A normal form for the fold bifurcation and its discretization

Lajos Lóczi*
Department of Mathematics,
Budapest University of Technology and Economics,
H-1521 Budapest, Hungary

Abstract

In the first part of the paper, normal forms for the time-h-map of an ordinary differential equation and its discretization near a fold bifurcation point in one dimension are derived together with suitable closeness estimates. These steps will pave the way for an anticipated generalization of the results in [1]. The second, complementary part of the paper shows that implicit Runge-Kutta methods completely preserve the fold as well as the cusp bifurcation conditions in N dimension.

1 Introduction

Consider the ordinary differential equation

$$\dot{x} = f(x, \alpha) \tag{1}$$

together with its discretization

$$x_{n+1} := \varphi(h, x_n, \alpha), \qquad n = 0, 1, 2, \ldots, \tag{2}$$

where $\alpha \in \mathbb{R}$ is a scalar bifurcation parameter, $h > 0$ is the step-size of the sufficiently smooth one-step method $\varphi : \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ of order $p \geq 1$, and the function $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is of class C^{p+k+1} with $k \geq 5$ and uniformly bounded derivatives.

By the definition of the order of the method, we have that

$$|\Phi(h, x, \alpha) - \varphi(h, x, \alpha)| \leq const \cdot h^{p+1}, \quad \forall h \in [0, h_0], \forall |x| \leq x_0, \forall |\alpha| \leq \alpha_0, \tag{3}$$

where $\Phi(h, \cdot, \alpha) : \mathbb{R} \to \mathbb{R}$ is the time-h-map of the solution flow induced by (1) at parameter value α, further h_0, x_0 and α_0 are some (small) positive constants. (Throughout the paper, the symbols *const* will denote the general constants in the

*This research was supported by the DFG Research Group 'Spektrale Analysis, asymptotische Entwicklungen und stochastische Dynamik' at Bielefeld University, further the Hungarian Scientific Research Fund OTKA under Grant No. T037491.

estimates, with dependence only on f.)

Suppose that the origin $x = 0$, $\alpha = 0$ is an equilibrium as well as a *fold-bifurcation* point for (1), that is the following conditions hold

$$f^E = 0, \quad f_x^E = 0, \quad f_{xx}^E \neq 0, \quad f_\alpha^E \neq 0, \tag{4}$$

where—and throughout the paper also—subscripts h, x (or z), and α denote partial differentiation with respect to their corresponding variables, while superscript ^{E} denotes *evaluation at equilibrium*: that is, evaluation at $x = 0$ and $\alpha = 0$. (The evaluation operator is understood, of course, to have the lowest precedence, i.e., it is performed *after* taking all partial derivatives.)

Our central result is formulated at the end of Section 2: Lemma 2.2 and Theorem 2.5 provide the existence of some smooth invertible coordinate and parameter changes transforming the mappings $x \mapsto \Phi(h, x, \alpha)$ and $x \mapsto \varphi(h, x, \alpha)$ into their corresponding normal forms. As a consequence of (3), these coordinate and parameter changes as well as the normal forms themselves turn out to be $\mathcal{O}(h^p)$-close to each other. The normal forms and estimates will be needed in the construction of a *conjugacy* (sufficiently close to the identity) between the h-flow and the h-map near the fold point.

Section 3 is somewhat independent in character of the anticipated general conjugacy result mentioned above, and its intention is to demonstrate that in a center manifold reduction—required to carry over the results from the scalar case to the general N-dimensional case—shifting the equilibria into the origin is, at least with implicit Runge-Kutta methods, unnecessary, since it automatically takes place.

2 Derivation of the normal forms

In this section, we compute normal forms for the mappings

$$x \mapsto \Phi(h, x, \alpha) \tag{5}$$

and

$$x \mapsto \varphi(h, x, \alpha) \tag{6}$$

near the equilibrium, which is a fold-bifurcation point. Since now—as opposed to [1]—they both depend on h also, this extra parameter together with uniform estimates on $[0, h_0]$ should be built into the computations [4] we follow.

The properties of the solution flow together with (3) imply for $h \geq 0$, $|x| \leq x_0$ and $|\alpha| \leq \alpha_0$ that

$$\Phi(h, 0, 0) = 0, \tag{7}$$

$$\varphi(0, x, \alpha) = \Phi(0, x, \alpha) = x, \tag{8}$$

$$\Phi_h(h, x, \alpha) = f(\Phi(h, x, \alpha), \alpha), \tag{9}$$

$$\varphi_h(0, x, \alpha) = \Phi_h(0, x, \alpha). \tag{10}$$

Instead of (9), the shorter $\Phi_h = f \circ \Phi$ form will be used. We remark that the property $\varphi(h, 0, 0) = 0$ is *not* assumed here; nevertheless it often holds for discretizations, see Section 3.

Lemma 2.1 *Under the assumptions above and for $h \in [0, h_0]$, $|x| \le x_0$, $|\alpha| \le \alpha_0$, we have that*

$$\Phi(h, x, \alpha) = f_0(h, \alpha) + f_1(h, \alpha)x + f_2(h, \alpha)x^2 + \psi_3(h, x, \alpha)x^3,$$

where

$$\begin{aligned} f_0(h,\alpha) &= \Phi^E_{h\alpha} \cdot h\alpha + h\alpha^2 \cdot \psi_0(h,\alpha), \qquad \Phi^E_{h\alpha} \neq 0, \\ f_1(h,\alpha) &\equiv 1 + g(h,\alpha) = 1 + h\alpha \cdot \psi_1(h,\alpha), \\ f_2(h,\alpha) &= \frac{1}{2}\Phi^E_{hxx} \cdot h + h\alpha \cdot \psi_2(h,\alpha), \qquad \Phi^E_{hxx} \neq 0, \\ \psi_3(h,x,\alpha) &= h \cdot \widehat{\psi}_3(h,x,\alpha) \end{aligned}$$

hold with some smooth functions ψ_0, ψ_1, ψ_2 and $\widehat{\psi}_3$.

Proof. We expand Φ in a multivariate Taylor series at the equilibrium with the remainders in integral form. For f_0 we have that

$$f_0(h,\alpha) = \Phi^E + \alpha \cdot \mathrm{I}_{001}(\alpha) + h \cdot \mathrm{I}_{100}(h) + h\alpha \cdot \Phi^E_{h\alpha} +$$

$$+ h\alpha^2 \cdot \mathrm{I}_{102}(\alpha) + h^2\alpha \cdot \mathrm{I}_{201}(h) + h^2\alpha^2 \cdot \mathrm{I}_{202}(h,\alpha),$$

where—taking into account (4) and (7)–(9) repeatedly— we get that $\Phi^E = 0$, and

$$\mathrm{I}_{001}(\alpha) = \int_0^1 \Phi_\alpha(0,0,\tau\alpha)\mathrm{d}\tau \equiv 0,$$

$$\mathrm{I}_{100}(h) = \int_0^1 \Phi_h(\tau h,0,0)\mathrm{d}\tau \equiv 0.$$

Further, since

$$\Phi_{hh\alpha} = (f \circ \Phi)_{h\alpha} = ((f_x \circ \Phi) \cdot \Phi_h)_\alpha = (f_x \circ \Phi)_\alpha \cdot \Phi_h + (f_x \circ \Phi) \cdot \Phi_{h\alpha}$$

