Conjugacy in the discretized transcritical bifurcation

Lajos Lóczi*
Department of Numerical Analysis,
Faculty of Informatics,
Eötvös Loránd University,
Budapest, Pázmány P. sétány 1/C,
H-1117 Hungary

January 25, 2005

Abstract

The present work can be considered as another case study—analogous to our earlier preprint [1]—in the direction of discretizing one-dimensional ordinary differential equations near non-hyperbolic equilibria. This time the hyperbolicity condition is violated due to the presence of a *transcritical bifurcation point*. The main aim is to show that the dynamics induced by the time-h-map of the original continuous system and that of the discretized one are still locally topologically equivalent, meaning that there exists a conjugacy between the corresponding phase portraits in the vicinity of the equilibrium. Besides the construction of a conjugacy map $J(h,\cdot,\alpha)$, the important point is that we also estimate the distance between $J(h,\cdot,\alpha)$ and the one-dimensional identity map.

In the first part of the paper, we derive normal forms for the time-h-map of the ordinary differential equation and its discretization near a transcritical bifurcation point at bifurcation parameter $\alpha = 0$ in one dimension and with discretization stepsize $h > 0$. We assume that the discretization method preserves equilibria. We will see that it is sufficient to construct a conjugacy between these normal forms.

In the second part, $J(h,\cdot,\alpha)$ is constructed for $0 < h \le h_0$ and $-\alpha_0 \le \alpha \le \alpha_0$ with h_0 and α_0 sufficiently small. Then the quantity $|x - J(h,x,\alpha)|$ is proved to be $\mathcal{O}(h^p)$ small, uniformly in x and α, in a small $x \in [-\varepsilon_0, \varepsilon_0]$ neighbourhood of the origin, where p denotes the order of the one-step discretization method.

*This research was supported by the DAAD project "Dynamics of evolution equations under simultaneous time and space discretization", by the DFG Research Group "Spektrale Analysis, asymptotische Entwicklungen und stochastische Dynamik" at Bielefeld University, further by the Hungarian Scientific Research Fund OTKA under Grant No. T037491.

1 Introduction and notation

Suppose we have a one-dimensional ordinary differential equation

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

and its one-step 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.

Since the numerical method is of order p, we have that

$$|\Phi(h, x, \alpha) - \varphi(h, x, \alpha)| \leq const \cdot h^{p+1}, \quad \forall h \in [0, h_0], \forall |x| \leq \varepsilon_0, \forall |\alpha| \leq \alpha_0, \tag{3}$$

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, ε_0 and α_0 are some small positive constants. Throughout the paper, the symbols *const* will denote generic positive constants in the estimates, with dependence only on f. (These can have possibly different values at different occurrences.)

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

$$\begin{gathered} f(0, \alpha) = 0, \quad \forall |\alpha| \leq \alpha_0, \\ f_x^B = 0, \quad f_{xx}^B \neq 0, \quad f_{x\alpha}^B \neq 0, \end{gathered} \tag{4}$$

where subscripts x and α denote partial differentiation with respect to their corresponding variables, while superscript ^{B} abbreviates *evaluation at the bifurcation point*, that is, evaluation at $x = 0$ and $\alpha = 0$. (The evaluation is performed *after* taking all partial derivatives.)

The evaluation operator ^{B} will also be used for functions of three variables—h, x and α—when we evaluate a function at $h = 0$, $x = 0$ and $\alpha = 0$, as in $\Phi_{hx\alpha}^B$ abbreviating $\Phi_{hx\alpha}(0, 0, 0)$. (Here subscript h, of course, again stands for partial differentiation.)

For functions of three variables h, x and α, the evaluation operator ^{E} denotes *evaluation at general parameter values* h and α, where the dependence of ^{E} on h and α is suppressed. (Values of the parameters $h \in [0, h_0]$ and $\alpha \in [-\alpha_0, \alpha_0]$ can be arbitrary but fixed.) Thus, for example, the function $J(h, \cdot, \alpha)$ is abbreviated to J^E, if $J : \mathbb{R} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$.

Some more notation is introduced. The symbol $g^{[-1]}$ means the *inverse* of a real function g. Similarly, $g^{[k]}$ is the k^{th} *iterate* ($k \in \mathbb{Z}$) of $f : \mathbb{R} \to \mathbb{R}$. The symbol *id* denotes the identity function of $\mathbb{R}$. Symbols $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$, as usual, denote the *floor* and the *ceiling* functions, respectively. The set of nonnegative integers is denoted by $\mathbb{N}$. Finally, for any $a, b \in \mathbb{R}$, the symbol $[\{a, b\}]$ represents the *closed interval between* the elements of the set $\{a, b\}$, that is $[\{a, b\}] := [\min(a, b), \max(a, b)]$.

Remark 1.1 Notice that instead of assumption $f(0,\alpha)=0,\ \forall|\alpha|\le\alpha_0$ in (4), [2] simply assumes $f(0,0)=0$ when it determines conditions for transcritical bifurcation of fixed points of maps. However, this is insufficient as illustrated by the map $x_{n+1}:=f(x_n,\alpha)$ with

$$f(x,\alpha):=\alpha^2+(1+\alpha)x+x^2.$$

Since $(x,\alpha)=(0,0)$ is the only fixed point of the map, clearly no bifurcation of fixed points can occur here. (The same discrepancy is present in [2] in the case of the *pitchfork bifurcation.*)

We add that [3], for example, correctly uses $f(0,0)$ *and* a kind of discriminant condition to define transcritical bifurcation of fixed points of maps. Condition $f(0,\alpha)=0$ we have adopted is more "direct" and a bit simpler to work with.

2 Construction of the normal forms

In this section, we compute normal forms for the maps

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

and

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

near the equilibrium being also a transcritical bifurcation point.

