

Review

# What Do You Mean by “Nonlinear Eigenvalue Problems”?

Raffaele Chiappinelli

Dipartimento di Ingegneria dell’Informazione e Scienze Matematiche, Università di Siena, I-53100 Siena, Italy; raffaele.chiappinelli@unisi.it

Received: 30 March 2018; Accepted: 6 June 2018; Published: 9 June 2018



**Abstract:** A nonlinear eigenvalue problem is generally described by an equation of the form  $F(\lambda, x) = 0$ , where  $F(\lambda, 0) = 0$  for all  $\lambda$ , and contains by definition two unknowns: the eigenvalue parameter  $\lambda$  and the “nontrivial” vector(s)  $x$  corresponding to it. The nonlinear dependence of  $F$  can be in either of them (and of course in both), and also the research in this area seems to follow two quite different directions. In this review paper, we try to collect some points of possible common interest for both fields.

**Keywords:** eigenspace; algebraic multiplicity; bifurcation; bifurcation

**MSC:** Primary 47J10; Secondary 47A56

## 1. Introduction

Nonlinear eigenvalue problems are generally described by equations of the form

$$F(\lambda, x) = 0 \quad (\lambda \in \mathbb{K}, x \in E) \quad (1)$$

where  $\mathbb{K}$  ( $= \mathbb{R}$  or  $\mathbb{C}$ ) is the field of real or complex numbers, and  $E$  is a real or complex Banach space that can in particular be the  $n$ -space  $\mathbb{R}^n$  or  $\mathbb{C}^n$ . In Equation (1),  $F$  is a continuous map of  $\mathbb{K} \times E$  into  $E$ , and it is assumed that  $F(\lambda, 0) = 0$  for all scalars  $\lambda$ . That is to say,  $x = 0$  solves trivially Equation (1) for all  $\lambda$ ; and one looks therefore for those  $\lambda$ 's (the *eigenvalues* of  $F$ ) such that Equation (1) has a solution  $x \neq 0$  (an *eigenvector* of  $F$  corresponding to  $\lambda$ ).

Of course, Equation (1) contains as a (very) special case the proper eigenvalue-eigenvector equation of Linear Algebra and Linear Functional Analysis,

$$F(\lambda, x) = Ax - \lambda x = (A - \lambda I)x = 0 \quad (2)$$

in which  $A \in L(E)$ , the space of all bounded linear operators acting in  $E$  and  $I$  is the identity map; to stress the linearity of  $A$ , we write as usual  $Ax$  rather than  $A(x)$ . In addition to Equation (2), consider now the following special forms of Equation (1):

$$F(\lambda, x) = G(\lambda)x = 0, \quad (3)$$

and

$$F(\lambda, x) = A(x) - \lambda C(x) = 0. \quad (4)$$

Evidently, both Equations (3) and (4) encompass the classical case Equation (2). However, there is quite a difference between them: in Equation (3),  $F$  depends linearly on  $x$  and arbitrarily in  $\lambda$ , the latter dependence being driven by a map  $G : \mathbb{K} \rightarrow L(E)$ , while, in Equation (4), it is rather the opposite, for here it is the dependence on  $x$  that is (possibly) nonlinear as dictated by the continuous

maps  $A, C : E \rightarrow E$ . One first consequence is that the terms *eigenvalue/eigenvector/eigenspace* retain their usual significance in the case of Equation (3), while on the contrary they have in general a poor meaning in the case of Equation (4). On the other hand, in the latter case, assuming that  $C(x) \neq 0$  for  $x \neq 0$ , the eigenvalue associated with an eigenvector is uniquely determined, for  $A(\hat{x}) - \lambda_1 C(\hat{x}) = A(\hat{x}) - \lambda_2 C(\hat{x})$  with  $\hat{x} \neq 0$  implies that  $\lambda_1 = \lambda_2$ , while in the former case this is not necessarily true.

In fact, in the past decades both Equations (3) and (4) have been usually referred to as *Nonlinear Eigenvalue Problems*. Those of the type in Equation (3), especially with  $E = \mathbb{K}^n$ , have been extensively studied in Numerical Analysis and Matrix Analysis (see, for instance, the review paper [1], where the abbreviation NLEVP is used to designate them), while problems of the type in Equation (4) have formed a main subject in Nonlinear Functional Analysis and its applications to differential equations, and are at the basis of, among others, Bifurcation Theory; see, for instance, the nowadays classic book [2] or the most recent [3].

Historically, the study of Nonlinear Eigenvalue Problems can be dated back to nearly one century ago, if we look in particular to the work—inspired by D. Hilbert and E. Landau—of E. Schmidt and A. Hammerstein, and subsequently of M. Golomb, on parameter-dependent nonlinear integral equations of the form

$$\lambda u(x) = \int_{\Omega} k(x, y) f(y, u(y)) dy.$$

On the eastern side of Europe, the investigation of this kind of problems received a strong impulse in the former Soviet Union on behalf of M.A. Krasnosel'skii and I.T. Gohberg. They were both pupils of M.G. Krein, and their subsequent work during many decades of the second half of the last century seems to have developed mainly on problems of the type in Equation (3) by Gohberg, and mainly on problems of the type in Equation (4) by Krasnosel'skii. For this reason, and in honor of these two true giants of Nonlinear Functional Analysis, I will often refer in the sequel to Equation (3) as describing problems of *type G*, and to Equation (4) as describing problems of *type K*.

The present paper does not contain new results in either field. It is rather a tentative review, having as a prominent scope that of indicating some problems and methods followed in each of the two classes, with a look for possible future interactions between them. This is done in the main part of the paper (Section 2). In fact, as to problems of type G—of which I became aware only a short time ago—my presentation (Section 2.1) will be that of a beginner, and limited to a few historical remarks, accompanied by some indication for further study and some motivation from ODEs.

Something more the reader will find about problems of type K, for which I have focused on a short account of some basic results and methods from Bifurcation Theory on the one hand (Section 2.2), and to a brief description of a very special—maybe the “closest to linear”—nonlinear eigenvalue problem on the other (Section 2.3). The latter is the *p-Laplace equation*

$$\begin{cases} -\operatorname{div}(|\nabla u|^{p-2} \nabla u) = \mu |u|^{p-2} u & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \end{cases} \tag{5}$$

where  $p > 1$  and  $\Omega$  is a bounded domain in  $\mathbb{R}^n$ . Here, one can prove the existence—exactly as for the classical Laplace operator,  $p = 2$ —of countably many eigenvalues which can be naturally arranged in an increasing sequence

$$\mu_1 < \mu_2 \leq \dots \mu_k \leq \dots, \quad \mu_k \rightarrow +\infty \tag{6}$$

The importance of this example is also because it shows—via the Lusternik–Schnirelmann theory—the full strength of the variational methods and of Critical Point Theory in particular. As is well known, these consist in searching a solution of a given equation as a *critical point* of a functional (i.e., a point where the derivative of the functional vanishes), and are of the utmost importance both for equations of the form

$$A(x) = 0 \tag{7}$$

and for equations of the form

$$A(x) = \lambda C(x) \tag{8}$$

the latter being in fact the nonlinear eigenvalue problem in Equation (4) for the pair  $(A, C)$ . Indeed, if  $A = \nabla f$  and  $C = \nabla g$ , then solutions of Equation (7) are the *free* critical points of  $f$ , while solutions of Equation (8) are—modulo technicalities—the *constrained* critical points of  $f$  on the manifold  $M = \{x : g(x) = \text{const.}\}$ . As explained for instance in [4], the Lusternik–Schnirelmann theory not only guarantees, under appropriate assumptions on the nonlinear operators  $A$  and  $C$ , the existence of infinitely many distinct eigenvalues of Equation (8), and thus in particular of Equation (5), but also provides for them a “minimax” characterization of the form (when  $C = I$ )

$$\lambda_n = \sup_{K_n} \inf_K \langle A(u), u \rangle,$$

obtained via suitable families  $K_n$  of subsets  $K$  of the unit sphere. This realizes a conceptually beautiful (and also practically useful, see for instance [5]) extension of the Courant–Weyl principle for the eigenvalues of a linear compact self-adjoint operator. The variational characterization of the eigenvalues seems to be a main point of common interest for either type of nonlinear eigenvalue problems, see, e.g., [6] and the references quoted in Sections 2.1 and 2.3.

In the second part of the paper (Section 3), we return to Equation (1) and look at the case in which a small “perturbation parameter”  $\epsilon$  enters in the problem, originating an equation of the form

$$F(\epsilon, \lambda, x) = 0 \tag{9}$$

of which Equation (1) is seen as the unperturbed form for  $\epsilon = 0$ . We consider parameter-dependent forms of both Equations (3) and (4), precisely

$$G(\epsilon, \lambda)x = 0 \tag{10}$$

and, taking in Equation (4)  $C = I$  and adding to a linear  $A$  a nonlinear perturbation  $\epsilon B$ ,

$$Ax + \epsilon B(x) = \lambda x. \tag{11}$$

In both Equations (10) and (11), one common problem is—in the light of what is done for linear operators [7]—to see how the perturbed eigenvalues  $\lambda(\epsilon)$  (provided that they exist) will depend on  $\epsilon$  near a given unperturbed eigenvalue  $\lambda_0$ . To this purpose, we review the main points of the recent contributions [8,9], respectively, to Equation (10) and to Equation (11).

Two more points deserve to be mentioned before closing this Introduction. The first is that, for a better understanding, Nonlinear Eigenvalue Problems—both of type G and of type K—should be set in the more general respective context of Nonlinear Spectral Theory. References for this are [10,11], respectively. The interested reader might look at [12] for a recent contribution to the latter. The second fact, clear enough from this Introduction, is that we have not even attempted to mention the various numerical methods used for the practical solution of Equation (3) in the case  $E = \mathbb{K}^n$ . The reader interested in this rich and fundamental research field might look into the excellent and very recent survey paper [13].

Let me repeat in conclusion that the only reasonable scope of this paper is to possibly arouse the curiosity of some expert in either field towards the problems treated in the other, and to give a chance of possible inspiration for further study.

## 2. The Two Types of NLEVP

### 2.1. Problems of Type G: (Linear) Operator- and Matrix-Valued Functions

A good point to start a presentation of nonlinear eigenvalue problems of the type in Equation (3) is perhaps R.E.L. Turner’s paper [6] of 1968. Given a complex Hilbert space  $H$ , rather than considering the spectrum of a single bounded linear operator  $A$  acting in  $H$ , he considers for  $\lambda \in \mathbb{C}$  operators of the form

$$A - B(\lambda) \equiv A - \sum_1^N \lambda^k B_k \tag{12}$$

where  $B_k \in L(H)$  are given ( $k = 1, \dots, N$ ), with  $B_1 = I$ ; thus if  $N = 1$  we are back to the familiar  $A - \lambda I$  considered in linear spectral theory. The *spectrum* of  $A - B(\lambda)$  is defined as the set of those  $\lambda \in \mathbb{C}$  for which  $A - B(\lambda)$  fails to be a homeomorphism of  $H$  onto itself. In particular, a point  $\lambda_0 \in \mathbb{C}$  such that  $A - B(\lambda_0)$  is not injective, i.e., such that the nullspace  $\text{Ker}(A - B(\lambda_0)) \neq \{0\}$ , is an *eigenvalue* of  $A - B(\lambda)$ . Note that, in the case  $N = 1$ , these definitions of spectrum and eigenvalues of  $A - \lambda I$  yield what we usually call the spectrum and eigenvalues of  $A$ . The new point of view is that the spectrum is now attributed to the (polynomial) *function* of  $\mathbb{C}$  into  $L(H)$  defined putting

$$G(\lambda) = A - B(\lambda). \tag{13}$$

In the case  $H = \mathbb{C}$ , the spectrum so defined consists very simply of the zeroes of the polynomial  $G$  itself. Now recall (see e.g., [14] or [15]) that, if  $A$  is compact, self-adjoint and nonnegative, then:

- The spectrum of  $A$  is at most countable and consists of a finite or infinite decreasing sequence of non-negative eigenvalues  $(\lambda_n)$ :

$$\lambda_1 \geq \lambda_2 \geq \dots \lambda_k \geq \dots$$

If the sequence is infinite, then  $\lambda_n \rightarrow 0$ .

- The eigenvectors  $(u_n)$  associated with the eigenvalues  $(\lambda_n)$  form an orthonormal basis of  $H$ .

Turner first generalizes this to operators as in Equation (12) where  $A$  is compact, self-adjoint and nonnegative,  $B_k$  is self-adjoint and non-negative for  $k = 1, \dots, N$  and  $A$  belongs to the Schatten class  $C_r$  (i.e., its eigenvalues  $(\alpha_i)$  satisfy the condition  $\sum_i (\alpha_i)^r < \infty$ ) for some  $r < \frac{1}{2}$ . Another basic fact concerning the spectrum of an operator  $A$  as above is the *variational characterization* of its positive eigenvalues  $(\lambda_n)$ : indeed,

$$\lambda_n = \max_{x \perp u_1, \dots, u_{n-1}} \frac{\langle Ax, x \rangle}{\langle x, x \rangle} = \min_{x \in [u_1, \dots, u_n]} \frac{\langle Ax, x \rangle}{\langle x, x \rangle} \tag{14}$$

$$= \min_{V \in V_{n-1}} \max_{x \perp V} \frac{\langle Ax, x \rangle}{\langle x, x \rangle} = \max_{V \in V_n} \min_{x \in V} \frac{\langle Ax, x \rangle}{\langle x, x \rangle} \tag{15}$$

where  $V$  is a vector subspace of  $H$ , and  $V_n$  denotes the family of all vector subspaces of dimension  $n$ .

Turner generalizes the variational principle as follows. For  $x \in H, x \neq 0$ , let  $Z(x)$  be the unique non-negative zero of the polynomial  $\lambda \rightarrow \langle G(\lambda)x, x \rangle$  [6]. Note that, in the case  $N = 1$ , as  $\langle G(\lambda)x, x \rangle = \langle Ax, x \rangle - \lambda \langle x, x \rangle$ , we have

$$Z(x) = \frac{\langle Ax, x \rangle}{\langle x, x \rangle} \tag{16}$$

so that the function  $Z$  is the usual *Rayleigh quotient* of  $A$ , of which the eigenvalues are *extremal values* as shown by Equation (14). Then, under the stated assumptions on  $A$  and  $B_k$ , if moreover the eigenvectors of  $A - B(\lambda)$ , corresponding to non-negative eigenvalues, form a basis for  $H$ , then the variational principles in Equations (14) and (15) hold replacing  $\langle Ax, x \rangle / \langle x, x \rangle$  with  $Z(x)$ .