and

$$\Phi^E_{h\alpha} = (f \circ \Phi)^E_\alpha = (f_x \circ \Phi)^E \cdot \Phi^E_\alpha + (f_\alpha \circ \Phi)^E = 0 + f^E_\alpha \neq 0,$$

so we also have

$$\mathrm{I}_{201}(h) = \int_0^1 (1-\tau)\Phi_{hh\alpha}(\tau h,0,0)\mathrm{d}\tau \equiv 0,$$

and $\Phi^E_{h\alpha} \neq 0$. The explicit form of the smooth functions

$$\mathrm{I}_{102}(\alpha) = \int_0^1 (1-\tau)\Phi_{h\alpha\alpha}(0,0,\tau\alpha)\mathrm{d}\tau$$

and

$$\mathrm{I}_{202}(h,\alpha) = \int_0^1 \int_0^1 (1-\tau)(1-\sigma)\Phi_{hh\alpha\alpha}(\tau h,0,\sigma\alpha)\mathrm{d}\sigma\mathrm{d}\tau$$

will not play any role in the following, hence grouping together these remaining terms into ψ_0 gives the desired expression for f_0.

As for f_1, one gets that

$$f_1(h,\alpha) = \Phi^E_x + \alpha \cdot \mathrm{I}_{011}(\alpha) + h \cdot \mathrm{I}_{110}(h) + h\alpha \cdot \mathrm{I}_{111}(h,\alpha),$$

where $\Phi_x^E = 1$,

$$\mathrm{I}_{011}(\alpha) = \int_0^1 \Phi_{x\alpha}(0,0,\tau\alpha)\mathrm{d}\tau \equiv 0,$$

$$\mathrm{I}_{110}(h) = \int_0^1 \Phi_{hx}(\tau h,0,0)\mathrm{d}\tau \equiv 0,$$

because $\Phi_{hx} = (f\circ\Phi)_x = (f_x\circ\Phi)\cdot\Phi_x$. Finally,

$$\mathrm{I}_{111}(h,\alpha) = \int_0^1\int_0^1 \Phi_{hx\alpha}(\tau h,0,\sigma\alpha)\mathrm{d}\sigma\mathrm{d}\tau.$$

In the case of f_2, we see that

$$f_2(h,\alpha) = \frac{1}{2}\left(\Phi_{xx}^E + \alpha\cdot\mathrm{I}_{021}(\alpha) + h\cdot\Phi_{hxx}^E + h^2\cdot\mathrm{I}_{220}(h) + h\alpha\cdot\mathrm{I}_{121}(h,\alpha)\right),$$

where $\Phi_{xx}^E = 0$ and

$$\mathrm{I}_{021}(\alpha) = \int_0^1 \Phi_{xx\alpha}(0,0,\tau\alpha)\mathrm{d}\tau \equiv 0.$$

However,

$$\Phi_{hxx}^E = (f\circ\Phi)_{xx}^E = (f_{xx}\circ\Phi)^E\cdot\left((\Phi_x)^2\right)^E + (f_x\circ\Phi)^E\cdot\Phi_{xx}^E = f_{xx}^E\cdot 1 + 0 \neq 0.$$

Further,

$$\Phi_{hhxx} = (f_x\circ\Phi)_{xx}\cdot\Phi_h + 2(f_x\circ\Phi)_x\cdot\Phi_{hx} + (f_x\circ\Phi)\cdot\Phi_{hxx},$$

thus

$$\mathrm{I}_{220}(h) = \int_0^1 (1-\tau)\Phi_{hhxx}(\tau h,0,0)\mathrm{d}\tau \equiv 0.$$

Finally,

$$\mathrm{I}_{121}(h,\alpha) = \int_0^1\int_0^1 \Phi_{hxx\alpha}(\tau h,0,\sigma\alpha)\mathrm{d}\sigma\mathrm{d}\tau.$$

For the remainder ψ_3, the integral formula gives

$$\psi_3(h,x,\alpha) = \frac{1}{2}\int_0^1 (1-\tau)^2\Phi_{xxx}(h,\tau x,\alpha)\mathrm{d}\tau. \tag{11}$$

But

$$\Phi_{xxx}(h,\tau x,\alpha) = \Phi_{xxx}(0,\tau x,\alpha) + h\cdot\int_0^1 \Phi_{hxxx}(\sigma h,\tau x,\alpha)\mathrm{d}\sigma$$

and $\Phi_{xxx}(0,\tau x,\alpha)\equiv 0$, so the lemma is proved. ■

Now let us perform a coordinate shift by introducing a new variable

$$\xi := x + \delta_0,$$

where $\delta_0 \equiv \delta_0(h,\alpha)$ will be defined soon via the implicit function theorem. This shift transforms (5) into $\xi \mapsto \Phi(h,\xi-\delta_0,\alpha) + \delta_0$, which—similarly, but more explicitly than in [4]—turns out to be equal to

$$\xi \mapsto \Big[f_0(h,\alpha) - g(h,\alpha)\delta_0(h,\alpha) + f_2(h,\alpha)\delta_0^2(h,\alpha) + h\cdot\delta_0^3(h,\alpha)\widehat{\psi}_{30}(h,\alpha,\delta_0)\Big] +$$

$$+\xi + \xi \cdot \Big[g(h,\alpha) - 2f_2(h,\alpha)\delta_0(h,\alpha) + h \cdot \delta_0^2(h,\alpha)\widehat{\psi}_{31}(h,\alpha,\delta_0) \Big] + \tag{12}$$
$$+\xi^2 \cdot \Big[f_2(h,\alpha) + h \cdot \delta_0(h,\alpha)\widehat{\psi}_{32}(h,\alpha,\delta_0) \Big] + h \cdot \widehat{\psi}_{33}(h,\xi,\alpha,\delta_0)\xi^3$$

with some smooth functions $\widehat{\psi}_{30}$, $\widehat{\psi}_{31}$, $\widehat{\psi}_{32}$, and $\widehat{\psi}_{33}$, where

$$\widehat{\psi}_{30}(h,\alpha,\delta) \equiv -\widehat{\psi}_3(h,-\delta,\alpha), \tag{13}$$

$$\widehat{\psi}_{31}(h,\alpha,\delta) \equiv 3\widehat{\psi}_3(h,-\delta,\alpha) - \delta \cdot \frac{\mathrm{d}}{\mathrm{d}x}\widehat{\psi}_3(h,-\delta,\alpha), \tag{14}$$

$$\widehat{\psi}_{32}(h,\alpha,\delta) \equiv -3\widehat{\psi}_3(h,-\delta,\alpha) + 3\delta \cdot \frac{\mathrm{d}}{\mathrm{d}x}\widehat{\psi}_3(h,-\delta,\alpha) - \frac{\delta^2}{2} \cdot \frac{\mathrm{d}^2}{\mathrm{d}x^2}\widehat{\psi}_3(h,-\delta,\alpha) \tag{15}$$

and

$$\widehat{\psi}_{33}(h,\xi,\alpha,\delta) \equiv \frac{1}{2}\int_0^1 (1-\tau)^2 \cdot \Phi_{xxx}(h,\tau\xi-\delta,\alpha)\mathrm{d}\tau.$$