The properties of the solution flow together with (3)–(4) imply for $h\ge0$, $|x|\le\varepsilon_0$ and $|\alpha|\le\alpha_0$ that

$$\Phi(h,0,\alpha)=0,\quad \forall|\alpha|\le\alpha_0, \tag{7}$$
$$\varphi(0,x,\alpha)=\Phi(0,x,\alpha)=x, \tag{8}$$
$$\Phi_h(h,x,\alpha)=f(\Phi(h,x,\alpha),\alpha), \tag{9}$$
$$\varphi_h(0,x,\alpha)=\Phi_h(0,x,\alpha). \tag{10}$$

Instead of (9), the shorter form $\Phi_h=f\circ\Phi$ will be used.

To ensure that *the origin $x=0$ is a fixed point also for the discretization map* (6), we assume that

$$\varphi(h,0,\alpha)=0 \tag{11}$$

holds for sufficiently small $h\ge0$ and $|\alpha|$, which is the case, for example, for all Runge-Kutta discretizations.

Lemma 2.1 *Under the assumptions above and for $h\in[0,h_0]$, $|x|\le\varepsilon_0$, $|\alpha|\le\alpha_0$, 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

$$\begin{aligned} f_0(h,\alpha) &\equiv 0,\\ f_1(h,\alpha) &\equiv 1+h\alpha\cdot f^B_{x\alpha}+h\alpha^2\cdot\psi_1(h,\alpha), \qquad f^B_{x\alpha}\ne0,\\ f_2(h,\alpha) &= \frac{1}{2}h\cdot f^B_{xx}+h\alpha\cdot\psi_2(h,\alpha), \qquad f^B_{xx}\ne0,\\ \psi_3(h,x,\alpha) &= h\cdot\widehat{\psi}_3(h,x,\alpha)\end{aligned}$$

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

Proof. We expand Φ in a multivariate Taylor series about the equilibrium with the remainders in integral form.

Since $f(0,\alpha)=0$ for all $|\alpha|$ sufficiently small, we have (7), hence $f_0(h,\alpha)$ should vanish.

As for f_1, we get that

$$f_1(h,\alpha)=\Phi_x^B+\alpha\cdot \mathrm{I}_{011}(\alpha)+h\cdot \mathrm{I}_{110}(h)+h\alpha\cdot\Phi_{hx\alpha}^B+$$

$$h\alpha^2\cdot \mathrm{I}_{112}(\alpha)+h^2\alpha\cdot \mathrm{I}_{211}(h)+h^2\alpha^2\cdot \mathrm{I}_{212}(h,\alpha),$$

where $\Phi_x^B=1$,

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

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

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

It is easy to verify that $\Phi_{hx\alpha}^B=f_{x\alpha}^B$. Indeed, we have that

$$\Phi_{hx\alpha}^B=(f\circ\Phi)_{x\alpha}^B=((f_x\circ\Phi)_\alpha\cdot\Phi_x+(f_x\circ\Phi)\cdot\Phi_{x\alpha})^B=(f_x\circ\Phi)_\alpha^B,$$

because $\Phi_{x\alpha}^B=0$ and $\Phi_x^B=1$. But

$$(f_x\circ\Phi)_\alpha^B=f_{xx}(\Phi^B,0)\cdot\Phi_\alpha^B+f_{x\alpha}(\Phi^B,0)=f_{x\alpha}^B,$$

since $\Phi_\alpha(0,x,\alpha)\equiv 0$.

The last three integrals read

$$\mathrm{I}_{112}(\alpha)=\int_0^1(1-\tau)\Phi_{hx\alpha\alpha}(0,0,\tau\alpha)\mathrm{d}\tau,$$

$$\mathrm{I}_{211}(h)=\int_0^1(1-\tau)\Phi_{hhx\alpha}(\tau h,0,0)\mathrm{d}\tau$$

and

$$\mathrm{I}_{212}(h,\alpha)=\int_0^1\int_0^1(1-\tau)(1-\sigma)\Phi_{hhx\alpha\alpha}(\tau h,0,\sigma\alpha)\mathrm{d}\tau\mathrm{d}\sigma.$$

We now show that $\mathrm{I}_{211}(h)$ vanishes, or, more precisely, that $\Phi_{hhx\alpha}(h,0,0)\equiv 0$ for every small $h\geq 0$. By direct differentiation we obtain that

$$\Phi_{hhx\alpha}=(f_{xx}\circ\Phi)_\alpha\cdot\Phi_x\cdot\Phi_h+(f_{xx}\circ\Phi)\cdot\Phi_{x\alpha}\cdot\Phi_h+$$

$$(f_{xx}\circ\Phi)\cdot\Phi_x\cdot\Phi_{h\alpha}+(f_x\circ\Phi)_\alpha\cdot\Phi_{hx}+(f_x\circ\Phi)\cdot\Phi_{hx\alpha}.$$

Here $\Phi_h(h,0,0)=f(\Phi(h,0,0),0)=f(0,0)=0$, so the first two terms above vanish. The third term is also zero, since

$$\Phi_{h\alpha}(h,0,0)=f_x(\Phi(h,0,0),0)\cdot\Phi_\alpha(h,0,0)+f_\alpha(\Phi(h,0,0),0)$$

but $\Phi(h,0,0)=0$ and $f_x(0,0)=0=f_\alpha(0,0)$. The fourth term is zero, because

$$\Phi_{hx}(h,0,0)=f_x(\Phi(h,0,0),0)\cdot\Phi_x(h,0,0)=0\cdot\Phi_x(h,0,0).$$

Finally, the fifth term vanishes due to the factor $f_x(\Phi(h,0,0),0)=0$.

By defining the smooth function $\psi_1(h,\alpha) := \mathrm{I}_{112}(\alpha) + h\cdot \mathrm{I}_{212}(h,\alpha)$, f_1 has the form stated above.

In the case of f_2, we have that

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

where $\Phi^B_{xx} = 0$ and

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

However,

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

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,

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

Thus, $\psi_2(h,\alpha) := \frac{1}{2}\mathrm{I}_{121}(h,\alpha)$ defines the desired smooth function.

