Proceedings of the Royal Society of Edinburgh, **144A**, 1027–1065, *2014*

Bifurcation at isolated singular points of the Hadamard derivative

C. A. Stuart
Section de Mathématiques, Station 8, EPFL, 1015 Lausanne, Switzerland (charles.stuart@epfl.ch)

(MS received 24 February 2013; accepted 15 July 2013)

For Banach spaces X and Y, we consider bifurcation from the line of trivial solutions for the equation $F(\lambda, u) = 0$, where $F\colon \mathbb{R} \times X \to Y$ with $F(\lambda, 0) = 0$ for all $\lambda \in \mathbb{R}$. The focus is on the situation where $F(\lambda, \cdot)$ is only Hadamard differentiable at 0 and Lipschitz continuous on some open neighbourhood of 0, without requiring any Fréchet differentiability. Applications of the results obtained here to some problems involving nonlinear elliptic equations on $\mathbb{R}^N$, where Fréchet differentiability is not available, are presented in some related papers, which shed light on the relevance of our hypotheses.

1. Introduction

Let X and Y be real Banach spaces with $X \subset Y$ and consider the equation $G(u) = \lambda u$ for a function $G\colon X \to Y$ such that $G(0) = 0$ and a scalar $\lambda \in \mathbb{R}$. For this equation, the bifurcation points on the line of trivial solutions $\{(\lambda, 0)\colon \lambda \in \mathbb{R}\}$ are those values of λ for which $(\lambda, 0)$ is an accumulation point of non-trivial solutions. The classical results deal with the case where G is Fréchet differentiable at 0. In that case, all bifurcation points belong to the spectrum of $G'(0)$ and the basic problem is to formulate criteria ensuring that a point λ_0 in this spectrum is indeed a bifurcation point and then to obtain some information about the nature of the set of non-trivial solutions in the vicinity of $(\lambda_0, 0)$.

Motivated by situations arising in the study of nonlinear elliptic partial differential equations, some recent work [4–7, 14, 16, 17] has investigated bifurcation when G is not Fréchet differentiable at 0 but is still differentiable in the weaker sense known as Hadamard differentiability [11]. For such cases, there may be bifurcation points not belonging to the spectrum of the Hadamard derivative. A new impetus for this line of research has been generated by the treatment of asymptotic bifurcation via the method of inversion. Bound states $u \in L^2(\mathbb{R}^N)$ of the elliptic equation

$$-\Delta u + Vu + g(u) = \lambda u \tag{1.1}$$

with $V \in L^\infty(\mathbb{R}^N)$ and $g \in C^1(\mathbb{R})$ asymptotically linear in the sense that

$$\lim_{|s|\to\infty} g(s)/s = \ell \in \mathbb{R}$$

are solutions of $G(u) = \lambda u$, where

$$G = -\Delta + V + g\colon X = H^2(\mathbb{R}^N) \to Y = L^2(\mathbb{R}^N)$$

provided that $g(0) = 0$. Using the inversion $u \mapsto v = u/\|u\|_X^2$ and $G^*(v) = \|v\|_X^2 G(v/\|v\|_X^2)$, asymptotic bifurcation (also called bifurcation from infinity) for $G(u) = \lambda u$ is equivalent to bifurcation for the equation $G^*(v) = \lambda v$. Using the spaces $X = H^2(\mathbb{R}^N)$ and $Y = L^2(\mathbb{R}^N)$, it is shown in [18] that $G^*\colon X \to Y$ is Hadamard differentiable at u under the above assumptions, whereas it is Fréchet differentiable if and only if $g(s) \equiv \ell s$ for all $s \in \mathbb{R}$. Furthermore, the existing results [4, 7, 16, 17] about bifurcation for non-Fréchet differentiable problems do not apply to (1.1) because requirements concerning concavity and compactness are not satisfied. The present paper aims to fill this gap by establishing conditions which are sufficient to ensure bifurcation for problems that are only Hadamard differentiable at 0 without any assumptions of this kind and without requiring any variational structure either. The abstract results proved here concern bifurcation at isolated eigenvalues of odd algebraic multiplicity of the Hadamard derivative and they are used to establish asymptotic bifurcation for (1.1) in [19]. Further applications of the present work are presented in [20], where we deal with the same kind of elliptic equations as in [7, 16, 17] but under hypotheses which do not imply the concavity and compactness properties required in those earlier treatments.

Rather than restricting our attention to the special form $G(u) = \lambda u$, it is preferable to treat the issue of bifurcation for the general case $F(\lambda, u) = 0$, where $F\colon \mathbb{R} \times X \to Y$ with $F(\lambda, 0) = 0$ and $F(\lambda, \cdot)\colon X \to Y$ Hadamard differentiable at 0 for all $\lambda \in \mathbb{R}$. In this context, the restriction $X \subset Y$ is no longer relevant. Since Hadamard and Fréchet differentiability are equivalent in finite dimensions, it is natural to try to reduce the problem of bifurcation for $F(\lambda, u) = 0$ to an equivalent equation in finite dimensions. To obtain sharp results by this approach, the finite-dimensional subspace of X must be chosen with some care and for this we introduce and explore a notion which we call the essential conditioning number $\gamma(D_u F(\lambda, 0))$ of the partial Hadamard derivative $D_u F(\lambda, 0)\colon X \to Y$ at points λ, where this is a Fredholm operator of index zero. Our approach covers bifurcation at points λ_0 such that

$$\gamma(D_u F(\lambda_0, 0)) L_X(R, \lambda_0) < 1, \tag{1.2}$$

where $L_X(R, \lambda_0)$ is the partial Lipschitz modulus of the remainder $R(\lambda, u) = F(\lambda, u) - D_u F(\lambda, 0)u$ at λ_0. Although it may appear that (1.2) is simply a technical device used to ensure a satisfactory reduction to finite dimensions, it plays a more fundamental role. Indeed, we provide examples showing that our conclusions about bifurcation can fail when this condition is not satisfied but all the other hypotheses of our results hold.

For the general form $F(\lambda, u) = 0$ the notion of isolated eigenvalue is replaced by that of isolated singular point λ_0 of the path of partial derivatives $D_u F(\lambda, 0)$. That is to say, $D_u F(\lambda_0, 0)\colon X \to Y$ is not necessarily an isomorphism, but, for some $\varepsilon > 0$, $D_u F(\lambda, 0)$ is an isomorphism for $0 < |\lambda - \lambda_0| < \varepsilon$. In the smooth case, there are several ways of formulating sufficient conditions for bifurcation in this context. An elegant approach, taking a global perspective, was laid down and perfected by Fitzpatrick and Pejsachowicz in [8–10] through their notion of parity of a path of Fredholm operators. Using a parametrix, this is ultimately based on the theory of compact perturbations of the identity. Much later, Benevieri *et al.*, in [1] and related papers, formulated similar results about bifurcation by using only the determinant of finite-rank perturbations of the identity without reference to an index like the parity. Since our concept of essential conditioning number is based on finite-rank perturbations of the identity it is close to the spirit of the work by Benevieri *et al.*, but we feel that the parity, in a local version sufficient for our context, is the appropriate way of expressing the criteria for bifurcation. Therefore, we present an abridged version of local parity across an isolated singular point of a path of Fredholm operators of index zero using only finite-rank perturbations of the identity, which fits in nicely with the quantities used for the essential conditioning number. Of course, everything we need concerning the local parity can be deduced from [1, 8, 10], but we give short self-contained proofs in Appendix A.

The rest of this paper is organized as follows. In § 2, we recall for convenience several types of derivative, including the Hadamard derivative, for mappings between Banach spaces. Section 3 is devoted to the Lipschitz modulus of functions, and we also establish the Hadamard differentiability of an implicit function defined using a parameter-dependent contraction mapping. Bifurcation and the more restrictive phenomenon which we call continuous bifurcation are defined in § 4, where we also formulate some results concerning bifurcation in finite dimensions, which will be used later. As mentioned above, in passing to infinite dimensions, extensive use is made of Fredholm operators of index zero and this is the subject of § 5. First we define a quantity called the essential conditioning number using regular pseudo-inverses. Then paths of regular pseudo-indices are used to define the local parity of a path of Fredholm operators across an isolated singular point. Section 6 begins with a list of the hypotheses which, in addition to (1.2), are used in our results. Using the reduction to finite dimensions set out in § 6.1, the main result concerning bifurcation is theorem 6.3. Part (i) gives conditions that are sufficient to ensure that continuous bifurcation occurs at an isolated singular point λ_0. Part (ii) gives conditions implying that λ_0 is not even a bifurcation point. Let us emphasize that in the context of Hadamard differentiability, bifurcation can occur at points where $D_uF(\lambda_0, 0)\colon X \to Y$ is an isomorphism and we give an example showing this at the end of § 6. In the case where $\ker D_uF(\lambda_0, 0)$ is one dimensional and a transversality condition is satisfied, we can obtain additional information about the form of solutions near $(\lambda_0, 0)$ and this is done in § 6.3. When X and Y are Hilbert spaces and $D_uF(\lambda_0, 0)$ is self-adjoint, our hypotheses can be expressed in a more intuitive form that is important for applications. This is explored in § 6.4, where we also return to the special case $F(\lambda, u) = G(u) - \lambda u$ in corollary 6.11, whose statement involves only well-known notions and standard terminology, rather than the more refined concepts required for the more general cases.

2. Various types of derivative

For a function $M\colon X \to Y$ between Banach spaces, there is a useful notion of differentiability, attributed by Fréchet to Hadamard, which is weaker than Fréchet differentiability, but equivalent to it when $\dim X < \infty$.

https://doi.org/10.1017/S0308210513000486

Definition 2.1. A function $M\colon X \to Y$ is Hadamard differentiable at $u \in X$ if there exists $T \in B(X, Y)$ such that

$$\left\| \frac{M(u + t_n v_n) - M(u)}{t_n} - Tv \right\| \to 0 \qquad \left(\iff \frac{\|M(u + t_n v_n) - M(u) - T(t_n v_n)\|}{|t_n|} \to 0 \right)$$

for all sequences $\{t_n\} \subset \mathbb{R} \setminus \{0\}$ and $\{v_n\} \subset X$ such that $t_n \to 0$ and $\|v_n - v\| \to 0$ for some $v \in X$.

In this case, T is unique and is denoted by $M'(u)$, the Hadamard derivative of M at u. Hadamard differentiability is discussed at length in [11, ch. 4]. See also [7] for some comments particularly relevant to our work on bifurcation. It is worth noting that M is Hadamard differentiable at $u \in X$ whenever it is Gâteaux differentiable at u and there is an open neighbourhood of u in X on which it is Lipschitz continuous. Here we refer to Gâteaux differentiability in the following form.

A function $M\colon X \to Y$ is Gâteaux differentiable at $u \in X$ (strong form) if there exists $T \in B(X, Y)$ such that

$$\lim_{t \to 0} \left\| \frac{M(u + tv) - M(u)}{t} - Tv \right\| = 0 \quad \text{for every } v \in X.$$

Then $M'(u) = T$.

For convenience, we now recall the usual notion of differentiability used in most of bifurcation theory.

A function $M\colon X \to Y$ is Fréchet differentiable at $u \in X$ if there exists $T \in B(X, Y)$ such that

$$\lim_{\|v\| \to 0} \frac{\|M(u + v) - M(u) - Tv\|}{\|v\|} = 0.$$

Clearly, Fréchet differentiability implies Hadamard differentiability, which implies Gâteaux differentiability. Finally, we recall a strengthened form of Fréchet differentiability which was used by Bartle and Graves to avoid requiring continuous differentiability in the inverse and implicit function theorems.

A function $M\colon X \to Y$ is strictly Fréchet differentiable at u if there exists $T \in B(X, Y)$ such that

$$\lim_{\substack{v, w \to u \\ v \neq w}} \frac{\|M(v) - M(w) - T(v - w)\|}{\|v - w\|} = 0.$$

This implies that M is Fréchet differentiable at u with $M'(u) = T$.

3. Lipschitz mappings

The Lipschitz modulus provides a quantity which will be used to formulate a fundamental restriction for our discussion of bifurcation.

https://doi.org/10.1017/S0308210513000486

3.1. The Lipschitz modulus

Consider a function $G\colon X \to Y$, where X and Y are real Banach spaces. The quantity $L(G)$ defined by

$$L(G) = \lim_{\delta\to 0} \sup_{\substack{u,v\in B(0,\delta)\\ u\neq v}} \frac{\|G(u)-G(v)\|}{\|u-v\|} = \limsup_{\substack{u,v\to 0\\ u\neq v}} \frac{\|G(u)-G(v)\|}{\|u-v\|}$$

is called the *Lipschitz modulus of* G *at* 0. Here $B(0,\delta) = \{u \in X\colon \|u\| < \delta\}$. Its properties are discussed thoroughly in [2]. The following observations are relevant for the context in which we use this concept.

(1) $L(G) < \infty$ iff G is Lipschitz continuous on some open neighbourhood of 0. Hence, $L(G) < \infty$ does not imply even Gâteaux differentiability of G at 0.

(2) However, $L(G) = 0$ iff G is strictly Fréchet differentiable at 0 with $G'(0) = 0$. Thus, G is strictly Fréchet differentiable at 0 iff there exists $T \in B(X,Y)$ such that $L(G-T) = 0$.

(3) If $F \in C^1(U,Y)$ for some open neighbourhood U of 0 in X, then $L(R) = 0$, where $R\colon X \to Y$ is the remainder, $R(u) = F(u) - \{F(0) + F'(0)u\}$. Also F is strictly differentiable at 0 and $L(F) = \|F'(0)\|$.

(4) G strictly Fréchet differentiable at 0 implies that $L(G) = \|G'(0)\|$. But G being Fréchet differentiable at 0 with $L(G) < \infty$ does not imply that G is strictly Fréchet differentiable at 0. Also, G being Fréchet differentiable at 0 with $G'(0) = 0$ does not imply that $L(G) < \infty$.

Example 3.1.

(i) The function $G\colon \mathbb{R} \to \mathbb{R}$ defined by $G(0) = 0$ and $G(t) = t^2 \sin(1/t)^2$ for $t \neq 0$ is differentiable at all points of $\mathbb{R}$ with $G'(0) = 0$ but $L(G) = \infty$. Consider $|G(u_n) - G(v_n)|/|u_n - v_n|$ for the sequences

$$u_n = \frac{1}{\sqrt{n\pi}} \quad \text{and} \quad v_n = \frac{1}{\sqrt{(n+\frac{1}{2})\pi}}.$$

Note that, for $t \neq 0$, $G'(t) = 2t\sin(1/t^2) - (2/t)\cos(1/t^2)$ so $G'(t)$ is unbounded as $t \to 0$.

(ii) The function $G\colon \mathbb{R} \to \mathbb{R}$ defined by $G(0) = 0$ and $G(t) = t^2 \sin(1/t)$ for $t \neq 0$ is differentiable at all points of $\mathbb{R}$ with $G'(0) = 0$ but $0 < L(G) \leqslant 1$. Note that, for $t \neq 0$, $G'(t) = 2t\sin(1/t) - \cos(1/t)$ so $G'(t)$ is bounded as $t \to 0$. For $u < v$, there exists $w \in (u,v)$ such that $G(u) - G(v) = G'(w)(u-v)$ and so $L(G) \leqslant \limsup_{w\to 0} |G'(w)| = 1$. Considering $|G(u_n) - G(v_n)|/|u_n - v_n|$ for the sequences $u_n = 1/n\pi$ and $v_n = 1/(n+\frac{1}{2})\pi$, we see that $L(G) \geqslant 2/\pi$, and so G is not strictly Fréchet differentiable at 0.

3.1.1. *The inverse function theorems of Bartle and Graves*

The implicit function theorem is a standard tool in bifurcation theory concerning C^1-mappings. With less smoothness it can fail, but strict Fréchet differentiability at a point is sufficient. It is enough to consider here the simpler setting of the inverse function theorem. Let X and Y be real Banach spaces and consider a mapping $F\colon X \to Y$ with $F(0) = 0$ and F strictly Fréchet differentiable at 0.

(1) The Bartle–Graves inverse function theorem (see, for example, [2, 5G3]) states that if $F'(0)X = Y$, there exist an open neighbourhood V of 0 in Y, a function $G \in C(V, X)$ and a constant $c > 0$ such that

$$F(G(y)) = y \quad \text{and} \quad \|G(y)\| \leqslant c\|y\| \quad \text{for all } y \in V.$$

(2) If $F'(0)\colon X \to Y$ is an isomorphism, then we also have that G is strictly Fréchet differentiable at 0 with $G'(0) = F'(0)^{-1}$. If X and Y are Hilbert spaces and $F'(0)X = Y$, we again have that G is strictly Fréchet differentiable at 0 with $G'(0) = A^*(AA^*)^{-1}$, where $A = F'(0)$. However, for Banach spaces we may have $L(G) = \infty$ and so it cannot be strictly differentiable (see [2, 5G1,5G2]).

As mentioned in § 1, we are concerned with cases where F is not strictly Fréchet differentiable, and the following example shows that the inverse function theorem can fail even when F is Fréchet differentiable at 0 and $L(G) < \infty$.

EXAMPLE 3.2. Consider $F\colon \mathbb{R} \to \mathbb{R}$ defined by $F(0) = 0$ and $F(t) = t - 2t^2 \sin(1/t)$ for $t \neq 0$. Clearly, however, F is differentiable on $\mathbb{R}$ with $F'(0) = 1$ and, for $t \neq 0$, $F'(t) = 1 - 4t\sin(1/t) + 2\cos(1/t)$. As in example 4.2(ii), we find that F is not strictly differentiable at 0 but $L(F) \leqslant \limsup_{t\to 0} |F'(t)| = 3$. For $t_n = 1/(n\pi)$ we get $F'(t_n) = 1 + 2(-1)^n = 3$ for n even, -1 for n odd, showing that F' changes sign in every neighbourhood of 0. Hence, there is no open neighbourhood of 0 on which a continuous inverse function can be defined.

3.1.2. *A partial Lipschitz modulus*

Consider a function $G\colon \mathbb{R} \times X \to Y$, where X and Y are real Banach spaces. For $\mu \in \mathbb{R}$, the quantity $L_X(G, \mu)$ defined by

$$\begin{aligned} L_X(G,\mu) &= \lim_{\delta\to 0} \sup_{\substack{(\lambda,u),(\lambda,v)\in B((\mu,0),\delta) \\ u\neq v}} \frac{\|G(\lambda,u) - G(\lambda,v)\|}{\|u-v\|} \\ &= \limsup_{\substack{(\lambda,u),(\lambda,v)\to(\mu,0) \\ u\neq v}} \frac{\|G(\lambda,u) - G(\lambda,v)\|}{\|u-v\|} \end{aligned}$$

is called the partial uniform Lipschitz modulus of G with respect to X at $(\mu, 0)$. Here $B((\mu, 0), \delta)$ denotes the open ball in $\mathbb{R} \times X$ with centre $(\mu, 0)$ and radius δ. Again, we refer the reader to [2] for a discussion of this quantity.

(1) $L_X(G,\mu) < \infty$ iff there exist $k > 0$ and $\delta > 0$ such that

$$\|G(\lambda,u) - G(\lambda,v)\| \leqslant k\|u-v\| \quad \text{for all } |\lambda-\mu| < \delta \text{ and } u,v \in B(0,\delta).$$

(2) $L_X(G,\mu) = 0$ implies that $G(\mu,\cdot)\colon X \to Y$ is Fréchet differentiable at 0 with $D_xG(\mu,0) = 0$. In fact, $G(\mu,\cdot)$ is strictly Fréchet differentiable at 0 with derivative 0. More generally, if there exists $L \in B(X,Y)$ such that $L_X(G-L,\mu) = 0$, then $G(\mu,\cdot)\colon X \to Y$ is strictly Fréchet differentiable at 0 with $D_xG(\mu,0) = L$.