In order to annihilate the parameter-dependent linear term in (12), define

$$F(h,\alpha,\delta) \equiv \frac{1}{h}\Big(g(h,\alpha) - 2f_2(h,\alpha)\delta + h \cdot \delta^2 \cdot \widehat{\psi}_{31}(h,\alpha,\delta) \Big),$$

where, in the case of $h = 0$, the continuous extension of F is used. Since we have that

$$F(h,0,0) = 0 \qquad \forall h \in [0,h_0],$$
$$\frac{\partial F}{\partial \delta}(h,0,0) = \frac{-2f_2(h,0)}{h} = -\Phi^E_{hxx} \neq 0 \qquad \forall h \in [0,h_0],$$

the implicit function theorem provides the local existence and uniqueness of a smooth function $\delta_0(h,\alpha)$, defined on $h \in [0,h_0]$ and $|\alpha| \leq \alpha_0$, for which

$$F(h,\alpha,\delta_0(h,\alpha)) \equiv 0$$

holds. From uniqueness, it is seen that this δ_0 also satisfies $\delta_0(h,0) = 0$ for $h \in [0,h_0]$, so

$$\delta_0(h,\alpha) = \alpha \cdot \psi_d(h,\alpha) \tag{16}$$

holds true for $h \in [0,h_0]$ and $|\alpha| \leq \alpha_0$ with some smooth function ψ_d.

As a next step, introduce a new parameter $\mu_0 \equiv \mu_0(h,\alpha)$ by

$$\mu_0(h,\alpha) := \frac{f_0(h,\alpha)}{h} - \frac{g(h,\alpha)\delta_0(h,\alpha)}{h} + \frac{f_2(h,\alpha)\delta_0^2(h,\alpha)}{h} + \delta_0^3(h,\alpha)\widehat{\psi}_{30}(h,\alpha,\delta_0),$$

i.e., as the ξ-independent term of (12) divided by h. Since $\mu_0(h,0) = 0$ and $\frac{\mathrm{d}}{\mathrm{d}\alpha}\mu_0(h,0) = \Phi^E_{h\alpha} \neq 0$ independently of $h \in [0,h_0]$, the inverse function theorem guarantees the local existence and uniqueness of a smooth inverse function $\overline{\alpha}_0 \equiv \overline{\alpha}_0(h,\mu)$ of $\alpha \mapsto \mu_0(h,\alpha)$. Moreover, the domain of definition of this inverse function is easily seen to contain a neighbourhood of the origin independent of $h \in [0,h_0]$. Further, $\overline{\alpha}_0(h,0) = 0$, hence

$$\overline{\alpha}_0(h,\mu) = \mu \cdot \psi_a(h,\mu) \tag{17}$$

holds for $h \in [0,h_0]$ and $|\mu|$ small with some smooth function ψ_a.

Therefore (12) now reads

$$\xi \mapsto h \cdot \mu_0 + \xi + h \cdot q(h,\mu_0) \cdot \xi^2 + h \cdot \xi^3 \cdot \widehat{\psi}_{m3}(h,\xi,\mu_0)$$

with $q(h,\mu_0) \equiv \frac{1}{2}\Phi^E_{hxx} + \widehat{\psi}_{m2}(h,\mu_0)$ and some smooth functions $\widehat{\psi}_{m2}$ and $\widehat{\psi}_{m3}$, where

$$\widehat{\psi}_{m2}(h,\mu_0) \equiv \overline{\alpha}_0 \cdot \psi_2(h,\overline{\alpha}_0) + \delta_0(h,\overline{\alpha}_0) \cdot \widehat{\psi}_{32}(h,\overline{\alpha}_0,\delta_0(h,\overline{\alpha}_0))$$

and

$$\widehat{\psi}_{m3}(h,\xi,\mu_0) \equiv \widehat{\psi}_{33}(h,\xi,\overline{\alpha}_0,\delta_0(h,\overline{\alpha}_0)).$$

A final scaling $\eta := |q(h,\mu_0)|\xi$ and $\beta := |q(h,\mu_0)|\mu_0$ with $s := \mathrm{sign}(q(h,0)) = \pm 1$ (being also independent of $h \in [0,h_0]$) yields the following normal form.

Lemma 2.2 *There are smooth invertible coordinate and parameter changes transforming the system*

$$x \mapsto \Phi(h,x,\alpha)$$

into

$$\eta \mapsto h\beta + \eta + s \cdot h\eta^2 + h\eta^3 \cdot \widehat{\eta}_3(h,\eta,\beta)$$

where $\widehat{\eta}_3(h,\eta,\beta) = \widehat{\psi}_{m3}(h,\xi,\mu_0) \cdot |q(h,\mu_0)|^{-2}$ *is a smooth function.*

Now let us consider the discretization map φ. We prove an analogous result to that of Lemma 2.1 first.

Lemma 2.3 *Under the assumptions of Lemma 2.1 and for* $h \in [0,h_0]$, $|x| \le x_0$, $|\alpha| \le \alpha_0$, *we have that*

$$\varphi(h,x,\alpha) = \widetilde{f}_0(h,\alpha) + \widetilde{f}_1(h,\alpha)x + \widetilde{f}_2(h,\alpha)x^2 + \chi_3(h,x,\alpha)x^3,$$

where

$$\begin{aligned}
\widetilde{f}_0(h,\alpha) &= h^{p+1} \cdot \chi_{00}(h) + \varphi^E_{h\alpha} \cdot h\alpha + h\alpha \cdot \chi_{01}(h,\alpha), \quad \varphi^E_{h\alpha} = \Phi^E_{h\alpha} \neq 0,\\
\widetilde{f}_1(h,\alpha) &\equiv 1 + \widetilde{g}(h,\alpha) = 1 + h^{p+1} \cdot \chi_{10}(h) + h\alpha \cdot \chi_{11}(h,\alpha),\\
\widetilde{f}_2(h,\alpha) &= h^{p+1} \cdot \chi_{20}(h) + \frac{1}{2}\varphi^E_{hxx} \cdot h + h\alpha \cdot \chi_{21}(h,\alpha), \quad \varphi^E_{hxx} = \Phi^E_{hxx} \neq 0,\\
\chi_3(h,x,\alpha) &= h \cdot \widehat{\chi}_3(h,x,\alpha)
\end{aligned}$$

hold with some smooth functions χ_{00}, χ_{01}, χ_{10}, χ_{11}, χ_{20}, χ_{21} *and* $\widehat{\chi}_3$. *Moreover, for* $h \in [0,h_0]$, $|x| \le x_0$ *and for* $|\alpha| \le \alpha_0$,

$$|\psi_3(h,x,\alpha) - \chi_3(h,x,\alpha)| \le const \cdot h^{p+1}. \tag{18}$$

Proof. Proceeding similarly as in Lemma 2.1, we get that

$$\begin{aligned}
\widetilde{f}_0(h,\alpha) = \varphi^E + \alpha \cdot \widetilde{\mathrm{I}}_{001}(\alpha) + h \cdot \widetilde{\mathrm{I}}_{100}(h) + h\alpha \cdot \varphi^E_{h\alpha} +\\
+h\alpha^2 \cdot \widetilde{\mathrm{I}}_{102}(\alpha) + h^2\alpha \cdot \widetilde{\mathrm{I}}_{201}(h) + h^2\alpha^2 \cdot \widetilde{\mathrm{I}}_{202}(h,\alpha),
\end{aligned} \tag{19}$$

where the integrals $\widetilde{\mathrm{I}}$'s are defined just as in the proof of Lemma 2.1, but with φ instead of Φ. Due to (4)–(9), here we also have $\varphi^E = 0$ and $\widetilde{\mathrm{I}}_{001}(\alpha) \equiv 0$. From (3) at $x = 0$ we infer that for $h \in [0,h_0]$ and for $|\alpha| \le \alpha_0$

$$\left| f_0(h,\alpha) - \widetilde{f}_0(h,\alpha) \right| \le const \cdot h^{p+1}. \tag{20}$$

Evaluating this at $\alpha = 0$ shows that $|h \cdot \widetilde{\mathrm{I}}_{100}(h)| \le const \cdot h^{p+1}$. Further, differentiating (10) yields that $\varphi^E_{h\alpha} = \Phi^E_{h\alpha}$.