For the remainder ψ_3, the integral formula gives

$$\psi_3(h,x,\alpha) = \frac{1}{2}\int_0^1 (1-\tau)^2\Phi_{xxx}(h,\tau x,\alpha)\mathrm{d}\tau. \tag{12}$$

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 we introduce a new parameter $\beta\equiv\beta(h,\alpha)$ by

$$\beta(h,\alpha) := \alpha\cdot f^B_{x\alpha} + \alpha^2\cdot \mathrm{I}_{112}(\alpha) + h\alpha^2\cdot \mathrm{I}_{212}(h,\alpha),$$

i.e., $\beta(h,\alpha) = \frac{f_1(h,\alpha)-1}{h}$.

We notice that $\beta(h,0)=0$ and $\frac{\mathrm{d}}{\mathrm{d}\alpha}\beta(h,0) = f^B_{x\alpha}\neq 0$ independently of $h\in[0,h_0]$, thus the inverse function theorem guarantees the local existence and uniqueness of a smooth inverse function $\overline{\alpha}_0\equiv\overline{\alpha}_0(h,\beta)$ of $\alpha\mapsto\beta(h,\alpha)$. Moreover, it is easy to see that the domain of definition of this inverse function contains a neighbourhood of the origin independent of $h\in[0,h_0]$. Further, $\overline{\alpha}_0(h,0)=0$, hence

$$\overline{\alpha}_0(h,\beta) = \beta\cdot\psi_a(h,\beta) \tag{13}$$

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

Therefore (5) is transformed into the map

$$x\mapsto(1+h\beta)x + h\cdot q(h,\beta)x^2 + h\cdot\widehat{\psi}_3(h,x,\overline{\alpha}_0(h,\beta))x^3$$

with $q(h,\beta)\equiv\frac{1}{2}f^B_{xx} + \frac{1}{2}\overline{\alpha}_0(h,\beta)\cdot \mathrm{I}_{121}(h,\overline{\alpha}_0(h,\beta))$.

A final scaling $\eta := |q(h,\beta)|x$ with $s := \operatorname{sign}(q(h,0)) = \pm1$ (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

$$\eta \mapsto (1 + h\beta)\eta + s \cdot h\eta^2 + h\eta^3 \cdot \widehat{\eta}_3(h, \eta, \beta)$$

where $\widehat{\eta}_3(h, \eta, \beta) = \widehat{\psi}_3(h, x, \overline{\alpha}_0(h, \beta)) \cdot |q(h, \beta)|^{-2}$ *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 together with (11) and for* $h \in [0, h_0]$, $|x| \leq \varepsilon_0$, $|\alpha| \leq \alpha_0$, *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

$$\begin{aligned}
\widetilde{f}_0(h, \alpha) &= 0, \\
\widetilde{f}_1(h, \alpha) &= 1 + h\alpha \cdot f^B_{x\alpha} + h^{p+1} \cdot \chi_{10}(h) + h\alpha \cdot \chi_{11}(h, \alpha), \\
\widetilde{f}_2(h, \alpha) &= \frac{1}{2}h \cdot f^B_{xx} + h^{p+1} \cdot \chi_{20}(h) + h\alpha \cdot \chi_{21}(h, \alpha), \\
\chi_3(h, x, \alpha) &= h \cdot \widetilde{\chi}_3(h, x, \alpha)
\end{aligned}$$

hold with some smooth functions χ_{10}, χ_{11}, χ_{20}, χ_{21} *and* $\widetilde{\chi}_3$. *Moreover, for* $h \in [0, h_0]$, $|x| \leq \varepsilon_0$ *and for* $|\alpha| \leq \alpha_0$,

$$|\psi_3(h, x, \alpha) - \chi_3(h, x, \alpha)| \leq const \cdot h^{p+1}. \tag{14}$$

Proof. By (11), we have that $\widetilde{f}_0(h, \alpha) \equiv 0$.

The remainders of the Taylor series are also represented by integrals and denoted—analogously to the proof of Lemma 2.1—by $\widetilde{\mathrm{I}}$'s. These integrals, of course, now always contain φ instead of Φ.

As for $\widetilde{f}_1$, by (8) one has that $\varphi^B_x = 1$ and $\widetilde{\mathrm{I}}_{011}(\alpha) \equiv 0$, further, we get that $\varphi^B_{hx\alpha} = \Phi^B_{hx\alpha} = f^B_{x\alpha} \neq 0$, hence

$$\widetilde{f}_1(h, \alpha) = 1 + h \cdot \widetilde{\mathrm{I}}_{110}(h) + h\alpha \cdot f^B_{x\alpha} +$$
$$h\alpha^2 \cdot \widetilde{\mathrm{I}}_{112}(\alpha) + h^2\alpha \cdot \widetilde{\mathrm{I}}_{211}(h) + h^2\alpha^2 \cdot \widetilde{\mathrm{I}}_{212}(h, \alpha).$$

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

$$\left| f_1(h, \alpha) - \widetilde{f}_1(h, \alpha) \right| \leq const \cdot h^{p+1}. \tag{15}$$

Evaluating this at $\alpha = 0$ yields $|h \cdot \widetilde{\mathrm{I}}_{110}(h)| \leq const \cdot h^{p+1}$. The smooth functions χ_{10} and χ_{11} are defined as

$$\chi_{10}(h) := \frac{h \cdot \widetilde{\mathrm{I}}_{110}(h)}{h^{p+1}}$$

and

$$\chi_{11}(h, \alpha) := \alpha \cdot \widetilde{\mathrm{I}}_{112}(\alpha) + h \cdot \widetilde{\mathrm{I}}_{211}(h) + h\alpha \cdot \widetilde{\mathrm{I}}_{212}(h, \alpha).$$

(It can be easily proved that $\widetilde{\mathrm{I}}_{112}(\alpha) \equiv \mathrm{I}_{112}(\alpha)$, but this property will not be needed later.)