Finally, we have by definition of  $Z(x)$  that

$$\langle G(Z(x))x, x \rangle = 0 \quad \text{for all } x \in H, x \neq 0 \tag{17}$$

Results similar to those of Turner, and practically at the same time, were obtained by K.P. Haderler in [16,17]. He considered several-parameter dependent operators of the form

$$A - (\lambda_1 B_1 + \lambda_2 B_2 + \dots + \lambda_N B_N)$$

with  $B_j$  bounded self-adjoint for  $j = 1, \dots, N$ , and in connection with the variational property of their eigenvalues introduced the general concept of *Rayleigh functional* of a matrix function as follows. Let  $\alpha \rightarrow T(\alpha)$  be a differentiable mapping of the real interval  $(a, b)$  to the set  $S_n$  of real symmetric matrices of order  $n$ . Then, a Rayleigh functional of  $T$  is a continuous real-valued function  $p$  on  $\mathbb{R}^n \setminus \{0\}$  such that  $p(x) \in (a, b)$  for all  $x \in \mathbb{R}^n \setminus \{0\}$  and

- $p(cx) = p(x)$  if  $c \neq 0$
- $\langle T(p(x))x, x \rangle = 0$
- $\langle T'(p(x))x, x \rangle > 0$ .

The last is a definiteness condition that can be replaced by  $\langle T'(p(x))x, x \rangle < 0$ , and is plainly satisfied in the basic case  $T(\alpha) = A - \alpha I$ , where  $T'(\alpha) = -I$ . Thus, looking at Equations (16) and (17), we see that this is a sensible extension of the definition and properties of the Rayleigh quotient.

The results of Turner and Haderler indicated above were developed and improved by, among others, H. Langer. For instance, in [18], studying combinations  $T(\lambda)$  of bounded self-adjoint operators of the form of Equation (12) considered by Turner, he assumed that, for each nonzero vector  $x$ , the polynomial  $p_x(L) \equiv \langle T(\lambda)x, x \rangle$  has only real roots

$$\lambda_1(x) \geq \lambda_2(x) \geq \dots \geq \lambda_n(x)$$

Under this assumption he showed that the ranges  $\Lambda_i$  of the functions  $\lambda_i$  are intervals, called *spectral zones*, whose interiors do not overlap.

A systematization of the spectral theory (that is, of the properties of eigenvalues and eigenvectors) of *polynomial operator pencils*, as had been named the families

$$A(\lambda) = A_0 + \lambda A_1 + \dots + \lambda^n A_n \tag{18}$$

where  $\lambda \in \mathbb{C}$  is a spectral parameter, and  $A_i, i = 1, \dots, n$ , are linear operators in a Hilbert space, was given by A.S. Markus in his book [19]. Among others, he considered in depth the problem of the *factorization* of pencils, which in the simplest case consists in representing a quadratic pencil  $A(\lambda) = \lambda^2 I + \lambda B + C$  in the form

$$A(\lambda) = (\lambda I - Y)(\lambda I - Z).$$

The importance of many results in [19] is due to the fact that they hold for the more general case of *holomorphic* (i.e., analytic) operator-valued functions, namely operators  $A(\lambda)$  expressed as the sum of convergent power series in  $L(E)$ :

$$A(\lambda) = \sum_0^\infty \lambda^n A_n. \tag{19}$$

For an updated reference reviewing the spectral properties of self-adjoint analytic operator functions, and in particular the factorization problem, see [10]. On the other hand, for further work on the variational characterization of eigenvalues as well as for the development of the theory of Rayleigh

functionals, the interested reader can look for instance into the quite recent papers by Binding, Eschwé and Langer [20], Hasanov [21], Voss [22], and Schwetlick and Schreiber [23], and the references therein.

Let us now add some more specific indication for the case in which  $E = \mathbb{K}^m$ , so that the function  $G$  appearing in Equation (3) takes its values in the space  $\mathbb{K}^{m \times m}$  of  $m \times m$  real or complex matrices. We shall stress the finite-dimensionality of the ambient space  $E$  using the letter  $M$  rather than  $G$ , and often the letter  $v$  rather than  $x$  for the vectors of  $E$ . A well known reference book for the matter is the one by Gohberg, Lancaster and Rodman [24], and the very Introduction to this book explains to us that problems of the form

$$M(\lambda)v = 0, \quad \lambda \in \mathbb{C}, \quad v \in \mathbb{C}^m, \quad v \neq 0 \tag{20}$$

where  $M(\lambda) \in \mathbb{C}^{m \times m}$  appear naturally when dealing with linear systems of higher order ordinary differential equations (ODE) with constant coefficients:

$$\frac{d^n u}{dt^n} + A_{n-1} \frac{d^{n-1} u}{dt^{n-1}} + \dots + A_1 \frac{du}{dt} + A_0 u = 0 \tag{21}$$

where  $A_i \in \mathbb{C}^{m \times m}$  for  $i = 0, 1, \dots, n - 1$ . Indeed, looking for solutions of the form  $u(t) = e^{\lambda t} v$  ( $\lambda \in \mathbb{C}, v \in \mathbb{C}^m$ ) of Equation (21) leads to the equation

$$e^{\lambda t} \{ \lambda^n + \lambda^{n-1} A_{n-1} + \dots + \lambda A_1 + A_0 \} v = 0 \tag{22}$$

which—as long as  $v \neq 0$ , and putting  $A_n = I$ —is equivalent to Equation (20) with

$$M(\lambda) = \sum_{i=0}^n \lambda^i A_i. \tag{23}$$

Thus,  $e^{\lambda_0 t} v_0$  is a nontrivial solution of Equation (21) if and only if  $\lambda_0$  is an eigenvalue of Equation (20), i.e., it is a zero of the *characteristic equation*

$$\det M(\lambda) = 0 \tag{24}$$

and  $v_0 \in \text{Ker } M(\lambda_0)$ . More generally, the function

$$u(t) = e^{\lambda_0 t} \left\{ \frac{t^k}{k!} v_0 + \dots + \frac{t}{1!} v_{k-1} + v_k \right\} \tag{25}$$

is a solution of Equation (21) if and only if the vectors  $v_0, v_1, \dots, v_k$  satisfy the relations

$$\sum_{j=0}^l \frac{1}{j!} \frac{d^j M}{d\lambda^j}(\lambda_0) v_{l-j} = 0, \quad l = 0, 1, \dots, k. \tag{26}$$

Such a set of vectors  $v_0, v_1, \dots, v_k$  is called a *Jordan chain* of length  $k + 1$  for the matrix function  $M(\lambda)$ , corresponding to the eigenvalue  $\lambda_0$  and *starting with* the eigenvector  $v_0$ . The above definitions extend from matrix polynomials as in Equation (23) to any analytic matrix function  $M(\lambda)$ . It is good to see the explicit form of Equation (26), which is

$$\begin{cases} l = 0 : & M(\lambda_0)v_0 = 0 \\ l = 1 : & M(\lambda_0)v_1 + \frac{1}{1!} \frac{dM}{d\lambda}(\lambda_0)v_0 = 0 \\ l = 2 : & M(\lambda_0)v_2 + \frac{1}{1!} \frac{dM}{d\lambda}(\lambda_0)v_1 + \frac{1}{2!} \frac{d^2M}{d\lambda^2}(\lambda_0)v_0 = 0 \\ & \dots \\ l = k : & M(\lambda_0)v_k + \frac{1}{1!} \frac{dM}{d\lambda}(\lambda_0)v_{k-1} + \dots + \frac{1}{k!} \frac{d^k M}{d\lambda^k}(\lambda_0)v_0 = 0. \end{cases} \tag{27}$$

If  $n = 1$  in Equation (23), we have  $M(\lambda) = \lambda A_1 + A_0$ ; and if moreover  $A_1 = -I$ , then  $M(\lambda) = A_0 - \lambda I$ . In this case,  $\frac{dM}{d\lambda}(\lambda_0) = -I$ , while  $\frac{d^j M}{d\lambda^j}(\lambda_0) = 0$  for all  $j > 1$ , so that the above equalities reduce to (putting  $A_0 = A$ )

$$\left\{ \begin{array}{l} (A - \lambda_0 I)v_0 = 0 \\ (A - \lambda_0 I)v_1 = v_0 \\ \dots \\ (A - \lambda_0 I)v_k = v_{k-1} \end{array} \right. \tag{28}$$

and are those defining an *ordinary* Jordan chain for the matrix  $A$  corresponding to  $\lambda_0$  and  $v_0$ , used to represent  $A$  in its Jordan canonical form and in particular to construct a basis of the *generalized eigenspace*  $E_{\lambda_0}(A)$  associated with  $\lambda_0$ . We recall that this is defined as

$$E_{\lambda_0}(A) = \text{Ker}((A - \lambda_0 I)^p) \tag{29}$$

where  $p$  is the least integer such that  $\text{Ker}((A - \lambda_0 I)^p) = \text{Ker}((A - \lambda_0 I)^{p+1})$ , and that the dimension  $\dim E_{\lambda_0}(A)$  of  $E_{\lambda_0}(A)$  is equal to the *algebraic multiplicity* of  $\lambda_0$ , that is, the multiplicity of the eigenvalue as a zero of the characteristic polynomial  $\det(A - \lambda I)$ . We say that  $\lambda_0$  is *semisimple* if  $p = 1$  in Equation (29)—that is, if the algebraic multiplicity coincides with the *geometric* multiplicity of  $\lambda_0$ , defined as  $\dim \text{Ker}(A - \lambda_0 I)$ —and that  $\lambda_0$  is *simple* if they are both equal to 1.

These familiar concepts from Linear Algebra, concerning the basic case  $M(\lambda) = A_0 - \lambda I$ , need to be extended to analytic matrix functions  $M(\lambda)$ . To this purpose, we quote from [25]; see also ([26], Chapter 7).

- Let  $x_0$  be an eigenvector corresponding to an eigenvalue  $\lambda_0$ . The maximal length of a Jordan chain starting at  $x_0$  is called the *multiplicity of  $x_0$*  and denoted by  $m(x_0)$ . An eigenvalue  $\lambda_0$  is said to be *normal* if it is an isolated eigenvalue and the multiplicity of each corresponding eigenvector is finite.
- Suppose that  $\lambda_0$  is a normal eigenvalue. Then, a corresponding *canonical system of Jordan chains*

$$x_0^k, x_1^k, \dots, x_{m_k-1}^k \quad (k = 1, \dots, N)$$

is defined by the following rules:

- (1) The vectors  $x_0^1, \dots, x_0^N$  form a basis of  $\text{Ker } M(\lambda_0)$  (and so  $N = \dim \text{Ker } M(\lambda_0)$ ).
  - (2)  $x_0^1, x_1^1, \dots, x_{m_1-1}^1$  is a Jordan chain of the maximal length  $m_1 \equiv m(x_0^1)$ .
  - (3) Once that the vectors  $x_0^1, x_0^2, \dots, x_0^{k-1}$  ( $1 \leq k < N$ ) have been chosen, then pick an eigenvector  $x_0^k$  linearly independent from  $x_0^1, x_0^2, \dots, x_0^{k-1}$  and form a Jordan chain  $x_0^k, x_1^k, \dots, x_{m_k-1}^k$  of the maximal length  $m_k \equiv m(x_0^k)$ .
- A canonical system is not defined uniquely; however, the numbers  $m_1, m_2, \dots, m_N$  do not depend on the choice of Jordan chains and are called *partial multiplicities* of the eigenvalue  $\lambda_0$ . The number  $Q(\lambda_0) \equiv m_1 + \dots + m_N$  is the *algebraic multiplicity of the eigenvalue  $\lambda_0$* .

The next statement—which is based on results found in [27]—proves that these definitions are a coherent generalization of the usual ones.

**Proposition 1.** *An eigenvalue  $\lambda_0$  is a zero of  $\det M(\lambda)$  of multiplicity  $Q(\lambda_0)$ .*

Based on Proposition 1, the definitions of simple and semisimple eigenvalue carry over to the case of matrix polynomials and more generally to analytic matrix functions. For instance, one may check that the matrix function

$$M_2(\lambda) = \begin{pmatrix} \lambda - 1 + e^{-\lambda} & 0 \\ 0 & \lambda + 1 \end{pmatrix} \tag{30}$$

considered in [8] has  $\lambda_0 = 0$  as a double (i.e., of algebraic multiplicity 2), nonsemisimple (i.e., of geometric multiplicity 1) eigenvalue, with Jordan chain

$$H_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad H_1 = \begin{pmatrix} \alpha \\ 0 \end{pmatrix} \tag{31}$$

for any  $\alpha \in \mathbb{R}$ . This example also shows that in the nonlinear case, generalized eigenvectors do not need to be linearly independent. Indeed, in the construction (and notation) recalled above, the generating vectors  $x_0^1, \dots, x_0^N$  of the system of Jordan chain are chosen to be linearly independent, but it is not necessarily so for the vectors in each corresponding chain, generated by the rules given by the system in Equation (27).

An especially important source for the study of NLEVP are the Delay Differential Equations (DDE), or systems of them. For instance, in [26] is considered the so-called Wright equation

$$x'(t) = -\alpha x(t-1)[1+x(t)] \tag{32}$$

where  $\alpha > 0$ . The objective is to determine the periodic orbits (if any) of Equation (32). To do this, one must first look at the linearized equation of Equation (32) near  $x \equiv 0$ , which is

$$x'(t) = -\alpha x(t-1) \tag{33}$$

Solutions  $e^{\lambda t}$  of this exist iff  $\lambda$  satisfies the characteristic equation

$$\lambda e^\lambda + \alpha = 0 \tag{34}$$

For  $\alpha = \pi/2$ , this has  $\lambda = i\frac{\pi}{2}$  as a simple purely imaginary root, corresponding to the periodic solution  $e^{i\frac{\pi}{2}t}$ . Studying the properties of these nonlinear eigenvalues, that is of the roots  $\lambda(\alpha)$  of Equation (34) as a function of  $\alpha$ , and using deep topological and functional-analytic results from [26], it is possible to demonstrate that Equation (32) has a Hopf bifurcation at  $\alpha = \pi/2$ , and that for every  $\alpha > \pi/2$  Equation (32) has a nonconstant periodic solution with period close to 4. Finally, the authors show that for  $p > 4$ , there is a periodic solution of Equation (32) of period  $p$ .

One can also consider systems of DDE, for instance

$$x'(t) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} x(t) + \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} x(t-1)$$

whose characteristic matrix is precisely that displayed in Equation (30). The general form of a system of  $N$  delay differential equations, with delays  $\tau_1, \dots, \tau_N$  is

$$x'(t) = A_0 x(t) + \sum_{i=1}^N A_i x(t - \tau_i) \tag{35}$$

with  $A_0, A_i \in \mathbb{C}^{N \times N}$ , and the corresponding characteristic matrix is

$$M(\lambda) = \lambda I - A_0 - \sum_{i=1}^N A_i e^{-\lambda \tau_i}.$$

More general forms of Equation (35) are considered in Section 3.