As for $\widetilde{f}_1$, one has that $\varphi_x^E = 1$ and $\widetilde{\mathrm{I}}_{011}(\alpha) \equiv 0$, hence

$$\widetilde{f}_1(h,\alpha) = 1 + h \cdot \widetilde{\mathrm{I}}_{110}(h) + h\alpha \cdot \widetilde{\mathrm{I}}_{111}(h,\alpha).$$

Since f is at least C^{p+4}, from [3] we obtain that

$$\left| f_1(h,\alpha) - \widetilde{f}_1(h,\alpha) \right| \leq const \cdot h^{p+1}. \tag{21}$$

Evaluation at $\alpha = 0$ yields $|h \cdot \widetilde{\mathrm{I}}_{110}(h)| \leq const \cdot h^{p+1}$.

Considering $\widetilde{f}_2$, we obtain that $\varphi_{xx}^E = 0$ and $\widetilde{\mathrm{I}}_{021}(\alpha) \equiv 0$, thus

$$\widetilde{f}_2(h,\alpha) = \frac{1}{2}\left(h \, \varphi_{hx}^E + h^2 \cdot \widetilde{\mathrm{I}}_{220}(h) + h\alpha \cdot \widetilde{\mathrm{I}}_{121}(h,\alpha) \right)$$

and again,

$$\left| f_2(h,\alpha) - \widetilde{f}_2(h,\alpha) \right| \leq const \cdot h^{p+1}. \tag{22}$$

Evaluating this at $\alpha = 0$, we see that $|h^2 \cdot \widetilde{\mathrm{I}}_{220}(h)| \leq const \cdot h^{p+1}$. Further, differentiating (10) again yields that $\varphi_{hxx}^E = \Phi_{hxx}^E$.

For the remainder χ_3, the same argument applies as in the proof of Lemma 2.1, together with the estimate

$$|\psi_3(h,x,\alpha) - \chi_3(h,x,\alpha)| \leq const \cdot h^{p+1} \cdot \frac{1}{2} \int_0^1 (1-\tau)^2 \mathrm{d}\tau,$$

which completes the proof of the lemma. ■

Now applying the corresponding coordinate shift with $\widetilde{\delta}$ instead of δ_0, we arrive at some formulae completely analogous to (12)–(15), where $\widetilde{\delta}$ is the implicit function defined by (the continuous extension at $h = 0$ of)

$$\widetilde{F}(h,\alpha,\delta) \equiv \frac{1}{h}\left(\widetilde{g}(h,\alpha) - 2\widetilde{f}_2(h,\alpha)\delta + h \cdot \delta^2 \cdot \widehat{\chi}_{31}(h,\alpha,\delta) \right).$$

However, for the $\mathcal{O}(h^p)$-estimates, we will need a *quantitative* (or parametrized) version of the *implicit function theorem*, see [8]. Instead of its full form (i.e. Banach space setting with more parameter-dependence), we restate that result in a simplified form tailored to our needs and using our notations.

Lemma 2.4 *Let $\widetilde{F} : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}$ be a C^j mapping. Assume there exist a function $\delta_0 : \mathbb{R}^2 \to \mathbb{R}$ and some constants $\kappa_1 > 0$, $\kappa_2 > 0$ such that for $|\delta - \delta_0(h,\alpha)| \leq r_1$ and $|h|, |\alpha| < r_2$ we have*

$$\left| \frac{\partial \widetilde{F}}{\partial \delta}(h,\alpha,\delta) - \frac{\partial \widetilde{F}}{\partial \delta}(h,\alpha,\delta_0(h,\alpha)) \right| \leq \kappa_2 < \kappa_1 \leq \left| \frac{\partial \widetilde{F}}{\partial \delta}(h,\alpha,\delta_0(h,\alpha)) \right|,$$

$$\left| \widetilde{F}(h,\alpha,\delta_0(h,\alpha)) \right| \leq (\kappa_1 - \kappa_2) \cdot r_1.$$

Then for any $|h|, |\alpha| < r_2$, $\widetilde{F}(h,\alpha,\cdot)$ has a unique C^j-smooth zero $\widetilde{\delta} \equiv \widetilde{\delta}(h,\alpha)$ near $\delta_0(h,\alpha)$, and the following estimate holds

$$\left| \widetilde{\delta}(h,\alpha) - \delta_0(h,\alpha) \right| \leq (\kappa_1 - \kappa_2)^{-1} \cdot |\widetilde{F}(h,\alpha,\delta_0(h,\alpha))|.$$

In order to verify the conditions of this lemma, define $\kappa_1 := \frac{1}{2}|\varphi^E_{hxx}|$ and $\kappa_2 := \frac{1}{2}\kappa_1$. The estimate

$$\left|\frac{\partial \widetilde{F}}{\partial \delta}(h,\alpha,\delta_0(h,\alpha))\right| =$$

$$= \left|\frac{-2\widetilde{f}_2(h,\alpha)}{h} + 2\delta_0(h,\alpha)\cdot\widehat{\chi}_{31}(h,\alpha,\delta_0) + \delta_0^2(h,\alpha)\cdot\frac{d}{d\delta}\widehat{\chi}_{31}(h,\alpha,\delta_0)\right| \geq \kappa_1$$

is seen to be valid—due to the form of $\widetilde{f}_2$ and (16)—provided that r_2 is small. On the other hand,

$$\left|\frac{\partial \widetilde{F}}{\partial \delta}(h,\alpha,\delta) - \frac{\partial \widetilde{F}}{\partial \delta}(h,\alpha,\delta_0(h,\alpha))\right| \leq$$

$$\leq |2[\delta-\delta_0(h,\alpha)]\widehat{\chi}_{31}(h,\alpha,\delta) + 2\delta_0(h,\alpha)\,[\widehat{\chi}_{31}(h,\alpha,\delta) - \widehat{\chi}_{31}(h,\alpha,\delta_0)]| +$$

$$+\left|\left[\delta^2-\delta_0^2(h,\alpha)\right]\frac{\mathrm{d}}{\mathrm{d}\delta}\widehat{\chi}_{31}(h,\alpha,\delta) + \delta_0^2(h,\alpha)\left[\frac{\mathrm{d}}{\mathrm{d}\delta}\widehat{\chi}_{31}(h,\alpha,\delta) - \frac{\mathrm{d}}{\mathrm{d}\delta}\widehat{\chi}_{31}(h,\alpha,\delta_0)\right]\right| \leq \kappa_2,$$