Considering $\widetilde{f}_2$, we obtain that $\varphi^B_{xx} = 0$ and $\widetilde{\mathrm{I}}_{021}(\alpha) \equiv 0$. By differentiating (10) we see that $\varphi^B_{hxx} = \Phi^B_{hxx} = f^B_{xx} \neq 0$, thus

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

and again, using $f \in C^{p+5}$ and [4]

$$\left| f_2(h,\alpha) - \widetilde{f}_2(h,\alpha) \right| \leq const \cdot h^{p+1}. \tag{16}$$

Evaluating this at $\alpha = 0$, we see that $|h^2 \cdot \widetilde{\mathrm{I}}_{220}(h)| \leq const \cdot h^{p+1}$, so we can set

$$\chi_{20}(h) := \frac{1}{2} \cdot \frac{h^2 \cdot \widetilde{\mathrm{I}}_{220}(h)}{h^{p+1}}$$

and

$$\chi_{21}(h,\alpha) := \frac{1}{2} \cdot \widetilde{\mathrm{I}}_{121}(h,\alpha)$$

to obtain two smooth functions.

To prove the product form of the remainder χ_3, we use the same argument as in (12). Finally, for (14) we take into account $f \in C^{p+6}$ and [4] again to get

$$|\psi_3(h,x,\alpha) - \chi_3(h,x,\alpha)| = \left| \frac{1}{2} \int_0^1 (1-\tau)^2 \left(\Phi_{xxx}(h,\tau x,\alpha) - \varphi_{xxx}(h,\tau x,\alpha)\right) \mathrm{d}\tau \right| \leq$$

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

completing the proof of the lemma. ■

Now we the introduce the analogue of parameter β. Set

$$\widetilde{\beta} \equiv \widetilde{\beta}(h,\alpha) := \widetilde{\mathrm{I}}_{110}(h) + \alpha \cdot f^B_{x\alpha} + \alpha^2 \cdot \widetilde{\mathrm{I}}_{112}(\alpha) + h\alpha \cdot \widetilde{\mathrm{I}}_{211}(h) + h\alpha^2 \cdot \widetilde{\mathrm{I}}_{212}(h,\alpha).$$

We will show that the function $\widetilde{\beta}(h,\cdot)$ is locally invertible at the origin for every $h \geq 0$ small enough, and its inverse function, $\widetilde{\alpha}(h,\cdot)$ is $\mathcal{O}(h^p)$-close to $\overline{\alpha}_0(h,\cdot)$, *i.e.* to the inverse of $\beta(h,\cdot)$. As in [5], we will use the same quantitative inverse function theorem, see Lemma 2.4 in [5]. (Now a letter G will play the role of $\widetilde{F}$ in that lemma.) We set

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

In order to check the conditions of the lemma, define $\kappa_1 := \frac{1}{2}|f^B_{x\alpha}| > 0$ and $\kappa_2 := \frac{1}{2}\kappa_1$. We have that

$$\frac{\partial G}{\partial \alpha}(h,\beta,\alpha) = f^B_{x\alpha} + 2\alpha \cdot \widetilde{\mathrm{I}}_{112}(\alpha) + \alpha^2 \frac{\mathrm{d}}{\mathrm{d}\alpha}\widetilde{\mathrm{I}}_{112}(\alpha) +$$

$$h \cdot \widetilde{\mathrm{I}}_{211}(h) + 2h\alpha \cdot \widetilde{\mathrm{I}}_{212}(h,\alpha) + h\alpha^2 \frac{\mathrm{d}}{\mathrm{d}\alpha}\widetilde{\mathrm{I}}_{212}(h,\alpha).$$

Thus

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

holds by smoothness of the functions $\widetilde{\mathrm{I}}$'s provided that $|\alpha - \overline{\alpha}_0(h,\beta)| \leq r_1$ and $h < r_2$ are small enough. It is also seen that

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

if $h, |\beta| < r_2$ are small enough, taking also into account (13). Finally, using that $\overline{\alpha}_0(h,\cdot)$ is the inverse function of $\beta(h,\cdot)$, we get that

$$|G(h,\beta,\overline{\alpha}_0(h,\beta))| = \left|\beta - \widetilde{\beta}(h,\overline{\alpha}_0(h,\beta))\right| = \left|\beta(h,\overline{\alpha}_0(h,\beta)) - \widetilde{\beta}(h,\overline{\alpha}_0(h,\beta))\right|.$$

But (15) implies that

$$|\beta(h,\alpha) - \widetilde{\beta}(h,\alpha)| \le const \cdot h^p, \tag{17}$$

hence $|G(h,\beta,\overline{\alpha}_0(h,\beta))| \le const \cdot h^p$ and also $|G(h,\beta,\overline{\alpha}_0(h,\beta))| \le (\kappa_1 - \kappa_2) \cdot r_1$ if $h < r_2$ is small enough.

Therefore, Lemma 2.4 in [5] is applicable in our situation and we get a unique zero $\widetilde{\alpha}(h,\beta)$ of $G(h,\beta,\cdot)$, which—by the construction of G—is the inverse function of $\alpha \mapsto \widetilde{\beta}(h,\alpha)$. Furthermore,

$$|\widetilde{\alpha}(h,\beta) - \overline{\alpha}_0(h,\beta)| \le const \cdot h^p \tag{18}$$

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

As a conclusion, (6) becomes

$$x \mapsto (1 + h\widetilde{\beta})x + h \cdot \widetilde{q}(h,\widetilde{\beta})x^2 + h \cdot \widetilde{\chi}_3(h,x,\widetilde{\alpha}(h,\widetilde{\beta}))x^3$$

with $\widetilde{q}(h,\widetilde{\beta}) \equiv \frac{1}{2}\left(f^B_{xx} + h \cdot \widetilde{\mathrm{I}}_{220}(h) + \widetilde{\alpha}(h,\widetilde{\beta}) \cdot \widetilde{\mathrm{I}}_{121}(h,\widetilde{\alpha}(h,\widetilde{\beta}))\right)$.