### 2.2. Problems of Type K: Nonlinear Operators and Bifurcation

Throughout this Section  $E$  will be a real Banach space, of finite or infinite dimension. Originally, bifurcation theory deals with the local study of Equation (1) near a point  $(\lambda_0, 0) \in \mathbb{R} \times E$ , and studies precisely the conditions under which from the given point  $(\lambda_0, 0)$  of the line  $\mathbb{R} \times \{0\} \subset \mathbb{R} \times E$

of the *trivial solutions* of Equation (1), there *bifurcates a branch* of nontrivial solutions, that is, of solutions  $(\lambda, x)$  with  $x \neq 0$ . Of course, the basic situation that comes to one’s mind is the case  $F(\lambda, x) = Ax - \lambda x$ , with  $\lambda_0$  an eigenvalue of the linear operator  $A$ , the “branch” being here the special subset  $\{\lambda_0\} \times (\text{Ker}(A - \lambda_0 I) \setminus \{0\})$  of  $\mathbb{R} \times E$ . The interesting case is when  $F$  depends in a less obvious way from  $\lambda$  and  $x$ ; an easy example of what we mean is given for instance by the equation

$$F(\lambda, x) \equiv ax + bx^3 - \lambda x = 0, \quad (\lambda, x) \in \mathbb{R}^2$$

in which the parabola  $\lambda = a + bx^2$  bifurcates at the point  $(a, 0)$  from the line of the trivial solutions. For a motivating introduction to the theory, and a discussion of some important physical problems that fall in this context, an excellent source is the old review paper by Stackgold [28].

The previous “naïf” idea of bifurcation needs to be made both more precise and more general, and this is done by saying that  $(\lambda_0, 0)$  is a *bifurcation point* for Equation (1) if any neighborhood of  $(\lambda_0, 0)$  in  $\mathbb{R} \times E$  contains nontrivial solutions of Equation (1). For this definition to make sense, it is enough that  $F$  be defined in an open set  $U \subset \mathbb{R} \times E$  with  $(\lambda_0, 0) \in U$ , and this is what we assume from now on. For the next step, we further assume that  $F$  is differentiable at the point  $(\lambda_0, 0)$ , so that  $F$  can be linearized near that point as

$$F(\lambda, x) = F(\lambda_0, 0) + D_\lambda F(\lambda_0, 0)(\lambda - \lambda_0) + D_x F(\lambda_0, 0)x + R(\lambda, x) = D_x F(\lambda_0, 0)x + R(\lambda, x) \quad (36)$$

where the remainder term  $R$  satisfies

$$R(\lambda, x) = o(\|(\lambda, x)\|) \quad \text{as } (\lambda, x) \rightarrow (\lambda_0, 0).$$

Some more regularity on  $F$  yields immediately a necessary condition for bifurcation:

**Theorem 1.** *Suppose that  $F$  is of class  $C^1$  in a neighborhood of  $(\lambda_0, 0)$ . If  $D_x F(\lambda_0, 0)$  is a homeomorphism of  $E$  onto itself, then  $(\lambda_0, 0)$  cannot be a bifurcation point for Equation (1).*

**Proof.** The assumption implies, via the Implicit Function Theorem, that there is a neighborhood  $I \times V$  of  $(\lambda_0, 0)$  such that, for any  $\lambda \in I$ , there is a *unique*  $x = x(\lambda) \in V$  such that  $F(\lambda, x) = 0$ . As by assumption  $F(\lambda, 0) = 0$  for any  $\lambda$ , we must have  $x(\lambda) = 0$  for  $\lambda \in I$ , so that there is no nontrivial solution to Equation (1) in the neighborhood  $I \times V$  of  $(\lambda_0, 0)$ .  $\square$

For simplicity, we shall henceforth consider only the special case

$$F(\lambda, x) = A(x) - \lambda x \quad (37)$$

where  $A(0) = 0$  and  $A$  is of class  $C^1$  near  $x = 0$ . Here,  $D_x F(\lambda_0, 0) = A'(0) - \lambda_0 I$ , and we have a more explicit form of the remainder term in the linearized form in Equation (36) of  $F$ : for we can write  $A(x) = A'(0)x + B(x)$  with  $B(x) = o(\|x\|)$  as  $\|x\| \rightarrow 0$ , so that Equation (37) yields

$$\begin{aligned} F(\lambda, x) &= A'(0)x + B(x) - \lambda x \\ &= (A'(0) - \lambda_0 I)x + B(x) - (\lambda - \lambda_0)x \end{aligned}$$

and comparing this with Equation (36) we see that  $R(\lambda, x) = B(x) - (\lambda - \lambda_0)x$  in this special case. Resuming, the equation we want to study is

$$A(x) - \lambda x = 0 \quad (38)$$

with  $A(0) = 0$  and  $A$  of class  $C^1$  near  $x = 0$ , and can be written as

$$Tx - \lambda_0 x + B(x) = (\lambda - \lambda_0)x \quad (39)$$

where  $T \equiv A'(0)$  and  $B(x) = o(\|x\|)$  as  $\|x\| \rightarrow 0$ . The necessary condition for bifurcation implicitly stated in Theorem 1 can now be rephrased as follows:

$$(\lambda_0, 0) \text{ bifurcation point of } A(x) - \lambda x = 0 \Rightarrow \lambda_0 \in \sigma(A'(0)).$$

The standard case considered in the literature is when  $\lambda_0$  is in the *point* spectrum of  $A'(0)$ , and we formalize this more precisely under the form of a basic assumption, which is plainly satisfied if  $\dim E < \infty$ :

**H0.**  $\lambda_0$  is an isolated eigenvalue of  $T \equiv A'(0)$  and  $T - \lambda_0 I$  is a Fredholm operator of index zero.

Let us recall (see, e.g., [14]) that a bounded linear operator  $L$  between two Banach spaces  $E$  and  $F$  is said to be a **Fredholm operator** if its nullspace  $\text{Ker } L$  has finite dimension and its range  $\text{Im } L$  is closed and has finite codimension; in this case, the *index* of  $L$ ,  $\text{ind } L$ , is defined as

$$\text{ind } L = \dim \text{Ker } L - \text{codim } \text{Im } L.$$

Thus, if  $\dim E = \dim F < \infty$ , then any linear operator is Fredholm of index zero. From the Riesz–Schauder theory of such operators (see e.g., [15]), it is known that also the nullspaces  $\text{Ker } L^j$  ( $j > 1$ ) are finitedimensional, and that they stabilize for  $j$  sufficiently large; with reference to the case  $L = T - \lambda_0 I$ , this means that there exists a least integer  $p$  such that  $\text{Ker}(T - \lambda_0 I)^p = \text{Ker}(T - \lambda_0 I)^{p+1}$ , and moreover one has

$$E = \text{Ker}(T - \lambda_0 I)^p \oplus \text{Im}(T - \lambda_0 I)^p.$$

It follows in particular that the *algebraic multiplicity* of  $\lambda_0$  is finite, where in general this is defined—consistently with the definition recalled in Section 2.1 for the case  $\dim E < \infty$ —as the dimension of the subspace

$$\bigcup_{j=1}^{\infty} \text{Ker}(T - \lambda_0 I)^j.$$

In the following, when speaking of *multiplicity* of an eigenvalue, we refer to the algebraic multiplicity. We recall that this coincides with the *geometric multiplicity*  $\dim \text{Ker}(T - \lambda_0 I)$  when  $T$  is a self-adjoint operator in a Hilbert space.

**Remark 1.** The assumption **H0** alone is not sufficient to guarantee that an eigenvalue of the “linear part”  $T$  of  $A$  at 0 is a bifurcation point for  $A$ . To see this, consider the example ([29], Chapter 11) given by the system

$$\begin{cases} x + y^3 = \lambda x \\ y - x^3 = \lambda y. \end{cases} \tag{40}$$

Here,  $E = \mathbb{R}^2$  and we have (in our notations)  $T = I$ ,  $\lambda_0 = 1$  and  $B(x, y) = (y^3, -x^3)$ . Multiplying the first equation by  $y$ , the second by  $x$  and subtracting the second from the first, we obtain  $x^4 + y^4 = 0$ , showing that Equation (40) has no nontrivial solution whatsoever. One way of seeing this is that the two-dimensional eigenspace associated with  $\lambda_0$  is completely destroyed by the addition of the perturbing term  $B$ .

Three typical situations are then considered, each of them guaranteeing bifurcation from  $\lambda_0$ , and described by the following assumptions, respectively:

**H1.**  $\lambda_0$  is a *simple* eigenvalue of  $A'(0)$ .

**H2.**  $A$  is *compact* and  $\lambda_0 \neq 0$  is an eigenvalue of *odd* multiplicity of  $A'(0)$ .

**H3.**  $A$  is a *gradient* operator in a Hilbert space and  $\lambda_0$  is an isolated eigenvalue of finite multiplicity of  $A'(0)$ .

These assumptions call immediately for some explanation. In fact, it could be noted at once that both **H2** and **H3** are a strengthening of **H0**. However, to proceed with some order, in the remaining part of this subsection, we shall give a precise statement for each of the three bifurcation results roughly indicated above, preceded by a comment on the respective assumption, and followed by an indication of the proof.

Thus, starting with **H2**, we recall that, if  $A$  is compact, then the linear operator  $A'(0)$  is a compact, too [30]. Therefore **H0** is redundant in this case, as it is a basic spectral property of any such operator [14].

**Theorem 2.** *If **H2** is satisfied, then  $\lambda_0$  is a bifurcation point for Equation (38). Moreover, it is a global bifurcation point in the following sense: if  $S$  denotes the closure in  $\mathbb{R} \times E$  of the set of nontrivial solutions of Equation (38), then  $S \cup \{(\lambda_0, 0)\}$  has a connected subset  $S_{\lambda_0}$  containing  $\{(\lambda_0, 0)\}$ , and which is either unbounded in  $\mathbb{R} \times E$  or contains a point  $\{(\lambda_1, 0)\}$  with  $\lambda_1$  an eigenvalue of odd multiplicity of  $T$ .*

**Proof.** The proof relies on the **Leray–Schauder degree**. Roughly speaking, this is a topological tool to detect the fixed points of a compact map and can be briefly introduced as follows (see, for instance, [3] (Part I) for a complete presentation). Suppose we have a continuous compact map  $C$  of  $E$  into itself, a bounded open set  $\Omega \subset E$  with  $0 \in \Omega$ , and suppose that  $C(x) \neq x$  for  $x \in \partial\Omega$ . Then, there exists an integer, denoted  $d(I - C, \Omega, 0)$  and called the (Leray–Schauder) *degree of  $I - C$  relative to the set  $\Omega$  and to the point 0*, having the following properties:

- (i) If  $d(I - C, \Omega, 0) \neq 0$ , then there exists an  $x \in \Omega$  such that  $x = C(x)$ .
- (ii)  $d(I, \Omega, 0) = 1$ .
- (iii) If  $\Omega_0 \subset \Omega$  and  $I - C$  has no zeroes in  $\Omega \setminus \overline{\Omega}_0$ , then  $d(I - C, \Omega, 0) = d(I - C, \Omega_0, 0)$ .
- (iv) Suppose  $C_1, C_2 : E \rightarrow E$  are compact maps, and put

$$H(t, x) = x - [C_1(x) - t(C_2(x) - C_1(x))], \quad t \in [0, 1], x \in E$$

If  $H(t, x) \neq 0$  for  $t \in [0, 1]$  and  $x \in \partial\Omega$ , then

$$d(I - C_1, \Omega, 0) = d(I - C_2, \Omega, 0).$$

- (v) If  $C$  is a *linear* compact map and  $I - C$  is injective, then

$$d(I - C, \Omega, 0) = (-1)^v$$

where  $v$  is the number of eigenvalues  $>1$  of  $C$ , each counted with its algebraic multiplicity.

To prove that  $\lambda_0$  is a bifurcation point, it is enough to show that for any sufficiently small  $r > 0$ , there exists a solution  $(\lambda_r, x_r)$  of Equation (38) with  $\lambda_r \in [\lambda_0 - r, \lambda_0 + r]$  and  $\|x_r\| = r$ . Thus, let  $B_r = \{x \in E : \|x_r\| < r\}$  be the open ball centered at  $x = 0$  and with radius  $r$ ; we consider the degree of various maps with respect to this neighborhood of 0. Precisely, assume for instance  $\lambda_0 > 0$  and write  $A(x) - \lambda x = -\lambda(x - \mu A(x))$ ,  $\mu = 1/\lambda$ , for  $\lambda$  near  $\lambda_0$ . Consider thus the equivalent equation

$$x - \mu A(x) = 0 \tag{41}$$

and let  $\mu$  vary in an interval  $[\underline{\mu}, \bar{\mu}]$  containing as interior point  $\mu_0 = 1/\lambda_0$  and no other *characteristic values* (as are named the reciprocals of the nonzero eigenvalues) of  $T \equiv A'(0)$  except  $\mu_0$ . Assume by way of contradiction that  $x - \mu A(x) \neq 0$  for  $\|x\| = r$  and  $\mu \in [\underline{\mu}, \bar{\mu}]$ ; then using the *Homotopy invariance Property* (iv) with  $C_1 = \underline{\mu}A, C_2 = \bar{\mu}A$  we would have

$$d(I - \underline{\mu}A, B_r, 0) = d(I - \bar{\mu}A, B_r, 0). \tag{42}$$

On the other hand, for small  $r > 0$ , using again Property (iv) we have

$$d(I - \underline{\mu}A, B_r, 0) = d(I - \underline{\mu}T, B_r, 0) \tag{43}$$

because  $I - \underline{\mu}A$  is homotopic to  $I - \underline{\mu}T$  on  $\partial B_r$ ; indeed, since the latter operator is a homeomorphism and since  $B(x) = o(\|x\|)$  as  $\|x\| \rightarrow 0$ , we have (diminishing  $r$  if necessary)

$$\|x - \underline{\mu}(Tx + tB(x))\| \geq \|x - \underline{\mu}Tx\| - \|\underline{\mu}tB(x)\| \geq k\|x\|$$

for some  $k > 0$  and for all  $(t, x) \in [0, 1] \times \bar{B}_r$ . Similarly,

$$d(I - \bar{\mu}A, B_r, 0) = d(I - \bar{\mu}T, B_r, 0). \tag{44}$$

However, using Property (v), we have

$$d(I - \underline{\mu}T, B_r, 0) = (-1)^\nu, \quad d(I - \bar{\mu}T, B_r, 0) = (-1)^{\bar{\nu}} \tag{45}$$

where  $\bar{\nu} = \nu + h$ ,  $h$  an odd integer (the algebraic multiplicity of  $\lambda_0$ ); therefore the two degrees in Equation (45) are different, contradicting the previous equalities in Equations (42)–(44). This proves that  $x_r - \mu_r A(x_r) = 0$  for some  $x_r \in \partial B_r$  and some  $\mu_r \in [\underline{\mu}, \bar{\mu}]$ , and therefore that there is bifurcation from  $(\mu_0, 0)$  for the Equation (41), or equivalently from  $(\lambda_0, 0)$  for Equation (38). The proof that under the stated assumptions the bifurcation has a *global* character, in the sense described by the statement of Theorem 2, requires the much deeper topological analysis performed by P.H. Rabinowitz in his famous paper [31].  $\square$

We now go on to comment assumption **H3**, and to briefly discuss the corresponding bifurcation result. For the next definition, and the statements following it, see for instance [2] or [30].