if r_1 and r_2 are sufficiently small. Finally,

$$\left|\widetilde{F}(h,\alpha,\delta_0(h,\alpha))\right| \leq |F(h,\alpha,\delta_0(h,\alpha))| + \left|\widetilde{F}(h,\alpha,\delta_0(h,\alpha)) - F(h,\alpha,\delta_0(h,\alpha))\right| \leq$$

$$\leq 0 + \frac{1}{h}\,|\widetilde{g}(h,\alpha) - g(h,\alpha)| + \frac{2|\delta_0(h,\alpha)|}{h}\left|\widetilde{f}_2(h,\alpha) - f_2(h,\alpha)\right| +$$

$$+\delta_0^2(h,\alpha)\left|\widehat{\chi}_{31}(h,\alpha,\delta_0) - \widehat{\psi}_{31}(h,\alpha,\delta_0)\right| \leq const\cdot h^p,$$

owing to (21), (22) and the estimate

$$\left|\widehat{\chi}_{31}(h,\alpha,\delta_0) - \widehat{\psi}_{31}(h,\alpha,\delta_0)\right| \leq \tag{23}$$

$$\leq 3\left|\widehat{\chi}_3(h,-\delta_0,\alpha) - \widehat{\psi}_3(h,-\delta_0,\alpha)\right| +$$

$$+|\delta_0(h,\alpha)|\left|\frac{1}{2h}\int_0^1 (1-\tau)^2\tau\cdot(\varphi_{xxxx}(h,-\tau\delta_0,\alpha) - \Phi_{xxxx}(h,-\tau\delta_0,\alpha))\,\mathrm{d}\tau\right| \leq$$

$$\leq const\cdot h^p + const\cdot h^p,$$

being valid due to (14), (11), (18) and the fact (see [3] again) that f is at least C^{p+5}.

Therefore, Lemma 2.4 proves the local existence and uniqueness of the function $\widetilde{\delta}$ such that

$$\widetilde{F}(h,\alpha,\widetilde{\delta}(h,\alpha)) \equiv 0,$$

and

$$\left|\widetilde{\delta}(h,\alpha) - \delta_0(h,\alpha)\right| \leq const\cdot h^p \tag{24}$$

holds for $h\in[0,h_0]$ and $|\alpha|\leq\alpha_0$.

Now let us define a new parameter $\widetilde{\mu}$ in an analogous way as we did before, i.e. as the ξ-independent term divided by h, that is

$$\widetilde{\mu}(h,\alpha) := \frac{\widetilde{f}_0(h,\alpha)}{h} - \frac{\widetilde{g}(h,\alpha)\widetilde{\delta}(h,\alpha)}{h} + \frac{\widetilde{f}_2(h,\alpha)\widetilde{\delta}^2(h,\alpha)}{h} + \widetilde{\delta}^3(h,\alpha)\widehat{\chi}_{30}(h,\alpha,\widetilde{\delta}).$$

We see from the analogous expression of (13) for $\widehat{\chi}_{30}$, from (18), (20)–(22) and (24) that

$$|\widetilde{\mu}(h,\alpha) - \mu_0(h,\alpha)| \le const \cdot h^p \tag{25}$$

holds for $h \in [0, h_0]$ and $|\alpha| \le \alpha_0$. In order to use a *quantitative inverse function theorem* for $\alpha \mapsto \widetilde{\mu}(h,\alpha)$, we apply Lemma 2.4 again, but this time with G instead of $\widetilde{F}$, and $\overline{\alpha}_0$ instead of δ_0, where

$$G(h,\mu,\alpha) := \mu - \widetilde{\mu}(h,\alpha).$$

To check the conditions of the lemma, define $\kappa_1 := \frac{1}{2}|\varphi^E_{h\alpha}|$ and $\kappa_2 := \frac{1}{2}\kappa_1$. We have that

$$\left|\frac{\partial G}{\partial \alpha}(h,\mu,\alpha)\right| =$$

$$= \left|\varphi^E_{h\alpha} + \chi_{01}(h,\alpha) + \alpha\frac{\mathrm{d}}{\mathrm{d}\alpha}\chi_{01}(h,\alpha) - \frac{\widetilde{g}(h,\alpha)}{h}\frac{\mathrm{d}}{\mathrm{d}\alpha}\widetilde{\delta}(h,\alpha) + \widetilde{\delta}(h,\alpha)\widetilde{\chi}_G(h,\alpha)\right|$$

holds with a suitable smooth function $\widetilde{\chi}_G$. By (24), (16) and (17) the expression $|\widetilde{\delta}(h,\overline{\alpha}_0(h,\mu))|$ can be made arbitrary small provided that $|h|, |\mu| < r_2$ are small enough. The same is true for $|\frac{1}{h}\widetilde{g}(h,\overline{\alpha}_0(h,\mu))|$. Moreover, the definition (19) of χ_{01} shows that $\chi_{01}(0,0) = 0$, so from these we can conclude that

$$\left|\frac{\partial G}{\partial \alpha}(h,\mu,\overline{\alpha}_0(h,\mu))\right| \ge \kappa_1,$$

provided that r_2 is sufficiently small. The other condition

$$\left|\frac{\partial G}{\partial \alpha}(h,\mu,\alpha) - \frac{\partial G}{\partial \alpha}(h,\mu,\overline{\alpha}_0(h,\mu))\right| \le \kappa_2$$

is seen to hold by continuity if $|\alpha - \overline{\alpha}_0(h,\mu)| \le r_1$ and r_2 are small enough. Finally, by (25),

$$|G(h,\mu,\overline{\alpha}_0(h,\mu))| = |\mu_0(h,\overline{\alpha}_0(h,\mu)) - \widetilde{\mu}(h,\overline{\alpha}_0(h,\mu))| \le const \cdot h^p.$$

Therefore, we get a unique zero $\widetilde{\alpha}(h,\mu)$ of $G(h,\mu,\cdot)$, which—by the construction of G—is just the inverse function of $\alpha \mapsto \widetilde{\mu}(h,\alpha)$. Furthermore,

$$|\widetilde{\alpha}(h,\mu) - \overline{\alpha}_0(h,\mu)| \le const \cdot h^p \tag{26}$$

holds for $h \in [0, h_0]$ and $|\mu|$ sufficiently small.

As a conclusion, (6) becomes

$$\widetilde{\xi} \mapsto h \cdot \widetilde{\mu} + \widetilde{\xi} + h \cdot \widetilde{q}(h,\widetilde{\mu}) \cdot \widetilde{\xi}^2 + h \cdot \widetilde{\xi}^3 \cdot \widehat{\chi}_{m3}(h,\widetilde{\xi},\widetilde{\mu})$$

with $\widetilde{q}(h,\widetilde{\mu}) \equiv \frac{1}{2}\varphi^E_{hxx} + \widehat{\chi}_{m2}(h,\widetilde{\mu})$ and some smooth functions $\widehat{\chi}_{m2}$ and $\widehat{\chi}_{m3}$. We claim that

$$|\widetilde{q}(h,\widetilde{\mu}) - q(h,\mu_0)| \le const \cdot h^p$$

also holds. Indeed, since

$$|\widetilde{q}(h,\widetilde{\mu}(h,\alpha)) - q(h,\mu_0(h,\alpha))| \le \left|\widetilde{f}_2(h,\alpha) - f_2(h,\alpha)\right| +$$

$$+h\left|\widetilde{\delta}(h,\alpha)\cdot\widehat{\chi}_{32}(h,\alpha,\widetilde{\delta}(h,\alpha)) - \delta_0(h,\alpha)\cdot\widehat{\psi}_{32}(h,\alpha,\delta_0(h,\alpha))\right|,$$

we deduce the desired estimate from the $\mathcal{O}(h^p)$-estimates obtained so far together with some standard (triangle) inequalities, and an estimate similar to (23) but with (15) and using that f is at least C^{p+6}.