We claim that

$$\left|\widetilde{q}(h,\widetilde{\beta}) - q(h,\beta)\right| \le const \cdot h^p \tag{19}$$

also holds. But this is a consequence of inequalities (18), (16) and the smoothness (and boundedness) of the functions I_{121} and $\widetilde{\mathrm{I}}_{121}$ when combined with standard triangle inequalities and the mean value theorem.

By applying a final scaling

$$\widetilde{\eta} := |\widetilde{q}(h,\widetilde{\beta})|x$$

with $s := \mathrm{sign}(\widetilde{q}(h,0)) = \pm 1$ (being independent of $h \in [0,h_0]$ for h_0 small enough, due to (18) evaluated at $\beta = 0$, (13) and the boundedness of the function $\widetilde{\mathrm{I}}_{121}$) and defining

$$\widetilde{\eta}_3(h,\widetilde{\eta},\widetilde{\beta}) := \widetilde{\chi}_3(h,x,\widetilde{\alpha}(h,\widetilde{\beta})) \cdot |\widetilde{q}(h,\widetilde{\beta})|^{-2},$$

we have derived a normal form for (6) in the theorem below.

For the closeness estimates in the theorem, we should only verify that

$$\left|\widehat{\eta}_3(h,\eta,\beta) - \widetilde{\eta}_3(h,\widetilde{\eta},\widetilde{\beta})\right| \le const \cdot h^p.$$

This estimate, however, is a simple consequence of (19) and the fact that

$$\left|\widehat{\psi}_3(h,x,\overline{\alpha}_0(h,\beta)) - \widetilde{\chi}_3(h,x,\widetilde{\alpha}(h,\widetilde{\beta}))\right| \le const \cdot h^p.$$

(For this last inequality, (14), the smoothness of $\widehat{\psi}_3$, a standard triangle inequality and the mean value theorem suffice.)

Theorem 2.4 *There are smooth invertible coordinate and parameter changes transforming the system*

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

into

$$\widetilde{\eta} \mapsto (1 + h\widetilde{\beta})\widetilde{\eta} + s \cdot h\widetilde{\eta}^2 + h\widetilde{\eta}^3 \cdot \widetilde{\eta}_3(h, \widetilde{\eta}, \widetilde{\beta})$$

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*

$$|\widehat{\eta}_3 - \widetilde{\eta}_3| \le const \cdot h^p \qquad \blacksquare$$

Finally, we apply a parameter shift $\widetilde{\beta} \mapsto \beta$ to the normal form in the theorem above, being $\mathcal{O}(h^p)$-close to the identity due to (17). So from now on we will use the bifurcation parameter α again instead of β and $\widetilde{\beta}$. To simplify our notations further, instead of η and $\widetilde{\eta}$ the letter x will be used.

3 Construction of the conjugacy

We have thus the following normal forms

$$\mathcal{N}_\Phi(h, x, \alpha) = (1 + h\alpha)x + s \cdot hx^2 + hx^3\,\widehat{\eta}_3(h, x, \alpha) \tag{20}$$

$$\mathcal{N}_\varphi(h, x, \alpha) = (1 + h\alpha)x + s \cdot hx^2 + hx^3\,\widetilde{\eta}_3(h, x, \alpha) \tag{21}$$

with $s = 1$ or $s = -1$, where $\widehat{\eta}_3$ and $\widetilde{\eta}_3$ are smooth functions. Let $K > 0$ denote a uniform bound on $\left|\frac{\mathrm{d}^i}{\mathrm{d}x^i}\eta(h, \cdot, \alpha)\right|$ $(i \in \{0, 1, 2\}, \eta \in \{\widehat{\eta}_3, \widetilde{\eta}_3\})$ in a neighbourhood of the origin for any small $h > 0$ and $|\alpha|$, as well as a uniform bound on $\left|\frac{\mathrm{d}}{\mathrm{d}\alpha}\eta(h, x, \cdot)\right|$ $(\eta \in \{\widehat{\eta}_3, \widetilde{\eta}_3\})$ in a neighbourhood of the origin for any small $h > 0$ and $|x|$. We also have that there exists a constant $c > 0$ such that

$$|\mathcal{N}_\Phi(h, x, \alpha) - \mathcal{N}_\varphi(h, x, \alpha)| \le c \cdot h^{p+1}|x|^3 \tag{22}$$

holds for all sufficiently small $h > 0$, $|x| \ge 0$ and $|\alpha| \ge 0$. Throughout the section, c will denote this particular positive constant. (Other generic constants, if needed, are denoted by *const*.)

We will consider the case $s = 1$, the other one is similar. Then it is easy to see that $\omega_{\Phi,0}(h, \alpha) \equiv 0$ is an attracting fixed point of the map $\mathcal{N}_\Phi(h, \cdot, \alpha)$ for $\alpha < 0$, and repelling for $\alpha > 0$. For any fixed $h > 0$ and $\alpha \in [-\alpha_0, \alpha_0] \setminus \{0\}$, this map possesses another fixed point, denoted by $\omega_{\Phi,+} \equiv \omega_{\Phi,+}(h, \alpha) > 0$ (if $\alpha < 0$) and $\omega_{\Phi,-} \equiv \omega_{\Phi,-}(h, \alpha) < 0$ (if $\alpha > 0$). It is seen that $\omega_{\Phi,+}$ is repelling and $\omega_{\Phi,-}$ is attracting. The two branches of fixed points, $\omega_{\Phi,0}(h, \alpha)$ and $\omega_{\Phi,\pm}(h, \alpha)$ merge at $\alpha = 0$.

Analogous results hold, of course, for the map $\mathcal{N}_\varphi(h, \cdot, \alpha)$. Its fixed points are denoted by $\omega_{\varphi,0}$ and $\omega_{\varphi,-}$ (or $\omega_{\varphi,+}$).