(3) If $G \in C^1(U,Y)$ for some open neighbourhood U of $(\mu,0)$ in $\mathbb{R} \times X$, then $L_X(G,\mu) = 0$.

The following more general result, which only requires continuity of the partial derivative at the point $(\mu,0)$, will be useful.

LEMMA 3.3. *Suppose that there exists an open neighbourhood, U, of $(\mu,0)$ in $\mathbb{R}\times X$ such that $G(\lambda,\cdot)$ is continuous and Gâteaux differentiable at u for all $(\lambda,u) \in U$ and that $\|D_xG(\lambda,u) - D_xG(\mu,0)\|_{B(X,Y)} \to 0$ as $(\lambda,u) \to (\mu,0)$. Then*

$$L_X(R,\mu) = 0, \quad \textit{where } R(\lambda,u) = G(\lambda,u) - D_xG(\lambda,0)u.$$

Proof. We can suppose that U is convex and then, for $(\lambda,u),(\lambda,v) \in U$ and $t \in [0,1]$, we have that $\phi(t) = R(\lambda, tu+(1-t)v)$ is continuous on $[0,1]$ and differentiable with $\phi'(t) = \{D_xG(\lambda,tu+(1-t)v) - D_xG(\lambda,0)\}(u-v)$. By [11, § 1.6.3, corollary], there exists $t_{(\lambda,u,v)} \in (0,1)$ such that

$$\|\phi(1) - \phi(0)\| \leqslant \|\phi'(t_{(\lambda,u,v)})\|.$$

Setting $w(\lambda,u,v) = t_{(\lambda,u,v)}u + (1-t_{(\lambda,u,v)})v$, this yields

$$\|R(\lambda,u) - R(\lambda,v)\| \leqslant \|D_xG(\lambda,w(\lambda,u,v)) - D_xG(\lambda,0)\|\|u-v\|,$$

where $\|w(\lambda,u,v)\| \to 0$, and hence $\|D_xG(\lambda,w(\lambda,u,v)) - D_xG(\lambda,0)\| \to 0$ as $(\lambda,u),(\lambda,v) \to (\mu,0)$. Hence, $L_X(R,\mu) = 0$. □

3.2. Parametrized contraction mappings

In this section, we make precise the way in which Hadamard differentiability is inherited by a function defined implicitly through the fixed points of a parameter-dependent contraction mapping. A result of this kind appears in some unpublished notes by Henry [12].

PROPOSITION 3.4. *Let X and Y be real Banach spaces, ω an open subset of X and M a closed subset of Y. Consider $C\colon \omega \times M \to M$ having the property that there exists $k \in [0,1)$ such that*

$$\|C(a,u) - C(a,v)\| \leqslant k\|u-v\| \quad \textit{for all } a \in \omega \textit{ and } u,v \in M.$$

https://doi.org/10.1017/S0308210513000486

(i) *For every* $a \in \omega$ *there exists a unique element* $\psi(a)$ *in* M *such that* $\psi(a) = C(a, \psi(a))$.

(ii) *If* $C \in C(\omega \times M, M)$, *then* $\psi \in C(\omega, M)$.

Suppose now that $X = \mathbb{R} \times E$, $\omega = J \times U$ *and* $M = \bar{N}$, *where* E *is a real Banach space,* J *is an open interval and* U, N *are open subsets of* E *and* Y, *respectively.*

(iii) *If* $(\lambda, v, \psi(\lambda, v)) \in \omega \times N$ *and* $C(\lambda, \cdot)$ *is Hadamard differentiable at* $(v, \psi(\lambda, v))$, *then* $\psi(\lambda, \cdot)$ *is Hadamard differentiable at* v *and*

$$D_E\psi(\lambda, v)z = D_E C(\lambda, v, \psi(\lambda, v))z + D_Y C(\lambda, v, \psi(\lambda, v)) D_E \psi(\lambda, v)z \quad \text{for all } z \in E.$$

Furthermore, $\|D_Y C(\lambda, v, \psi(\lambda, v))\| \leqslant k$ *and so* $I - D_Y C(\lambda, v, \psi(\lambda, v)) \colon Y \to Y$ *is an isomorphism.*

(iv) *If* $(\lambda_0, 0, 0) \in J \times U \times N$, $C(\lambda_0, 0, 0) = 0$ *and* $L_{E\times Y}(C, \lambda_0) < \infty$, *then* $\psi(\lambda_0, 0) = 0$ *and* $L_E(\psi, \lambda_0) \leqslant L_{E\times Y}(C, \lambda_0)/(1 - k)$.

Proof. (i) M is a complete metric space and $C(a, \cdot) \colon M \to M$ is a contraction for every $a \in \omega$.

(ii) For $a, b \in \omega$,

$$\begin{aligned}\|\psi(a) - \psi(b)\| &= \|C(a, \psi(a)) - C(b, \psi(b))\| \\ &\leqslant \|C(a, \psi(a)) - C(b, \psi(a))\| + \|C(b, \psi(a)) - C(b, \psi(b))\| \\ &\leqslant \|C(a, \psi(a)) - C(b, \psi(a))\| + k\|\psi(a) - \psi(b)\|\end{aligned}$$

and so $\|\psi(a) - \psi(b)\| \leqslant (1/(1 - k))\|C(a, \psi(a)) - C(b, \psi(a))\|$. The continuity of ψ at a now follows from the continuity of C at $(a, \psi(a))$.

(iii) For any $y \in Y$,

$$D_Y C(\lambda, v, \psi(\lambda, v))y = \lim_{t\to 0} \frac{C(\lambda, v, \psi(\lambda, v) + ty) - C(\lambda, v, \psi(\lambda, v))}{t}$$

and $\|C(\lambda, v, \psi(\lambda, v)+ty)-C(\lambda, v, \psi(\lambda, v))\| \leqslant k\|ty\|$. Hence, $\|D_Y C(\lambda, v, \psi(\lambda, v))\| \leqslant k$.

For the Hadamard differentiability of $\psi(\lambda, \cdot)$ at v, consider sequences $\{t_n\} \subset \mathbb{R} \setminus \{0\}$ and $\{z_n\} \subset E$ such that $t_n \to 0$ and $\|z_n - z\| \to 0$ for some $z \in E$. We should show that there exists $L \in B(E, Y)$ such that

$$\frac{\psi(\lambda, v + t_n z_n) - \psi(\lambda, v)}{t_n} \to Lz$$

or, equivalently, that

$$\frac{\psi(\lambda, v + t_n z_n) - \psi(\lambda, v) - t_n L(z_n)}{t_n} \to 0 \text{ in } Y,$$

since $Lz_n \to Lz$.

Let $L_E = D_E C(\lambda, v, \psi(\lambda, v))$, $L_Y = D_Y C(\lambda, v, \psi(\lambda, v))$ and $L = (I_Y - L_Y)^{-1} L_E$. Then $L \in B(E, Y)$ and

$$(I - L_Y)L = L_E, \quad \text{so } L = L_E + L_Y L.$$

Now,

$$\begin{aligned}
&\psi(\lambda, v + t_n z_n) - \psi(\lambda, v) - t_n L z_n \\
&\quad = C(\lambda, v + t_n z_n, \psi(\lambda, v + t_n z_n)) - C(\lambda, v, \psi(\lambda, v)) - t_n\{L_E z_n + L_Y L z_n\} \\
&\quad = C(\lambda, v + t_n z_n, \psi(\lambda, v + t_n z_n)) - C(\lambda, v + t_n z_n, \psi(\lambda, v) + t_n L z_n) \\
&\qquad\quad + C(\lambda, v + t_n z_n, \psi(\lambda, v) + t_n L z_n) - C(\lambda, v, \psi(\lambda, v)) - t_n\{L_E z_n + L_Y L z_n\},
\end{aligned}$$

where

$$\begin{aligned}
&\|C(\lambda, v + t_n z_n, \psi(\lambda, v + t_n z_n)) - C(\lambda, v + t_n z_n, \psi(\lambda, v) + t_n L z_n)\| \\
&\qquad\qquad\qquad \leqslant k\|\psi(\lambda, v + t_n z_n) - \psi(\lambda, v) - t_n L z_n\|
\end{aligned}$$

and so

$$\begin{aligned}
&\|\psi(\lambda, v + t_n z_n) - \psi(\lambda, v) - t_n L z_n\| \\
&\qquad \leqslant \frac{1}{1-k}\|C(\lambda, v + t_n z_n, \psi(\lambda, v) + t_n L z_n) \\
&\qquad\qquad - C(\lambda, v, \psi(\lambda, v)) - t_n\{L_E z_n + L_Y L z_n\}\|.
\end{aligned}$$

But,

$$\begin{aligned}
&C(\lambda, v + t_n z_n, \psi(\lambda, v) + t_n L z_n) - C(\lambda, v, \psi(\lambda, v)) - t_n\{L_E z_n + L_Y L z_n\} \\
&\quad = C(\lambda, (v, \psi(\lambda, v)) + t_n(z_n, L z_n)) - C(\lambda, v, \psi(\lambda, v)) \\
&\qquad - t_n D_{E\times Y} C(\lambda, v, \psi(\lambda, v))(z_n, L z_n)
\end{aligned}$$

and so

$$\frac{C(\lambda, v + t_n z_n, \psi(\lambda, v) + t_n L z_n) - C(\lambda, v, \psi(\lambda, v)) - t_n\{L_E z_n + L_Y L z_n\}}{t_n} \to 0$$

by the Hadamard differentiability of $C(\lambda, \cdot)$ at $(v, \psi(\lambda, v))$. This proves that $\psi(\lambda, \cdot)$ is Hadamard differentiable at v with $D_E \psi(\lambda, v) = L$.

(iv) Since 0 is a fixed point of $C(\lambda_0, 0, \cdot)$, $\psi(\lambda_0, 0) = 0$. For $\lambda \in J$ and $v, z \in U$,

$$\|\psi(\lambda, v) - \psi(\lambda, z)\| \leqslant \frac{1}{1-k}\|C(\lambda, v, \psi(\lambda, v)) - C(\lambda, z, \psi(\lambda, v))\|$$

as in part (ii). Given $\varepsilon > 0$, there exists an open ball B with centre $(\lambda_0, 0)$ in $X \times Y$ such that $\|C(\lambda, p) - C(\lambda, q)\| \leqslant [L_{E\times Y}(C, \lambda_0) + \varepsilon]\|p - q\|$ for all $(\lambda, p), (\lambda, q) \in B$. By part (ii), there exists $\delta > 0$ such that $(\lambda, v, \psi(\lambda, v)), (\lambda, z, \psi(\lambda, v)) \in B$ for $|\lambda - \lambda_0| < \delta$ and $v, z \in B(0, \delta) \subset E$ and, consequently,

$$\|\psi(\lambda, v) - \psi(\lambda, z)\| \leqslant \frac{1}{1-k}[L_{E\times Y}(C, \lambda_0) + \varepsilon]\|v - z\|.$$

□

https://doi.org/10.1017/S0308210513000486

4. Bifurcation and continuous bifurcation

We begin by formulating the basic definitions concerning bifurcation points, which will be used. Then these notions are explored in the case where the underlying Banach space has finite dimension, but avoiding assumptions involving continuous differentiability.

Notation

For real Banach spaces X and Y,

- $B(X, Y) = \{L\colon X \to Y \mid L \text{ is linear and bounded}\}$,
- $\mathrm{Iso}(X, Y) = \{L \in B(X, Y) \mid L\colon X \to Y \text{ is an isomorphism}\}$,
- for $L \in B(X, Y)$, $\ker L = \{u \in X\colon Lu = 0\}$ and $\operatorname{rge} L = LX$.

4.1. Bifurcation and continuous bifurcation

Let X and Y be real Banach spaces and consider a function $F\colon \Omega \to Y$, where Ω is an open subset of $\mathbb{R} \times X$ and $F(\lambda, 0) = 0$ for all $(\lambda, 0) \in \Omega$. If $(\lambda_0, 0) \in \Omega$, $\lambda_0 \in \mathbb{R}$ is called a *bifurcation point* for the equation $F(\lambda, 0) = 0$, provided that there exists a sequence

$$\{(\lambda_n, u_n)\} \subset \mathcal{S} \equiv \{(\lambda, u) \in \Omega\colon F(\lambda, u) = 0 \text{ and } u \neq 0\}$$

such that $\lambda_n \to \lambda_0$ and $\|u_n\| \to 0$ as $n \to \infty$. There is *continuous bifurcation* at λ_0 if there is a connected subset $\mathcal{C}$ of $\mathcal{S}$ such that $\bar{\mathcal{C}} \cap [\mathbb{R} \times \{0\}] = \{(\lambda_0, 0)\}$. In this case, $\bar{\mathcal{C}}$ is a connected subset of $\mathcal{S}' \equiv \mathcal{S} \cup \{(\lambda_0, 0)\}$ but, even if $\mathcal{C}$ is a maximal connected subset of $\mathcal{S}$, it is not necessarily the connected component $\mathcal{C}'$ of $\mathcal{S}'$ containing $\{(\lambda_0, 0)\}$ since, in general, $\mathcal{C}' \setminus \{(\lambda_0, 0)\}$ is not a connected subset of $\mathcal{S}$. We have chosen to deal with $\mathcal{C}$ rather than $\mathcal{C}'$ in this work because it has the advantage that, in applications to asymptotic bifurcation via inversion [19], its image under inversion is connected.

For $\lambda \in \mathbb{R}$, $\Omega_\lambda = \{u \in X\colon (\lambda, u) \in \Omega\}$. If $F(\lambda, \cdot)\colon \Omega_\lambda \to Y$ is Fréchet differentiable at 0 for all $(\lambda, 0) \in \Omega$ with Fréchet partial derivative $D_x F(\cdot, 0) \in B(X, Y)$ continuous at λ_0 and, for some $\varepsilon > 0$, $D_x F(\lambda, 0) \in \mathrm{Iso}(X, Y)$ for $0 < |\lambda - \lambda_0| < \varepsilon$, then λ_0 is said to be an *isolated singular point* of the path $\lambda \mapsto D_x F(\lambda, 0)$. Note that $D_x F(\lambda_0, 0)\colon X \to Y$ may or may not be an isomorphism. If $X = Y$ with $\dim X < \infty$ and λ_0 is an isolated singular point of $D_x F(\cdot, 0)$, there are only two cases:

Case 1. $\det D_x F(\lambda, 0) \det D_x F(\mu, 0) > 0$ for $\lambda_0 - \varepsilon < \lambda < \lambda_0 < \mu < \lambda_0 + \varepsilon$.

Case 2. $\det D_x F(\lambda, 0) \det D_x F(\mu, 0) < 0$ for $\lambda_0 - \varepsilon < \lambda < \lambda_0 < \mu < \lambda_0 + \varepsilon$.

In case 2, $\det D_x F(\lambda, 0)$ changes sign as λ crosses λ_0. The local parity across λ_0, which is introduced in § 5.2, provides an extension of this dichotomy to the infinite-dimensional setting.

In order to formulate hypotheses ensuring that $D_x F(\lambda_0, 0) \notin \mathrm{Iso}(X, Y)$ is a necessary condition for bifurcation at λ_0, without requiring F to be C^1 in a neighbourhood of $(\lambda_0, 0)$, we introduce the following condition.

https://doi.org/10.1017/S0308210513000486

Consider a function $F\colon \Omega \to Y$, where X and Y are real Banach spaces and Ω is an open subset of $\mathbb{R}\times X$. For a point $(\lambda, 0) \in \Omega$ and $\delta > 0$ such that $B((\lambda,0),\delta) \subset \Omega$ set

$$\Delta_\delta(F,\lambda) = \sup_{\substack{(\mu,u)\in B((\lambda,0),\delta)\\ u\neq 0}} \frac{\|F(\mu,u)-F(\lambda,u)\|}{\|u\|}. \tag{4.1}$$

(C3) $\lim_{\delta\to 0} \Delta_\delta(F,\lambda) = 0$.

Noting that in example 4.3 $F(\lambda, 2\lambda^{1/2})/\lambda = -9/17$ for all $\lambda > 0$ and $F(0,0) = 0$, we see that $\Delta_\delta(F,0) \geqslant 9/17$ for all $\delta > 0$ and so (C3) is not satisfied at $\lambda_0 = 0$ in that example. However, there are several different situations in which (C3) does hold.

(1) If $F \in C^1(\Omega, Y)$, condition (C3) is satisfied at every $(\lambda,0) \in \Omega$. More generally, if D_xF exists on Ω and $D_xF \in C(\Omega, B(X,Y))$, then (C3) holds at every $(\lambda, 0) \in \Omega$ since, for $(\mu,u) \in B((\lambda,0),\delta) \subset \Omega$,

$$\begin{aligned}\|F(\mu,u)-F(\lambda,u)\| &= \left\| \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}\{F(\mu,tu)-F(\lambda,tu)\}\,\mathrm{d}t \right\| \\ &\leqslant \|u\| \int_0^1 \|D_xF(\mu,tu)-D_xF(\lambda,tu)\|\,\mathrm{d}t,\end{aligned}$$

from which it follows easily that $\Delta_\delta(F,\lambda) \to 0$ as $\delta \to 0$.

(2) Alternatively, if $D_\lambda F$ exists on Ω and there exists $L(\lambda,u) \in B(X,Y)$ such that $D_\lambda F(\lambda,u) = L(\lambda,u)u$ for $(\lambda,u) \in \Omega$, then (C3) at λ_0 is satisfied provided that $\|L(\lambda,u)\|_{B(X,Y)}$ is bounded on some open neighbourhood of $(\lambda_0, 0)$, since for $(\lambda,u) \in B((\lambda_0,0),\delta) \subset \Omega$,

$$\|F(\lambda,u)-F(\lambda_0,u)\| \leqslant |\lambda-\lambda_0| \sup_{B((\lambda_0,0),\delta)} \|D_\lambda F\| \leqslant |\lambda-\lambda_0|\|u\| \sup_{B((\lambda_0,0),\delta)} \|L\|$$

in this case. An important special case of this situation occurs when F has the form $F(\lambda,u) = G(u) - \lambda u$ for some $G\colon V \to Y$, where V is an open subset of X.

(3) If $F(\lambda,u) = G(u) - L(\lambda)u$, where $L\colon J \subset \mathbb{R} \to B(X,Y)$, condition (C3) is satisfied at points of continuity of L.

4.2. Bifurcation in the finite-dimensional case

The first result is a more or less standard application of Brouwer degree, but uses rather weaker hypotheses than usual. We give two examples illustrating its limitations before formulating a result that is free from the deficiencies shown by the examples, but still requiring only a minimal amount of smoothness.

Proposition 4.1. *Let X be a real Banach space with $\dim X < \infty$ and Ω an open subset of $\mathbb{R}\times X$. Consider $F\colon \Omega \to X$ having the following properties.*

https://doi.org/10.1017/S0308210513000486

(C1) *$F \in C(\Omega, X)$ and $F(\lambda, 0) = 0$ for all $\lambda \in \mathbb{R}$ such that $(\lambda, 0) \in \Omega$.*

(C2) *$(\lambda_0, 0) \in \Omega$ and there exist sequences $\{\mu_n^-\}$ and $\{\mu_n^+\}$ such that*

- *$(\mu_n^{\pm}, 0) \in \Omega$ and $\mu_n^- < \lambda_0 < \mu_n^+$ for all n,*
- $\lim_{n\to\infty} \mu_n^- = \lim_{n\to\infty} \mu_n^+ = \lambda_0$,
- *$F(\mu_n^{\pm}, \cdot)\colon \Omega_{\mu_n^{\pm}} \subset X \to Y$ is Fréchet differentiable at 0 with $D_x F(\mu_n^{\pm}, 0)$ invertible,*
- $\det D_x F(\mu_n^-, 0) \det D_x F(\mu_n^+, 0) < 0$.