By applying a final scaling

$$\widetilde{\eta} := |\widetilde{q}(h,\widetilde{\mu})|\widetilde{\xi} \quad \text{and} \quad \widetilde{\beta} := |\widetilde{q}(h,\widetilde{\mu})|\widetilde{\mu}$$

with $s := \text{sign}(\widetilde{q}(h,0)) = \pm 1$ (being independent of $h \in [0,h_0]$), further taking into account the fact that $|\xi - \widetilde{\xi}|$, $|\eta - \widetilde{\eta}|$ and $|\beta - \widetilde{\beta}|$ are all $\mathcal{O}(h^p)$-small, we have derived the following normal form together with the desired closeness estimates.

Theorem 2.5 *There are smooth invertible coordinate and parameter changes transforming the system*

$$x \mapsto \varphi(h,x,\alpha)$$

into

$$\widetilde{\eta} \mapsto h\widetilde{\beta} + \widetilde{\eta} + s \cdot h\widetilde{\eta}^2 + h\widetilde{\eta}^3 \cdot \widetilde{\eta}_3(h,\widetilde{\eta},\widetilde{\beta})$$

where $\widetilde{\eta}_3$ is a smooth function.

Moreover, the smooth invertible coordinate and parameter changes above and those in Lemma 2.2 are $\mathcal{O}(h^p)$-close to each other, further

$$|\widehat{\eta}_3 - \widetilde{\eta}_3| \leq const \cdot h^p.$$

3 Preservation of bifurcations under Runge-Kutta methods

In this section we show that the conditions for the *fold bifurcation* and for the *cusp bifurcation* in N dimension are preserved by *implicit Runge-Kutta methods.*

Throughout the section, $\text{null}(A)$ and $\text{ran}(A)$ denote the *null space* and the *range* of the linear operator A, respectively. The evaluation operator ^{E} again evaluates functions at $z = 0$, $\alpha = 0$, and also at $h > 0$, when it applies. Finally, the $N \times N$ identity matrix is denoted by I_N.

Consider the ordinary differential equation

$$\dot{z} = f(z,\alpha) \tag{27}$$

depending on a parameter $\alpha \in \mathbb{R}$. Suppose that the smooth function $f : \mathbb{R}^N \times \mathbb{R} \to \mathbb{R}^N$ has a fold bifurcation [5] at the equilibrium $z = 0$, $\alpha = 0$, that is the following conditions are satisfied:

- $f^E = 0$,
- $\dim \text{null}(f_z^E) = 1$,
- $f_\alpha^E \notin \text{ran}(f_z^E)$,
- $f_{zz}^E(v,v) \notin \text{ran}(f_z^E)$, where $\text{null}(f_z^E)$=span(v).

Now consider a discretization $\varphi(h,z,\alpha)$ of the above equation, with the function $\varphi : \mathbb{R}^+ \times \mathbb{R}^N \times \mathbb{R} \to \mathbb{R}^N$ coming from an *s-stage implicit Runge-Kutta method with step-size* $h > 0$, that is

$$z_{n+1} := \varphi(h, z_n, \alpha), \qquad n = 0, 1, 2, \ldots \tag{28}$$

where

$$\varphi(h,z,\alpha) \equiv z + h \sum_{i=1}^{s} \gamma_i \cdot k_i(h,z,\alpha),$$

and every function k_i $(i = 1, 2, \ldots, s)$ satisfies the implicit equation

$$k_i(h,z,\alpha) = f(z + h \sum_{j=1}^{s} \beta_{ij} \cdot k_j(h,z,\alpha), \alpha) \tag{29}$$

with some γ_i, β_{ij} $(i,j = 1, 2, \ldots, s)$ given real constants.

The origin $z = 0$, $\alpha = 0$ is a fold bifurcation point for the map $\varphi(h,\cdot,\cdot)$, if the following conditions hold:

- $\varphi^E \equiv \varphi(h,0,0) = 0$,
- $\dim \text{null}(\varphi_z^E - I_N) = 1$,
- $\varphi_\alpha^E \notin \text{ran}(\varphi_z^E - I_N)$,
- $\varphi_{zz}^E(v,v) \notin \text{ran}(\varphi_z^E - I_N)$, where $\text{null}(\varphi_z^E - I_N)$=span$(v)$.

Proposition 3.1 *Suppose that the equation (27) has a fold bifurcation at the equilibrium* $z = 0$, $\alpha = 0$, *and* $\Gamma := \sum_{i=1}^{s} \gamma_i \neq 0$. *Then the map (28) also has a fold bifurcation at* $z = 0$, $\alpha = 0$ *for* $h > 0$ *sufficiently small.*

Remark. It is well known that the condition $\Gamma = 1$ is necessary for a Runge-Kutta method to be of order at least one, hence the above assumption on Γ is natural.

Proof of the proposition. Step 1. The first, well-known property follows from the fact [7] that for h small enough, there is a locally unique solution to the defining system of equations (29) for the functions k_i, which is seen to be $k_i^E \equiv k_i(h,0,0) = f^E = 0$ for every $i = 1, 2, \ldots, s$.

Step 2. Next we show that $\text{null}(f_z^E) \subset \text{null}\big((k_i)_z^E\big)$ for all $i = 1, 2, \ldots, s$. To this end, choose $0 \neq v \in \text{null}(f_z^E)$ and use (29) to obtain for every i that

$$(k_i)_z^E v = f_z^E \left(I_N + h \sum_{j=1}^{s} \beta_{ij} \cdot (k_j)_z^E \right) v = f_z^E h \sum_{j=1}^{s} \beta_{ij} \cdot (k_j)_z^E v,$$

that is

$$(k_i)_z^E v - h \sum_{j=1}^{s} \beta_{ij} \cdot f_z^E (k_j)_z^E v = 0, \qquad i = 1, 2, \ldots, s.$$

Notice that these s equations can be represented by a single matrix equation as

$$\left(I_{N\cdot s} - h\cdot\beta\otimes f_z^E\right)\begin{pmatrix}(k_1)_z^E v\\(k_2)_z^E v\\\vdots\\(k_s)_z^E v\end{pmatrix} = 0\in\mathbb{R}^{N\cdot s},$$

where we have used the Kronecker product $\otimes$ of the matrices $\beta := [\beta_{ij}]\in\mathbb{R}^{s\times s}$ and f_z^E. However, for small h, the matrix $I_{N\cdot s} - h\cdot\beta\otimes f_z^E$ is invertible, hence $(k_i)_z^E v = 0$ for every $i = 1, 2, \ldots, s$, and the assertion follows.