We will construct a conjugacy in a natural way and prove optimal closeness estimates in the $x \le 0$ region—the $x > 0$ case is similar due to symmetry.

In what follows, we suppose that

$$0 < h \le h_0 := \frac{1}{5},$$

$$|x| \le \varepsilon_0 := \min\left(\frac{1}{25}, \frac{1}{25K}\right) \text{ and} \tag{23}$$

$$|\alpha| \le \alpha_0 := \min\left(\frac{1}{51}, \frac{1}{51K}\right).$$

With these values of h_0, ε_0 and α_0, all constructions and proofs below can be carried out. (There is only one constraint which has not been taken into account explicitly: if the domain of definition of the functions $\widehat{\eta}_3$ and $\widetilde{\eta}_3$ is smaller than $(0, h_0] \times [-\varepsilon_0, \varepsilon_0] \times [-\alpha_0, \alpha_0]$ given above, then h_0, ε_0 or α_0 should be decreased further suitably.)

Lemma 3.1 *For every* $0 < h \le h_0$ *and* $0 < \alpha \le \alpha_0$ *we have that*

$$\{\omega_{\varphi,-}, \omega_{\Phi,-}\} \subset \left(-\frac{3}{2}\alpha, -\frac{6}{7}\alpha\right).$$

Proof. By definition, $\omega_{\varphi,-}$ solves $\alpha + x + x^2 \cdot \widetilde{\eta}_3(h, x, \alpha) = 0$. But $|x| \le \frac{1}{6K}$ implies $\frac{2}{3} \le 1 + x\,\widetilde{\eta}_3 \le \frac{7}{6}$, so

$$-\frac{3\alpha}{2} \le \omega_{\varphi,-} = \frac{-\alpha}{1 + \omega_{\varphi,-} \cdot \widetilde{\eta}_3(h, \omega_{\varphi,-}, \alpha)} \le -\frac{6\alpha}{7}.$$

The proof for $\omega_{\Phi,-}$ is similar. ■

By iterating one of the normal forms, say $\mathcal{N}_\varphi(h, \cdot, \alpha)$, let us define three sequences x_n, y_n and z_n. For $\alpha > 0$, let $x_n \equiv x_n(h, \alpha)$ be defined as

$$x_{n+1} := \mathcal{N}_\varphi(h, x_n, \alpha), \quad n = 0, 1, 2, \ldots$$

with $x_0 := -\frac{\alpha}{3}$, further, let $y_n \equiv y_n(h, \alpha)$ be defined as

$$y_n := \left(\mathcal{N}_\varphi^E\right)^{[-n]}(x_0), \quad n = 0, 1, 2, \ldots,$$

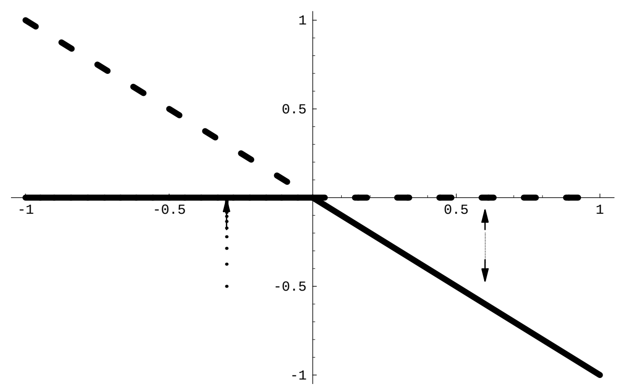
so $y_0 := x_0$, and set $y_{-1} := x_1$. Finally, for all $\alpha \in [-\alpha_0, \alpha_0]$ define $z_n \equiv z_n(h, \alpha)$ as

$$z_n := \left(\mathcal{N}_\varphi^E\right)^{[n]}(z_0), \quad n = 0, 1, 2, \ldots,$$

with $z_0 < 0$ being independent of h and α such that $2\alpha_0 < |z_0| < \frac{1}{2K}$ holds. An appropriate choice for z_0 is, *e.g.*, $z_0 := -\varepsilon_0$.

Simple calculations show that, for example, under conditions (23), both $\mathcal{N}_\varphi^E$ and $\mathcal{N}_\Phi^E$ (together with their inverses) are monotone increasing, further $|\alpha| < \frac{6}{K}$ implies $x_0(\alpha) > x_1(h, \alpha)$ and $2\alpha_0 < |z_0| < \frac{1}{2K}$ implies $z_0 < z_1(h, \alpha)$. This means that x_n is monotone decreasing, y_n is monotone increasing (if $\alpha > 0$ and $n \ge 0$), and $\lim_{n\to\infty} x_n(h, \alpha) = \omega_{\varphi,-}$, while $\lim_{n\to\infty} y_n(h, \alpha) = \omega_{\varphi,0}$. Moreover, z_n is monotone increasing, further, for $\alpha > 0$, $\lim_{n\to\infty} z_n(h, \alpha) = \omega_{\varphi,-}$ and for $\alpha \le 0$, $\lim_{n\to\infty} z_n(h, \alpha) = \omega_{\varphi,0}$.

The following figure shows the branch of stable and unstable fixed points of $\mathcal{N}_\varphi^E$ in the (α, x)-plane together with the first few terms of the inner sequences ($x_n(h,\alpha)$ and $y_n(h,\alpha)$), and the outer sequence $z_n(h,\alpha)$ with some $h > 0$ and α fixed. The arrows indicate the direction of the sequences.



A homeomorphism J^E satisfying the conjugacy equation

$$J^E \circ \mathcal{N}_\varphi^E = \mathcal{N}_\Phi^E \circ J^E \tag{24}$$

is now piecewise defined on the fundamental domains, *i.e.* on $[x_{n+1}, x_n]$, $[y_n, y_{n+1}]$ and $[z_n, z_{n+1}]$ ($n \in \mathbb{N}$), for any fixed $0 < h \le h_0$ and $-\alpha_0 \le \alpha \le \alpha_0$.