**Definition 1.** Let  $H$  be a real Hilbert space with scalar product denoted  $\langle \cdot, \cdot \rangle$ . An operator  $A : H \rightarrow H$  is said to be a gradient (or potential) operator if there exists a differentiable functional  $a : H \rightarrow \mathbb{R}$  such that

$$\langle A(x), y \rangle = a'(x)y \quad \text{for all } x, y \in H. \tag{46}$$

One then writes  $A = \nabla a$ ; the functional  $a$ —the *potential* of  $A$ —is uniquely determined by the requirement that  $a(0) = 0$ , and is explicitly given by the formula

$$a(x) = \int_0^1 \langle A(tx), x \rangle dt. \tag{47}$$

A bounded linear operator is a gradient if and only if it is self-adjoint. Moreover, if a gradient operator  $A$  is differentiable at a point  $x_0$ , then  $A'(x_0)$  is self-adjoint.

**Theorem 3.** If **H3** is satisfied, then  $\lambda_0$  is a bifurcation point of Equation (38). Moreover, for each  $r > 0$  sufficiently small, Equation (38) has at least two distinct solutions  $(\lambda_r, x_r)$  such that  $\|x_r\| = r$ .

**Proof.** The proof makes use of the **Lyapunov–Schmidt method** (see, for instance, ([3], Chapter 2) or ([29]), Chapter 11) which allows to reduce the infinite-dimensional problem in Equation (38) to a problem in the finite-dimensional space  $\text{Ker}(T - \lambda_0 I)$ . Indeed, consider the equivalent form Equation (39) of Equation (38), and rewrite it as

$$Lx + B(x) = \delta x \tag{48}$$

where  $L = T - \lambda_0 I$  and  $\delta = \lambda - \lambda_0$ . Now recalling that  $T = A'(0)$ , the assumption **H3** implies that  $H$  is the orthogonal sum

$$H = \text{Ker } L \oplus \text{Im } L. \tag{49}$$

Then, letting  $P, Q$  denote the orthogonal projections of  $H$  onto  $\text{Ker } L$  and  $\text{Im } L$ , respectively, we have

$$x = Px + Qx \equiv v + w \tag{50}$$

and using this in Equation (48), we obtain the equivalent system

$$PB(v + w) = \delta v \tag{51}$$

$$Lw + QB(v + w) = \delta w. \tag{52}$$

The restriction  $L|_{\text{Im } L}$  of  $L$  to  $\text{Im } L$  is a homeomorphism of  $\text{Im } L$  onto itself. A standard application of the Implicit Function Theorem, together with the condition  $B(x) = o(\|x\|)$  as  $\|x\| \rightarrow 0$ , then allows to solve the complementary equation, Equation (52) in the form

$$w = w(\delta, v) \tag{53}$$

with  $w(0, 0) = 0$ , where  $\delta$  and  $v$  belong to suitably small neighborhoods  $J$  and  $V$  of  $\delta = 0$  and  $v = 0$ , respectively in  $\mathbb{R}$  and in  $\text{Ker } L$ . Replacing this in Equation (51) first yields

$$\langle PB(v + w(\delta, v)), v \rangle = \delta \|v\|^2 \tag{54}$$

whence, applying once more the Implicit Function Theorem, one can recover  $\delta$  as a function of  $v$ ,

$$\delta = \delta(v), \quad \delta(0) = 0, \tag{55}$$

for  $v$  in a neighborhood  $V_0 \subset V$  of 0 in  $\text{Ker } L$ . Finally, putting

$$\phi(v) = w(\delta(v), v), \quad v \in V_0 \tag{56}$$

and replacing this in Equation (51), one is left with the finite-dimensional equation (the bifurcation equation)

$$F_0(v) \equiv PB(v + \phi(v)) = \delta(v)v. \tag{57}$$

Any solution  $v \in V_0, v \neq 0$ , of this equation will give rise to a solution  $(\delta, x), x \neq 0$ ,

$$(\delta, x) = (\delta(v), v + \phi(v))$$

of the original Equation (48), and the continuity (in fact,  $C^1$  regularity) of the maps  $\delta = \delta(v), w = w(\delta, v)$  will ensure that this solution  $(\delta, x)$  stays into a given small neighborhood of  $(0, 0)$  in  $\mathbb{R} \times H$  provided that  $v$  is small enough. Thus, proving bifurcation from  $\lambda_0$  for Equation (38)—or equivalently, bifurcation from  $\delta = 0$  for Equation (48)—reduces to prove that Equation (57) has solutions  $v \neq 0$  of arbitrarily small norm.

**Remark 2.** The Lyapunov–Schmidt reduction can be applied more generally, and with minor modifications, in a Banach space  $E$  whenever the basic assumption **H0** (i.e., that  $L = T - \lambda_0 I$  is Freholm of index zero) holds and is supplemented by the transversality condition

$$\text{Ker } L \cap \text{Im } L = \{0\} \tag{58}$$

which is plainly satisfied when  $T$  is self-adjoint, as the two subspaces in Equation (58) are then orthogonal. Note that Equation (58) is in general equivalent to  $\text{Ker } L = \text{Ker } L^2$ , and thus to the fact that the algebraic and geometric multiplicities of  $\lambda_0$  coincide. **H0** and Equation (58) imply the direct decomposition of  $E$  into (closed) subspaces as in Equation (49), and therefore allow for the same reduction on taking for  $P, Q$  the (continuous) projections associated with Equation (49).

Returning to the proof of Theorem 3, we let now come in the assumption that the whole  $A$ , and therefore also its “nonlinear part”  $B$ , is a gradient. Here, we bound ourselves to give the main idea of the particularly clear demonstration provided by C. Stuart [32]. Thus, let  $f$  be such that  $\nabla f = L + B$ , and consider the reduced functional  $f_0 : V_0 \subset \text{Ker } L \rightarrow \mathbb{R}$  defined putting

$$f_0(v) \equiv f(v + \phi(v)). \tag{59}$$

Moreover, for small  $r > 0$  put

$$M_r = \{v \in V_0 : g(v) \equiv \|v\|^2 + \|\phi(v)\|^2 = r^2\}. \tag{60}$$

Then,  $M_r$  is the level set of the  $C^1$  functional  $g$ , and is compact because it is a closed and bounded subset of the finite dimensional space  $\text{Ker } L$ . Thus,  $f_0$  attains its minimum and its maximum on  $M_r$ , and if  $v_0 \in M_r$  is such an extremal point we have, by the Lagrange multiplier’s rule,

$$\nabla f_0(v_0) = \lambda \nabla g(v_0). \tag{61}$$

Performing the computations of  $\nabla f_0(v_0)$  and  $\nabla g(v_0)$  by the definitions in Equations (59) and (60), and using the fact that  $w(\delta, v)$  satisfies the complementary equation, Equation (52), one checks that  $\lambda = \delta(v_0)$  and that Equation (57) is satisfied.  $\square$

We finally come to **H1**. Unlike **H2** and **H3**, in general **H1** is independent from **H0**, and must be supplemented with it to guarantee bifurcation. Of course, when  $E$  is finite dimensional, **H0** does not play any role, and indeed **H1** can in this case be viewed as a special case of **H2**, because any continuous map is then compact.

**Theorem 4.** *If **H0** and **H1** are satisfied, then  $\lambda_0$  is a bifurcation point of Equation (38). Moreover, if  $A$  is of class  $C^2$  in a neighborhood of  $x = 0$ , then near  $(\lambda_0, 0)$  the solution set of Equation (38) consists of the trivial solutions  $\{(\lambda, 0)\}$  and of a  $C^1$  curve*

$$\gamma(t) = (\lambda(t), x(t)), \quad t \in ]-\delta, \delta[$$

with  $\gamma(0) = (\lambda_0, 0)$  and  $x(t) \neq 0$  for  $t \neq 0$ . Finally, if  $\text{Ker } L = [\phi]$ , then as  $t \rightarrow 0$

$$\begin{cases} x(t) &= t\phi + o(|t|) \\ \lambda(t) &= \lambda_0 + o(1). \end{cases} \tag{62}$$

The statement of Theorem 4 means that near  $(\lambda_0, 0)$ , the solution set of Equation (38) is topologically equivalent to the “cross”

$$(-1, 1[ \times \{0\}) \cup (\{0\} \times ]-1, 1[).$$

As to the proof, this goes for a first part along the same lines used to prove the previous Theorem 3, that is, using the Lyapounov–Schmidt decomposition in the sense indicated in Remark 2. What is specific here is that, since  $\dim \text{Ker } L = 1$ , one ends with an equation in  $\mathbb{R}$ ; a further nontrivial application of the Implicit Function Theorem then leads to the result: see, for instance, ([3], Chapter 2).

### 2.3. A Very Special Nonlinear Problem: The $p$ -Laplace Equation

Let  $\Omega$  be a bounded open set in  $\mathbb{R}^n$ , let  $p > 1$ , and let  $E$  be the Sobolev space  $W_0^{1,p}(\Omega)$ , equipped with the norm

$$\|v\|_{W_0^{1,p}} = \left( \int_{\Omega} |\nabla v|^p dx \right)^{\frac{1}{p}}. \tag{63}$$

That this is actually a norm in  $W_0^{1,p}(\Omega)$ , equivalent to the standard one of  $W^{1,p}(\Omega)$ , is a consequence of *Poincaré’s inequality* (see e.g., [14]), stating that

$$\int_{\Omega} |v|^p dx \leq C \int_{\Omega} |\nabla v|^p dx \tag{64}$$

for some  $C > 0$  and for all  $v \in W_0^{1,p}(\Omega)$ . Let  $E' = W^{-1,p'}(\Omega)$  be the dual space of  $E$ . A (weak) solution of the  $p$ -Laplace Equation (5) is a function  $u \in E$  such that

$$A_p(u) = \lambda B_p(u) \tag{65}$$

where  $\lambda = \mu^{-1}$  (it will soon be clear that  $\mu = 0$  is not an eigenvalue of Equation (5)) and  $A_p, B_p : E \rightarrow E'$  are defined by duality via the equations

$$\langle A_p(u), v \rangle = \int_{\Omega} |u|^{p-2} uv dx, \quad \langle B_p(u), v \rangle = \int_{\Omega} |\nabla u|^{p-2} \nabla u \nabla v dx \tag{66}$$

where  $u, v \in E$  and  $\langle \cdot, \cdot \rangle$  denotes the duality pairing between  $E$  and  $E'$ .

The proof of the existence of countably many eigenvalues and eigenfunctions of Equation (65) relies on the Lusternik–Schnirelmann (LS) theory of critical points for an even functional on a symmetric manifold. Complete presentations of this theory, in both finite and infinite dimensional spaces, can be found, among others, in [3,4,29,33,34]. Theorem 5 below is essentially a simplified version of Theorem A in [35], save that with respect to [35] we have for expository convenience interchanged the roles of the operators  $A$  and  $B$ . Thus, let  $E$  be a real, infinite dimensional, uniformly convex Banach space with dual  $E'$ , and consider the problem

$$A(u) = \lambda B(u) \tag{67}$$

where  $A, B : E \rightarrow E'$  are continuous gradient operators with potentials  $a, b$ , respectively:  $A = \nabla a, B = \nabla b$ . Definition 1 of gradient operator extends of course to mappings of  $E$  into  $E'$  replacing the scalar product with the duality pairing.

Suppose that  $b(u) > 0$  for  $u \neq 0$ ; then, the eigenvectors of Equation (67) satisfying a normalization condition  $b(u) = r$  ( $r > 0$ ), are precisely the constrained critical points of  $a$  on the level set

$$M_r = \{u \in E : b(u) = r\}.$$

The additional key assumptions that we make on  $A$  and  $B$  are as follows:

- $A, B$  are *odd* (that is,  $A(-u) = -A(u)$  for  $u \in E$ , and similarly for  $B$ ).
- $A$  is *non-negative* (that is,  $\langle A(u), u \rangle \geq 0$  for  $u \in E$ ) and *strongly sequentially continuous* (that is, if  $(u_n) \subset E$  converges weakly to  $u_0 \in E$ , then  $A(u_n)$  converges strongly to  $A(u_0)$  in  $E'$ ).
- $B$  is *strongly monotone* in the following sense: there exist constants  $k > 0$  and  $p > 1$  such that, for all  $u, v \in E$ ,

$$\langle B(u) - B(v), u - v \rangle \geq k \|u - v\|^p. \tag{68}$$

By the above assumptions on  $B$ ,  $M_r$  is *symmetric* (that is,  $u \in M_r \Rightarrow -u \in M_r$ ) and *sphere-like*, in the sense that each ray through the origin hits  $M_r$  in exactly one point. If  $K \subset M_r$  is compact and symmetric, then the *genus* of  $K$ , denoted  $\gamma(K)$ , is defined as

$$\gamma(K) = \inf\{n \in \mathbb{N} : \text{there exists a continuous odd map of } K \text{ into } \mathbb{R}^n \setminus \{0\}\}.$$

If  $V$  is a subspace of  $E$  with  $\dim V = n$ , then  $\gamma(M_r \cap V) = n$ . For  $n \in \mathbb{N}$  put

$$K_n(r) = \{K \subset M_r : K \text{ compact and symmetric, } \gamma(K) \geq n\}. \tag{69}$$

**Theorem 5.** Let  $A, B : E \rightarrow E'$  be as above. Suppose moreover that  $a(u) \neq 0$  implies  $A(u) \neq 0$ . For  $n \in \mathbb{N}$  and  $r > 0$ , put

$$C_n(r) \equiv \sup_{K_n(r)} \inf_K a(u) \tag{70}$$

where  $K_n(r)$  is as in Equation (69). Then

$$\sup_{M_r} a(u) = C_1(r) \geq \dots \geq C_n(r) \geq C_{n+1}(r) \geq \dots \geq 0. \tag{71}$$

Moreover,  $C_n(r) \rightarrow 0$  as  $n \rightarrow \infty$ , and if  $C_n(r) > 0$  then  $C_n(r)$  is attained and is critical value of  $a$  on  $M_r$ ; thus, there exist  $u_n(r) \in M_r$  and  $\lambda_n(r) \in \mathbb{R}$  such that

$$C_n(r) = a(u_n(r)) \tag{72}$$

and

$$A(u_n(r)) = \lambda_n(r)B(u_n(r)). \tag{73}$$

Here are a few indications for the Proof of Theorem 5:

(i) The sequence  $(C_n(r))$  is non-decreasing because, for any  $n \in \mathbb{N}$ , we have  $K_n(r) \supset K_{n+1}(r)$  as shown by Equation (69). (ii) In addition,  $C_1(r) = \sup_{M_r} a(u)$  because  $K_1(r)$  contains all sets of the form  $\{x\} \cup \{-x\}$ ,  $x \in M_r$ . (iii) The proof that  $C_n(r) \rightarrow 0$  as  $n \rightarrow \infty$ , together with a lot of related information, can be found for instance in [34]. (iv) Finally, the assumption that  $A(u) \neq 0$  whenever  $a(u) \neq 0$ , together with the stated continuity properties of  $A$  and  $B$ , ensures that  $a$  satisfies the crucial Palais–Smale (PS) condition on  $M_r$  at any level  $C > 0$ , needed to prove the final (and most important) assertion of the Theorem via the standard deformation methods of Critical Point Theory; see for this any of the above cited references.