Then λ_0 is a bifurcation point for the equation $F(\lambda, u) = 0$.

Proof. If λ_0 is not a bifurcation point, there exists $\delta > 0$ such that $[\lambda_0 - \delta, \lambda_0 + \delta] \times B_X(0, \delta) \subset \Omega$ and $u = 0$ for all (λ, u) with $F(\lambda, u) = 0$, $|\lambda - \lambda_0| \leqslant \delta$ and $\|u\| \leqslant \delta$. In particular, the Brouwer degree $d(F(\lambda, \cdot), B_X(0, \delta), 0)$ is defined for all $|\lambda - \lambda_0| \leqslant \delta$ and it is independent of λ. Choose some n so that $|\mu_n^{\pm} - \lambda_0| \leqslant \delta$. Then $d(D_x F(\mu_n^{\pm}, 0), B_X(0, r), 0)$ is defined for all $r > 0$ and is equal to

$$\frac{\det D_x F(\mu_n^{\pm}, 0)}{|\det D_x F(\mu_n^{\pm}, 0)|}.$$

Furthermore, there exist $r_n^{\pm} \in (0, \delta]$ such that

$$d(F(\mu_n^{\pm}, \cdot), B_X(0, r_n^{\pm}), 0) = d(D_x F(\mu_n^{\pm}, 0), B_X(0, r_n^{\pm}), 0),$$

whereas $d(F(\mu_n^{\pm}, \cdot), B_X(0, r_n^{\pm}), 0) = d(F(\mu_n^{\pm}, \cdot), B_X(0, \delta), 0)$ since $F(\mu_n^{\pm}, u) \neq 0$ for $r_n^{\pm} \leqslant \|u\| \leqslant \delta$. Thus,

$$\begin{aligned} d(F(\mu_n^-, \cdot), B_X(0, \delta), 0)&d(F(\mu_n^+, \cdot), B_X(0, \delta), 0) \\ &= \frac{\det D_x F(\mu_n^-, 0)}{|\det D_x F(\mu_n^-, 0)|} \frac{\det D_x F(\mu_n^+, 0)}{|\det D_x F(\mu_n^+, 0)|} \\ &< 0, \end{aligned}$$

contradicting the earlier conclusion that $d(F(\lambda, \cdot), B_X(0, \delta), 0)$ is independent of λ in $[\lambda_0 - \delta, \lambda_0 + \delta]$. □

Since the proof of proposition 4.1 is an easy consequence of the properties of Brouwer degree, one might hope to establish continuous bifurcation under the same hypotheses by a more careful use of degree theory in a way that leads to the usual results about global bifurcation, inspired by Rabinowitz's original paper [15]. The following example shows that this cannot be done without strengthening condition (C2).

Example 4.2. For some $n \in \mathbb{N} \setminus \{0\}$, define $\alpha\colon \mathbb{R} \to \mathbb{R}$ by $\alpha(0) = 0$ and $\alpha(\lambda) = \lambda^n \sin(1/\lambda)$ for $\lambda \neq 0$. Then $\alpha \in C(\mathbb{R}) \cap C^1(\mathbb{R} \setminus \{0\})$. Furthermore, α is differentiable at 0 if and only if $n \geqslant 2$ and then $\alpha'(0) = 0$. Hence, $\alpha \in C^1(\mathbb{R})$ if and only if $n \geqslant 3$. Now consider $F\colon \mathbb{R}^2 \to \mathbb{R}$ defined by $F(\lambda, x) = x\alpha(\lambda) - x^2$. Then F and $\partial_x F \in C(\mathbb{R}^2)$ with $\partial_x F(\lambda, 0) = \alpha(\lambda)$, so the conditions (C1) and (C2) are satisfied at $\lambda_0 = 0$. However,

$$\mathcal{S} = \{(\lambda, x)\colon F(\lambda, x) = 0 \text{ and } x \neq 0\} = \{(\lambda, \alpha(\lambda))\colon \lambda \neq 0 \text{ and } 1/(\lambda\pi) \notin \mathbb{Z}\}$$

https://doi.org/10.1017/S0308210513000486

and the connected components of $\mathcal{S}$ are the sets $\{(\lambda, \alpha(\lambda)) \colon \lambda \in J\}$, where J is one of the intervals

$$\left(-\infty, -\frac{1}{\pi}\right), \left(-\frac{1}{n\pi}, -\frac{1}{(n+1)\pi}\right), \left(\frac{1}{(n+1)\pi}, \frac{1}{n\pi}\right), \left(\frac{1}{\pi}, \infty\right) \quad \text{for } n \in \mathbb{N} \setminus \{0\}.$$

No component has $(0,0)$ in its closure and therefore there is not continuous bifurcation at $\lambda_0 = 0$. Note that $F \in C^1(\mathbb{R}^2)$ for $n \geqslant 3$.

In Example 4.2, $\lambda_0 = 1/(k\pi)$ is an isolated singular point of $\partial_x F(\cdot, 0)$ for every $k \in \mathbb{Z} \setminus \{0\}$ and the second case occurs. From the explicit form of $\mathcal{S}$ we see that there is continuous bifurcation at all such points.

At an isolated singular point where the second case occurs we must have that $\det D_x F(\lambda_0, 0) = 0$, and when $F \in C^1(\Omega, X)$ it is well known that this is a necessary condition for bifurcation to occur at λ_0. However, with less smoothness of F this fails to hold.

Example 4.3. Consider $F(\lambda, x) = g(\lambda, x) - x$, where $g \colon \mathbb{R}^2 \to \mathbb{R}$ is defined by

$$g(0,0) = 0 \quad \text{and} \quad g(\lambda, x) = \frac{2\lambda x^3}{\lambda^2 + x^4} \text{ for } (\lambda, x) \neq (0,0).$$

Then $g \in C(\mathbb{R}^2)$ and $g(\lambda, x) = 0$ when $\lambda x = 0$. Hence, $\partial_\lambda g(\lambda, 0) = \partial_x g(0, x) = 0$ for all $\lambda, x \in \mathbb{R}$. Thus, $F \in C(\mathbb{R}^2)$ and $\partial_\lambda F(\lambda, x)$, $\partial_x F(\lambda, x)$ exist for all $(\lambda, x) \in \mathbb{R}^2$. In fact, $\partial_\lambda F, \partial_x F \in C(\mathbb{R}^2 \setminus \{(0,0)\})$ but are discontinuous at $(0,0)$. In particular, $\partial_x F(\lambda, 0) = -1$ for all $\lambda \in \mathbb{R}$ so $\lambda_0 = 0$ is an isolated singular point at which case 1 occurs and $\partial_x F(0,0) = -1$. However,

$$\mathcal{S} = \{(\lambda, x) \colon F(\lambda, x) = 0 \text{ and } x \neq 0\} = \{(\lambda, \pm\sqrt{\lambda}) \colon \lambda > 0\},$$

showing that continuous bifurcation occurs at $\lambda_0 = 0$.

Part (a) of the next result shows that our condition (C3) is sufficient to exclude bifurcation occurring at a point where $D_x F(\lambda_0, 0) \in \mathrm{Iso}(X, X)$. Part (b) shows that continuous bifurcation will occur at an isolated singular point in case 2, provided that (C3) is satisfied.

Proposition 4.4. *Let X be a real Banach space with $\dim X < \infty$ and Ω an open subset of $\mathbb{R} \times X$. Consider $F \colon \Omega \to X$ satisfying condition (C1).*

(a) *If (C3) is satisfied at $(\lambda_0, 0) \in \Omega$ and $F(\lambda_0, \cdot) \colon \Omega_{\lambda_0} \to X$ is Fréchet differentiable at 0 with $\det D_x F(\lambda_0, 0) \neq 0$, then λ_0 is not a bifurcation point for the equation $F(\lambda, u) = 0$.*

(b) *Suppose also that $F(\lambda, \cdot) \colon \Omega_\lambda \to Y$ is Fréchet differentiable at 0 for all $(\lambda, 0) \in \Omega$ with $\lambda \mapsto D_x F(\lambda, 0) \in B(X, Y)$ continuous at λ_0. If λ_0 is an isolated singular point and $\det D_x F(\lambda, 0)$ changes sign as λ crosses λ_0, then λ_0 is a bifurcation point. If in addition (C3) holds for all $(\lambda, 0) \in \Omega$, there is continuous bifurcation at λ_0.*

Proof. (a) Set $T = D_x F(\lambda_0, 0)$ and then $\varepsilon = 1/(4\|T^{-1}\|)$. Choose $\delta > 0$ such that $B((\lambda_0, 0), \delta) \subset \Omega$, $\Delta_\delta(F, \lambda_0) < \varepsilon$ and $\|R(u)\| \leqslant \varepsilon \|u\|$ for $\|u\| < \delta$, where

https://doi.org/10.1017/S0308210513000486

$R(u) = F(\lambda_0, u) - Tu$. Then, for any $(\lambda, u) \in B((\lambda_0, 0), \delta)$ such that $F(\lambda, u) = 0$, we have

$$\|u\| \leqslant \|T^{-1}\|\|Tu\| = \|T^{-1}\|\|F(\lambda_0, u) - R(u)\| \leqslant \tfrac{1}{2}\|u\|$$

since

$$\|F(\lambda_0, u)\| = \|F(\lambda, u) - F(\lambda_0, u)\| \leqslant \|u\|\Delta_\delta(F, \lambda_0).$$

Thus, $u = 0$ and λ_0 cannot be a bifurcation point.

(b) That λ_0 is a bifurcation point follows immediately from proposition 4.1.

Suppose now that (C3) is satisfied for all $(\lambda, 0) \in \Omega$. Choose any $\varepsilon > 0$ such that $B = \overline{B((\lambda_0, 0), \varepsilon)} \subset \Omega$ and $\det D_x F(\lambda, 0) \neq 0$ for $0 < |\lambda - \lambda_0| \leqslant \varepsilon$. By part (a), this implies that there are no bifurcation points in $[\lambda_0 - \varepsilon, \lambda_0 + \varepsilon]$ except for λ_0. Hence, for any $\rho \in (0, \varepsilon)$, there exists $r(\rho) > 0$ such that, for $\rho \leqslant |\lambda - \lambda_0| \leqslant \varepsilon$, $u = 0$ is the only solution of $F(\lambda, u) = 0$ with $\|u\| \leqslant r(\rho)$. Furthermore, for

$$\mathcal{S} = \{(\lambda, u) \in B \colon u \neq 0 \text{ and } F(\lambda, 0) = 0\} \text{ and } \mathcal{S}' = \mathcal{S} \cup \{(\lambda_0, 0)\},$$

$\mathcal{S}'$ is compact. Let $\mathcal{C}'$ be the connected component of $\mathcal{S}'$ containing $\{(\lambda_0, 0)\}$ and suppose that $\mathcal{C}' \cap \partial B = \emptyset$. Since no connected subset of $\mathcal{S}'$ intersects $\{(\lambda_0, 0)\}$ and $\partial B \cap \mathcal{S}'$, it follows from Whyburn's lemma (see [21, ch. 1, § 9.3]) that there exist compact subsets K_1 and K_2 of $\mathcal{S}'$ such that

$$\{(\lambda_0, 0)\} \subset K_1, \quad \partial B \cap \mathcal{S}' \subset K_2, \quad K_1 \cap K_2 = \emptyset, \quad K_1 \cup K_2 = \mathcal{S}'.$$

It follows that $d_1 = \operatorname{dist}(K_1, K_2) > 0$ and $d_2 = \operatorname{dist}(K_1, \partial B) > 0$. Set $\delta = \frac{1}{2}\min\{d_1, d_2\}$ and consider the open set

$$U = \{(\lambda, u) \in \mathbb{R} \times X \colon \operatorname{dist}((\lambda, u), K_1) < \delta\}.$$

We have $\bar{U} \subset B((\lambda_0, 0), \varepsilon)$, $\partial U \cap \mathcal{S} = \emptyset$ and there exists $\rho \in (0, \varepsilon/2)$ such that $(\lambda, 0) \in U$ for $0 \leqslant |\lambda - \lambda_0| \leqslant 2\rho$. Let

$$V = \{(\lambda, u) \in U \colon \|u\| > r(\rho)\}.$$

Then V is open and $F(\lambda, u) \neq 0$ for all $(\lambda, u) \in \partial V$ with $|\lambda - \lambda_0| \geqslant \rho$. It follows that

$$d(F(\lambda_0 - 2\rho, \cdot), V_{\lambda_0 - 2\rho}, 0) = d(F(\lambda_0 + 2\rho, \cdot), V_{\lambda_0 + 2\rho}, 0) = 0$$

and hence that

$$\begin{aligned} d(F(\lambda_0 \pm 2\rho, \cdot), U_{\lambda_0 \pm 2\rho}, 0) &= d(F(\lambda_0 \pm 2\rho, \cdot), B_X(0, r(\rho)), 0) \\ &= d(D_x F(\lambda_0 \pm 2\rho, \cdot), B(0, r(\rho)), 0) \\ &= \frac{\det D_x F(\lambda_0 \pm 2\rho, 0)}{|\det D_x F(\lambda_0 \pm 2\rho, 0)|}. \end{aligned}$$

But we also have that $F(\lambda, u) \neq 0$ for all $(\lambda, u) \in \partial U$ with $|\lambda - \lambda_0| \leqslant 2\rho$ and so

$$d(F(\lambda_0 - 2\rho, \cdot), U_{\lambda_0 - 2\rho}, 0) = d(F(\lambda_0 + 2\rho, \cdot), U_{\lambda_0 + 2\rho}, 0).$$

These calculations show that $\det D_x F(\lambda_0 - 2\rho, 0) \det D_x F(\lambda_0 + 2\rho, 0) > 0$ which contradicts the hypothesis that $\det D_x F(\lambda, 0)$ changes sign as λ crosses λ_0. Hence, $\mathcal{C}' \cap \partial B \neq \emptyset$. Choosing any point $(\mu, v) \in \mathcal{C}' \cap \partial B$ and noting that $\mathcal{C}'$ is a compact

https://doi.org/10.1017/S0308210513000486

connected set, it follows from [13, ch. V, § 48I, theorem 1] that there is an irreducible continuum $\mathcal{D}$ in $\mathcal{C}'$ such that $(\lambda_0, 0)$ and $(\mu, v) \in \mathcal{D}$. By [13, ch. V, § 48II, theorem 4], $\mathcal{C} \equiv \mathcal{D} \setminus \{(\lambda_0, 0), (\mu, v)\} \subset \mathcal{S}$ is connected and $(\lambda_0, 0) \in \bar{\mathcal{C}}$. Since $\bar{\mathcal{C}} \subset B$ and λ_0 is the only bifurcation point in $[\lambda_0 - \varepsilon, \lambda_0 + \varepsilon]$, we have that $\bar{\mathcal{C}} \cap [\mathbb{R} \times \{0\}] = \{(\lambda_0, 0)\}$, proving that there is continuous bifurcation at λ_0. □

5. Fredholm operators of index zero

Let X and Y be real Banach spaces. In addition to the notation introduced at the beginning of § 4, we set

- $\Phi_0(X, Y) = \{L \in B(X, Y) \colon L \text{ is a Fredholm operator of index } 0\}$,
- $F(X, Y) = \{L \in B(X, Y) \colon \dim \operatorname{rge} L < \infty\}$,
- $\Psi(X) = \{L \in B(X, X) \colon I - L \in F(X, X)\}$,
- $\operatorname{Pr}(X) = \{P \in B(X, X) \colon P^2 = P\}$.

5.1. The essential conditioning number

For $L \in B(X, Y)$, set $[L] = \{T \colon T - L \in F(X, Y)\}$ and $[L]_r = [L] \cap \operatorname{Iso}(X, Y)$.

LEMMA 5.1. *Let $L \in B(X, Y)$. Then $[L]_r \neq \emptyset \iff L \in \Phi_0(X, Y)$.*

Proof. Suppose that $T \in [L]_r$. Then there exists $A \in F(X, Y)$ such that $L + A \in \operatorname{Iso}(X, Y)$ and so $L = (L + A) - A \in \Phi_0(X, Y)$.

Conversely, suppose that $L \in \Phi_0(X, Y)$. Choose closed subspaces W and Z so that $X = \ker L \oplus W$ and $Y = Z \oplus \operatorname{rge} L$. Let Π be the projection of X onto $\ker L$ and Q the projection of Y onto $\operatorname{rge} L$. Then $\dim Z = \dim \ker L$ and $V = L|_W \in \operatorname{Iso}(W, \operatorname{rge} L)$. Choose some $M \in \operatorname{Iso}(\ker L, Z)$. It is easy to check that $L + M\Pi \in [L]_r$ with $(L + M\Pi)^{-1} = M^{-1}(I_Y - Q) + V^{-1}Q$. □

REMARK 5.2. It follows from this lemma that $\Phi_0(X, Y) \neq \emptyset$ iff X and Y are isomorphic.

For $L \in B(X, Y)$, a *regular pseudo-inverse* for L is an operator $S \in \operatorname{Iso}(Y, X)$ such that $SL \in \Psi(X)$ and $LS \in \Psi(Y)$. Note that $S \in \operatorname{Iso}(Y, X)$ is a regular pseudo-inverse for $L \iff SL \in \Psi(X) \iff LS \in \Psi(Y)$.

LEMMA 5.3. *Let $L \in B(X, Y)$.*

(i) *S is a regular pseudo-inverse for $L \iff S^{-1} \in [L]_r$.*

(ii) *If S is a regular pseudo-inverse for L, $[S]_r$ is the set of all regular pseudo-inverses of L.*

Proof. (i) If S is a regular pseudo-inverse for L, $S \in \operatorname{Iso}(Y, X)$ and there exists $A \in F(X, X)$ such that $SL = I_X + A$. Hence, $S^{-1} = L - S^{-1}A \in [L]_r$.

Conversely, if $S = T^{-1}$ for some $T \in [L]_r$, there exists $A \in F(X, X)$ such that $L + A \in \operatorname{Iso}(X, Y)$ and $S = (L + A)^{-1}$. Then $SL = I_X - SA \in \Psi(X)$.

(ii) If $T \in [S]_r$, then $T \in \text{Iso}(Y,X)$ and there exists $C \in F(Y,X)$ such that $T = S + C$. Hence, $CL \in F(X,X)$ and $LC \in F(Y,Y)$, so that $TL = SL + CL \in \Psi(X)$, showing the T is also a regular pseudo-inverse for L.

Conversely, if S and T are both pseudo-inverses for L, then $S,T \in \text{Iso}(Y,X)$ and there exist $A \in F(X,X)$ and $B \in F(Y,Y)$ such that $SL = I_X + A$ and $LT = I_Y + B$. Therefore, $SLT = T + AT = S + SB$ and so $T - S = SB - AT \in F(Y,X)$, showing that $T \in [S]_r$. □

It follows from these lemmas that $L \in B(X,Y)$ has a regular pseudo-inverse if and only if $L \in \Phi_0(X,Y)$. Furthermore, $[L]_r^{-1} \equiv \{T^{-1} \colon T \in [L]_r\}$ is the set of all regular pseudo-inverses of L.

For $L \in \Phi_0(X,Y)$, the quantity

$$\begin{aligned}\gamma(L)&\\ &= \inf\{\|(I_X - P)S\|_{B(Y,X)} \colon S \in [L]_r^{-1},\ P \in \text{Pr}(X) \text{ with } \operatorname{rge} P = \operatorname{rge}(I_X - SL)\}\\ &= \inf\{\|(I_X - P)T^{-1}\| \colon T \in [L]_r,\ P \in \text{Pr}(X) \text{ with } \operatorname{rge} P = \operatorname{rge}(I_X - T^{-1}L)\}\end{aligned}$$

will be referred to as the *essential conditioning number* of L.