Step 3. The previous step also proves that $\mathrm{null}(f_z^E)\subset\mathrm{null}(\varphi_z^E - I_N)$, since for any $v\in\mathrm{null}(f_z^E)$ we have that

$$(\varphi_z^E - I_N)v = \left(h\sum_{i=1}^{s}\gamma_i\cdot f_z^E\left(I_N + h\sum_{j=1}^{s}\beta_{ij}\cdot(k_j)_z^E\right)\right)v =$$

$$= h\Gamma\cdot f_z^E v + h^2\cdot f_z^E\sum_{i=1}^{s}\sum_{j=1}^{s}\gamma_i\beta_{ij}\cdot(k_j)_z^E v = 0.$$

Step 4. In order to prove that $\mathrm{null}(\varphi_z^E - I_N)$ is in fact one dimensional, choose an arbitrary nonzero vector w from this subspace. A similar rearrangement as in the previous step shows that

$$0 = (\varphi_z^E - I_N)w = h\cdot f_z^E A w,$$

where we have used the abbreviation

$$A\equiv\Gamma\cdot I_N + h\sum_{i=1}^{s}\sum_{j=1}^{s}\gamma_i\beta_{ij}\cdot(k_j)_z^E.$$

Therefore, $Aw\in\mathrm{null}(f_z^E)\subset\mathrm{null}\left((k_j)_z^E\right)$ for all $j = 1, 2, \ldots, s$, which implies that

$$AAw = \Gamma\cdot Aw + h\sum_{i=1}^{s}\sum_{j=1}^{s}\gamma_i\beta_{ij}\cdot(k_j)_z^E Aw = \Gamma\cdot Aw.$$

But A is invertible, because $\Gamma\neq 0$ and h is small, so we have

$$Aw = \Gamma\cdot w,$$

which shows that $w\in\mathrm{null}(f_z^E)$, and also that $\mathrm{null}(\varphi_z^E - I_N) = \mathrm{null}(f_z^E)$.

Step 5. As for the first range condition $\varphi_\alpha^E\notin\mathrm{ran}(\varphi_z^E - I_N)$, suppose to the contrary that there exists a vector $w\in\mathbb{R}^N$ such that $\varphi_\alpha^E = (\varphi_z^E - I_N)w$ holds. This is equivalent to saying that

$$h\sum_{i=1}^{s}\gamma_i\cdot\left(f_z^E\left(h\sum_{j=1}^{s}\beta_{ij}\cdot(k_j)_\alpha^E\right) + f_\alpha^E\right) = h\cdot f_z^E A w,$$

which is just

$$f_\alpha^E = \frac{1}{\Gamma}\cdot f_z^E\left(Aw - h\sum_{i=1}^{s}\sum_{j=1}^{s}\gamma_i\beta_{ij}\cdot(k_j)_\alpha^E\right).$$

This means, however, that $f_\alpha^E \in \mathrm{ran}(f_z^E)$, a contradiction.

Step 6. Finally, to prove the second range condition, one has to work with the bilinear forms representing the second derivatives. Suppose again, to the contrary, that there exists a vector $w \in \mathbb{R}^N$ such that $\varphi_{zz}^E(v,v) = (\varphi_z^E - I_N)w$, where $v \in \mathrm{null}(f_z^E) = \mathrm{null}(\varphi_z^E - I_N)$. Since

$$\varphi_{zz}^E(v,v) = h\sum_{i=1}^{s} \gamma_i \cdot (k_i)_{zz}^E(v,v),$$

we first need to compute $(k_i)_{zz}^E(v,v)$. To accomplish this, introduce the functions

$$F(z) \equiv f(z,\alpha)$$

and for any $i = 1, 2, \ldots, s$

$$G_i(z) \equiv z + h\sum_{j=1}^{s} \beta_{ij} \cdot k_j(h,z,\alpha).$$

Now $k_i = F \circ G_i$, so according to the higher-order chain rule [6], we get that

$$(k_i)_{zz}^E(v,v) = (F_{zz} \circ G_i)^E\left((G_i)_z^E v, (G_i)_z^E v\right) + (F_z \circ G_i)^E\left((G_i)_{zz}^E(v,v)\right) =$$

$$f_{zz}^E\left(v + h\sum_{j=1}^{s} \beta_{ij} \cdot (k_j)_z^E v, v + h\sum_{j=1}^{s} \beta_{ij} \cdot (k_j)_z^E v\right) + f_z^E\left((G_i)_{zz}^E(v,v)\right).$$

But $v \in \mathrm{null}(f_z^E) \subset \mathrm{null}\left((k_j)_z^E\right)$ for every j, hence

$$(k_i)_{zz}^E(v,v) = f_{zz}^E(v,v) + f_z^E\left((G_i)_{zz}^E(v,v)\right).$$

If $\varphi_{zz}^E(v,v) = (\varphi_z^E - I_N)w$ were true, then

$$\Gamma \cdot f_{zz}^E(v,v) + f_z^E\left(\sum_{i=1}^{s} \gamma_i \cdot (G_i)_{zz}^E(v,v)\right) = f_z^E A w$$

would hold, in other words

$$f_{zz}^E(v,v) = \frac{1}{\Gamma} \cdot f_z^E\left(Aw - \sum_{i=1}^{s} \gamma_i \cdot (G_i)_{zz}^E(v,v)\right),$$

which would clearly violate our original assumption $f_{zz}^E(v,v) \notin \mathrm{ran}(f_z^E)$. ■

As for the cusp case, consider (27) again, but this time with $\alpha \in \mathbb{R}^2$. The smooth function $f : \mathbb{R}^N \times \mathbb{R}^2 \to \mathbb{R}^N$ has a cusp bifurcation [5] at the equilibrium $z = 0$, $\alpha = 0$, if

- $f^E = 0$,
- $\dim \mathrm{null}(f_z^E) = 1$,
- $f_{zz}^E(v,v) \in \mathrm{ran}(f_z^E)$, where $\mathrm{null}(f_z^E)$=span(v),

• $f^E_{zzz}(v,v,v) + 3f^E_{zz}(v,x) \notin \mathrm{ran}(f^E_z)$, where v is as above and x is any solution to the equation $f^E_z x = -f^E_{zz}(v,v)$.

Remark. One can make x unique assuming one extra condition, but we will not make use of this property.

Consider the corresponding Runge-Kutta discretization $\varphi : \mathbb{R}^+ \times \mathbb{R}^N \times \mathbb{R}^2 \to \mathbb{R}^N$. The equilibrium $z = 0$, $\alpha = 0$ is a cusp bifurcation point for the map $\varphi(h,\cdot,\cdot)$, if the following conditions hold:

• $\varphi(h,0,0) = 0$,

• $\dim \mathrm{null}(\varphi^E_z - I_N) = 1$,

• $\varphi^E_{zz}(v,v) \in \mathrm{ran}(\varphi^E_z - I_N)$, where $\mathrm{null}(\varphi^E_z - I_N)$=span$(v)$,

• $\varphi^E_{zzz}(v,v,v) + 3\varphi^E_{zz}(v,y) \notin \mathrm{ran}(\varphi^E_z - I_N)$, where v is as above and y is any solution to the equation $(\varphi^E_z - I_N)y = -\varphi^E_{zz}(v,v)$.