Of special importance—with reference to the the  $p$ -Laplace equation—is the case in which  $A$  and  $B$  have the additional property of being **positively homogeneous** of the same degree  $p - 1 > 0$ , meaning that  $A(tu) = t^{p-1}A(u)$  for  $u \in E$  and  $t > 0$ , and similarly for  $B$ . In this case, we have from Equation (47)

$$a(u) = \frac{\langle A(u), u \rangle}{p}, \quad b(u) = \frac{\langle B(u), u \rangle}{p} \tag{74}$$

so that  $a(u) \neq 0$  implies  $A(u) \neq 0$ . Moreover, the use of Equation (74) in Equations (72) and (73) yields at once the relation  $C_n(r) = \lambda_n(r)r$ . In fact, here,  $\lambda_n(r)$  is independent of  $r > 0$ : to see this, it is convenient to re-parameterize the level sets on putting for  $R > 0$

$$M_R = \{u \in E : b(u) = \frac{R^p}{p}\} = \{u \in E : \langle B(u), u \rangle = R^p\}.$$

As  $a$  and  $b$  are  $p$ -homogeneous, it follows that  $M_R = RM_1$ , that each  $K \in K_n(R)$  is the image of the corresponding set in  $K_n(1)$  under the map  $u \rightarrow Ru$ , and that  $C_n(R) = R^p C_n(1)$ . By these remarks, we thus have the equalities

$$\lambda_n(R) \frac{R^p}{p} = C_n(R) = R^p C_n(1)$$

showing as expected that  $\lambda_n(R)$  is independent of  $R$ , and precisely that

$$\lambda_n(R) = pC_n(1) = \sup_{K_n} \inf_K \langle A(u), u \rangle \equiv \lambda_n \tag{75}$$

where  $K_n \equiv K_n(1)$ . From Theorem 5, we then get immediately the following statements about  $\lambda_n$ :

- $\sup_{M_1} \langle A(u), u \rangle = \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq \dots$ ;
- $\lambda_n \rightarrow 0$  as  $n \rightarrow \infty$ ; and

- if  $\lambda_n > 0$ , then there exists  $u_n \in M_1$  (that is,  $\langle B(u_n), u_n \rangle = 1$ ) such that  $Au_n = \lambda_n B(u_n)$ ; in particular,  $\lambda_n = \langle A(u_n), u_n \rangle$ .

**Remark 3.** The situation just described contains as a more special case that of two linear operators  $A$  and  $B$ , in which the above formulae hold with  $p = 2$ . Suppose in particular that  $A$  acts in a real Hilbert space  $H$  and  $B = I$ ; then  $M_1 = \{u \in H : \|u\| = 1\}$  is the unit sphere in  $H$ , while  $A$  is a compact, self-adjoint, non-negative linear operator (strong sequential continuity and compactness are equivalent properties for a linear operator acting in a reflexive Banach space, see e.g., [15]). Then, Equation (75) and the statements following this formula yield a good part of the familiar spectral properties of such operators: indeed it is not hard to see that the LS variational characterization in Equation (75) of  $\lambda_n$  reduces in this case to the classical Courant’s minimax principle expressed by Equation (15), so that the sequence in Equation (75) of the LS eigenvalues of  $A$  coincides with the decreasing sequence of all the eigenvalues of  $A$ , each counted with its multiplicity.

Returning finally to the  $p$ -Laplacian, it is now a matter of applying the above information to the operators  $A_p, B_p$  defined in Equation (66). One can check (see [36,37], for instance) that they satisfy all the requirements for the application of Theorem 5. Moreover, they are evidently positively homogeneous of degree  $p - 1$ , and finally  $A_p$  is (strictly) positive, for

$$\langle A_p(v), v \rangle = \int_{\Omega} |v|^p dx > 0 \quad \text{for } v \in E, v \neq 0.$$

This implies that each of the numbers  $\lambda_n$  defined in Equation (75) for the pair  $A_p, B_p$  is strictly positive, whence it follows—using the last statement of Theorem 5—that the eigenvalue problem in Equation (65) for the  $p$ -Laplacian possesses an infinite sequence of eigenvalues  $\lambda_n > 0$ , each given by

$$\lambda_n = \sup_{K_n} \inf_K \int_{\Omega} |v|^p \quad (n = 1, 2, \dots)$$

where

$$K_n = \{K \subset \{v \in W_0^{1,p} : \int_{\Omega} |\nabla v|^p = 1\}, K \text{ compact and symmetric, } \gamma(K) \geq n\}.$$

Setting  $\mu_n = \lambda_n^{-1}$ , this finally proves the properties of Equation (5) stated in the Introduction, and in particular Equation (6).

**Remark 4.** For the very special properties owned by the first eigenvalue  $\mu_1$  in the sequence in Equation (6) and by the associated eigenfunctions, see for instance [37]. Anyway, it follows by our discussion that  $\lambda_1 = \mu_1^{-1}$  is the best constant in Poincaré’s inequality, Equation (64):

$$\lambda_1 = \sup_{v \in W_0^{1,p}, v \neq 0} \frac{\int_{\Omega} |v|^p}{\int_{\Omega} |\nabla v|^p}.$$

To conclude this section, let us remark that the study and research in problems related to the  $p$ -Laplacian has grown enormously in the last decades, and even remaining in the strict context of a “spectral theory” for Equation (5), one should at least mention the following relevant points: (i) the problem of the asymptotic distribution of the LS eigenvalues (along the classical Weyl’s law for the Laplacian); (ii) the question of the existence of other eigenvalues outside the LS sequence; and (iii) the Fredholm alternative for perturbed non-homogeneous versions of Equation (5). For information on these issues, we refer the reader to [37–39] and to the recent and very clear review paper [36]. Related material can be found in [40].

### 3. Nonlinear Perturbation of an Isolated Eigenvalue

As a way to introduce and motivate the more specific content of this section, let me start recalling a famous and beautiful result of F. Rellich in perturbation theory of linear eigenvalue problems:

**Theorem 6.** ([41], Theorem 1). Let  $A(\epsilon)$  be a family of Hermitian  $n \times n$  matrices depending analytically on the real parameter  $\epsilon$  for  $\epsilon$  near 0. Let  $\lambda_0$  be an eigenvalue of multiplicity  $m > 1$  of  $A = A(0)$ . Then, for  $\epsilon$  near 0,  $A(\epsilon)$  possesses  $m$  eigenvalues

$$\lambda_1(\epsilon), \dots, \lambda_m(\epsilon)$$

and corresponding orthonormal eigenvectors  $u_1(\epsilon), \dots, u_m(\epsilon)$ ; that is, for all sufficiently small  $\epsilon$ , we have

$$A(\epsilon)u_i(\epsilon) = \lambda_i(\epsilon)u_i(\epsilon) \quad (i = 1, \dots, m). \tag{76}$$

Moreover,  $\lambda_i(0) = \lambda_0$  for all  $i = 1, \dots, m$  and the functions  $\lambda_i$  and  $u_i$  depend analytically on  $\epsilon$  near  $\epsilon = 0$ .

As is well known, the “ideal” situation described by Theorem 6 for the **splitting of the multiple eigenvalue** does not hold in general. In Rellich’s words, “...our question about the eigenvalues reduces to asking whether or not the zeroes of a polynomial [in the case, the characteristic polynomial of a matrix whose elements depend analytically on a parameter  $\epsilon$ ] are themselves regular analytic functions of  $\epsilon$  for small  $\epsilon$ . In general the answer is no; a counterexample is  $\lambda^2 + \epsilon$ . What is true is that if  $\lambda = \lambda(0)$  is a zero for  $\epsilon = 0$ , then the zero  $\lambda(\epsilon)$  can be written as a convergent (for small  $\epsilon$ ) power series in  $\epsilon^{\frac{1}{h}}$  (**Puiseux series**) where  $h$  is the multiplicity of  $\lambda = \lambda(0)$ .” The example indicated by Rellich can be displayed as

$$A(\epsilon) \equiv \begin{pmatrix} 0 & 1 \\ -\epsilon & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + \epsilon \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix} \equiv A + \epsilon B \tag{77}$$

and shows the unperturbed eigenvalue  $\lambda_0 = 0$  of  $A$ , of multiplicity  $h = 2$ , splitting into the two simple eigenvalues  $\lambda(\epsilon) = \pm(-\epsilon)^{\frac{1}{2}}$  of  $A(\epsilon)$ . In general, if  $\lambda_0$  has multiplicity  $h$ , the perturbed eigenvalue(s)  $\lambda(\epsilon)$  will admit an expansion such as

$$\lambda(\epsilon) = \lambda_0 + \epsilon^{\frac{1}{h}}\lambda_1 + \epsilon^{\frac{2}{h}}\lambda_2 + \dots + \epsilon\lambda_h + \dots = \lambda_0 + \sum_{i=1}^{\infty} \epsilon^{\frac{i}{h}}\lambda_i \tag{78}$$

For the special case that  $A(\epsilon)$  is Hermitian, using the reality of  $\lambda(\epsilon)$  Rellich showed in [41] that only integral powers of  $\epsilon$  can have non-zero coefficients in the expansion of Equation (78), thus proving the analytic dependence on  $\epsilon$  of the perturbed eigenvalues as stated in Theorem 6. Rellich’s work was a main starting point for the very vast literature concerning the systematic analysis of the perturbation of eigenvalues of linear operators, both in finite and infinite dimensional spaces; see Kato’s book [7] and the references therein. Our aim in this section is to indicate some partial results about similar questions for *nonlinear* eigenvalue problems, both of type  $G$  and of type  $K$ , recently appearing in [8,9], respectively.

#### 3.1. A Perturbation Problem of Type $G$

In the paper [8], the authors study the splitting of a multiple eigenvalue of the nonlinear eigenvalue problem, depending on the real parameter  $\epsilon$ ,

$$M(\lambda; \epsilon)v = 0, \quad \lambda \in \mathbb{C}, \quad v \in \mathbb{C}^n, \quad v \neq 0. \tag{79}$$

Here,  $M(\lambda; \epsilon)$  is an  $n \times n$  complex matrix having an eigenvalue  $\lambda_0$  for  $\epsilon = 0$  (i.e.,  $\det M(\lambda_0; 0) = 0$ ). As in the linear case, a perturbation theory for the eigenvalue  $\lambda_0$  consists in the study of the eigenvalues of Equation (79)—and of the corresponding eigenvectors—in the vicinity of  $\lambda_0$ , and will focus precisely on the behaviour of such eigenvalues/eigenvectors as functions  $\lambda(\epsilon), v(\epsilon)$  of the parameter  $\epsilon$  for  $\epsilon$

near 0; one assumes to know the solutions of Equation (79) for  $\epsilon = 0$ , i.e., to know the nullspace of  $M(\lambda_0; 0)$ . In the linear case, we have

$$M(\lambda; \epsilon) = A(\epsilon) - \lambda I$$

for some assigned function  $A$  of  $\epsilon$  into  $\mathbb{C}^{n \times n}$ , and Rellich’s theorem can be rephrased on saying that if this function is analytic and with Hermitian values, and if

$$\dim \text{Ker}(A(0) - \lambda_0 I) = m$$

then there exist  $m$  pairs of analytic functions  $\lambda_i(\epsilon), u_i(\epsilon)$  such that  $\lambda_i(0) = \lambda_0, u_i(0) \in \text{Ker}(A(0) - \lambda_0 I)$ , each pair satisfying identically Equation (76) for  $\epsilon$  near 0.

For the study of Equation (79), it is assumed that  $M(\lambda; \epsilon)$  depends regularly on  $\lambda$  and  $\epsilon$  in the following sense: there exists an open set  $\Omega \subset \mathbb{C}$  containing  $\lambda_0$ , and an open interval  $I \subset \mathbb{R}$  containing zero, such that for all  $\epsilon \in I$  the entries of  $M$  are analytic functions of  $\lambda$  in  $\Omega$ , and for all  $\lambda \in \Omega$  the entries of  $M$  are smooth functions of  $\epsilon$  in  $I$ . In the first part of [8], the authors develop previous work on the subject and consider the case in which the geometric multiplicity of  $\lambda_0$  (that is, the dimension of the nullspace of  $M(\lambda_0, 0)$ ) is one, while its algebraic multiplicity (that is, the multiplicity of  $\lambda$  as a root of the characteristic equation  $\det M(\lambda; 0) = 0$ ) is  $m > 1$ . Thus,  $\lambda_0$  is a multiple, nonsemisimple eigenvalue of  $M$  for  $\epsilon = 0$ . The following notations are used in the sequel:

$$M_\epsilon \equiv \frac{\partial M}{\partial \epsilon}(\lambda_0, 0); M_\lambda \equiv \frac{\partial M}{\partial \lambda}(\lambda_0, 0); M_{\lambda\lambda} \equiv \frac{\partial^2 M}{\partial \lambda^2}(\lambda_0, 0); \dots M_{\lambda^m} \equiv \frac{\partial^m M}{\partial \lambda^m}(\lambda_0, 0).$$

**Theorem 7.** [8]. *Let  $\lambda_0$  be an eigenvalue of Equation (79) for  $\epsilon = 0$ , with algebraic multiplicity equal to  $m$  and geometric multiplicity one, with Jordan chain  $(H_0, \dots, H_{m-1})$ . Let  $U_0$  be the corresponding left eigenvector. Assume that the condition  $U_0^* M_\epsilon H_0 \neq 0$  holds. Then, around  $\epsilon = 0$ , the eigenvalues in the vicinity of  $\lambda_0$  can be expanded as the branches of the Puiseux series in Equation (78), where*

$$\lambda_1^m = - \frac{U_0^* M_\epsilon H_0}{U_0^* (\frac{1}{1!} M_\lambda H_{m-1} + \frac{1}{2!} M_{\lambda\lambda} H_{m-2} + \dots + \frac{1}{m!} M_{\lambda^m} H_0)}. \tag{80}$$

**Remark 5.** *In the classical terminology of Numerical Analysis (see e.g., ([42], p. 137), a (column) vector  $U \in \mathbb{C}^n = \mathbb{C}^{n \times 1}$  is a left eigenvector of a matrix  $M$  if  $U^* M = 0$ , where  $U^*$  denotes the transpose of the conjugate. “Starring both sides”, this is equivalent to  $M^* U = 0$ , that is  $U$  is a (“right”) eigenvector of the adjoint matrix  $M^*$ . With the same notations, for the scalar product in  $\mathbb{C}^n$  we have*

$$\langle x, y \rangle = \sum_{i=1}^n x_i \bar{y}_i = y^* x$$

(the last product being the matrix product between  $y^* \in \mathbb{C}^{1 \times n}$  and  $x \in \mathbb{C}^{n \times 1}$ ), and therefore Equation (80) reads

$$\lambda_1^m = - \frac{\langle M_\epsilon H_0, U_0 \rangle}{\langle Z, U_0 \rangle}$$

with  $Z = (\frac{1}{1!} M_\lambda H_{m-1} + \frac{1}{2!} M_{\lambda\lambda} H_{m-2} + \dots + \frac{1}{m!} M_{\lambda^m} H_0)$ .