For any $L \in \Phi_0(X,Y)$, the operator $S_0 = (L + M\Pi)^{-1} = M^{-1}(I_Y - Q) + V^{-1}Q$ defined in the proof of lemma 5.1 is a regular pseudo-inverse for L since $S_0 L = I_X - \Pi$. Hence, $\operatorname{rge}(I_X - S_0 L) = \ker L$ and $\gamma(L) \leqslant \|(I_X - \Pi)S_0\|$. But $(I_X - \Pi)S_0 = V^{-1}Q$ and we obtain the estimate $\gamma(L) \leqslant \|V^{-1}Q\|$. However, it is clear from the following discussion that this estimate is rarely optimal.

Consider first the case where either X or Y has finite dimension. Then for $\Phi_0(X,Y) \neq \emptyset$ we must have $\dim X = \dim Y$. In this case, $B(X,Y) = \Phi_0(X,Y) = F(X,Y)$ and, for any $L \in B(X,Y)$, $[L] = B(X,Y)$ and $[L]_r = \text{Iso}(X,Y)$. Let $J \in \text{Iso}(Y,X)$ and note that $S_n = J/n$ is a regular pseudo-inverse of L for all $n \neq 0$. For $n > \|JL\|$, we have that $\|S_n L\| < 1$ and so $I_X - S_n L \in \text{Iso}(X,X)$. Then I_X is the only projection onto $\operatorname{rge}(I_X - S_n L)$ and $\gamma(L) = 0$.

In the general situation, $\|(I_X - P)S\| \leqslant \|I_X - P\|\|S\|$, but in the absence of a bound for $\dim \operatorname{rge}(I_X - SL)$ as S ranges over $[L]_r^{-1}$, we do not have an upper bound for $\|I_X - P\|$. An exception is the case where X (and hence Y) is isomorphic to a Hilbert space, since in this case we can use an orthogonal projection. Hence, when X is isomorphic to a Hilbert space,

$$\gamma(L) \leqslant \inf_{S \in [L]_r^{-1}} \|S\| = \inf_{T \in [L]_r} \|T^{-1}\| \quad \text{for all } L \in \Phi_0(X,Y).$$

For self-adjoint operators, we can calculate $\gamma(L)$ precisely.

5.1.1. *Self-adjoint operators in Hilbert space*

Let $(H, \langle\cdot,\cdot\rangle, \|\cdot\|)$ be a real Hilbert space. For a self-adjoint operator $S \colon D(S) \subset H \to H$ acting in H, the graph norm of S on $D(S)$ is defined by

$$\|u\|_S = \{\|u\|^2 + \|Su\|^2\}^{1/2} \quad \text{for } u \in D(S).$$

Recall that since S is closed, the graph space $(D(S), \|\cdot\|_S)$ is a Hilbert space.

Proposition 5.4. *Let $S\colon D(S)\subset H\to H$ and $T\colon D(T)\subset H\to H$ be two self-adjoint operators having the same domain $X=D(S)=D(T)$. Then $\|\cdot\|_S$ and $\|\cdot\|_T$ are equivalent norms on the subspace X and $S,T\in B(X,H)$ for such norms.*

Proof. Since S is a closed operator acting in H and $\|\cdot\|\leqslant\|\cdot\|_T$, $S\colon(X,\|\cdot\|_T)\to(H,\|\cdot\|)$ is also closed. But $(X,\|\cdot\|_T)$ is complete and so, by the closed graph theorem, $S\in B((X,\|\cdot\|_T),(H,\|\cdot\|))$. Hence, there exists a constant $A>0$ such that $\|\cdot\|_S\leqslant A\|\cdot\|_T$. The proof is completed by interchanging the roles of S and T. □

For a self-adjoint operator S, its spectrum and essential spectrum are defined by

$$\sigma(S)=\{\lambda\in\mathbb{R}\colon S-\lambda I\colon X\to H \text{ is not an isomorphism}\}$$

and

$$\sigma_e(S)=\{\lambda\in\mathbb{R}\colon S-\lambda I\notin\Phi_0(X,H)\},$$

where $X=D(S)$ with the graph norm of S. It is well known (see, for example, [3]) that $\lambda\in\sigma(S)\setminus\sigma_e(S)$ if and only if λ is an isolated eigenvalue of finite multiplicity of S.

Theorem 5.5. *Let $(H,\langle\cdot,\cdot\rangle,\|\cdot\|)$ be a real Hilbert space and $L\colon D(L)\subset H\to H$ a self-adjoint operator.*

(i) *Suppose that $0\notin\sigma_e(L)$ and let $T=f(L)$ for some continuous function $f\colon\mathbb{R}\to\mathbb{R}$ such that*

$$0<\liminf_{|\lambda|\to\infty}\left|\frac{f(\lambda)}{\lambda}\right|\leqslant\limsup_{|\lambda|\to\infty}\left|\frac{f(\lambda)}{\lambda}\right|<\infty.$$

Then

$$\gamma(L)\leqslant\frac{1}{d(0,\sigma_e(L))}+b(T),$$

where $\gamma(L)$ is the essential conditioning number of $L\in\Phi_0(X,H)$ for $X=D(L)=D(T)$ with the graph norm $\|\cdot\|_T$ and

$$b(T)=\sup\left\{\left|\frac{f(\lambda)}{\lambda}\right|\colon\lambda\notin J\right\}$$

for the maximal interval J of $\mathbb{R}\setminus\sigma_e(L)$ containing 0. If $\sigma_e(L)=\emptyset$, then $\gamma(L)\leqslant b(T)$.

(ii) *If $D(L)=H$, $L\in B(H,H)$ and, if $0\notin\sigma_e(L)$, $\gamma(L)\leqslant 1/d(0,\sigma_e(L))$ for the norm $\|\cdot\|$. Note that when $L\in B(H,H)$, $\sigma_e(L)=\emptyset$ if and only if $\dim H<\infty$ and that $\gamma(L)=0$ for all norms on $D(L)$ in this case.*

Proof. (i) Consider an open interval (a,b) such that $0\in(a,b)$, where $(a,b)\cap\sigma_e(L)=\emptyset$ and $a,b\notin\sigma(L)$. Set $m=\min\{|a|,b\}$ and $M=\max\{|a|,b\}$.

Let $\{E(\lambda)\colon\lambda\in\mathbb{R}\}$ be the resolution of the identity for L, as in [22, ch. XI, § 6, theorem 1]. Then

$$X=D(L)=\left\{u\in H\colon\int\lambda^2\,\mathrm{d}\langle E(\lambda)u,u\rangle<\infty\right\}\quad\text{and}\quad L=\int\lambda\,\mathrm{d}E(\lambda).$$

The set $[a,b]\cap\sigma(L)$ consists of a finite number of eigenvalues of L, all having finite multiplicity. Let $P=E(b)-E(a)$ and $Q=I-P$ and set $U=P(H)$ and $V=Q(H)$. Then $P\in B(H,H)$ is the orthogonal projection onto U, $\dim U<\infty$ and $U^{\perp}=V$. Furthermore, $U\subset D(L)$ and $Qu\in D(L)$ if and only if $u\in D(L)$. Setting $V_1=V\cap D(L)$, we have that $L(U)\subset U$, $L(V_1)\subset V$ and $X=U\oplus V_1$, where this is an orthogonal direct sum for the graph scalar product

$$\langle u,v\rangle_L=\langle u,v\rangle+\langle Lu,Lv\rangle.$$

Setting $C=L|_U$ and $D=L|_{V_1}$ we have that $C(U)\subset U$ and $D(V_1)\subset V$.

For $u\in X$,

$$Lu=\int_{-\infty}^{\infty}\lambda\,\mathrm{d}E(\lambda)u \quad\text{and}\quad E(\lambda)E(\mu)=E(\min\{\lambda,\mu\}).$$

It follows easily that, for $u\in U$,

$$Cu=Lu=L[E(b)-E(a)]u=\int_a^b\lambda\,\mathrm{d}E(\lambda)u$$

and then, for $u\in V_1$,

$$Du=Lu=LQu=Lu-LPu=\int_{-\infty}^{\infty}\lambda\{1-\chi_{[a,b]}(\lambda)\}\,\mathrm{d}E(\lambda)u.$$

Furthermore,

$$\|Cu\|^2=\int_a^b\lambda^2\,\mathrm{d}\langle E(\lambda)u,u\rangle\leqslant M^2\int_a^b\mathrm{d}\langle E(\lambda)u,u\rangle\leqslant M^2\|u\|^2\quad\text{for } u\in U$$

and hence

$$\|(C-\mu I_U)u\|\geqslant(\mu-M)\|u\|\quad\text{for } u\in U \text{ and } \mu>M.$$

On the other hand,

$$\|Du\|^2=\int_{-\infty}^{\infty}\lambda^2\{1-\chi_{[a,b]}(\lambda)\}^2\,\mathrm{d}\langle E(\lambda)u,u\rangle\geqslant m^2\|u\|^2\quad\text{for } u\in V_1.$$

Hence, for $\mu>M$, $C-\mu I_U\colon U\to U$ and $D\colon V_1\subset V\to V$ are invertible self-adjoint operators in the Hilbert spaces U and V, respectively, with

$$\|(C-\mu I_U)^{-1}u\|\leqslant\frac{1}{\mu-M}\|u\|\quad\text{for } u\in U$$

and

$$\|D^{-1}u\|\leqslant\frac{1}{m}\|u\|\qquad\text{for } u\in V_1.$$

Hence, we can define $S\colon H\to X$ by

$$Su=(C-\mu I_U)^{-1}Pu+D^{-1}Qu\quad\text{for } u\in H.$$

Then, for $u \in H$, we have that

$$\begin{aligned} LSu &= C(C-\mu I_U)^{-1}Pu + Qu \\ &= Pu + \mu(C-\mu I_U)^{-1}Pu + Qu \\ &= u + \mu(C-\mu I_U)^{-1}Pu \end{aligned}$$

and, for $u \in X$,

$$\begin{aligned} SLu &= (C-\mu I_U)^{-1}PLu + D^{-1}QLu \\ &= (C-\mu I_U)^{-1}LPu + D^{-1}LQu \\ &= (C-\mu I_U)^{-1}CPu + D^{-1}DQu \\ &= u + \mu(C-\mu I_U)^{-1}Pu. \end{aligned}$$

Note that $\operatorname{rge}(C-\mu I_U)^{-1} = U$ and $\dim U < \infty$. To complete the proof that S is a regular pseudo-inverse for L we must show that $S \in B(H,X)$ and then that $S \in \operatorname{Iso}(H,X)$. In the first step, we shall estimate the norm of S using the graph norm of T on X and this will yield the required estimate for $\gamma(L)$.

By [22, § XI.5, theorem 2], $T = f(L)$ is a self-adjoint operator acting in H with domain

$$D(T) = \left\{ u \in H \colon \int_{-\infty}^{\infty} f(\lambda)^2 \, \mathrm{d}\langle E(\lambda)u, u\rangle < \infty \right\}$$

and, by the hypotheses on f,

$$\int_{-\infty}^{\infty} f(\lambda)^2 \, \mathrm{d}\langle E(\lambda)u, u\rangle < \infty \iff \int_{-\infty}^{\infty} \lambda^2 \, \mathrm{d}\langle E(\lambda)u, u\rangle < \infty$$

so $D(T) = D(L) = X$. The result in [22] also ensures that $E(\lambda)T \subset TE(\lambda)$ and so $PTu = TPu$ for all $u \in X$. It follows that for $u \in V_1$, $QTu = Tu - PTu = Tu - TPu = TQu$ and so $X = U \oplus V_1$ is an orthogonal direct sum for the scalar product $\langle\cdot,\cdot\rangle_T$ associated with the graph norm of T on X. Therefore, for $u \in X$,

$$\begin{aligned} \|TSu\|^2 &= \|T(C-\mu I_U)^{-1}Pu\|^2 + \|TD^{-1}Qu\|^2 \\ &= \int_a^b \left[\frac{f(\lambda)}{\lambda-\mu}\right]^2 \mathrm{d}\langle E(\lambda)Pu, Pu\rangle \\ &\quad + \int_{-\infty}^{\infty} \left[\frac{f(\lambda)}{\lambda}\right]^2 \{1-\chi_{[a,b]}(\lambda)\}^2 \, \mathrm{d}\langle E(\lambda)Qu, Qu\rangle \\ &\leqslant \frac{A^2}{(\mu-b)^2} \int_a^b \mathrm{d}\langle E(\lambda)Pu, Pu\rangle + B^2 \int_{-\infty}^{\infty} \{1-\chi_{[a,b]}(\lambda)\}^2 \, \mathrm{d}\langle E(\lambda)Qu, Qu\rangle \\ &\leqslant \frac{A^2}{(\mu-b)^2}\|Pu\|^2 + B^2\|Qu\|^2 \leqslant B^2\|u\|^2, \end{aligned}$$

where $A = \max_{\lambda\in[a,b]} |f(\lambda)|$, $B = \sup_{\lambda\notin[a,b]} |f(\lambda)/\lambda|$ and $\mu \geqslant \max\{M, b + A/B\} \equiv K$ so that $A/(\mu-b) \leqslant B$.

For $u \in X$ and $\mu \geqslant K$, we now have that

$$\begin{aligned}\|Su\|_T^2 &= \|Su\|^2 + \|TSu\|^2 \leqslant \|Su\|^2 + B^2\|u\|^2 \\ &\leqslant \frac{1}{(\mu - M)^2}\|Pu\|^2 + \frac{1}{m^2}\|Qu\|^2 + B^2\|u\|^2 \\ &\leqslant \left(\frac{1}{m^2} + B^2\right)\|u\|^2,\end{aligned}$$

provided that $\mu \geqslant \max\{K, M + m\}$. This proves that $S \in B(H, X)$ and $\|S\| \leqslant 1/m + B$ if we use the norm $\|\cdot\|_T$ on X. It is easy to see that $S \in \mathrm{Iso}(H, X)$ and so $\gamma(L) \leqslant 1/m + B$. Expanding (a, b) so that $m \to d(0, \sigma_e(L))$ yields the estimate

$$\gamma(L) \leqslant \frac{1}{d(0, \sigma_e(L))} + \sup\left\{\left|\frac{f(\lambda)}{\lambda}\right| : |\lambda| \geqslant d(0, \sigma_e(L))\right\}$$

for the norm $\|\cdot\|_T$ on X.

(ii) From the proof of part (i), we have that

$$\|Su\|^2 \leqslant \frac{1}{(\mu - M)^2}\|Pu\|^2 + \frac{1}{m^2}\|Qu\|^2 \leqslant \frac{1}{m^2}\|u\|^2$$

for all $\mu > M + m$ and the conclusion follows. □

Corollary 5.6. *Let $S\colon D(S) \subset H \to H$ be a self-adjoint operator acting in a real Hilbert space $(H; \langle\cdot,\cdot\rangle, \|\cdot\|)$ and consider $\lambda_0 \notin \sigma_e(S)$. For any $\varepsilon > 0$ and $\mu \in \mathbb{R}$,*

$$\gamma(S - \lambda_0 I) \leqslant \frac{1}{d(\lambda_0, \sigma_e(S))} + \varepsilon K(\mu)$$

provided that we consider $S - \lambda_0 I \in B(X, H)$ with $X = D(S)$ equipped with the graph norm $\|\cdot\|_{\varepsilon(S-\mu I)}$, where

$$K(\mu) = \max\left\{1, \frac{|\alpha - \mu|}{\lambda_0 - \alpha}, \frac{|\beta - \mu|}{\beta - \lambda_0}\right\}$$

and (α, β) is the maximal interval in $\mathbb{R}\backslash\sigma_e(L)$ containing λ_0. If $\alpha = -\infty$ or $\beta = \infty$, replace the corresponding ratio by 1.

Proof. Setting $L = S - \lambda_0 I$, we have that $0 \notin \sigma_e(L)$ and $\varepsilon(S - \mu I) = f(L)$, where $f(\lambda) = \varepsilon(\lambda + \lambda_0 - \mu)$. Since $d(0, \sigma_e(L)) = d(\lambda_0, \sigma_e(S))$, the conclusion follows from theorem 5.5. □

5.2. Local parity of a Fredholm path

We begin by recalling the definition and properties of the determinant of a finite-rank perturbation of the identity acting on a real Banach space X.

If $T \in \Psi(X)$ and E is a subspace of X such that $\mathrm{rge}(I - T) \subset E$ and $\dim E < \infty$, then $T(E) \subset E$ and so $\det T|_E$ is defined. Furthermore, $\det T|_E$ does not depend on the choice of E and so we set

$$\det T = \det T|_E$$

for any finite-dimensional subspace E of X with $\mathrm{rge}(I - T) \subset E$.

The mapping $\det\colon \Psi(X) \to \mathbb{R}$ has the following properties (see, for example, [1]).

(d1) If $T \in \Psi(X)$ and $S \in \mathrm{Iso}(X,Y)$, then $STS^{-1} \in \Psi(Y)$ and $\det STS^{-1} = \det T$.

(d2) If $S, T \in \Psi(X)$, then $ST \in \Psi(X)$ and $\det ST = \det S \det T$.

(d3) If $T \in \Psi(X)$, then $\det T \neq 0 \iff T \in \mathrm{Iso}(X,X)$.

(d4) $\mathrm{sgn}(\det) = \det/|\det|$ is continuous on the set $\Psi(X) \cap \mathrm{Iso}(X,X)$, but if $\dim X = \infty$, det is not continuous on $\Psi(X) \cap \mathrm{Iso}(X,X)$.

A path of Fredholm operators of index zero is a function $L \in C(J, \Phi_0(X,Y))$, where J is an interval and $\Phi_0(X,Y)$ has the metric inherited from $B(X,Y)$. A *local parametrix* for L at $\lambda_0 \in J$ is a path $S \in C(J_0, B(Y,X))$, where $J_0 \subset J$ is some open interval containing λ_0, such that $S(\lambda) \in [L(\lambda)]_r^{-1}$ for all $\lambda \in J_0$. A local parametrix exists at every interior point λ_0.