Proposition 3.2 *Suppose that the equation (27) has a cusp bifurcation at the equilibrium $z = 0$, $\alpha = 0$, and $\Gamma := \sum_{i=1}^s \gamma_i \neq 0$. Then the corresponding Runge-Kutta discretization map also has a cusp bifurcation at $z = 0$, $\alpha = 0$ for $h > 0$ sufficiently small.*

Proof. Due to the previous proposition, only the last two conditions have to be checked.
Step 1. Suppose that $f^E_{zz}(v,v) = f^E_z u$ holds with some $u \in \mathbb{R}^N$ and $0 \neq v \in \mathrm{null}(f^E_z)$. Set

$$w := A^{-1}\left(\Gamma u + \sum_{i=1}^s \gamma_i \cdot (G_i)^E_{zz}(v,v)\right),$$

where the linear operator A and the functions G_i $(i = 1,2,\ldots,s)$ are as in the previous proof, see Step 4 and 6 there. Then we have that

$$(\varphi^E_z - I_N)w = h \cdot f^E_z A w = h \cdot f^E_z \left(\Gamma u + \sum_{i=1}^s \gamma_i \cdot (G_i)^E_{zz}(v,v)\right) =$$

$$= h\sum_{i=1}^s \gamma_i \cdot f^E_z u + h\sum_{i=1}^s \gamma_i \cdot f^E_z\left((G_i)^E_{zz}(v,v)\right) =$$

$$= h\sum_{i=1}^s \gamma_i \left(f^E_{zz}(v,v) + f^E_z\left((G_i)^E_{zz}(v,v)\right)\right) = h\sum_{i=1}^s \gamma_i \cdot (k_i)^E_{zz}(v,v) = \varphi^E_{zz}(v,v).$$

Step 2. Suppose to the contrary that there exists a vector $w \in \mathbb{R}^N$ such that

$$\varphi^E_{zzz}(v,v,v) + 3\varphi^E_{zz}(v,y) = (\varphi^E_z - I_N)w \tag{30}$$

holds with $0 \neq v \in \mathrm{null}(\varphi^E_z - I_N) = \mathrm{null}(f^E_z)$ and y being any solution to the equation $(\varphi^E_z - I_N)y = -\varphi^E_{zz}(v,v)$.
In order to compute the trilinear and the bilinear forms here, we appeal again

to the higher-order chain rule [6] (with the same notations as in Step 6 in the proof of the previous proposition) to get for every $i = 1, 2, \ldots, s$ that

$$(k_i)^E_{zzz}(v,v,v) = (F_{zzz} \circ G_i)^E \left((G_i)^E_z v, (G_i)^E_z v, (G_i)^E_z v\right) +$$

$$+3\,(F_{zz} \circ G_i)^E \left((G_i)^E_{zz}(v,v), (G_i)^E_z v\right) + (F_z \circ G_i)^E \left((G_i)^E_{zzz}(v,v,v)\right),$$

where symmetry of the bilinear forms has also been taken into account. Performing some of the evaluations, we arrive at the following formula

$$(k_i)^E_{zzz}(v,v,v) = f^E_{zzz}(v,v,v) + 3f^E_{zz}\left((G_i)^E_{zz}(v,v), v\right) + f^E_z\left((G_i)^E_{zzz}(v,v,v)\right).$$

In a similar manner, we have that

$$(k_i)^E_{zz}(v,y) = f^E_{zz}\left(v, y + h\sum_{j=1}^{s}\beta_{ij}\cdot(k_j)^E_z y\right) + f^E_z\left((G_i)^E_{zz}(v,y)\right).$$

Now (30) is equivalent to the following

$$h\sum_{i=1}^{s}\gamma_i\{f^E_{zzz}(v,v,v) + 3f^E_{zz}\left((G_i)^E_{zz}(v,v), v\right) + f^E_z\left((G_i)^E_{zzz}(v,v,v)\right) +$$

$$+3f^E_{zz}\left(v, y + h\sum_{j=1}^{s}\beta_{ij}\cdot(k_j)^E_z y\right) + 3f^E_z\left((G_i)^E_{zz}(v,y)\right)\} = h\cdot f^E_z Aw,$$

that is

$$f^E_{zzz}(v,v,v) + 3f^E_{zz}\left(v, \frac{1}{\Gamma}\left(\sum_{i=1}^{s}\gamma_i\cdot(G_i)^E_{zz}(v,v) + \Gamma y + h\sum_{i=1}^{s}\sum_{j=1}^{s}\gamma_i\beta_{ij}\cdot(k_i)^E_z y\right)\right) =$$

$$= \frac{1}{\Gamma}\cdot f^E_z\left(Aw - \sum_{i=1}^{s}\gamma_i\cdot(G_i)^E_{zzz}(v,v,v) - 3\sum_{i=1}^{s}\gamma_i\cdot(G_i)^E_{zz}(v,y)\right),$$

using again the symmetry of the bilinear forms.

The desired contradiction will immediately follow as soon as we have shown that

$$x := \frac{1}{\Gamma}\left(\sum_{i=1}^{s}\gamma_i\cdot(G_i)^E_{zz}(v,v) + \Gamma y + h\sum_{i=1}^{s}\sum_{j=1}^{s}\gamma_i\beta_{ij}\cdot(k_i)^E_z y\right)$$

solves

$$f^E_z x = -f^E_{zz}(v,v).$$

But we know that y satisfies

$$(\varphi^E_z - I_N)(-y) = \varphi^E_{zz}(v,v),$$

which—by the last part of Step 6 in the proof of the previous proposition—implies that

$$f^E_{zz}(v,v) = \frac{1}{\Gamma}\cdot f^E_z\left(A(-y) - \sum_{i=1}^{s}\gamma_i\cdot(G_i)^E_{zz}(v,v)\right).$$

By the definition of x and A, the right hand side is just $f^E_z(-x)$, so the proof is complete. ■

Acknowledgement

The author wishes to thank Professor Wolf-Jürgen Beyn for his helpful remarks and consultations about the paper.

References

[1] G. Farkas, *Conjugacy in the discretized fold bifurcation*, Computers & Mathematics with Applications 43 (2002) pp. 1027-1033.

[2] B. M. Garay, *Discretization and some qualitative properties of ordinary differential equations about equilibria*, Acta Math. Univ. Comenianae, Vol. LXII, 2 (1993), pp. 249-275

[3] B. M. Garay, *On C^j-closeness Between the Solution Flow and its Numerical Approximation*, Journal of Difference Eq. and Appl., 1996, Vol. 2, pp. 67-86

[4] Y. A. Kuznetsov, *Elements of Applied Bifurcation Theory*, Springer-Verlag, New York, 1998.

[5] W.-J. Beyn, A. Champneys, E. Doedel, W. Govaerts, Y. A. Kuznetsov, B. Sandstede, *Numerical continuation, and computation of normal forms*, In: Handbook of Dynamical Systems (Ed.: B. Fiedler), Vol. 2, Elsevier, 2002.

[6] W.-J. Beyn, W. Kless, *Numerical Taylor expansions of invariant manifolds in large dynamical systems*, Numer. Math. (1998) 80: 1-38

[7] E. Hairer, S.P. Nørsett, G. Wanner, *Solving Ordinary Differential Equations I.*, 2nd Edition, Springer-Verlag, Berlin, Heidelberg, New York, 1993.

[8] Y.-K. Zou, W.-J. Beyn, *On manifolds of connecting orbits in discretizations of dynamical systems*, Nonlinear Analysis TMA (to appear)

Lajos Lóczi, Budapest University of Technology and Economics
E-mail address: `lloczi@math.bme.hu`