To some extent, the proof of Theorem 7 relies on previous work by Lancaster et al. [25,43] on the perturbation of analytic matrix functions. To indicate the main idea followed to obtain Equation (80), consider that by definition the perturbed  $\lambda(\epsilon), v(\epsilon)$  have to satisfy for all  $\epsilon$  the condition

$$M(\lambda(\epsilon), \epsilon) v(\epsilon) = 0. \tag{81}$$

Now use the Taylor expansion of  $M(\lambda, \epsilon)$  around  $(\lambda_0, 0)$ ,

$$M(\lambda, \epsilon) = M(\lambda_0, 0) + M_\epsilon \epsilon + \frac{1}{1!} M_\lambda (\lambda - \lambda_0) + \frac{1}{2!} M_{\lambda\lambda} (\lambda - \lambda_0)^2 + \dots$$

and replace  $\lambda$  with the expansion of Equation (78) for  $\lambda(\epsilon)$ . Using a similar expansion for  $v(\epsilon)$ ,

$$v(\epsilon) = V_0 + \sum_{i=1}^{\infty} \epsilon^{\frac{i}{m}} V_i$$

starting with an eigenvector  $V_0$  associated with  $\lambda_0$ , and putting all this in Equation (81) yields (equalling to zero the coefficients of the increasing powers of  $\epsilon^{\frac{1}{m}}$ )  $m + 1$  recursive equations that contain the elements of a Jordan chain built upon the unknown vectors  $V_1, \dots, V_m$ . Solving these equations with the help of a technical lemma (Lemma 2.1 in [8]) that relates all possible Jordan chains corresponding to the same eigenvalue, one returns to the original chain  $(H_0, \dots, H_{m-1})$  and finally obtains Equation (80).

**Example 1.** [8]. Consider the perturbed matrix

$$M_2(\lambda, \epsilon) = \begin{pmatrix} \lambda - 1 + e^{-\lambda} + \epsilon & 0 \\ 0 & \lambda + 1 + \epsilon \end{pmatrix} \tag{82}$$

that for  $\epsilon = 0$  reduces to Equation (30). For the unperturbed eigenvalue  $\lambda_0 = 0$ , in addition to the Jordan chain in Equation (31), one has

$$U_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad M_\epsilon = I, \quad M_\lambda = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad M_{\lambda\lambda} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \tag{83}$$

Substituting these values in Equation (80), one obtains

$$\lambda_1 = \sqrt{-2}.$$

In the second part of [8], the authors consider general linear functional differential equations of the form

$$x'(t) = \int_{-\tau_{max}}^0 d\mu(\theta)x(t + \theta), \quad x(t) \in \mathbb{C}^n \tag{84}$$

where  $\mu : [-\tau_{max}, 0] \rightarrow \mathbb{C}^{n \times n}$  is a function of bounded variation such that  $\mu(0) = 0$ . We refer to Chapter 7 of [26] for a thorough discussion of this kind of problems. Note that Equation (84) contains as special cases both equations with discrete delay and equations with continuous delay: indeed, Equation (84) takes the form of Equation (35) if one lets

$$0 = \tau_0 < \tau_1 < \dots < \tau_k \leq \tau_{max}$$

and defines  $\mu : [-\tau_{max}, 0] \rightarrow \mathbb{C}^{n \times n}$  as follows:

$$\begin{cases} \mu(0) &= 0 \\ \mu(\theta) &= -\sum_{i=0, -\tau_i > \theta}^k A_i, \quad \theta \in (-\tau_{max}, 0) \\ \mu(-\tau_{max}) &= -\sum_{i=0}^k A_i. \end{cases} \tag{85}$$

On the other hand, taking in Equation (84)  $\mu(\theta) = -\int_\theta^0 A(s) ds$  for  $\theta \in [-\tau_{max}, 0]$ , with  $A$  a continuous function from  $[-\tau_{max}, 0]$  to  $\mathbb{C}^{n \times n}$ , yields the system with distributed delay

$$x'(t) = \int_{-\tau_{max}}^0 A(\theta)x(t + \theta)d\theta.$$

The relation of Equation (84) with Equation (79) is as follows: looking for solutions  $x(t) = e^{\lambda t}v$  ( $v \in \mathbb{C}^n$ ) of Equation (84) yields the equation

$$\lambda v = \int_{-\tau_{max}}^0 d\mu(\theta)e^{\lambda\theta}v \equiv N(\lambda)v \tag{86}$$

that is a non-parametric form of Equation (79) with  $M(\lambda) = \lambda I - N(\lambda)$ . The authors then consider the infinite-dimensional vector space  $X = \mathbb{C}^n \times L^2([-\tau_{max}, 0], \mathbb{C}^n)$  and a suitably defined linear operator  $A$  acting in  $X$  and having the property that Equation (84) can be rewritten as the abstract ordinary differential equation  $z'(t) = Az(t)$  in  $X$ . Moreover,  $\lambda_0$  is an eigenvalue of the nonlinear eigenvalue problem in Equation (86) if and only if it is an eigenvalue of the linear operator  $A$ , and in Theorem 3.1 of [8] it is shown how to build an (ordinary) Jordan chain for  $A$  corresponding to  $\lambda_0$  starting from a Jordan chain for  $\lambda_0$  as an eigenvalue of the NLEVP in Equation (86), and vice versa; for this matter, see also ([26], Chapter 7, Theorem 4.2). Further exploiting this functional-analytic point of view, the authors are then able to deal with parameter-dependent forms of Equation (84)—that is, with functions  $\mu = \mu(\theta, \epsilon)$ —and to reformulate the sensitivity formula in Equation (80) for the eigenvalues  $\lambda_\epsilon$  of the perturbed matrix  $N(\lambda; \epsilon)$ , corresponding to  $\mu = \mu(\theta, \epsilon)$  as in Equation (86), in terms of eigenvectors and generalized eigenvectors of the linear operator  $A(\epsilon)$  acting in  $X$ . This produces a more readable formula, given in Theorem 3.2 of [8], for the coefficient  $\lambda_1$  of the leading term in the expansion in Equation (78).

The concluding section of [8] shows applications of the theory to some numerical examples, that deal in particular with a planar time-delay system containing an uncertain delay  $\tau + \epsilon$  and with a model problem for spectral abscissa optimization.

### 3.2. A Perturbation Problem of Type K

In [9] we have considered the following parameter-dependent version of Equation (39),

$$Tx + \epsilon B(x) = \lambda x, \quad x \in S \tag{87}$$

where—as in Section 2.2— $T$  is a self-adjoint bounded linear operator acting in a real Hilbert space  $H$  and having  $\lambda_0 \in \mathbb{R}$  as an isolated eigenvalue of finite multiplicity. In Equation (87),  $S$  stands for the unit sphere in  $H$ , so that  $S \cap \text{Ker}(T - \lambda_0 I)$  is the unit sphere in some  $\mathbb{R}^n$ . As to the nonlinear term  $B$ , we shall soon give precise assumptions, but roughly speaking can say that the  $\epsilon$  term appearing before it in Equation (87) replaces the condition  $B(x) = o(\|x\|)$  as  $\|x\| \rightarrow 0$  previously considered for bifurcation in Equation (39). Indeed, rather than looking for solutions of small norm as in Equation (39), we now look for *normalized* eigenvectors of the perturbed eigenvalue problem  $Tx + \epsilon B(x) = \lambda x$ . Here, is our result for Equation (87):

**Theorem 8.** *Let  $T$  be a self-adjoint bounded linear operator acting in a real Hilbert space  $H$ , and having  $\lambda_0$  as an isolated eigenvalue of finite multiplicity. Suppose that  $B$  is a  $C^1$  map of  $H$  into itself, and suppose moreover that at least one of the following conditions is satisfied: either*

- (a) *the dimension of the nullspace  $N \equiv \text{Ker}(T - \lambda_0 I)$  is odd; or*
- (b)  *$B$  is a gradient operator.*

*Then, there exist  $\epsilon_0 > 0, \delta_0 > 0$  such that for any  $\epsilon \in [-\epsilon_0, \epsilon_0]$ , there exist  $\lambda_\epsilon \in [\lambda_0 - \delta_0, \lambda_0 + \delta_0]$  and  $x_\epsilon \in S$  such that*

$$Tx_\epsilon + \epsilon B(x_\epsilon) = \lambda_\epsilon x_\epsilon. \tag{88}$$

*If moreover  $B$  is bounded on  $S$ , then  $\lambda_\epsilon \rightarrow \lambda_0$  as  $\epsilon \rightarrow 0$ . Finally, if we suppose in addition that  $B(0) = 0$  and that  $B$  is Lipschitz continuous in the unit ball  $U = \{x \in H : \|x\| \leq 1\}$  of  $H$ , i.e., that there exist  $k > 0$  such that*

$$\|B(x) - B(y)\| \leq k\|x - y\| \tag{89}$$

for  $x, y \in U$ , then putting

$$C = \inf_{0 < \|v\| \leq 1, v \in N} \frac{\langle B(v), v \rangle}{\|v\|^2}, \quad D = \sup_{0 < \|v\| \leq 1, v \in N} \frac{\langle B(v), v \rangle}{\|v\|^2} \tag{90}$$

the following asymptotic estimate for  $\lambda_\epsilon$  hold as  $\epsilon \rightarrow 0^+$  :

$$\lambda_0 + \epsilon C + O(\epsilon^2) \leq \lambda_\epsilon \leq \lambda_0 + \epsilon D + O(\epsilon^2). \tag{91}$$

The same estimate, with reversed inequalities, holds for  $\epsilon \rightarrow 0^-$ .

**Remark 6.** The bounds in Equation (91) are sharp in the sense that there exist perturbing operators  $B$  satisfying all the assumptions of the Theorem, and perturbed eigenvalues  $\lambda^+(\epsilon), \lambda^-(\epsilon)$  of  $T + \epsilon B$  that satisfy at least one of the inequalities in Equation (91) with the equality sign. To see this, just consider a linear operator  $B_0$  acting in the finite-dimensional subspace  $N$ , and then extend it to all of  $H$  on putting  $B(x) = B_0(v)$  for all  $x \in H$ , with  $v$  the orthogonal projection of  $x$  onto  $N$ . If  $B_0 : N \rightarrow N$  is taken to be self-adjoint, then it has  $n$  eigenvalues (counting multiplicities)  $\mu_0^1 \leq \dots \leq \mu_0^n$  with normalized eigenvectors  $v_1, \dots, v_n$ , say; that is,  $B_0 v_i = \mu_0^i v_i$  and  $\|v_i\| = 1$ . Then, putting for each  $i = 1, \dots, n$

$$\lambda_\epsilon = \lambda_0 + \epsilon \mu_0^i, \quad x_\epsilon = v_i$$

we have  $n$  families of eigenvalues/eigenvectors satisfying Equation (87) for all  $\epsilon \in \mathbb{R}$ . We have  $\mu_0^i = \langle B_0 v_i, v_i \rangle = \langle B v_i, v_i \rangle$  for each  $i$ , and the variational characterization of the eigenvalues of  $B_0$  gives in particular

$$\mu_0^1 = \inf_{0 < \|v\| \leq 1, v \in N} \frac{\langle B(v), v \rangle}{\|v\|^2} = C$$

and similarly  $\mu_0^n = D$ . Therefore taking  $\lambda^-(\epsilon) = \lambda_0 + \epsilon \mu_0^1$  (respectively,  $\lambda^+(\epsilon) = \lambda_0 + \epsilon \mu_0^n$ ), the left-hand side (respectively, the right-hand side) of Equation (91) is satisfied with equality sign and  $O(\epsilon^2) = 0$ .

The first part of Theorem 8 is proved following the track indicated in Section 2.2, that is performing the Lyapounov–Schmidt reduction of Equation (87). One non-trivial difference is that here a global version of the Implicit Function Theorem is employed in order to obtain a mapping

$$(\delta, \epsilon, v) \rightarrow w(\delta, \epsilon, v)$$

defined in an open neighborhood  $Y_1 = I_1 \times J_1 \times V_1 \subset \mathbb{R} \times \mathbb{R} \times N$  of  $\{0\} \times \{0\} \times S$  by the rule that the  $\epsilon$ -dependent complementary equation (see Equation (52)) can be solved uniquely with respect to  $w$  for each given  $(\delta, \epsilon, v) \in Y_1$ . Moreover,  $w(0, 0, v) = 0$  for any  $v \in S$ , and the mapping  $(\delta, \epsilon, v) \rightarrow w(\delta, \epsilon, v)$  of  $Y_1$  into  $W$  is of class  $C^1$ . Next, expressing  $\delta$  as a  $C^1$  function  $\delta(\epsilon, v)$  of  $(\epsilon, v)$  in a possibly smaller neighborhood  $J \times V \subset J_1 \times V_1$ , and putting for convenience

$$\phi(\epsilon, v) \equiv w(\delta(\epsilon, v), \epsilon, v), \quad (\epsilon, v) \in J \times V$$

we arrive at the  $\epsilon$ -dependent form of the bifurcation Equation (57), namely

$$\epsilon P B(v + \phi(\epsilon, v)) = \delta(\epsilon, v)v \tag{92}$$

that is here accompanied by the norm constraint

$$\|v + \phi(\epsilon, v)\|^2 = \|v\|^2 + \|\phi(\epsilon, v)\|^2 = 1. \tag{93}$$

A solution  $(\lambda, x)$  of the original problem in Equation (87) will then be given by the formulae

$$\lambda = \lambda_0 + \delta(\epsilon, v), \quad x = v + \phi(\epsilon, v). \tag{94}$$

In any of the two Cases (a) and (b) listed in Theorem 8, using as needed either of the methods (topological or variational) recalled in general in Section 2.2, we find for each  $\epsilon$  small a solution  $v_\epsilon$  of Equations (92) and (93). Therefore, making an appropriate choice of  $\delta_0, \epsilon_0$  for the intervals  $I_0 \equiv [-\delta_0, \delta_0], J_0 \equiv [-\epsilon_0, \epsilon_0]$  and putting

$$\delta_\epsilon = \delta(\epsilon, v_\epsilon) \quad \text{and} \quad w_\epsilon = \phi_\epsilon(v_\epsilon) = w(\delta_\epsilon, \epsilon, v_\epsilon) \tag{95}$$

the first part of Theorem 8, asserting the existence of at least one solutions  $(\lambda_\epsilon, x_\epsilon) \in I_0 \times S$  of Equation (87) for each  $\epsilon \in J_0$ , is proved with  $\lambda_\epsilon = \lambda_0 + \delta_\epsilon$  and  $x_\epsilon = v_\epsilon + w_\epsilon$ .

