_n) = (1, \ldots, 1)$, as desired. □

<u>Remark 5</u>: More generally, the proof of Theorem 7.3 can be repeated to show that bifurcation is guaranteed if the nondegeneracy condition (5.1) (ensuring

$\kappa_1 = \cdots = \kappa_n = 1$), namely

$$(A'(0))^{-1}(\text{Range } A(0)) \cap \text{Ker } A(0) = \{0\} ,$$

is satisfied and if the (nondegenerate) operator $PA'(0)P \in \mathcal{L}(\text{Ker } A(0))$ where P denotes the orthogonal projection onto Ker A(0), has nonzero Morse index. This is of course the case if n is odd, a particular case that remains valid when F is not a gradient as was seen in Section 5. □

Remark 6: Theorem 7.3 is a very particular case among the bifurcation theorems that can be deduced through the combination of Theorem 7.2 and Proposition 7.1. Still, it trivially contains the example when

$$A(\mu) = (\lambda_0 + \mu)I - L ,$$

and L is selfadjoint, λ_0 is an isolated eigenvalue of L with finite multiplicity. Apparently, this was the most general situation to be found in the literature (e.g. Section 4.11 in Chow and Hale [3]). □

It should not be inferred from Theorem 7.3 that the partial multiplicities of $\mu = 0$ in $A(\mu)$ and the sign characteristic(s) can always be determined from the restriction of the quadratic form $(A(\mu)x, x)$ for $x \in \text{Ker } A(0)$. Consider for instance the following counter example: Let $X = \mathbf{R}^4$ and

$$A(\mu) = \frac{1}{1 + \mu^2}\begin{pmatrix} \mu^3 + \mu^2 & 0 & -\mu^4 + \mu & 0 \\ 0 & \mu^3 - \mu^2 & 0 & -\mu^4 - \mu \\ -\mu^4 + \mu & 0 & \mu^5 + 1 & 0 \\ 0 & -\mu^4 - \mu & 0 & \mu^5 - 1 \end{pmatrix} .$$

Then, $A(\mu)$ has eigenvalues μ^3 (double) with eigenvectors $(1, 0, -\mu, 0)$ and $(0, 1, 0, -\mu)$, 1 (simple) with eigenvector $(\mu, 0, 1, 0)$ and -1 (simple) with eigenvector $(0, \mu, 0, 1)$. Obviously, $A(\mu)$ verifies the conditions required in

Theorem 7.2 since $n = \dim \operatorname{Ker} A(0) = 2$, $\kappa_1 = \kappa_2 = 3$ and $\sigma_1 = \sigma_2 = 1$. Now, with $x = (x_1, x_2, x_3, x_4)$, Ker A(0) consists of those vectors of the form $(x_1, x_2, 0, 0)$. For such a choice of x, one has

$$(A(\mu)x, x) = \frac{\mu^2}{1+\mu^2}[(\mu+1)x_1^2 + (\mu-1)x_2^2] .$$

This quadratic form is represented by the matrix

$$\frac{\mu^2}{1+\mu^2}\begin{pmatrix} \mu+1 & 0 \\ 0 & \mu-1 \end{pmatrix}$$

with eigenvalues $\mu^2(\mu \pm 1)/(1+\mu^2)$ yielding the wrong values $\kappa_1 = \kappa_2 = 2$, $\sigma_1 = 1$, $\sigma_2 = -1$.