Lemma 5.7. *Consider a path $L \in C(J, \Phi_0(X,Y))$, where J is an open interval, $\lambda_0 \in J$ and $S_0 \in [L(\lambda_0)]_r^{-1}$. There exists a local parametrix $S \in C(J_0, B(Y,X))$ for L at λ_0 such that $S(\lambda_0) = S_0$.*

Proof. Indeed, since $L(\lambda_0) \in \Phi_0(X,Y)$, $[L(\lambda_0)]_r^{-1} \neq \emptyset$ by lemmas 5.1 and 5.3. For any $S_0 \in [L(\lambda_0)]_r^{-1}$, set $K = S_0^{-1} \in [L(\lambda_0)]_r$ and then $T(\lambda) = K + L(\lambda) - L(\lambda_0)$ for $\lambda \in J$. Clearly, $T \in C(J, B(X,Y))$ and $T(\lambda_0) = K \in \mathrm{Iso}(X,Y)$, so there exists $\delta > 0$ such that $T(\lambda) \in \mathrm{Iso}(X,Y)$ for all $\lambda \in J_0 = (\lambda_0 - \delta, \lambda_0 + \delta)$. Setting $S(\lambda) = T(\lambda)^{-1}$ for $\lambda \in J_0$, we have that

$$S(\lambda)L(\lambda) = T(\lambda)^{-1}\{T(\lambda) + L(\lambda_0) - K\} = I_X + T(\lambda)^{-1}\{L(\lambda_0) - K\} \in \Psi(X),$$

showing that $S(\lambda)$ is a regular pseudo-inverse for $L(\lambda)$. □

Lemma 5.8. *Let $S_1 \in C(J_1, B(Y,X))$ and $S_2 \in C(J_2, B(Y,X))$ be local parametrices for a path $L \in C(J, \Phi_0(X,Y))$ at $\lambda_0 \in J$. Then $S_1(\lambda)S_2(\lambda)^{-1} \in \Psi(X) \cap \mathrm{Iso}(X,X)$ for all $\lambda \in J_1 \cap J_2$ and $\mathrm{sgn}(\det S_1 S_2^{-1})$ is constant on $J_1 \cap J_2$. Furthermore, setting $M_i(\lambda) = S_i(\lambda)L(\lambda)$, we have that*

$$\det M_1(\lambda) = \det[S_1(\lambda)S_2(\lambda)^{-1}] \det M_2(\lambda) \quad \textit{for all } \lambda \in J_1 \cap J_2.$$

Proof. We have that $S_i^{-1} \in C(J_i, B(X,Y))$ and $S_i(\lambda)^{-1} \in [L(\lambda)]_r$ for $\lambda \in J_i$ so $A_i(\lambda) \equiv S_i(\lambda)^{-1} - L(\lambda) \in F(X,Y)$. Hence, for $\lambda \in J_1 \cap J_2$,

$$S_1(\lambda)S_2(\lambda)^{-1} = \{L(\lambda) + A_1(\lambda)\}^{-1}\{L(\lambda) + A_2(\lambda)\} = I_X + S_1(\lambda)\{A_2(\lambda) - A_1(\lambda)\},$$

showing that $S_1(\lambda)S_2(\lambda)^{-1} \in \Psi(X)$ for all $\lambda \in J_1 \cap J_2$ and by property (d3) that $\det S_1(\lambda)S_2(\lambda)^{-1} \neq 0$. Since $\mathrm{sgn}(\det S_1 S_2^{-1})$ is continuous on $J_1 \cap J_2$ by (d4), it must be constant on $J_1 \cap J_2$. Also,

$$\det M_1(\lambda) = \det[S_1(\lambda)S_2(\lambda)^{-1}M_2(\lambda)] = \det[S_1(\lambda)S_2(\lambda)^{-1}] \det M_2(\lambda)$$

by (d2). □

https://doi.org/10.1017/S0308210513000486

A number λ_0 is an *isolated singular point* of a path $L \in C(J, \Phi_0(X,Y))$ if there exists $\delta > 0$ such that $J_{\lambda_0,\delta} \equiv (\lambda_0 - \delta, \lambda_0 + \delta) \subset J$ and $L(\lambda) \in \mathrm{Iso}(X,Y)$ for $0 < |\lambda - \lambda_0| < \delta$. The *local parity* $\sigma(L, \lambda_0)$ of the path L at an isolated singularity λ_0 is defined as follows. Let $S \in C(J_0, B(Y,Y))$ be a local parametrix for L at λ_0 and choose $\delta > 0$ so that $J_{\lambda_0,\delta} \subset J_0$. Now consider $M(\lambda) \equiv S(\lambda)L(\lambda)$. We have that $M \in C(J_{\lambda_0,\delta}, B(X,X)), M(\lambda) \in \Psi(X)$ for all $\lambda \in J_{\lambda_0,\delta}$ and $M(\lambda) \in \mathrm{Iso}(X,X)$ for $\lambda \neq \lambda_0$. Hence, $\mathrm{sgn}(\det M(\lambda))$ is constant on $(\lambda_0 - \delta, \lambda_0)$ and on $(\lambda_0, \lambda_0 + \delta)$ by properties (d3) and (d4). Now set

$$\sigma(L, \lambda_0) = \frac{\det M(\lambda_l)}{|\det M(\lambda_l)|}\frac{\det M(\lambda_r)}{|\det M(\lambda_r)|}, \quad \text{where } \lambda_0 - \delta < \lambda_l < \lambda_0 < \lambda_r < \lambda_0 + \delta,$$

and observe that lemma 5.8 shows that this quantity does not depend on the choice of local parametrix used to construct M. Thus, $\sigma(L, \lambda_0)$ depends only on the path $L \in C(J, \Phi_0(X,Y))$ and the isolated singular point $\lambda_0 \in J$. It will be referred to as the local parity of L at λ_0. That is, $\sigma(L, \lambda_0) = -1$ if $\det M(\lambda)$ changes sign as λ crosses λ_0 and $\sigma(L, \lambda_0) = 1$ if it does not change sign. It has the following useful properties.

(a) If $L(\lambda_0) \in \mathrm{Iso}(X,Y)$, then $\sigma(L, \lambda_0) = 1$.

(b) If $X = Y$ and $L(\lambda) \in \Psi(X)$ for all $\lambda \in J$, then $S(\lambda) \equiv I$ is a local parametrix for L at λ_0 and so $\sigma(L, \lambda_0) = -1$ if and only if $\det L(\lambda)$ changes sign as λ crosses λ_0.

(c) Let X, Y and Z be real Banach spaces. If $U \in C(J, \Phi_0(Y,Z))$ and λ_0 is an isolated singular point of U, then $\sigma(UL, \lambda_0) = \sigma(U, \lambda_0)\sigma(L, \lambda_0)$. To check this, let S and T be local parametrices for L and U at λ_0 with $S(\lambda)L(\lambda) = I_X + A(\lambda)$ and $T(\lambda)U(\lambda) = I_Y + B(\lambda)$ for $\lambda \in J_0$, where $A(\lambda) \in F(X,X)$ and $B(\lambda) \in F(Y,Y)$. Then

$$STUL = STUS^{-1}SL = \{I_X + SBS^{-1}\}\{I_X + A\} \in \Psi(X),$$

showing that ST is a local parametrix for UL at λ_0 and that

$$\begin{aligned}\det STUL &= \det(I_X + SBS^{-1})\det(I_X + A) = \det(I_Y + B)\det SL \\ &= \det TU \det SL\end{aligned}$$

by properties (d1) and (d2) of det.

There are some well-known situations in which the parity can be calculated. See conditions (T) and (C) in [10, corollary of theorem 1]. To keep our presentation self-contained, proofs are given in an appendix.

Criterion 1. Consider $L \in C(J, \Phi_0(X,Y))$, where J is an open interval and L is differentiable at λ_0 with

$$\{u \in \ker L(\lambda_0) \colon L'(\lambda_0)u \in \mathrm{rge}\, L(\lambda_0)\} = \{0\}. \tag{5.1}$$

Then λ_0 is an isolated singular point of L and $\sigma(L, \lambda_0) = (-1)^n$, where $n = \dim\ker L(\lambda_0)$.

https://doi.org/10.1017/S0308210513000486

CRITERION 2. Let $K \in B(X,X)$ be compact and $\lambda_0 \in \sigma(K) \setminus \{0\}$. Setting $L(\lambda) = K - \lambda I$, we have that λ_0 is an isolated singular point of L and $\sigma(L,\lambda_0) = (-1)^n$, where n is the algebraic multiplicity of λ_0 as an eigenvalue of K.

For a more general version of criterion 2, see condition (C) in [9].

6. The main results about bifurcation

We shall use the following hypotheses and notation.

Let X and Y be real Banach spaces and consider a mapping $F\colon U \subset \mathbb{R} \times X \to Y$, where U is an open subset of $\mathbb{R} \times X$, having the following properties.

(B1) $F \in C(U,Y)$ and $F(\lambda,0) = 0$ for all $(\lambda,0) \in U$.

(B2) $F(\lambda,\cdot)\colon X \to Y$ is Hadamard differentiable at 0 for all $(\lambda,0) \in U$ and $\lambda \mapsto D_xF(\lambda,0) \in B(X,Y)$ is continuous.

(B3) $(\lambda_0,0) \in U$ and $D_xF(\lambda_0,0) \in \Phi_0(X,Y)$.

(B4) $L_x(R,\lambda_0) < \infty$, where $R(\lambda,x) = F(\lambda,x) - D_xF(\lambda,0)x$ for $(\lambda,x),(\lambda,0) \in U$.

Notation and conventions

Let $L(\lambda) = D_xF(\lambda,0)$. Assumption (B3) and lemma 5.1 ensure that there exist $S_0 \in [L(\lambda_0)]_r^{-1}$ and a projection $P \in B(X,X)$ of X onto the finite-dimensional subspace $\operatorname{rge}(I_X - S_0L(\lambda_0))$. For a choice of S_0 and P the following notation will be used:

$$E = \operatorname{rge}\{I_X - S_0L(\lambda_0)\}, \quad Q = I - P \quad \text{and} \quad F = \operatorname{rge} Q.$$

Then $\dim E < \infty$, F is a closed subspace of X and $X = E \oplus F$. Clearly, we always have that $\ker L(\lambda_0) \subset E$. If X is a Hilbert space, we choose P to be the orthogonal projection of X onto E. Then $\|I - P\| = 1$ and $F = E^{\perp}$.

Since L is continuous at λ_0 and $\Phi_0(X,Y)$ is an open subset of $B(X,Y)$, the hypotheses (B2) and (B3) imply that there exists an open interval J_ε containing λ_0 such that $L(\lambda) \in \Phi_0(X,Y)$ for all $\lambda \in J_\varepsilon \equiv (\lambda_0 - \varepsilon, \lambda_0 + \varepsilon)$.

For $S_0 \in [L(\lambda_0)]_r^{-1}$, we know that, for ε small enough, there exists a local parametrix $S \in C(J_\varepsilon, B(Y,X))$ for the path $L(\lambda) = D_xF(\lambda,0)$ at λ_0 such that $S(\lambda_0) = S_0$. Setting $K(\lambda) = S(\lambda)L(\lambda) - I_X$, we have that $K \in C(J_\varepsilon, B(X,X))$ and $K(\lambda) \in F(X,X)$ for all $\lambda \in J_\varepsilon$. We note that

$$E = \operatorname{rge} K(\lambda_0) \quad \text{and} \quad QK(\lambda_0) = 0.$$

For $\lambda \in J_\varepsilon$, let $V_\lambda = \{(v,w) \in E \times F\colon (\lambda, v+w) \in U\}$ and then define $C(\lambda,\cdot)\colon V_\lambda \subset E \times F \to F$ by

$$C(\lambda,v,w) = -QK(\lambda)(v+w) - QS(\lambda)R(\lambda,v+w),$$

where $R(\lambda,u) = F(\lambda,u) - L(\lambda)u$.

Lemma 6.1. *Under the hypotheses (B1)–(B4), suppose that*

$$\gamma(D_xF(\lambda_0,0))L_x(R,\lambda_0)<1. \tag{6.1}$$

Then S_0 and P can be chosen such that such that

$$\|(I-P)S_0\|L_x(R,\lambda_0)<1. \tag{6.2}$$

In this case, there exist positive constants ε, ρ and η such that, for all $(\lambda,v)\in J_\varepsilon\times B_E(0,\rho)$,

(i) *$J_\varepsilon\times B_E(0,\rho)\times M\subset U$ and $C(\lambda,v,M)\subset M$, where $M=\overline{B_F(0,\eta)}$,*

(ii) *there exists $k\in[0,1)$ such that*

$$\|C(\lambda,v,w)-C(\lambda,v,z)\|\leqslant k\|w-z\| \quad \textit{for all } w,z\in M.$$

Proof. By (B4), (6.2) follows from (6.1) and the definition of γ. By (6.2), there exists $\xi>0$ such that $c\equiv[\|QS(\lambda_0)\|+\xi][L_x(R,\lambda_0)+\xi]<1$. The continuity of S ensures that we can assume that $\|QS(\lambda)\|\leqslant\|QS(\lambda_0)\|+\xi$ for all $\lambda\in J_\varepsilon$, by making ε smaller if necessary. In the same way, there exists $\eta>0$ such that $J_\varepsilon\times B_X(0,2\eta)\subset U$ and

$$\|R(\lambda,x)-R(\lambda,y)\|\leqslant[L_x(R,\lambda_0)+\xi]\|x-y\|$$

for all $\lambda\in J_\varepsilon$ and $x,y\in B_X(0,2\eta)$.

Choose $k\in(c,1)$. Since $QK(\lambda_0)=0$, ε can be further reduced so as to ensure that $c+\|QK(\lambda)\|\leqslant k$ for all $\lambda\in J_\varepsilon$. It follows that

$$\|QK(\lambda)\|+\|QS(\lambda)\|[L_x(R,\lambda_0)+\xi]\leqslant\|QK(\lambda)\|+c\leqslant k \quad \text{for all } \lambda\in J_\varepsilon.$$

Now set $\rho=\min\{\eta,\eta(1-k)/k\}$.

For (i), note that $\|R(\lambda,u)\|=\|R(\lambda,u)-R(\lambda,0)\|\leqslant[L_x(R,\lambda_0)+\xi]\|u\|$ for $\lambda\in J_\varepsilon$ and $u\in B_X(0,2\eta)$. For $v\in B_E(0,\rho)$ and $w\in M$, $\|v+w\|<\rho+\eta\leqslant2\eta$ and so $B_E(0,\rho)\times M\subset V_\lambda$ for $\lambda\in J_\varepsilon$ and

$$\begin{aligned}\|C(\lambda,v,w)\|&\leqslant\|QK(\lambda)\|\|v+w\|+\|QS(\lambda)\|\|R(\lambda,v+w)\|\leqslant k\|v+w\|\\&\leqslant k(\rho+\eta)\leqslant\eta.\end{aligned} \tag{6.3}$$

For (ii), we have, with the same restrictions on (λ,v,w),

$$\begin{aligned}&\|C(\lambda,v,w)-C(\lambda,v,z)\|\\&\quad\leqslant\|QK(\lambda)(w-z)\|+\|QS(\lambda)\|\|R(\lambda,v+w)-R(\lambda,v+z)\|\\&\quad\leqslant\{\|QK(\lambda)\|+\|QS(\lambda)\|[L_x(R,\lambda_0)+\xi]\}\|w-z\|\leqslant k\|w-z\|.\end{aligned} \tag{6.4}$$

□

6.1. Reduction to finite dimensions

Using (6.1) and the contraction mapping principle, we are able to replace the study of equation $F(\lambda,u)=0$ near $(\lambda_0,0)$ by an equation $f(\lambda,v)=0$, where v lies in a finite-dimensional space.

https://doi.org/10.1017/S0308210513000486

THEOREM 6.2. *Under the hypotheses (B1)–(B4), consider* λ_0 *such that (6.1) holds and choose* S_0 *and* P *such that (6.2) is satisfied. Then we have the following situation.*

(i) *There exist* $\varepsilon > 0$, $\rho > 0$ *and* $\psi \in C(J_\varepsilon \times B_E(0,\rho), F)$ *such that* $(\lambda, v + \psi(\lambda, v)) \in U$,

$$QS(\lambda)F(\lambda, v + \psi(\lambda, v)) = 0 \quad \textit{and} \quad \psi(\lambda, 0) = 0 \tag{6.5}$$

for all $\lambda \in J_\varepsilon$ *and* $v \in B_E(0,\rho)$. *Thus, for* $\lambda \in J_\varepsilon$ *and* $v \in B_E(0,\rho)$,

$$F(\lambda, v + \psi(\lambda, v)) = 0 \iff PS(\lambda)F(\lambda, v + \psi(\lambda, v)) = 0.$$

The implicit function ψ *has some additional properties:*

(a) *for all* $\lambda \in J_\varepsilon$, $\psi(\lambda, \cdot)\colon B_E(0,\rho) \to F$ *is Fréchet differentiable at* 0 *with* $D_v\psi(\cdot, 0) \in C(J_\varepsilon, B(E, F))$ *and* $D_v\psi(\lambda_0, 0) = 0$;

(b) *we have that*

$$\|\psi(\lambda, v)\| \leqslant \frac{k}{1-k}\|v\| \quad \textit{and} \quad \|D_v\psi(\lambda, 0)\| \leqslant \frac{\|Q\|}{1-k}\|K(\lambda) - K(\lambda_0)\|$$

for all $\lambda \in J_\varepsilon$ *and* $v \in B_E(0,\rho)$.

(ii) *There exists* $r > 0$ *such that, if* $F(\lambda, u) = 0$ *and* $(\lambda, u) \in B((\lambda_0, 0), r)$, *then* $\lambda \in J_\varepsilon$, $Pu \in B_E(0,\rho)$ *and* $Qu = \psi(\lambda, Pu)$.

(iii) *If* $\lim_{\delta\to 0} \Delta_\delta(F, \lambda_0) = 0$, *then* $\lim_{\delta\to 0} \Delta_\delta(\psi, \lambda_0) = 0$. *If* $\lim_{\delta\to 0} \Delta_\delta(F, \lambda) = 0$ *for all* λ *in an open neighbourhood of* λ_0, *then* $\lim_{\delta\to 0} \Delta_\delta(\psi, \lambda) = 0$ *for all* $\lambda \in J_\varepsilon$, *by making* ε *small enough.*

Proof. (i) With ε, ρ and η as in lemma 6.1, let $\Omega = J_\varepsilon \times B_E(0,\rho)$.

First of all we note that, for $\lambda \in J_\varepsilon$ and $(v, w) \in V_\lambda$,

$$\begin{aligned} QS(\lambda)F(\lambda, v + w) &= Q\{S(\lambda)L(\lambda)(v + w) + S(\lambda)R(\lambda, v + w)\} \\ &= Q\{[I + K(\lambda)](v + w) + S(\lambda)R(\lambda, v + w)\} \\ &= w - C(\lambda, v, w), \end{aligned} \tag{6.6}$$

where C is the contraction mapping discussed in lemma 6.1. It follows from proposition 3.4(i) that, for each $(\lambda, v) \in \Omega$, there exists a unique element $\psi(\lambda, v) \in M$ such that $\psi(\lambda, v) = C(\lambda, v, \psi(\lambda, v))$. Since $C(\lambda, 0, 0) = 0$, we also have that $\psi(\lambda, 0) = 0$ for all $\lambda \in J_\varepsilon$.

We now expose some additional properties of C and their implications for $\psi(\lambda, v)$.

The continuity of $F\colon U \subset \mathbb{R} \times X \to Y$ and $S\colon J_\varepsilon \to B(Y, X)$ imply that $C\colon \Omega \times M \to M$ is continuous. By proposition 3.4(ii) this implies that $\psi\colon \Omega \to F$ is continuous, proving part (i).

Furthermore, for each $\lambda \in J_\varepsilon$, $R(\lambda, \cdot)$ is Hadamard differentiable at 0 with $D_x R(\lambda, 0) = 0$. Hence, we see that, for each $\lambda \in J_\varepsilon$, $C(\lambda, \cdot, \cdot)\colon B_E(0,\rho) \times M \to F$ is Hadamard differentiable at $(0, 0)$ with $D_{(v,w)}C(\lambda, 0, 0)(e, f) = -QK(\lambda)(e + f)$

https://doi.org/10.1017/S0308210513000486

for $(e,f) \in E \times F$. By proposition 3.4(iii) this implies that $\psi(\lambda,\cdot)\colon B_E(0,\rho) \to F$ is Hadamard differentiable at 0 with

$$D_v\psi(\lambda,0) = D_E C(\lambda,0,0) + D_F C(\lambda,0,0) D_v\psi(\lambda,0) = -QK(\lambda)[I_E + D_v\psi(\lambda,0)]$$

and hence

$$[I_F + QK(\lambda)]D_v\psi(\lambda,0) = -QK(\lambda)|_E \quad \text{for all } \lambda \in J. \tag{6.7}$$

But $\dim E < \infty$, so $\psi(\lambda,\cdot)$ is Fréchet differentiable at 0. Recalling that $QK(\lambda_0) = 0$, we have that $D_v\psi(\lambda_0,0) = 0$. Furthermore, since $S \in C(J_\varepsilon, B(Y,X))$, we also have that $K \in C(J_\varepsilon, B(X,X))$ and $\|QK(\lambda)\| \leqslant k$ for all $\lambda \in J_\varepsilon$. Thus, $D_v\psi(\cdot,0) = -[I_F + QK(\cdot)|_F]^{-1}QK(\cdot)|_E \in C(J_\varepsilon, B(E,F))$, proving part (ii).