Some words are now in order to explain the estimates in Equation (91). One first shows that the component  $w_\epsilon$  of  $x_\epsilon$  (as defined in Equation (95)) satisfies  $w_\epsilon \rightarrow 0$  as  $\epsilon \rightarrow 0$ , uniformly with respect to  $v_\epsilon$ , and consequently with respect to  $x_\epsilon$ . This in turn implies that  $\lambda_\epsilon \rightarrow \lambda_0$  as  $\epsilon \rightarrow 0$ , uniformly with respect to  $x_\epsilon$ . Indeed, using Equation (92), we have

$$\epsilon PB(v_\epsilon + w_\epsilon) = \delta_\epsilon v_\epsilon$$

for all  $\epsilon$ , whence taking the scalar product with  $v_\epsilon$  of both members, we obtain

$$\epsilon \langle B(v_\epsilon + w_\epsilon), v_\epsilon \rangle = \delta_\epsilon \|v_\epsilon\|^2. \tag{96}$$

Therefore,

$$\delta_\epsilon = \epsilon \frac{\langle B(x_\epsilon), v_\epsilon \rangle}{\|v_\epsilon\|^2}. \tag{97}$$

Moreover, as  $x_\epsilon = v_\epsilon + w_\epsilon \in S$  and  $w_\epsilon \rightarrow 0$  as indicated above, then necessarily  $\|v_\epsilon\| \rightarrow 1$  as  $\epsilon \rightarrow 0$ . Therefore, since

$$\frac{|\langle B(x_\epsilon), v_\epsilon \rangle|}{\|v_\epsilon\|^2} \leq \frac{\|B(x_\epsilon)\|}{\|v_\epsilon\|}$$

it follows, by the boundedness assumption on  $B$ , that the term multiplying  $\epsilon$  in Equation (97) remains bounded as  $\epsilon \rightarrow 0$ , implying that  $\delta_\epsilon = O(\epsilon)$  as  $\epsilon \rightarrow 0$ , uniformly with respect to  $x_\epsilon$ .

We can now prove the asymptotic formula in Equation (91) on  $\lambda_\epsilon$  if  $B$  satisfies Equation (89). In this respect, the utility of Equation (89) is twofold. First, it permits to improve significantly the information on  $w_\epsilon$  as it yields by means of straightforward computations the estimate

$$\|w(\delta, \epsilon, v)\| \leq C_1 |\epsilon| \|v\| \tag{98}$$

holding for some constant  $C_1 > 0$  and all  $(\delta, \epsilon, v) \in [-\delta_0, \delta_0] \times [-\epsilon_0, \epsilon_0] \times (U \cap N)$ . Moreover, Equation (89) implies via the Schwarz' inequality that, for any  $v$  and  $w$  such that  $v, v + w \in U$ , one has

$$|\langle B(v + w), v \rangle - \langle B(v), v \rangle| \leq k \|v\| \|w\|.$$

Writing this for  $w(\delta, \epsilon, v)$  and using Equation (98), we then get the inequality

$$|\langle B(v + w(\delta, \epsilon, v)), v \rangle - \langle B(v), v \rangle| \leq C_2 |\epsilon| \|v\|^2 \tag{99}$$

with  $C_2 = kC_1$ , valid for all the (possible) solutions  $(\delta, x = v + w(\delta, \epsilon, v))$  of Equation (87) having sufficiently small  $\epsilon$ . Using in turn this estimate in Equation (96) for the actual solutions  $(\delta_\epsilon, x_\epsilon = v_\epsilon + w_\epsilon)$  we see that as  $\epsilon \rightarrow 0$

$$\delta_\epsilon \|v_\epsilon\|^2 = \epsilon \langle B(v_\epsilon + w_\epsilon), v_\epsilon \rangle = \epsilon \langle B(v_\epsilon), v_\epsilon \rangle + \|v_\epsilon\|^2 O(\epsilon^2). \tag{100}$$

This implies that

$$\delta_\epsilon = \epsilon \frac{\langle B(v_\epsilon), v_\epsilon \rangle}{\|v_\epsilon\|^2} + O(\epsilon^2) \tag{101}$$

as  $\epsilon \rightarrow 0$ , and thus for  $\epsilon > 0$  yields immediately the estimate of Equation (91) in view of the definition in Equation (90) of  $C$  and  $D$ .

**Example 2.** Theorem 8 can be used to evaluate the convergence rate as  $\epsilon \rightarrow 0$  of the eigenvalues  $\mu_\epsilon$  of the nonlinear elliptic problem

$$\begin{cases} -\Delta u = \mu(u + \epsilon f(x, u)) & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \end{cases} \tag{102}$$

near an eigenvalue  $\mu_0$  of the unperturbed linear problem  $-\Delta u = \mu u$  in  $\Omega$ ,  $u = 0$  on  $\partial\Omega$ . Here,  $\Omega$  is a bounded domain in  $\mathbb{R}^N$  ( $N \geq 1$ ) with boundary  $\partial\Omega$ , and  $\Delta = \sum_{i=1}^N \frac{\partial^2}{\partial x_i^2}$  is the familiar Laplace operator acting on sufficiently smooth real functions  $u$  defined in  $\Omega$ . Under appropriate hypotheses on  $f$ , and assuming in particular that

$$mt^2 \leq f(x, t)t \leq Mt^2 \quad (x \in \Omega, t \in \mathbb{R}) \tag{103}$$

for some real constants  $0 \leq m \leq M$ , one proves that as  $\epsilon \rightarrow 0^+$

$$\mu_0 - \epsilon\mu_0 M + O(\epsilon^2) \leq \mu_\epsilon \leq \mu_0 - \epsilon\mu_0 m + O(\epsilon^2). \tag{104}$$

These inequalities can be used for actual computation, once an efficient approximation of the linear eigenvalue  $\mu_0$  is available and the bounds in Equation (103) for  $f$  are known with accuracy. For instance, if  $f(x, t) = f(t) = \frac{t}{1+t^2}$ , just put in Equation (104)

$$m = \inf_{t \neq 0} \frac{f(t)t}{t^2} = 0, \quad M = \sup_{t \neq 0} \frac{f(t)t}{t^2} = 1. \tag{105}$$

#### 4. Concluding Remarks, Open Problems and Applicability

To summarize and motivate again the content of this paper, let me define it as an attempt to identify and logically re-connect (or at least give a common frame to) two important and presently distinct research areas in Mathematical Analysis and its applications that appear in the current literature under the same name of *Nonlinear Eigenvalue Problems*. As better explained in the Introduction, problems in these two areas are described (in abstract operator form) by the two equations

$$G(\lambda)x = 0 \tag{106}$$

(“problems of type G”) and

$$A(x) - \lambda C(x) = 0 \tag{107}$$

(“problems of type K”). Some basic facts and solution methods about each of the two equations are reported in Section 2. Section 3 is devoted to discuss two specific problems, one for each type, with the scope of giving samples of very recent research in either field.

While for the problem discussed in Section 3.1 there are already concrete numerical examples [8], these are still missing for the problem presented in Section 3.2 [9]. The main aim of this final section is to partially fill this gap by further commenting (in Section 4.1) on the formula

$$\lambda_0 + \epsilon C + O(\epsilon^2) \leq \lambda_\epsilon \leq \lambda_0 + \epsilon D + O(\epsilon^2) \tag{108}$$

proved in Theorem 8 and by finally providing, at least in a special case, a recipe ready for use in numerical simulation (Section 4.2).

### 4.1. Open Problems

The basic idea standing behind the formula in Equation (108) is that if the (algebraic and geometric) multiplicity  $m(\lambda_0)$  of the unperturbed eigenvalue  $\lambda_0$  of the linear operator  $T$  in Equation (87) is equal to  $m$ , then upon perturbation by a small term  $\epsilon B$  there are potentially  $m$  eigenvalue functions  $\lambda_1(\epsilon), \dots, \lambda_m(\epsilon)$  of  $T + \epsilon B$  satisfying Equation (108). This is what actually happens for a *linear* operator  $B$  (essentially under the further assumption that  $T$  and  $B$  are self-adjoint) as described by Rellich’s Theorem 6. Our idea is that something of this persists also for *nonlinear* operators that have good similarity with the linear ones: the class of Lipschitz continuous maps considered in Theorem 8 is apparently quite close to that of bounded linear maps, and in fact contains properly the latter. Indeed, the conclusions of Theorem 8 point in in this direction. However, many problems remain open and we describe here three of them (in increasing order of interest and difficulty):

- Verify on specific examples of nonlinear ODE/PDE/Equations in  $\mathbb{R}^n$ —by means of explicit computation or by means of a numerical analysis—the existence of at least one “eigenvalue branch”  $\lambda(\epsilon)$  satisfying the bounds in Equation (108) as predicted by the theory.
- Verify by the same means that the bounds in Equation (108) are optimal by producing examples of nonlinear problems (in the same fields as above) where at least one eigenvalue function exists that satisfies the RHS (LHS) bound in Equation (108) with equality sign.
- Exhibit examples of “nonlinear splitting of the multiple eigenvalue”, that is, of nonlinear problems in which starting from an unperturbed eigenvalue  $\lambda_0$  of multiplicity  $\geq 2$  there exist **two** different families  $\lambda_+(\epsilon), \lambda_-(\epsilon)$  respecting Equation (108), and possibly each satisfying the RHS (LHS) bound of it with equality sign.

Here, is a very simple example that highlights the above issues, and the last in particular. For the linear case, these questions were answered by Remark 6.

**Example 3.** Consider the system

$$\begin{cases} x + \epsilon x^3 = \lambda x \\ y + \epsilon y^3 = \lambda y. \end{cases} \tag{109}$$

In the notations of Equation (87) and of Theorem 8, we have here  $H = \mathbb{R}^2, T = I, \lambda_0 = 1$  and  $B(x, y) = (x^3, y^3)$ . Solving Equation (109) with the constraint  $x^2 + y^2 = 1$  gives the solutions

$$\begin{cases} (x, y) = (0, \pm 1), & (x, y) = (\pm 1, 0), & \lambda = 1 + \epsilon \\ (x, y) = (\pm \frac{1}{\sqrt{2}}, \pm \frac{1}{\sqrt{2}}), & & \lambda = 1 + \epsilon/2. \end{cases} \tag{110}$$

Therefore,  $\lambda_0 = 1$  splits into the two eigenvalue functions (each carrying four distinct eigenvectors)

$$\lambda_+(\epsilon) = 1 + \epsilon, \quad \lambda_-(\epsilon) = 1 + \epsilon/2.$$

This is in full agreement with Equation (108), for

$$D = \sup_{0 < \|v\| \leq 1, v \in N} \frac{\langle B(v), v \rangle}{\|v\|^2} = \sup_{0 < x^2 + y^2 \leq 1} \frac{x^4 + y^4}{x^2 + y^2} = 1 \tag{111}$$

and similarly, replacing “sup” with “inf”, we find that  $C = 1/2$ .

### 4.2. Applicability

As indicated in Example 2, the bounds in Equation (108) proved in Theorem 8 for the operator Equation (87) can be used for concrete nonlinear elliptic problems in a bounded domain  $\Omega \subset \mathbb{R}^N (N \geq 1)$  such as Equation (102), or more generalized forms of it in which  $-\Delta$  is replaced by

a uniformly elliptic second order operator in divergence form. In particular for  $N = 1$  this applies to the Sturm–Liouville problem

$$\begin{cases} -(p(x)u')' + q(x)u = \mu(u + \epsilon f(x, u)) & \text{in } ]a, b[ \\ u(a) = u(b) = 0 \end{cases} \tag{112}$$

where  $p \in C^1([a, b]), p > 0$  and  $q \in C([a, b])$ . The first remark on the applicability of Theorem 8 to such kind of problems, in one or more space variables, is that one must not worry of the multiplicity of the unperturbed eigenvalue because—as recalled for instance in [9]—the operator  $B$  corresponding to the nonlinear term  $f$  is a gradient operator, so that the assumption  $b$ ) of Theorem 8 is satisfied. In [9] it is previously recalled that these problems can be cast in the operator form of Equation (87) on taking as Hilbert space the Sobolev space  $H_0^1(\Omega) \equiv W_0^{1,2}(\Omega)$ , equipped with the scalar product

$$\langle u, v \rangle = \int_{\Omega} \nabla u(x) \nabla v(x) \, dx \tag{113}$$

and that the operator  $B$  mentioned above is defined via the duality relation

$$\langle B(u), v \rangle = \int_{\Omega} f(x, u(x))v(x) \, dx. \tag{114}$$

The representation in Equation (114) of  $B$  is the key formula to be used to gain information on the constants  $C, D$  appearing in Equation (108), for these are defined by the formulae in Equation (90) that involve precisely the nonlinear Rayleigh quotient of  $B$ . Indeed, using Equation (114) and the fact (see e.g., [9]) that for  $v \in N$  we have

$$\|v\|^2 = \int_{\Omega} \nabla v^2(x) \, dx = \mu_0 \int_{\Omega} v^2(x) \, dx \tag{115}$$

where  $\mu_0$  is the unperturbed eigenvalue and  $N$  the corresponding eigenspace, yields the following quite readable expression for  $D$ :

$$D = \sup_{0 < \|v\| \leq 1, v \in N} \frac{\langle B(v), v \rangle}{\|v\|^2} = \sup_{0 < \|v\| \leq 1, v \in N} \frac{\int_{\Omega} f(x, v(x))v(x) \, dx}{\mu_0 \int_{\Omega} v^2(x) \, dx} \tag{116}$$

Thus, essentially, in the applications of the theory to elliptic PDE or ODE, estimating the Rayleigh quotient of  $B$  reduces to estimating the ratio appearing in the RHS of Equation (116). In turn, this can be easily obtained by **pointwise bounds** on  $f$ : for clearly if  $f$  satisfies Equation (103), then it follows that for every  $v \in H$  (and in fact for every  $v \in L^2(\Omega)$ )

$$m \leq \frac{\int_{\Omega} f(x, v(x))v(x) \, dx}{\int_{\Omega} v^2(x) \, dx} \leq M.$$

We conclude by Equation (116) and the dual formula for  $C$  that

$$\frac{m}{\mu_0} \leq C, \quad D \leq \frac{M}{\mu_0}. \tag{117}$$

Using the inequalities in Equations (117) and (108) and putting  $\lambda_{\epsilon} = 1/\mu_{\epsilon}$  yield bounds on the perturbed eigenvalues  $\mu_{\epsilon}$  of Equation (102). Considering for instance the right-hand side of in Equation (108), we obtain

$$\frac{1}{\mu_{\epsilon}} \leq \frac{1}{\mu_0} + \frac{\epsilon M}{\mu_0} + O(\epsilon^2) = \frac{1}{\mu_0} (1 + \epsilon M + O(\epsilon^2)) \tag{118}$$

and doing the same with the lower bound thus yields

$$\frac{\mu_0}{1 + \epsilon m + O(\epsilon^2)} \geq \mu_\epsilon \geq \frac{\mu_0}{1 + \epsilon M + O(\epsilon^2)}. \tag{119}$$

**Remark 7.** Note that Equation (119) contains—as due—the equality

$$\mu_\epsilon = \frac{\mu_0}{1 + \epsilon a}$$

which plainly holds for the eigenvalues of Equation (102) in the linear case  $f(x, s) = as, a = \text{const}$ .