We first consider the region between the fixed points for $0 < \alpha \le \alpha_0$.

Let $J^E(x_0) := x_0$ and $J^E(x_1) := \mathcal{N}_\Phi^E(x_0)$. For $x \in [x_1, x_0]$ extend J^E linearly. For $n \ge 1$ and $x \in [x_{n+1}, x_n]$, we recursively set

$$J^E(x) := \left(\mathcal{N}_\Phi^E \circ J^E \circ \left(\mathcal{N}_\varphi^E\right)^{[-1]}\right)(x),$$

while for $n \ge 0$ and $x \in [y_n, y_{n+1}]$, we let

$$J^E(x) := \left(\left(\mathcal{N}_\Phi^E\right)^{[-1]} \circ J^E \circ \mathcal{N}_\varphi^E\right)(x).$$

(Since $[y_{-1}, y_0] \equiv [x_1, x_0]$, these two definitions are compatible.) Finally, set

$$J^E(\omega_{\varphi,-}) := \omega_{\Phi,-}$$

and

$$J^E(\omega_{\varphi,0}) := \omega_{\Phi,0}.$$

Then J^E is continuous, strictly monotone increasing on $[\omega_{\varphi,-}, 0]$, since it is a composition of three such functions, and satisfies (24).

In the outer region, *i.e.* below the fixed points, fix $z_0 < 0$ ($2\alpha_0 < |z_0| < \frac{1}{2K}$), then for $\alpha \in [-\alpha_0, \alpha_0]$ the construction of J^E is analogous to the construction above with the sequence x_n: this time z_n plays the role of x_n. (Of course, now the counterpart of the sequence y_n is not needed.) Then the function J^E becomes continuous, strictly monotone increasing on $[z_0, \omega_{\varphi,-}]$ ($0 < \alpha \le \alpha_0$) and $[z_0, \omega_{\varphi,0}]$ (for $-\alpha_0 \le \alpha \le 0$), and satisfies (24).

The construction of J^E—with the appropriate and natural modifications—in the upper half-plane $x > 0$ is analogous to the one presented above.

4 The closeness estimate for the conjugacy

4.1 Optimality at the fixed points

We first prove that the constructed conjugacy J^E is $\mathcal{O}(h^p\,\alpha^2)$-close to the identity at the fixed points $\omega_{\varphi,-}(h,\alpha)$, further, an explicit example will show that this estimate is optimal in h and α.

Since fixed points must be mapped into nearby fixed points by the conjugacy and we are going to prove $\mathcal{O}(h^p)$-closeness in the whole domain, the result above means that our estimates of $|\,id - J^E|$ near a transcritical bifurcation point are optimal in h.

The following auxiliary estimate will frequently be used.

Lemma 4.1 *For any $0 < h \le h_0$, $-\varepsilon_0 \le x < 0$ and $-\alpha_0 \le \alpha \le \alpha_0$, we have that*

$$(\mathcal{N}_\Phi^E)'(x) \le 1 + h\alpha + \frac{7}{4}hx.$$

Proof. The conditions in (23) have been set up to imply this inequality, too. ∎

Lemma 4.2 *For any $0 < h \le h_0$ and $0 < \alpha \le \alpha_0$ (satisfying (23)), we have that*

$$|\,\omega_{\varphi,-} - \omega_{\Phi,-}| \le \frac{27}{4}c \cdot h^p\,\alpha^2.$$

Proof.

$$|\,id - J^E|(\omega_{\varphi,-}(h,\alpha)) \le |\mathcal{N}_\varphi^E(\omega_{\varphi,-}) - \mathcal{N}_\Phi^E(\omega_{\varphi,-})| + |\mathcal{N}_\Phi^E(\omega_{\varphi,-}) - \mathcal{N}_\Phi^E(\omega_{\Phi,-})| \le$$

$$c \cdot h^{p+1}|\omega_{\varphi,-}|^3 + \left(\sup_{[\{\,\omega_{\varphi,-},\omega_{\Phi,-}\}]}(\mathcal{N}_\Phi^E)'\right)|\,\omega_{\varphi,-} - \omega_{\Phi,-}| \le$$

$$\frac{27}{8}c \cdot h^{p+1}\alpha^3 + \left(1 - \frac{h\alpha}{2}\right)|\,\omega_{\varphi,-} - \omega_{\Phi,-}|,$$

by Lemma 3.1, (22) and Lemma 4.1. Solving the above inequality for $|\,\omega_{\varphi,-} - \omega_{\Phi,-}| \equiv |\,id - J^E|(\omega_{\varphi,-})$ yields the desired result. ∎

Remark 4.1 on optimality. The next example shows that the distance of fixed points of normal forms satisfying (22) can be bounded from *below* by $\mathcal{O}(h^p)$ $(h \to 0)$.

Indeed, set $\mathcal{N}_\Phi(h,x,\alpha) := (1+h\alpha)x + hx^2$ and $\mathcal{N}_\varphi(h,x,\alpha) := (1+h\alpha)x + hx^2 + h^{p+1}x^3$. Then these maps satisfy (22) in a neighbourhood of the origin, further, $\omega_{\Phi,-} = -\alpha$ and $\omega_{\varphi,-} = \frac{-1+\sqrt{1-4h^p\,\alpha}}{2h^p}$. Using inequality $1+\frac{t}{2}-\frac{t^2}{4} \le \sqrt{1+t} \le 1+\frac{t}{2}-\frac{t^2}{8}$ for $-\frac{1}{2} \le t \le 0$, one sees that

$$|\,\omega_{\varphi,-} - \omega_{\Phi,-}| \ge h^p\,\alpha^2,$$

if, for example, $h \le 1$ and $\alpha \le \frac{1}{8}$.

4.2 The inner region

Now the closeness estimate in $(\omega_{\varphi,-},x_0]$ is proved for any fixed $0 < h \le h_0$ and $0 < \alpha \le \alpha_0$. It is clear that $\sup_{(\omega_{\varphi,-},x_0]} |\, id - J^E| = \sup_{n\in\mathbb{N}} \sup_{[x_{n+1},x_n]} |\, id - J^E|$.