(iii) For $(\lambda,v) \in \Omega$, it follows from (6.3) that $\|\psi(\lambda,v)\| = \|C(\lambda,v,\psi(\lambda,v))\| \leqslant k(\|v\| + \|\psi(\lambda,v)\|)$ and hence that

$$\|\psi(\lambda,v)\| \leqslant \frac{k}{1-k}\|v\| \quad \text{and} \quad \|v + \psi(\lambda,v)\| \leqslant \frac{1}{1-k}\|v\| \quad \text{for all } (\lambda,v) \in \Omega. \tag{6.8}$$

Returning to (6.7) we have that

$$D_v\psi(\lambda,0)z = -QK(\lambda)\{z + D_v\psi(\lambda,0)z\}, \quad \text{where } \|QK(\lambda)\| \leqslant k$$

for $\lambda \in J_\varepsilon$ and $z \in E$. It follows that

$$\begin{aligned}\|D_v\psi(\lambda,0)\| &\leqslant \frac{1}{1-k}\|QK(\lambda)\| \\ &= \frac{1}{1-k}\|Q[K(\lambda) - K(\lambda_0)]\| \\ &\leqslant \frac{\|Q\|}{1-k}\|K(\lambda) - K(\lambda_0)\|,\end{aligned}$$

proving (iii).

(iv) For $0 < r < \min\{\varepsilon, \rho/\|P\|, \eta/\|Q\|\}$, we have that $\lambda \in J_\varepsilon, Pu \in B_E(0,\rho)$ and $Qu \in M$ for all $(\lambda,u) \in B((\lambda_0,0),r)$. Since $F(\lambda,u) = 0$ implies that

$$0 = QS(\lambda)F(\lambda, Pu + Qu) = Qu - C(\lambda, Pu, Qu),$$

the uniqueness of the fixed point of $C(\lambda, Pu, \cdot)$ in M means that $Qu = \psi(\lambda, Pu)$, proving (iv).

(v) For $(\lambda,v),(\mu,v) \in \Omega$, we have

$$\begin{aligned}\|\psi(\lambda,v) - \psi(\mu,v)\| &= \|C(\lambda,v,\psi(\lambda,v)) - C(\lambda,v,\psi(\mu,v)) \\ &\qquad + C(\lambda,v,\psi(\mu,v)) - C(\mu,v,\psi(\mu,v))\| \\ &\leqslant k\|\psi(\lambda,v) - \psi(\mu,v)\| + \|C(\lambda,v,\psi(\mu,v)) - C(\mu,v,\psi(\mu,v))\|\end{aligned}$$

by (6.4) and so

$$\|\psi(\lambda,v) - \psi(\mu,v)\| \leqslant \frac{1}{1-k}\|C(\lambda,v,\psi(\mu,v)) - C(\mu,v,\psi(\mu,v))\|. \tag{6.9}$$

https://doi.org/10.1017/S0308210513000486

By the continuity of ψ, the values of ε and ρ can be reduced so as to ensure that, for all $(\lambda, v) \in \Omega = J_\varepsilon \times B_E(\rho)$, we have

$$\|F(\lambda, u)\| = \|F(\lambda, u) - F(\lambda, 0)\| \leqslant [L_x(F, \lambda_0) + 1]\|u\| \quad \text{for all } (\lambda, v) \in \Omega,$$

where $u = v + \psi(\lambda, v)$.

Consider $(\mu, v) \in B((\lambda, 0), \delta) \subset \Omega \subset \mathbb{R}\times E$. It follows from (6.8) that $v+\psi(\mu, v) \in B((\lambda, 0), \delta/(1-k)) \subset \mathbb{R} \times X$.

Now $C(\lambda, v, w) = w - QS(\lambda)F(\lambda, v + w)$ by (6.6) and so, with $u = v + \psi(\mu, v)$, we have that

$$\begin{aligned}
&\|C(\lambda, v, \psi(\mu, v)) - C(\mu, v, \psi(\mu, v))\| \\
&\qquad = \|QS(\lambda)F(\lambda, u) - QS(\mu)F(\mu, u)\| \\
&\qquad \leqslant \|Q\|\{\|S(\lambda) - S(\mu)\|\|F(\lambda, u)\| + \|S(\mu)\|\|F(\lambda, u) - F(\mu, u)\|\} \\
&\qquad \leqslant \|Q\|\{\|S(\lambda) - S(\mu)\|[L_x(F, \lambda_0) + 1]\|u\| + \|S(\mu)\|\Delta_{\delta/(1-k)}(F, \lambda)\|u\|\} \\
&\qquad \leqslant C\{\|S(\lambda) - S(\mu)\| + \Delta_{\delta/(1-k)}(F, \lambda)\}\|v + \psi(\mu, v)\|.
\end{aligned} \tag{6.10}$$

From (6.8)–(6.10) it follows that

$$\|\psi(\lambda, v) - \psi(\mu, v)\| \leqslant \frac{C}{(1-k)^2}\{\|S(\lambda) - S(\mu)\| + \Delta_{\delta/(1-k)}(F, \lambda)\}\|v\|$$

for all $(\mu, v) \in B((\lambda, 0), \delta) \subset J_\varepsilon \times B_E(\rho)$, showing that

$$\Delta_\delta(\psi, \lambda) \leqslant \frac{C}{(1-k)^2}\Big\{\sup_{|\lambda-\mu|<\delta} \|S(\lambda) - S(\mu)\| + \Delta_{\delta/(1-k)}(F, \lambda)\Big\},$$

from which (v) follows. □

6.2. Bifurcation theorems

We now come to the main results about bifurcation.

THEOREM 6.3. *Under hypotheses (B1)–(B4), consider λ_0 such that (6.1) is satisfied.*

(i) *If λ_0 is an isolated singular point of $D_xF(\cdot, 0)$ with $\sigma(D_xF(\cdot, 0), \lambda_0) = -1$, then λ_0 is a bifurcation point for $F(\lambda, u) = 0$. If, in addition,*

$$\lim_{\delta\to 0} \Delta_\delta(F, \lambda) = 0 \quad \textit{for all } \lambda \textit{ in an open neighbourhood of } \lambda_0, \tag{6.11}$$

there is continuous bifurcation at λ_0.

(ii) *If $\ker D_xF(\lambda_0, 0) = \{0\}$ and $\lim_{\delta\to 0} \Delta_\delta(F, \lambda_0) = 0$, then λ_0 is not a bifurcation point for $F(\lambda, u) = 0$.*

REMARK 6.4.

(1) The hypothesis (6.1) plays a crucial role in this result. We have examples where (B1)–(B4) and (6.11) are satisfied and there is an isolated singular point λ_0 of $L(\lambda_0)$ with $\sigma(L, \lambda_0) = -1$ which is not a bifurcation point. There are also examples where (B1)–(B4) and (6.11) hold with $\ker L(\lambda_0) = \{0\}$ but λ_0 is a bifurcation point for the equation $F(\lambda, u) = 0$.

https://doi.org/10.1017/S0308210513000486

(2) The condition (6.1) is satisfied if $L_x(R, \lambda_0) = 0$ and we have shown that this occurs if there exists an open neighbourhood, U, of $(\lambda_0, 0)$ in $\mathbb{R} \times X$ such that $F(\lambda, \cdot)$ is continuous and Gâteaux differentiable at u for all $(\lambda, u) \in U$ and $\|D_x F(\lambda, u) - D_x F(\lambda_0, 0)\|_{B(X,Y)} \to 0$ as $(\lambda, u) \to (\lambda_0, 0)$. However, in this case, $F(\lambda_0, \cdot)$ is strictly Fréchet differentiable, and hence Fréchet differentiable, at $u = 0$.

(3) If $F(\lambda_0, \cdot)$ is Hadamard differentiable, but not Fréchet differentiable at 0, then $L_x(R, \lambda_0) > 0$ and so checking (6.1) requires estimates for $\gamma(L(\lambda_0))$. If $\dim X = \dim Y < \infty$, then $\gamma(L(\lambda_0)) = 0$ and (6.1) is satisfied. In this case (B2) implies that $F(\lambda, \cdot)$ is Fréchet differentiable at 0 but it may not be strictly differentiable at 0 and so, even in finite dimensions, we may have $L_x(R, \lambda_0) > 0$. Thus, proposition 4.4 is a special case of theorem 6.3. When X and Y are infinite-dimensional, by using the standard construction of a regular pseudo-inverse from the proof of lemma 5.1, we see that (6.1) is satisfied provided that

$$\|V^{-1}Q\| L_x(R, \lambda_0) < 1, \tag{6.12}$$

where $E = \ker L(\lambda_0)$, $X = E \oplus W$, $Y = \operatorname{rge} L(\lambda_0) \oplus Z$, $V = L(\lambda_0)|_W$ and Q is the projection of Y onto $\operatorname{rge} L(\lambda_0)$. We have shown in § 5.1.1 how to get better estimates for $\gamma(L(\lambda_0))$ in the Hilbert space case with $L(\lambda_0)$ self-adjoint. We exploit this in § 6.4, where we formulate some useful corollaries of theorem 6.3. Of course, if $F \in C^1(U, Y)$, we have $L_x(R, \lambda_0) = 0$ so (6.1) is satisfied, as is (6.11), and we recover the usual results about bifurcation using either criterion 1 or criterion 2 from § 5.2. For example, see [9, theorem 1] for a C^1-version of theorem 6.3. Furthermore, [9, rem. 2, p. 1001] refers to a non-C^1 situation provided that $L_x(R, \lambda_0)$ is sufficiently small. Our introduction of the essential conditioning number provides a way of quantifying this smallness requirement and as we show in several contexts the condition (6.1) if sharp. The proof of [9, theorem 1], which does not use a Lyapunov–Schmidt reduction, cannot yield a sharp criterion.

(4) Using the assumption (B2), the property (6.11) is equivalent to requiring that

$$\lim_{\delta \to 0} \sup_{\substack{(\mu,u) \in B((\lambda,0),\delta) \\ u \neq 0}} \frac{\|R(\mu, u) - R(\lambda, u)\|}{\|u\|} = 0 \tag{6.13}$$

for all λ in an open neighbourhood of λ_0 and so it is trivially satisfied when R is independent of λ.

Proof of theorem 6.3. We use the notation introduced at the beginning of this section. Let ε, ρ and ψ be given by theorem 6.2 and set $\Omega = J_\varepsilon \times B_E(0, \rho)$. Since

$$QS(\lambda) F(\lambda, v + \psi(\lambda, v)) = 0 \quad \text{for all } (\lambda, v) \in \Omega$$

and the equation (6.7) can be expressed as

$$QS(\lambda) D_x F(\lambda, 0)[z + D_v \psi(\lambda, 0) z] = 0. \tag{6.14}$$

(i) Consider the mapping $f\colon \Omega \subset \mathbb{R} \times E \to E$ defined by

$$f(\lambda, v) = PS(\lambda) F(\lambda, v + \psi(\lambda, v)).$$

From theorem 6.2 we obtain the following properties of f:

(1) $f \in C(\Omega, E)$;

(2) $f(\lambda, 0) = PS(\lambda)F(\lambda, 0) = 0$ for all $\lambda \in J$;

(3) for each $\lambda \in J_\varepsilon$, $f(\lambda, \cdot)\colon B_E(0, \rho) \to E$ is Hadamard differentiable at 0 with $D_v f(\lambda, 0)z = PS(\lambda)D_x F(\lambda, 0)[z + D_v\psi(\lambda, 0)z]$ for all $z \in E$.

Since $\dim E < \infty$, $f(\lambda, \cdot)\colon B_E(0, \rho) \to E$ is Fréchet differentiable at 0 and furthermore, $D_v f(\cdot, 0) \in C(J_\varepsilon, B(E, E))$.

To prove that λ_0 is a bifurcation point it is enough to show that

(a) λ_0 is an isolated singular point of $D_v f(\cdot, 0)$ and

(b) $\sigma(D_v f(\cdot, 0), \lambda_0) = -1$.

It then follows from proposition 4.1 that λ_0 is a bifurcation point for the equation $f(\lambda, v) = 0$. But $f(\lambda, v) = 0$ implies that $S(\lambda)F(\lambda, v + \psi(\lambda, v)) = 0$ and hence that $F(\lambda, v + \psi(\lambda, v)) = 0$ since $S(\lambda) \in \mathrm{Iso}(Y, X)$ for all $\lambda \in J_\varepsilon$. Furthermore, setting $u_n = v_n + \psi(\lambda_n, v_n)$, we have that $v_n \neq 0$ and $(\lambda_n, v_n) \to (\lambda_0, 0)$ imply that $u_n \neq 0$ and $(\lambda_n, u_n) \to (\lambda_0, 0)$.

We now prove statements (a) and (b).

Proof of (a): since λ_0 is an isolated singular point of $D_x F(\cdot, 0)$, there exists $\delta \in (0, \varepsilon]$ such that $D_x F(\cdot, 0) \in \mathrm{Iso}(X, Y)$ for $0 < |\lambda - \lambda_0| < \delta$. Also $S(\lambda) \in \mathrm{Iso}(Y, X)$ for all $\lambda \in J_\varepsilon$. Now if $D_v f(\lambda, 0)z = 0$ for some $\lambda \in J_\varepsilon$ and $z \in E$, we have $PS(\lambda)D_u F(\lambda, 0)[z + D_v\psi(\lambda, 0)z] = 0$ and so, by (6.14), this means that

$$S(\lambda)D_u F(\lambda, 0)[z + D_v\psi(\lambda, 0)z] = 0.$$

But $S(\lambda)D_x F(\lambda, 0) \in \mathrm{Iso}(X, X)$ for $0 < |\lambda - \lambda_0| < \delta$ so $z + D_v\psi(\lambda, 0)z = 0$ in this case. Therefore, $\ker D_v f(\lambda, 0) = \{0\}$ for $0 < |\lambda - \lambda_0| < \delta$ since $E \cap F = \{0\}$. This proves that λ_0 is a isolated singular point of $D_v f(\cdot, 0)$.

Proof of (b): to calculate the local parity at λ_0, consider the following linear operators. Let $M(\lambda) = S(\lambda)D_x F(\lambda, 0) = I + K(\lambda)$ for $\lambda \in J_\varepsilon$. Then $D_v f(\lambda, 0) = PM(\lambda)A(\lambda)$, where

$$A(\lambda) \in B(E, X) \text{ is defined by } A(\lambda)z = z + D_v\psi(\lambda, 0)z.$$

Setting $T(\lambda) = QM(\lambda)|_F$, we have that $T \in C(J_\varepsilon, B(F, F))$ and $T(\lambda_0) = Q(I + K(\lambda_0))|_F = Q|_F = I_F$ since $QK(\lambda_0) = 0$. Hence, $\delta \in (0, \varepsilon]$ can be chosen so that $T(\lambda) \in \mathrm{Iso}(F, F)$ for $\lambda \in J_\delta = \{\lambda\colon |\lambda - \lambda_0| < \delta\} \subset J$.

Now define $H(\lambda) \in B(X, X)$ by

$$H(\lambda)u = Pu + T(\lambda)^{-1}Qu.$$

Then $H \in C(J_\delta, B(X, X))$ and it is easily seen that $H(\lambda) \in \mathrm{Iso}(X, X)$ for all $\lambda \in J_\delta$.

Finally, we define $B(\lambda) \in B(X, X)$ by $B(\lambda)u = A(\lambda)Pu + Qu$. Again $B \in C(J_\varepsilon, B(X, X))$ and $B(\lambda_0) = I_X$ since $D_v\psi(\lambda_0, 0) = 0$. Hence, by adjusting the value of δ we have that $B(\lambda) \in \mathrm{Iso}(X, X)$ for all $\lambda \in J_\delta$.

Note that

$$H(\lambda)M(\lambda)B(\lambda) = H(\lambda)S(\lambda)D_xF(\lambda,0)B(\lambda),$$

where $H, B \in C(J_\delta, \mathrm{Iso}(X,X))$ and $S \in C(J_\delta, \mathrm{Iso}(Y,X))$. Thus, λ_0 is an isolated singular point of $HMB \in C(J_\delta, \Phi_0(X,X))$ and $\sigma(HMB,\lambda_0) = \sigma(D_xF(\cdot,0),\lambda_0) = -1$.

However, for $\lambda \in J_\delta$, we have that

$$H(\lambda)M(\lambda)B(\lambda) = PM(\lambda)A(\lambda)P + PM(\lambda)Q + T(\lambda)^{-1}QM(\lambda)A(\lambda)P + T(\lambda)^{-1}QM(\lambda)Q.$$

But (6.14) means that $QM(\lambda)A(\lambda)P = 0$ and $T(\lambda)^{-1}QM(\lambda)Q = Q = I_X - P$, so in fact

$$H(\lambda)M(\lambda)B(\lambda) = I_X - P + PM(\lambda)A(\lambda)P + PM(\lambda)Q.$$

Hence, for $\lambda \in J_\delta$, we have that $H(\lambda)M(\lambda)B(\lambda) \in \Psi(X)$ and

$$\begin{aligned}\det[H(\lambda)M(\lambda)B(\lambda)] &= \det[H(\lambda)M(\lambda)B(\lambda)|_{\mathrm{rge}\,P=E}] \\ &= \det[PM(\lambda)A(\lambda)] \\ &= \det D_vf(\lambda,0).\end{aligned}$$

Therefore, $\sigma(D_vf(\cdot,0),\lambda_0) = \sigma(HMB,\lambda_0) = -1$.

Continuous bifurcation. Using the additional assumption, it follows from part (iii) of theorem 6.2 that f satisfies the hypotheses of proposition 4.4(b) and so there is continuous bifurcation at λ_0 for the equation $f(\lambda,v) = 0$. Thus, there is a connected subset $\mathcal{D}$ of $\{(\lambda,v) \in \Omega\colon f(\lambda,v) = 0 \text{ and } v \neq 0\}$ such that $\bar{\mathcal{D}} \cap [\mathbb{R}\times\{0\}] = \{(\lambda_0,0)\}$ and, as before, $f(\lambda,v) = 0$ implies that $F(\lambda, v + \psi(\lambda,v)) = 0$.

Setting $\mathcal{C} = \{(\lambda, v + \psi(\lambda,v))\colon (\lambda,v) \in \mathcal{D}\}$, the continuity of $(\lambda,v) \mapsto (\lambda, v + \psi(\lambda,v))$ ensures that $\mathcal{C}$ is a connected subset of $\mathbb{R}\times X$. Since $v + \psi(\lambda,v) = 0$ if and only if $v = 0$, we have that $\mathcal{C} \subset \mathcal{S} = \{(\lambda,u) \in U\colon F(\lambda,u) = 0 \text{ and } u \neq 0\}$ for all $(\lambda,v) \in \mathcal{D}$. There exists a sequence $(\lambda_n,v_n) \in \mathcal{D}$ such that $\lambda_n \to \lambda_0$ and $\|v_n\| \to 0$, from which it follows that $(\lambda_n, v_n + \psi(\lambda_n,v_n)) \to (\lambda_0,0)$ and $(\lambda_0,0) \in \bar{\mathcal{C}} \cap [\mathbb{R}\times\{0\}]$. On the other hand, if $(\lambda_n,u_n) \in \mathcal{C}$ and $(\lambda_n,u_n) \to (\mu,0)$, then $Pu_n \to 0$ and $u_n = Pu_n + \psi(\lambda_n,Pu_n)$, where $(\lambda_n,Pu_n) \in \mathcal{D}$. Thus, $(\mu,0) \in \bar{\mathcal{D}}$ and so $\mu = \lambda_0$ proving that there is continuous bifurcation at λ_0 for the equation $F(\lambda,u) = 0$.