Finally, using in Equation (119) the asymptotic relation  $(1 + x)^{-1} = 1 - x + O(x^2)$  for  $x \rightarrow 0$ , we obtain as  $\epsilon \rightarrow 0^+$  the formula in Equation (104), which as remarked is ready for use in numerical experiments once  $\mu_0, m$  and  $M$  are known. For instance, taking  $f(x, s) = f(s) = \frac{s}{1+s^2}$  and using the bounds  $m = 0, M = 1$  (see Equation (105)) yields, for  $\epsilon \rightarrow 0^+$ ,

$$\mu_0 - \epsilon\mu_0 + O(\epsilon^2) \leq \mu_\epsilon \leq \mu_0 + O(\epsilon^2). \tag{120}$$

**The case of the simple eigenvalue.** More information can be gained in the case that  $\dim N = 1$ , so that  $N = \{t\phi, t \in \mathbb{R}\}$  for some  $\phi$  that we normalize taking  $\|\phi\| = 1$ . Then, the Rayleigh quotient of  $B$  simplifies as

$$\frac{\langle B(v), v \rangle}{\|v\|^2} = \frac{\langle B(t\phi), t\phi \rangle}{t^2\|\phi\|^2} = \frac{1}{t} \langle B(t\phi), \phi \rangle \equiv h(t), \quad 0 < |t| \leq 1. \tag{121}$$

Note that  $h$  is bounded since  $B$  is sublinear (that is,  $\|B(u)\| \leq k\|u\|$  for all  $u \in H$ ) as follows from Equation (89) and the condition  $B(0) = 0$ . It follows by Equations (114) and (121) that

$$h(t) = \frac{1}{t} \int_{\Omega} f(x, t\phi(x))\phi(x) \, dx. \tag{122}$$

Considering as above the example  $f(x, s) = f(s) = \frac{s}{1+s^2}$ , we obtain

$$h(t) = \int_{\Omega} \frac{\phi^2(x)}{1 + t^2\phi^2(x)} \, dx$$

showing that  $h$  can be extended continuously to  $t = 0$  and that it is an even function of  $t$  in  $[-1, 1]$ . Therefore, using also Equation (115) and the condition  $\|\phi\| = 1$ , we get

$$D = \sup_{-1 \leq t \leq 1} h(t) = \sup_{0 \leq t \leq 1} h(t) = \int_{\Omega} \phi^2(x) \, dx = \frac{1}{\mu_0} \tag{123}$$

while

$$C = \inf_{-1 \leq t \leq 1} h(t) = \inf_{0 \leq t \leq 1} h(t) = \int_{\Omega} \frac{\phi^2(x)}{1 + \phi^2(x)} \, dx \equiv K. \tag{124}$$

These computations show that in the present case:

- The upper bound  $\frac{M}{\mu_0} = \frac{1}{\mu_0}$  given by Equation (117) for  $D$  is optimal.
- The lower bound 0 given by Equation (117) for  $C$  can be improved to  $C = K > 0$ .

Proceeding in the same way as before (see Equations (118) and (119)) and using as before the asymptotic expansion for  $\frac{1}{1+x}$  as  $x \rightarrow 0$ , we see that, as  $\epsilon \rightarrow 0^+$ , Equation (120) can be replaced by the more precise formula

$$\mu_0 - \epsilon\mu_0 + O(\epsilon^2) \leq \mu_\epsilon \leq \mu_0 - \epsilon\mu_0^2 K + O(\epsilon^2). \tag{125}$$

The practical use of Equation (125) for numerical purposes requires explicit knowledge of the eigenvalue  $\mu_0$  and of the corresponding eigenfunction  $\phi$ . Typical cases in which these data are available are:

- $N = 2$ ,  $\Omega$  a rectangle or a circle, and  $\mu_0$  the first eigenvalue of the Dirichlet Laplacian in  $\Omega$  (see e.g., [44]).
- $N = 1$ ,  $\Omega = ]a, b[$  and  $\mu_0 = \mu_n$  any eigenvalue of the Sturm–Liouville problem in Equation (112) with simple forms of the coefficients  $p$  and  $q$ . For instance, if  $]a, b[ = ]0, \pi[$ ,  $p \equiv 1$  and  $q \equiv 0$  we have

$$\mu_n = n^2, \quad \phi_n(x) = \frac{1}{n} \sqrt{\frac{2}{\pi}} \sin nx \quad (0 \leq x \leq \pi).$$

As to the expression of  $\phi_n$ , recall that we have normed  $H_0^1(a, b)$  via the formula in Equation (113), which in this case reduces to  $(u, v) = \int_0^\pi u'(x)v'(x) dx$ .

**Acknowledgments:** The author wishes to thank the Referees for their valuable suggestions that have helped to improve the quality of the present paper.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Betcke, T.; Higham, N.J.; Mehrmann, V.; Schroder, C.; Tisseur, F. NLEVP: A collection of nonlinear eigenvalue problems. *ACM Trans. Math. Softw.* **2013**, *39*, 28. [CrossRef]
2. Berger, M.S. *Nonlinearity and Functional Analysis*; Academic Press: Cambridge, MA, USA, 1977.
3. Ambrosetti, A.; Malchiodi, A. *Nonlinear Analysis and Semilinear Elliptic Problems*; Cambridge University Press: Cambridge, UK, 2007.
4. Palais, R.S. Critical point theory and the minimax principle. In Proceedings of the Symposia in Pure Mathematics, Berkeley, CA, USA, 1–26 July 1968; American Mathematics Society: Providence, RI, USA, 1970; Volume XV, pp. 185–212.
5. Chabrowski, J. On nonlinear eigenvalue problems. *Forum Math.* **1992**, *4*, 359–375. [CrossRef]
6. Turner, R.E.L. A class of nonlinear eigenvalue problems. *J. Funct. Anal.* **1968**, *2*, 297–322. [CrossRef]
7. Kato, T. *Perturbation Theory for Linear Operators*, 2nd ed.; Springer: Berlin, Germany; New York, NY, USA, 1976.
8. Michiels, W.; Boussaada, I.; Niculescu, S.I. An explicit formula for the splitting of multiple eigenvalues for nonlinear eigenvalue problems and connections with the linearization for the delay eigenvalue problem. *SIAM J. Matrix Anal. Appl.* **2017**, *38*, 599–620. [CrossRef]
9. Chiappinelli, R. Approximation and convergence rate of nonlinear eigenvalues: Lipschitz perturbations of a bounded self-adjoint operator. *J. Math. Anal. Appl.* **2017**, *455*, 1720–1732. [CrossRef]
10. Langer, H.; Markus, A.; Matsaev, V. Linearization, factorization, and the spectral compression of a self-adjoint analytic operator function under the condition (VM). In *A Panorama of Modern Operator Theory and Related Topics*; Birkhauser/Springer Basel AG: Basel, Switzerland, 2012; pp. 445–463.
11. Appell, J.; De Pascale, E.; Vignoli, A. *Nonlinear Spectral Theory*; de Gruyter: Berlin, Germany, 2004.
12. Chiappinelli, R. Surjectivity of coercive gradient operators in Hilbert space and Nonlinear Spectral Theory. *Ann. Funct. Anal.* **2018**, *9*, in press.
13. Guttel, S.; Tisseur, F. The nonlinear eigenvalue problem. *Acta Numer.* **2017**, *26*, 1–94. [CrossRef]
14. Brezis, H. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*; Springer: Berlin/Heidelberg, Germany, 2011.
15. Taylor, A.; Lay, D. *Introduction to Functional Analysis*; Wiley: Hoboken, NJ, USA, 1980.
16. Hädeler, K.P. Mehrparametrische und nichtlineare Eigenwertaufgaben. *Arch. Ration. Mech. Anal.* **1967**, *27*, 306–328. [CrossRef]
17. Hädeler, K.P. Variationsprinzipien bei nichtlinearen Eigenwertaufgaben. *Arch. Ration. Mech. Anal.* **1968**, *30*, 297–307. [CrossRef]

18. Langer, H. Über eine Klasse polynomialer Scharen selbstadjungierter Operatoren im Hilbertraum. *J. Funct. Anal.* **1973**, *12*, 13–29. [[CrossRef](#)]
19. Markus, A.S. *Introduction to the Spectral Theory of Polynomial Operator Pencils*; Translations of Mathematical Monographs, 71; American Mathematical Society: Providence, RI, USA, 1988.
20. Binding, P.; Eschwé, D.; Langer, H. Variational principles for real eigenvalues of self-adjoint operator pencils. *Integral Equ. Oper. Theory* **2000**, *38*, 190–206. [[CrossRef](#)]
21. Hasanov, M. An approximation method in the variational theory of the spectrum of operator pencils. *Acta Appl. Math.* **2002**, *71*, 117–126. [[CrossRef](#)]
22. Voss, H. A minmax principle for nonlinear eigenproblems depending continuously on the eigenparameter. *Numer. Linear Algebra Appl.* **2009**, *16*, 899–913. [[CrossRef](#)]
23. Schwetlick, H.; Schreiber, K. Nonlinear Rayleigh functionals. *Linear Algebra Appl.* **2012**, *436*, 3991–4016. [[CrossRef](#)]
24. Gohberg, I.; Lancaster, P.; Rodman, L. *Matrix Polynomials*; Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers]: New York, NY, USA; London, UK, 1982.
25. Hryniv, R.; Lancaster, P. On the perturbation of analytic matrix functions. *Integral Equ. Oper. Theory* **1999**, *34*, 325–338. [[CrossRef](#)]
26. Hale, J.K.; Verduyn Lunel, S.M. *Introduction to Functional Differential Equations*; Applied Mathematical Sciences, 99; Springer: New York, NY, USA, 1993.
27. Gohberg, I.; Lancaster, P.; Rodman, L. *Invariant Subspaces of Matrices With Applications*; A Wiley-Interscience Publication; John Wiley and Sons, Inc.: New York, NY, USA, 1986.
28. Stackgold, I. Branching of solutions of non-linear equations. *SIAM Rev.* **1971**, *13*, 289–332. [[CrossRef](#)]
29. Rabinowitz, P.H. *Minimax Methods in Critical Point Theory With Applications to Differential Equations*; CBMS Regional Conference Series Mathematics; American Mathematics Society: Providence, RI, USA, 1986; Volume 65.
30. Krasnoselskii, M.A. *Topological Methods in the Theory of Nonlinear Integral Equations*; Pergamon Press: Oxford, UK, 1964.
31. Rabinowitz, P.H. Some global results for nonlinear eigenvalue problems. *J. Funct. Anal.* **1971**, *7*, 487–513. [[CrossRef](#)]
32. Stuart, C.A. An introduction to bifurcation theory based on differential calculus. In *Nonlinear Analysis and Mechanics: Heriot-Watt Symposium*; Research Notes in Mathematics; Pitman: Totowa, NJ, USA 1979; Volume IV, pp. 76–135.
33. Mawhin, J.; Willem, M. *Critical Point Theory and Hamiltonian Systems*; Applied Mathematical Sciences, 74; Springer: New York, NY, USA, 1989.
34. Zeidler, E. *Nonlinear Functional Analysis and Its Applications. III. Variational Methods and Optimization*; Springer: New York, NY, USA, 1985.
35. Amann, H. Liusternik-Schnirelman theory and non-linear eigenvalue problems. *Math. Ann.* **1972**, *199*, 55–72. [[CrossRef](#)]
36. Fernandez Bonder, J.; Pinasco, J.P.; Salort, A.M. Quasilinear eigenvalues. *Rev. Union Mat. Argent.* **2015**, *56*, 1–25.
37. Lindqvist, P. A nonlinear eigenvalue problem. In *Topics in Mathematical Analysis*; World Scientific Publishing: Hackensack, NJ, USA, 2008; pp. 175–203.
38. Drábek, P. On the variational eigenvalues which are not of Ljusternik-Schnirelmann type. *Abstr. Appl. Anal.* **2012**, 434631. [[CrossRef](#)]
39. Drábek, P.; Robinson, S.B. Resonance problems for the  $p$ -Laplacian. *J. Funct. Anal.* **1999**, *169*, 189–200. [[CrossRef](#)]
40. Appell, J.; Drábek, P.; Chiappinelli, R. (Eds.) *Mini-Workshop: Nonlinear Spectral and Eigenvalue Theory with Applications to the  $p$ -Laplace Operator*; Abstracts from the Mini-Workshop held 15–21 February 2004; Oberwolfach Report; Mathematisches Forschungsinstitut Oberwolfach: Oberwolfach, Germany, 2004; pp. 407–437.
41. Rellich, F. *Perturbation Theory of Eigenvalue Problems*; Gordon and Breach Science Publishers: New York, NY, USA; London, UK; Paris, France, 1969.
42. Isaacson, E.; Keller, H.B. *Analysis of Numerical Methods*; John Wiley and Sons, Inc.: New York, NY, USA; London, UK; Sydney, Australia, 1966.

43. Lancaster, P.; Markus, A.S.; Zhou, F. Perturbation theory for analytic matrix functions: The semisimple case. *SIAM J. Matrix Anal. Appl.* **2003**, *25*, 606–626. [[CrossRef](#)]
44. Courant, R.; Hilbert, D. *Methods of Mathematical Physics*; Wiley: Hoboken, NJ, USA, 1953; Volume I.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).