As a conclusion, we shall now show that the condition $\sum_{j=1}^{n}(1-(-1)^{\kappa_j})\sigma_j \neq 0$ in Theorem 7.2 is sharp. Take $X = \mathbf{R}^2$ and, with $x = (x_1, x_2)$

$$F(\mu, x) = (\mu x_1 + x_2^3 + 3x_1^2x_2, \ -\mu x_2 + x_1^3 + 3x_1x_2^2) .$$

Then, $F = \nabla J$ with

$$J(\mu, x) = \tfrac{1}{2}\mu(x_1^2 - x_2^2) + x_1^3x_2 + x_1x_2^3 .$$

Clearly,

$$A(\mu) = \begin{pmatrix} \mu & 0 \\ 0 & -\mu \end{pmatrix} ,$$

so that $\operatorname{Ker} A(0) = \mathbf{R}^2$, $\kappa_1 = \kappa_2 = 1$, $\sigma_1 = 1$ and $\sigma_2 = -1$. Thus, here, $\sigma_1 + \sigma_2 = 0$ and, actually, no bifurcation occurs. Indeed, $F(\mu, x) = 0$ requires $F(\mu, x)\cdot(x_2, x_1) = 0$, namely $x_1^4 + 6x_1^2x_2^2 + x_2^4 = 0$, forcing $x_1 = x_2 = 0$.

The results of this section can be extended to the case when bifurcation is studied from an arbitrary smooth curve $(\mu, x(\mu))$ since changing x into

$x + x(\mu)$ does not affect the variational character of the problem and allows for a reduction to the case when bifurcation is studied from the trivial branch.

[1] Beyn, W. J., Private communication.

[2] Bohme, R., Die Losung der Verwegungsgleichungen fur nichtlineare Eigenwertprobleme, Math. Z., 127, pp. 105-126 (1972).

[3] Chow, S. N. and Hale, J. K., Methoas of bifurcation theory, Grund. Math. Wiss., 251, Springer (1982).

[4] Conley, C., Isolated invariant sets and the Morse index, Conf. Board Math. Sci., no. 38, AMS publications (1978).

[5] Crandall, M. and Rabinowitz, P., Bifurcation from simple eigenvalues, J. Funct. Anal., 8, pp. 321-340 (1971).

[6] Deimling, K., Nonlinear functional analysis, Springer (1985).

[7] Gohberg, I., Lancaster, P. and Rodman, L., Matrix polynomials, Comp. Sc. Appl. Math., Academic Press (1982).

[8] Golubitsky, M. and Schaeffer, D., A theory for imperfect bifurcation via singularity theory, Comm. Pure Appl. Math., 32, p. 21-98 (1979).

[9] Ize, J., Bifurcation theory for Fredholm operators, AMS Memoirs, 174 (1974).

[10] Kato, T., Perturbation theory for linear operators, 2nd ed., Grund. Math. Wiss., 132, Springer (1976).

[11] Krasnoselskii, M. A., Topological methods in the theory of nonlinear integral equations, Pergamon Press (1964).

[12] Lyusternik, L. A. and Sobolev, V. J., Elements of functional analysis, Gordon and Breach (1961).

[13] Magnus, R. J., A generalization of multiplicity and the problem of bifurcation, Proc. London Math. Soc., 32, 251-278 (1976).

[14] Marino, A., La biforcazione nel caso variazionale, Conf. Sem. Mat. Univ. Bari, 132 (1977).

[15] Rabier, P. J., Topics in one-parameter bifurcation problems, Tata Inst. Lec. Notes, 76, Springer (1985).

[16] Rabier, P. J., Generalized Jordan chains and bifurcation with one-dimensional null-space (preprint).

[17] Rabinowitz, P., A bifurcation theorem for potential operators, J. Funct. Anal., 25, pp. 412-424 (1977).

[18] Rabinowitz, P., Minimax methods in critical point theory with applications to differential equations, Conf. Board Math. Sci., no. 65, AMS Publications (1986).

[19] Smoller, J., Shock Waves and Reaction-diffusion Equations, Grund. Math.

Wiss., 258, Springer (1983).

[20] Westreich, D., Bifurcation at eigenvalues of odd multiplicity, Proc. Am. Math. Soc., 41, pp. 609-614 (1973).

[21] Yosida, K., Functional analysis, Springer (1965).

END
DATE
FILMED
6-88
DTIC