(ii) Since $D_xF(\lambda_0,0) \in \Phi_0(X,Y)$ and $\ker D_xF(\lambda_0,0) = \{0\}$, it follows that

$$D_xF(\lambda_0,0) \in \mathrm{Iso}(X,Y).$$

If $z \in \ker D_vf(\lambda_0,0)$, the proof of (a) in part (i) shows that $S(\lambda_0)D_xF(\lambda_0,0)[z + D_v\psi(\lambda_0,0)z] = 0$, from which it now follows that $z + D_v\psi(\lambda_0,0)z = 0$ and then $z = 0$. Using part (iii) of theorem 6.2, we now have that $f\colon \Omega \to E$ satisfies the hypotheses of proposition 4.4(a) and so λ_0 is not a bifurcation point for the equation $f(\lambda,v) = 0$.

Consider a sequence $\{(\lambda_n,u_n)\} \subset U$ such that $F(\lambda_n,u_n) = 0$ and $(\lambda_n,u_n) \to (\lambda_0,0)$ in $\mathbb{R}\times X$. By part (ii) of theorem 6.2 we can suppose that $Qu_n = \psi(\lambda_n,Pu_n)$ and hence that $f(\lambda_n,v_n) = 0$, where $v_n \equiv Pu_n \to 0$. Since λ_0 is not a bifurcation point for the equation $f(\lambda,v) = 0$, there exists n_0 such that $v_n = 0$ for all $n \geqslant n_0$.

https://doi.org/10.1017/S0308210513000486

This implies that $u_n = 0$ for all $n \geqslant n_0$, showing that λ_0 is not a bifurcation point for the equation $F(\lambda, u) = 0$. □

6.3. Bifurcation from a simple eigenvalue

In the case where $\dim\ker D_x F(\lambda_0, 0) = 1$, a transversality condition implies that the parity is -1 and some extra information can be obtained concerning the form of non-trivial solutions near the bifurcation point.

THEOREM 6.5. *Under the hypotheses (B1)–(B4) and (6.11), suppose that (6.1) is satisfied. Suppose also that* $L(\lambda) = D_x F(\lambda, 0)$ *is differentiable at* λ_0 *and that*

$$\ker L(\lambda_0) = \operatorname{span}\{\phi\}, \quad \textit{where } L'(\lambda_0)\phi \notin \operatorname{rge} L(\lambda_0). \tag{6.15}$$

Then λ_0 *is a bifurcation point for* $F(\lambda, u) = 0$ *and there exists a connected subset* $\mathcal{C}$ *of* $\mathcal{S} = \{(\lambda, u) \in U : F(\lambda, u) = 0$ *and* $u \neq 0\}$ *such that* $\bar{\mathcal{C}} \cap [\mathbb{R} \times \{0\}] = \{(\lambda_0, 0)\}$.

Choose Z *such that* $E = \operatorname{span}\{\phi\} \oplus Z$ *and set* $W = Z \oplus F$. *Let* $\Pi \in B(X, X)$ *be the projection onto* $\operatorname{span}\{\phi\}$ *associated with the decomposition* $X = \operatorname{span}\{\phi\} \oplus W$. *Then*

$$\lim_{\substack{(\lambda,u)\to(\lambda_0,0)\\ F(\lambda,u)=0,\ u\neq 0}} \frac{\|u - \Pi u\|}{\|u\|} = 0. \tag{6.16}$$

REMARK 6.6. Since

$$\frac{\|u\| - \|u - \Pi u\|}{\|u\|} \leqslant \frac{\|\Pi u\|}{\|u\|} \leqslant \frac{\|u\| + \|u - \Pi u\|}{\|u\|},$$

it follows from (6.16) that

$$\lim_{\substack{(\lambda,u)\to(\lambda_0,0)\\ F(\lambda,u)=0,\ u\neq 0}} \frac{\|\Pi u\|}{\|u\|} = 1, \tag{6.17}$$

and the property (6.16) can be expressed as follows. There exists $t(u) \in \mathbb{R}$ such that $\Pi u = t(u)\phi$ and $t(u) \to 0$ as $\|u\| \to 0$. It follows from (6.17) that there exists $\delta > 0$ such that $t(u) \neq 0$ if $F(\lambda, u) = 0$, $u \neq 0$ and $|\lambda - \lambda_0| + \|u\| < \delta$. For such u, we set

$$w(u) = \frac{u - t(u)\phi}{t(u)} \text{ so that } u = t(u)\{\phi + w(u)\}, \quad \text{where } w(u) \in W, \tag{6.18}$$

and

$$\lim_{\substack{(\lambda,u)\to(\lambda_0,0)\\ F(\lambda,u)=0,\ u\neq 0}} \|w(u)\| = \lim_{\substack{(\lambda,u)\to(\lambda_0,0)\\ F(\lambda,u)=0, u\neq 0}} \frac{\|u - \Pi u\|}{\|u\|} \frac{\|u\|\|\phi\|}{\|\Pi u\|} = 0.$$

Although the development (6.18) resembles the usual result about bifurcation at a simple eigenvalue in the case where $F \in C^1(U, Y)$ and (6.15) is satisfied, the conclusion of theorem 6.5 is significantly weaker than what is known in the smooth case. There is no claim that $\mathcal{C}$ is a curve and no information is given about the number of other connected sets having the same properties as $\mathcal{C}$.

https://doi.org/10.1017/S0308210513000486

Remark 6.7. In the case where X is a Hilbert space, $F = E^{\perp}$ and we can choose $Z = E \cap [\ker L(\lambda_0)]^{\perp}$ so that $W = [\ker L(\lambda_0)]^{\perp}$. Then $t(u) = \langle u, \phi \rangle$, provided that $\|\phi\| = 1$.

Proof. From (6.15) and criterion 1 in § 5.2 it follows that λ_0 is an isolated singular point of the path L and that $\sigma(L, \lambda_0) = -1$. Hence, the existence of the connected set $\mathcal{C}$ of solutions is ensured by theorem 6.3.

By theorem 6.2(ii), there exists $r > 0$ such that $Qu = \psi(\lambda, Pu)$ for all $(\lambda, u) \in B((\lambda_0, 0), r)$ such that $F(\lambda, u) = 0$ and so we have that

$$u = Pu + \psi(\lambda, Pu) = \Pi Pu + (I - \Pi)Pu + \psi(\lambda, Pu),$$

where $(I - \Pi)Pu + \psi(\lambda, Pu) \in Z \oplus F = W$. Setting $z(u) = (I - \Pi)Pu$, this shows that

$$\Pi Pu = \Pi u, \quad Pu = \Pi u + z(u) \quad \text{and} \quad u = \Pi u + z(u) + \psi(\lambda, Pu).$$

Since $\ker L(\lambda_0) \cap Z = \{0\}$ and $\dim Z < \infty$, there exists a constant $\beta > 0$ such that $\|L(\lambda_0)z\| \geqslant \beta\|z\|$ for all $z \in Z$. For $(\lambda, u) \in \mathcal{C} \cap B((\lambda_0, 0), \delta)$ with $\delta \in (0, r)$, we have

$$\begin{aligned} L(\lambda_0)z(u) &= L(\lambda_0)z(u) - F(\lambda_0, u) + F(\lambda_0, u) - F(\lambda, u) \\ &= L(\lambda_0)z(u) - L(\lambda_0)u - R(\lambda_0, u) + \{F(\lambda_0, u) - F(\lambda, u)\} \\ &= -L(\lambda_0)\psi(\lambda, Pu) - R(\lambda_0, u) + \{F(\lambda_0, u) - F(\lambda, u)\}, \end{aligned}$$

and hence

$$\beta\|z(u)\| \leqslant \|L(\lambda_0)\psi(\lambda, Pu)\| + \|R(\lambda_0, u)\| + \Delta_\delta(F, \lambda)\|u\|.$$

Now, for δ small enough,

$$\begin{aligned} \|R(\lambda_0, u)\| &\leqslant \|R(\lambda_0, u) - R(\lambda_0, Pu)\| + \|R(\lambda_0, Pu)\| \\ &\leqslant [L_x(R, \lambda_0) + 1]\|\psi(\lambda, Pu)\| + \|R(\lambda_0, Pu)\| \end{aligned}$$

and

$$\begin{aligned} \|\psi(\lambda, Pu)\| &\leqslant \|\psi(\lambda, Pu) - \psi(\lambda_0, Pu)\| + \|\psi(\lambda_0, Pu)\| \\ &\leqslant \Delta_{\delta\|P\|}(\psi, \lambda)\|Pu\| + \|\psi(\lambda_0, Pu)\|. \end{aligned} \tag{6.19}$$

Thus, we obtain

$$\begin{aligned} \beta\|z(u)\| &\leqslant \Delta_\delta(F, \lambda)\|Pu\| \\ &\qquad + \{\|L(\lambda_0)\| + L_x(R, \lambda_0) + 1 + \Delta_\delta(F, \lambda)\}\|\psi(\lambda, Pu)\| + \|R(\lambda_0, Pu)\| \\ &\leqslant \Delta_\delta(F, \lambda)\|Pu\| + C\|\psi(\lambda, Pu)\| + \|R(\lambda_0, Pu)\| \\ &\leqslant \{\Delta_\delta(F, \lambda) + C\Delta_{\delta\|P\|}(\psi, \lambda)\}\|Pu\| + C\|\psi(\lambda_0, Pu)\| + \|R(\lambda_0, Pu)\| \end{aligned}$$

with $C = \|L(\lambda_0)\| + L_x(R, \lambda_0) + 2$ since we may assume by (6.11) that $\Delta_\delta(F, \lambda) \leqslant 1$ for all $\lambda \in J_\varepsilon$ and $\delta \leqslant r$. Recalling that both $\psi(\lambda_0, \cdot)\colon B_E(0, \rho) \to F$ and $R(\lambda_0, \cdot)\colon B_E(0, \rho) \to Y$ are Fréchet differentiable at 0 with $D_v\psi(\lambda_0, 0) = 0$ and $D_x R(\lambda_0, 0) = 0$, it follows that

$$\beta\|z(u)\| \leqslant g(\lambda, \|u\|)\|Pu\| \leqslant g(\lambda, \|u\|)\{\|\Pi u\| + \|z(u)\|\},$$

where $g(\lambda, s) \to 0$ as $(\lambda, s) \to (\lambda_0, 0)$. Hence, for $(\lambda, u) \in B((\lambda_0, 0), r)$ with r sufficiently small and $F(\lambda, u) = 0$,

$$\|z(u)\| \leqslant \frac{2}{\beta} g(\lambda, \|u\|)\|\Pi u\|.$$

But, from (6.19) and $D_v\psi(\lambda_0, 0) = 0$, we also have that

$$\|\psi(\lambda, Pu)\| \leqslant h(\lambda, \|u\|)\|Pu\| \leqslant h(\lambda, \|u\|)\{\|\Pi u\| + \|z(u)\|\}$$

where $h(\lambda, s) \to 0$ as $(\lambda, s) \to (\lambda_0, 0)$ and so we obtain

$$\|z(u) + \psi(\lambda, u)\| \leqslant \left\{\frac{2}{\beta} g(\lambda, \|u\|) + h(\lambda, \|u\|) + \frac{2}{\beta} g(\lambda, \|u\|) h(\lambda, \|u\|)\right\}\|\Pi u\|,$$

proving (6.16). □

6.4. Bifurcation in Hilbert space

In the case where X and Y are Hilbert spaces and the linearization is self-adjoint, the condition (6.1) can be expressed in a more transparent form. Finally, for equations of the form $G(u) = \lambda u$, we can formulate a special case of the main result using only standard notions. We refer the reader to § 5.1.1 for the notation and results concerning graph norms.

COROLLARY 6.8. *Let $(Y, \langle\cdot,\cdot\rangle, \|\cdot\|)$ be a real Hilbert space and let X be the graph space of some self-adjoint operator acting in Y. Suppose that the hypotheses (B1)–(B4) are satisfied and that $L(\lambda_0) = D_x F(\lambda_0, 0)\colon X \subset Y \to Y$ is a self-adjoint operator acting in Y. Suppose also that*

$$\sup_{\varepsilon \in (0,1]} L_x^\varepsilon(R, \lambda_0) < d(0, \sigma_e(L(\lambda_0))), \tag{6.20}$$

where $L_x^\varepsilon(R, \lambda_0)$ denotes the partial uniform Lipschitz modulus of R with respect to the graph norm $\|\cdot\|_{\varepsilon L(\lambda_0)}$ on X.

(i) *If λ_0 is an isolated singular point of $D_x F(\cdot, 0)$ with $\sigma(D_x F(\cdot, 0), \lambda_0) = -1$, then λ_0 is a bifurcation point for $F(\lambda, u) = 0$. If (6.11) is also satisfied, there exists a connected subset $\mathcal{C}$ of $\mathcal{S} = \{(\lambda, u) \in U\colon F(\lambda, u) = 0$ and $u \neq 0\}$ such that $\bar{\mathcal{C}} \cap [\mathbb{R} \times \{0\}] = \{(\lambda_0, 0)\}$.*

(ii) *Suppose that (6.11) is also satisfied, that $L(\lambda) = D_x F(\lambda, 0)$ is differentiable at λ_0 and that*

$$\ker L(\lambda_0) = \operatorname{span}\{\phi\}, \quad \text{where } \|\phi\| = 1 \text{ and } \langle L'(\lambda_0)\phi, \phi\rangle \neq 0. \tag{6.21}$$

Then the hypotheses of part (i) are satisfied and there exists $\delta > 0$ such that, for $(\lambda, u) \in B((\lambda_0, 0), \delta)$ and $F(\lambda, u) = 0$,

$$u = \langle u, \phi\rangle\{\phi + w(u)\}, \quad \text{where } \langle w(u), \phi\rangle = 0 \text{ and } \lim_{\|u\|_X \to 0} \|w(u)\|_X = 0.$$

(iii) *If $\ker D_x F(\lambda_0, 0) = \{0\}$ and $\lim_{\delta \to 0} \Delta_\delta(F, \lambda_0) = 0$, then λ_0 is not a bifurcation point for $F(\lambda, u) = 0$.*

https://doi.org/10.1017/S0308210513000486

Remark 6.9. If $\sigma_e(L(\lambda_0)) = \emptyset$, (6.20) means that $M \equiv \sup_{\varepsilon\in(0,1]} L_x^{\varepsilon}(R, \lambda_0) < \infty$.

Remark 6.10. The condition (6.20) is satisfied when $R = R_1 + R_2$, where

(1) there exists an open neighbourhood, U, of $(\lambda_0, 0)$ in $\mathbb{R} \times X$ such that $R_1(\lambda, \cdot)$ is Gâteaux differentiable at u for all $(\lambda, u) \in U$ and

$$\|D_x R_1(\lambda, u) - D_x R_1(\lambda_0, 0)\|_{B(X,Y)} \to 0 \quad \text{as } (\lambda, u) \to (\lambda_0, 0),$$

and

(2) $L_x^Y(R_2, \lambda_0) < d(0, \sigma_e(L(\lambda_0)))$, where $L_x^Y(R_2, \lambda_0)$ is the partial uniform Lipschitz modulus of R_2 with respect to the norm $\|\cdot\|_Y$ on both spaces X and Y.

Indeed, (B1) and (2) imply that $R_1(\lambda, \cdot)$ is continuous on a neighbourhood of 0. Noting that (1) ensures the same property for all equivalent norms on X, lemma 3.3 implies that $L_x^{\varepsilon}(R_1, \lambda_0) = 0$ for all $\varepsilon > 0$. Since $\|\cdot\|_Y \leqslant \|\cdot\|_{\varepsilon L(\lambda_0)}$ for all $\varepsilon > 0$, $L_x^{\varepsilon}(R_2, \lambda_0) \leqslant L_x^Y(R_2, \lambda_0)$, and hence (2) implies that

$$L_x^{\varepsilon}(R, \lambda_0) \leqslant L_x^{\varepsilon}(R_1, \lambda_0) + L_x^{\varepsilon}(R_2, \lambda_0) < d(0, \sigma_e(L(\lambda_0)))$$

for all $\varepsilon > 0$, as required.

Proof of corollary 6.8. By corollary 5.6, there exists a constant $K > 0$ such that

$$\gamma^{\varepsilon}(L(\lambda_0)) \leqslant \frac{1}{d(0, \sigma_e(L(\lambda_0)))} + \varepsilon K \quad \text{for all } \varepsilon > 0,$$

where γ^{ε} denotes the essential conditioning number of $L(\lambda_0)\colon X \to Y$ calculated using the graph norm $\|\cdot\|_{\varepsilon L(\lambda_0)}$ on X.

Setting $M = \sup_{\varepsilon\in(0,1]} L_x^{\varepsilon}(R, \lambda_0)$, choose $\varepsilon \in (0, 1]$ such that

$$\varepsilon K < \frac{1}{M} - \frac{1}{d(0, \sigma_e(L(\lambda_0)))}.$$

Then

$$\gamma^{\varepsilon}(L(\lambda_0)) L_x^{\varepsilon}(R, \lambda_0) \leqslant \left\{\frac{1}{d(0, \sigma_e(L(\lambda_0)))} + \varepsilon K\right\} M < 1.$$

Thus, (6.1) is satisfied and the conclusions of parts (i) and (iii) follow from theorem 6.3.

For part (ii), we note that the scalar product for the norm $\|\cdot\|_{\varepsilon L(\lambda_0)}$ is

$$\langle u, v\rangle_{\varepsilon} = \langle u, v\rangle + \varepsilon^2 \langle L(\lambda_0)u, L(\lambda_o)v\rangle \text{ for } u, v \in X$$

and that $L(\lambda_0)\phi = 0$. Hence, $\langle u, \phi\rangle_{\varepsilon} = \langle u, \phi\rangle$ for all $u \in X$ and $\|\phi\|_{\varepsilon L(\lambda_0)} = \|\phi\| = 1$. Thus, for any $\varepsilon > 0$, the orthogonal projection of X onto span $\{\phi\}$ for the norm $\|\cdot\|_{\varepsilon L(\lambda_0)}$ is given by $\mathit{\Pi} u = \langle u, \phi\rangle\phi$ and the conclusion (ii) follows from Remark 6.6. □

Recalling that (6.11) is trivially satisfied when F has the form $F(\lambda, u) = M(u) - \lambda u$, we obtain the following special case.

Corollary 6.11. *Let $(Y, \langle\cdot,\cdot\rangle, \|\cdot\|_Y)$ be a real Hilbert space and let X be the graph space of some self-adjoint operator acting in Y. Let $B_X(0,\delta) = \{u \in X : \|u\|_X < \delta\}$. Consider the equation $M(u) = \lambda u$, where the function $M : X \to Y$ has the following properties.*

(H1) $M(0) = 0$.

(H2) *M is Gâteaux differentiable at 0 and $M'(0) \in B(X, Y)$ is a self-adjoint operator acting in Y with domain X.*

(H3) *For some $\delta > 0$, $M = M_1 + M_2$, where $M_1 \in C^1(B_X(0,\delta), Y)$ with $M_1'(0) = M'(0)$ and there exists a constant L such that $\|M_2(u) - M_2(v)\|_Y \leqslant L\|u - v\|_Y$ for all $u, v \in B_X(0,\delta)$. Let*

$$L^Y(M_2) = \lim_{\delta\to 0} \sup_{\substack{u,v\in B_X(0,\delta)\\ u\neq v}} \frac{\|M_2(u) - M_2(v)\|_Y}{\|u - v\|_Y} < \infty.$$

Then, for λ_0 such that $d(\lambda_0, \sigma_e(M'(0))) > L^Y(M_2)$ we have the following conclusions.

(i) *If $\ker\{M'(0) - \lambda_0 I\} = \{0\}$, λ_0 is not a bifurcation point.*

(ii) *If $\dim\ker\{M'(0) - \lambda_0 I\}$ is odd, λ_0 is a bifurcation point. Indeed, there exists a connected subset $\mathcal{C}$ of $\mathcal{S} = \{(\lambda, u) \in \mathbb{R} \times B_X(0,\delta) : M(u) = \lambda u \text{ and } u \neq 0\}$, with the metric inherited from $\mathbb{R} \times X$, such that $\bar{\mathcal{C}} \cap [\mathbb{R} \times \{0\}] = \{(\lambda_0, 0)\}$.*

(iii) *If $\ker\{M'(0) - \lambda_0 I\} = \operatorname{span}\{\phi\}$, where $\|\phi\|_Y = 1$, λ_0 is a bifurcation point and, for any sequence of solutions $\{(\lambda_n, u_n)\} \subset \mathbb{R} \times X$ such that $\lambda_n \to \lambda_0$ and $\|u_n\|_X \to 0$, we have that $u_n = \langle u, \phi\rangle\{\phi + w_n\}$, where $\langle w_n, \phi\rangle = 0$ and $\|w_n\|_X \to 0$.*

Proof. Setting $F(\lambda, u) = M(u) - \lambda u$, it follows from (H3) that $F(\lambda, \cdot) : B_X(0,\delta) \to Y$ is Lipschitz continuous and so (H2) and (H3) imply that $F(\lambda, \cdot) : B_X(0,\delta) \subset X \to Y$ is Hadamard differentiable at 0 with $D_x F(\lambda, 0) = M'(0) - \lambda I$ for all $\lambda \in \mathbb{R}$. Setting $R_1 = M_1 - M'(0)$, we have that

$$R(\lambda, u) = F(\lambda, u) - D_x F(\lambda, 0)u = M(u) - M'(0) = R_1 + M_2,$$

where $R_1 \in C^1(B_X(0,\delta), Y)$ with $R_1'(0) = 0$. Using the graph norm $\|\cdot\|_{\varepsilon L(\lambda_0)}$ with $L(\lambda_0) = D_x F(\lambda_0, 0) = M'(0) - \lambda_0 I$ on X, it follows that $L_X^{\varepsilon}(R, \lambda_0) \leqslant L^Y(M_2) < d(\lambda_0, \sigma_e(M'(0))) = d(0, \sigma_e(L(\lambda_0)))$. Note that (6.13) is trivially satisfied since R does not depend on λ. The conclusions now follow from corollary 6.8. □

To end this section we give a simple example showing that, given hypotheses (H1)–(H3) of corollary 6.11, the condition $d(\lambda_0, \sigma_e(M'(0))) > L^Y(M_2)$ may be necessary and sufficient for conclusion (i) to hold. A similar situation occurs with respect to conclusions (ii) and (iii), but we shall illustrate this later in the more interesting context of elliptic equations on $\mathbb{R}^N$ (see [20, corollary 5.2]).

Example 6.12. Let $X = Y = L^2([0,1])$ and consider the Nemytskii operator $G\colon X \to X$ defined by $G(u)(x) = g(u(x))$ for $u \in X$, where $g(s) = s^2/(1+|s|)$ for $s \in \mathbb{R}$. Noting that $g \in C^1(\mathbb{R})$ with $g(0) = g'(0) = 0$ and that $|g'(s)| \leqslant 1$ for all $s \in \mathbb{R}$, we see that $G(u) \in X$ for all $u \in X$ with $\|G(u) - G(v)\| \leqslant \|u - v\|$ for all $u, v \in X$ and the usual norm on X. Furthermore, for all $v \in X$ and $t \in \mathbb{R} \setminus \{0\}$, $|G(tv)/t| \leqslant |v|$ and it follows by dominated convergence that $G\colon X \to X$ is Gâteaux differentiable at 0 with $G'(0) = 0$. Hence, the hypotheses (H1)–(H3) of corollary 6.11 are satisfied with $G_1 = 0$, $G_2 = G$ and $L^X(G_2) \leqslant 1$. In fact, $L^X(G_2) = 1$. To justify this assertion, consider the functions $u_n^k = n\chi_k$ for $n, k \in \mathbb{N}$, where χ_k denotes the characteristic function of the interval $(0, 1/k)$ for $k \geqslant 1$. Then $\|u_n^k\| = n/\sqrt{k}$, $\|u_n^k - u_{n-1}^k\| = 1/\sqrt{k}$ and

$$g(u_n^k) - g(u_{n-1}^k) = \{g(n) - g(n-1)\}\chi_k = \left\{1 - \frac{1}{n(n+1)}\right\}\chi_k,$$

so

$$\frac{\|G(u_n^k) - G(u_{n-1}^k)\|}{\|u_n^k - u_{n-1}^k\|} = 1 - \frac{1}{n(n+1)} \quad \text{for } n, k \geqslant 1.$$

Given $\varepsilon \in (0,1)$ and $\delta > 0$, we choose $n_\varepsilon \in \mathbb{N}$ such that $n_\varepsilon \geqslant 1/\sqrt{\varepsilon}$ and then we choose $k \in \mathbb{N}$ such that $k > (n_\varepsilon/\delta)^2$. Then $u_{n_\varepsilon}^k$ and $u_{n_\varepsilon - 1}^k \in B_X(0,\delta)$ and

$$\frac{\|G(u_{n_\varepsilon}^k) - G(u_{n_\varepsilon-1}^k)\|}{\|u_{n_\varepsilon}^k - u_{n\varepsilon-1}^k\|} \geqslant 1 - \varepsilon,$$

showing that $L^X(G_2) = L^X(G) \geqslant 1$, as required.

We now consider the equation $G(u) = \lambda u$ for $u \in X$ and we shall show that

(1) λ_0 is a bifurcation point iff $|\lambda_0| \leqslant 1$,

(2) $\sigma(G'(0)) = \sigma_e(G'(0)) = \{0\}$ so that $d(\lambda_0, \sigma_e(G'(0))) = |\lambda_0|$.

Thus, hypotheses (H1)–(H3) of corollary 6.11 are satisfied and $\ker(G'(0) - \lambda_0 I) = \{0\}$ for all $\lambda_0 \neq 0$. This shows that the condition $d(\lambda_0, \sigma_e(G'(0))) > L^X(G_2) = 1$ is sharp for conclusion (i) of corollary 6.11. The statement (2) is trivial since $G'(0) = 0$. For (1), we consider first a point λ_0 with $0 < |\lambda_0| < 1$ and we set $u_k = (\lambda_0/(1 - |\lambda_0|))\chi_k$ for $k \in \mathbb{N}$ using the notation introduced earlier. Then

$$\|u_k\| = \frac{|\lambda_0|}{1 - |\lambda_0|}\frac{1}{\sqrt{k}} \to 0$$

and $G(u_k) = \lambda_0 u_k$ for all k, showing that λ_0 is indeed a bifurcation point. It follows from this that all points in $[-1,1]$ are bifurcation points. On the other hand, if $G(u) = \lambda u$, then $|\lambda|\|u\| = \|G(u)\| \leqslant \|u\|$ and so either $|\lambda| \leqslant 1$ or $\|u\| = 0$. Therefore, the equation has no solutions with $u \not\equiv 0$ and $|\lambda| > 1$ and consequently all bifurcation points must belong to the interval $[-1,1]$, completing the justification of (1).

Acknowledgements

The work for this paper was started during a visit to the Oxford Centre for Nonlinear PDE (OxPDE), Oxford University in autumn 2011. It was pursued in the spring of 2012 at the Wuhan Institute of Physics and Mathematics at the Chinese Academy of Sciences while I was supported by the Chinese Academy of Sciences Visiting Professorship for Senior International Scientists (Grant no. Z010T1J10). It is with great pleasure that I record my sincere thanks to these institutions for their hospitality and support.

Appendix A.

The relation of parity to other criteria for bifurcation is thoroughly discussed in [10]. In our restricted setting, the following result is sufficient to establish the usual conditions ensuring that the local parity is -1. See also [1, § 4].

PROPOSITION A.1. *Let* $L \in C(J, \Phi_0(X,Y))$ *and let* λ_0 *be an isolated singular point of* L*. Let* $Y = F \oplus F_0$*, where* $\dim F < \infty$ *and* $F_0 \subset \operatorname{rge} L(\lambda_0)$ *with* P *the associated projection onto* F *and* $Q = I - P$*. Set* $E = \ker QL(\lambda_0)$*.*

(a) *Then* $\dim E = \dim F$ *and, for any* E_0 *such that* $X = E \oplus E_0$*, there exists* $\delta > 0$ *such that* $QL(\lambda)|_{E_0} \in \operatorname{Iso}(E_0, F_0)$ *for all* $\lambda \in J_\delta = (\lambda_0 - \delta, \lambda_0 + \delta)$*.*

For $\lambda \in J_\delta$*, define* $A(\lambda)\colon E \to X$ *by* $A(\lambda)z = z - V(\lambda)QL(\lambda)z$*, where* $V(\lambda) = [QL(\lambda)|_{E_0}]^{-1}\colon F_0 \to E_0$*.*

(b) *Choose any* $U \in \operatorname{Iso}(F,E)$*. Then* λ_0 *is an isolated singular point of the path* $UPLA \in C(J_\delta, B(E,E))$ *and* $\sigma(L,\lambda_0) = \sigma(UPLA, \lambda_0)$*. Hence,* $\sigma(L,\lambda_0) = -1$ *if and only if* $\det[UPL(\lambda)A(\lambda)]$ *changes sign as* λ *passes through* λ_0*.*

Proof. (a) The projection $Q\colon Y \to F_0$ is a Fredholm operator of index

$$n = \dim \ker Q = \dim F.$$

Hence, $QL(\lambda)\colon X \to F_0$ is a Fredholm operator with $\operatorname{ind}(QL(\lambda)) = \operatorname{ind} Q + \operatorname{ind} L(\lambda) = n$ for all $\lambda \in J$. Since $\operatorname{rge} QL(\lambda_0) = F_0$, it follows that $\dim\ker QL(\lambda_0) = n$ and we can choose E_0 such that $X = E \oplus E_0$. Then $QL(\lambda_0)|_{E_0} \to F_0$ is an isomorphism and so by continuity of $\lambda \mapsto QL(\lambda)|_{E_0}$, there exists $\delta > 0$ such that $QL(\lambda)|_{E_0} \in \operatorname{Iso}(E_0, F_0)$ for all $\lambda \in J_\delta$.

(b) Choose some $U \in \operatorname{Iso}(F,E)$ and consider the operator $H(\lambda)\colon Y \to X$ defined by $H(\lambda)y = UPy + V(\lambda)Qy$ for $\lambda \in J_\delta$. Clearly, $H \in C(J_\delta, B(Y,X))$ and it is easy to check that $H(\lambda) \in \operatorname{Iso}(Y,X)$ for all $\lambda \in J_\delta$.

Next define $B(\lambda)\colon X \to X$ by $B(\lambda) = A(\lambda)\Pi + (I - \Pi)$, where Π is the projection onto E associated with the decomposition $X = E \oplus E_0$. Clearly, $A \in C(J_\delta, B(E,X))$ and $B \in C(J_\delta, B(X,X))$. Furthermore, since $E = \ker QL(\lambda_0)$, $A(\lambda_0)z = z$ for all $z \in E$ and, consequently, $B(\lambda_0) = I_X$. Hence, by reducing δ if necessary, we can assume that $B(\lambda) \in \operatorname{Iso}(X,X)$ for all $\lambda \in J_\delta$. It follows by property (c) of the local parity that λ_0 is an isolated singular point of $HLB \in C(J_\delta, \Phi_0(X,X))$ with

$$\sigma(HLB, \lambda_0) = \sigma(H,\lambda_0)\sigma(L,\lambda_0)\sigma(B,\lambda_0) = \sigma(L,\lambda_0)$$

since $H(\lambda_0) \in \mathrm{Iso}(Y, X)$ and $B(\lambda_0) \in \mathrm{Iso}(X, X)$. Note also that, for $z \in E$, $QL(\lambda)A(\lambda)z = 0$ for all $\lambda \in J_\delta$ so that

$$\begin{aligned} H(\lambda)L(\lambda)B(\lambda) &= UPL(\lambda)A(\lambda)\Pi + UPL(\lambda)(I - \Pi) + V(\lambda)QL(\lambda)(I - \Pi) \\ &= UPL(\lambda)A(\lambda)\Pi + UPL(\lambda)(I - \Pi) + I - \Pi. \end{aligned}$$

Hence, $\mathrm{rge}(I - H(\lambda)L(\lambda)B(\lambda)) \subset E$ and

$$\det H(\lambda)L(\lambda)B(\lambda) = \det H(\lambda)L(\lambda)B(\lambda)|_E = \det UPL(\lambda)A(\lambda).$$

Furthermore, for $z \in E$, $UPL(\lambda)A(\lambda)z = 0$ implies that $PL(\lambda)A(\lambda)z = 0$ and, recalling that $QL(\lambda)A(\lambda)z = 0$ for all $z \in E$, we get $L(\lambda)A(\lambda)z = 0$. Since λ_0 is an isolated singular point of L, we may deduce that $A(\lambda)z = 0$ for $0 < |\lambda - \lambda_0| < \delta$, by reducing the size of δ if necessary. But $A(\lambda)z = 0$ implies that $z = V(\lambda)QL(\lambda)z \in E \cap E_0 = \{0\}$ and we have shown that $UPL(\lambda)A(\lambda) \in B(E, E)$ is an isomorphism for $0 < |\lambda - \lambda_0| < \delta$. Hence, $\sigma(HLB, \lambda_0) = -1$ if and only if $\det UPL(\lambda)A(\lambda)$ changes sign as λ passes through λ_0. □

Proof of criterion 1. We can write

$$L(\lambda) = L(\lambda_0) + (\lambda - \lambda_0)\rho(\lambda),$$

where $\rho \in C(J, B(X, Y))$ and $\rho(\lambda_0) = L'(\lambda_0)$. Using proposition A.1 with $F_0 = \mathrm{rge}\, L(\lambda_0)$, we have $E = \ker L(\lambda_0)$ and $\sigma(L, \lambda_0) = \sigma(UPLA, \lambda_0)$, where $PL(\lambda_0) = 0$ since $Q = I - P$ is now a projection onto $\mathrm{rge}\, L(\lambda_0)$. Hence,

$$UPL(\lambda)A(\lambda) = (\lambda - \lambda_0)UP\rho(\lambda)A(\lambda)$$

and

$$\det UPL(\lambda)A(\lambda) = (\lambda - \lambda_0)^n \det UP\rho(\lambda)A(\lambda).$$

But $A(\lambda_0) = I_E$ and so, for $z \in E$, $UP\rho(\lambda_0)A(\lambda_0)z = 0$ implies that $PL'(\lambda_0)z = 0$. That is, $L'(\lambda_0)z \in \mathrm{rge}\, L(\lambda_0)$ and therefore $z = 0$ by (1). Hence, $UP\rho(\lambda_0)A(\lambda_0) \in \mathrm{Iso}(E, E)$. Consequently, $\det UP\rho(\lambda)A(\lambda)$ does not change sign as λ passes through λ_0 and the result follows. □

Proof of criterion 2. From the spectral theory of compact linear operators we have that $L(\lambda) \in \Phi_0(X, X)$ for all $\lambda \neq 0$. Setting

$$N = \bigcup_{j=1}^{\infty} \ker(K - \lambda_0 I)^j \quad \text{and} \quad R = \bigcap_{j=1}^{\infty} \mathrm{rge}(K - \lambda_0 I)^j$$

we have $\dim N = n$, $K(N) \subset N$, $K(R) \subset R$, $X = N \oplus R$, $L(\lambda_0)|_N$ is nilpotent and $L(\lambda_0)|_R \in \mathrm{Iso}(R, R)$.

We can use proposition A.1 with $X = Y$ and $F_0 = R$. Then we can choose $F = N$ and we find that $E = N$ so we can also choose $E_0 = R$ and $U = I_E$. Furthermore, for $z \in E$, we have $L(\lambda)z \in E$ for all λ and consequently $A(\lambda)z = z$ since $QL(\lambda)z = 0$. Hence,

$$UPL(\lambda)A(\lambda) = PL(\lambda)|_E = (K - \lambda I)|_N.$$

Since $\sigma((K - \lambda_0 I)|_N) = \{0\}$, $\sigma((K - \lambda I)|_N) = \{\lambda_0 - \lambda\}$ and $\det UPL(\lambda)A(\lambda) = (\lambda_0 - \lambda)^n$ from which the conclusion follows. □

References

1 P. Benevieri, M. Furi, M. P. Pera and M. Spadini. About the sign of oriented Fredholm operators between Banach spaces. *Z. Analysis Anwend.* **22** (2003), 619–645.

2 A. L. Dontchev and R. T. Rockafellar. *Implicit functions and solution mappings* (Springer, 2009).

3 D. E. Edmunds and W. D. Evans. *Spectral theory and differential operators* (Oxford University Press, 1987).

4 G. Evéquoz and C. A. Stuart. On differentiability and bifurcation. *Adv. Math. Econ.* **8** (2006), 155–184.

5 G. Evéquoz and C. A. Stuart. Bifurcation points of a degenerate elliptic boundary-value problem. *Rend. Lincei Mat. Appl.* **17** (2006), 309–334.

6 G. Evéquoz and C. A. Stuart. Bifurcation and concentration of radial solutions of a nonlinear degenerate elliptic eigenvalue problem. *Adv. Nonlin. Studies* **6** (2006), 215–232.

7 G. Evéquoz and C. A. Stuart. Hadamard differentiability and bifurcation. *Proc. R. Soc. Edinb.* A **137** (2007), 1249–1285.

8 P. M. Fitzpatrick. Homotopy, linearization and bifurcation. *Nonlin. Analysis TMA* **12** (1988), 171–184.

9 P. M. Fitzpatrick and J. Pejsachowicz. A local bifurcation theorem for C^1-Fredholm maps. *Proc. Am. Math. Soc.* **109** (1990), 995–1002.

10 P. M. Fitzpatrick and J. Pejsachowicz. Parity and generalized multiplicity. *Trans. Am. Math. Soc.* **326** (1991), 281–305.

11 T. M. Flett. *Differential analysis* (Cambridge University Press, 1980).

12 D. Henry. Differential calculus in Banach space. Lecture notes, ch. 3. (Available at http://www.ime.usp.br/map/dhenry/danhenry/main.htm.)

13 K. Kuratowski. *Topology*, vol. 2 (New York: Academic Press, 1968).

14 P. J. Rabier. Bifurcation in weighted spaces. *Nonlinearity* **21** (2008), 841–856.

15 P. H. Rabinowitz. Some global results for nonlinear eigenvalue problems. *J. Funct. Analysis* **7** (1971), 487–513.

16 C. A. Stuart. Bifurcation for some non-Fréchet differentiable problems. *Nonlin. Analysis TMA* **69** (2008), 1011–1024.

17 C. A. Stuart. Bifurcation and decay of solutions for a class of elliptic equations on $\mathbb{R}^N$. *Contemp. Math.* **540** (2011), 203–230.

18 C. A. Stuart. Asymptotic linearity and Hadamard differentiability. *Nonlin. Analysis* **75** (2012), 4699–4710.

19 C. A. Stuart. Asymptotic bifurcation and second order elliptic equations on $\mathbb{R}^N$. Preprint, 2012.

20 C. A. Stuart. Bifurcation at isolated eigenvalues for some elliptic equations on $\mathbb{R}^N$. In *Progress in nonlinear differential equations*, vol. 85, pp. 423–443 (Springer, 2014).

21 G. T. Whyburn. *Topological analysis* (Princeton University Press, 1958).

22 K. Yosida. *Functional analysis* (Springer, 1966).

(*Issued 3 October 2014*)