Since $x_0 = J^E(x_0)$, we have that

$$\sup_{[x_1,x_0]} |\, id - J^E| = |x_1 - J^E(x_1)| = |\mathcal{N}_\varphi^E(x_0) - \mathcal{N}_\Phi^E(x_0)|$$

$$\le c\cdot h^{p+1}|x_0|^3 = \frac{c}{27}h^{p+1}\alpha^3,$$

while for $n \ge 1$

$$\sup_{[x_{n+1},x_n]} |\, id - J^E| \le \sup_{[x_{n+1},x_n]} \left|\mathcal{N}_\varphi^E \circ (\mathcal{N}_\varphi^E)^{[-1]} - \mathcal{N}_\Phi^E \circ (\mathcal{N}_\varphi^E)^{[-1]}\right| +$$

$$+ \sup_{[x_{n+1},x_n]} \left|\mathcal{N}_\Phi^E \circ (\mathcal{N}_\varphi^E)^{[-1]} - \mathcal{N}_\Phi^E \circ J^E \circ (\mathcal{N}_\varphi^E)^{[-1]}\right| =$$

$$= \sup_{[x_n,x_{n-1}]} \left|\mathcal{N}_\varphi^E - \mathcal{N}_\Phi^E\right| + \sup_{[x_n,x_{n-1}]} \left|\mathcal{N}_\Phi^E - \mathcal{N}_\Phi^E \circ J^E\right| \le$$

$$\le \sup_{[x_n,x_{n-1}]} \left|\mathcal{N}_\varphi^E - \mathcal{N}_\Phi^E\right| + \sup_{x\in[x_n,x_{n-1}]} \left(\left(\sup_{[\{x,J^E(x)\}]} (\mathcal{N}_\Phi^E)'\right) |\, x - J^E(x)|\right) \le$$

$$\le c\cdot h^{p+1}|x_n|^3 + \left(1 + h\alpha + \frac{7}{4}h \max\left(x_{n-1}, J^E(x_{n-1})\right)\right) \sup_{[x_n,x_{n-1}]} |\, id - J^E|,$$

the last inequality being true due to

$$\sup_{[\{x,J^E(x)\}]} (\mathcal{N}_\Phi^E)' \le \sup_{[\{x,J^E(x)\}]} (1 + h\alpha + \frac{7}{4}h\cdot id) \le 1 + h\alpha + \frac{7}{4}h \max\left(x, J^E(x)\right)$$

taking into account Lemma 4.1, then using the fact that the functions id and J^E are increasing.

From these we have for $n \ge 1$ that

$$\sup_{[x_{n+1},x_n]} |\, id - J^E| \le c\cdot h^{p+1} \sum_{i=0}^{n} |x_i|^3 \prod_{j=i}^{n-1} \left(1 + h\alpha + \frac{7}{4}h \max\left(x_j, J^E(x_j)\right)\right),$$

where $\prod_{j=n}^{n-1}$ is understood to be 1.

So in order to prove that the conjugacy J^E is $\mathcal{O}(h^p)$-close to the identity on the interval $(\omega_{\varphi,-},x_0]$ for any $h \in (0,h_0]$ and $\alpha \in (0,\alpha_0]$, it is enough to show that

$$\sup_{h\in(0,h_0]} \sup_{\alpha\in(0,\alpha_0]} \sup_{n\in\mathbb{N}} h \sum_{i=0}^{n} |x_i|^3 \prod_{j=i}^{n-1} \left(1 + h\alpha + \frac{7}{4}h \max\left(x_j, J^E(x_j)\right)\right) \le const \qquad (25)$$

holds with a suitable $const \ge 0$.

First an explicit estimate of the sequence $\max\left(x_n, J^E(x_n)\right)$ is given.

Lemma 4.3 *For $n \geq 0$, set*

$$a_n(h,\alpha) := -\frac{3}{4}\alpha \cdot \frac{(1+h\alpha)^{n+1}}{2+(1+h\alpha)^n},$$

then we have that $x_n \in (\omega_{\varphi,-}, a_n)$ and $J^E(x_n) \in (\omega_{\Phi,-}, a_n)$.

Proof. It is easily checked that, due to assumptions (23),

$$\max\left(\omega_{\varphi,-}, \omega_{\Phi,-}\right) < a_n$$

for $n \geq 0$, so the intervals in the lemma are non-degenerate. We proceed by induction.

$a_0 = -\frac{\alpha}{4}(1+h\alpha) > x_0 \equiv J^E(x_0) \equiv -\frac{\alpha}{3}$ is equivalent to $h\alpha < \frac{1}{3}$, being true by assumptions (23) on h_0 and α_0.

So suppose that the statement is true for some $n \geq 0$. Since $\mathcal{N}_\varphi^E(x) < (1+h\alpha)x + \frac{6}{5}h\,x^2$ is implied by $|x| \leq \varepsilon_0 < \frac{1}{5K}$, and $\mathcal{N}_\varphi^E$ is monotone increasing, we get that

$$x_{n+1} = \mathcal{N}_\varphi^E(x_n) < \mathcal{N}_\varphi^E(a_n) < (1+h\alpha)a_n + \frac{6}{5}h\,a_n^2,$$

thus it is enough to prove that the right-hand side above is smaller than a_{n+1}. But

$$a_{n+1} - \left((1+h\alpha)a_n + \frac{6}{5}h\,a_n^2\right) =$$

$$-\frac{3h\alpha^2(1+h\alpha)^{2+2n}\left(-2+(1+h\alpha)^n(-1+9h\alpha)\right)}{40\left(2+(1+h\alpha)^n\right)^2\left(2+(1+h\alpha)^{n+1}\right)} > 0$$

is equivalent to $-2+(1+h\alpha)^n(-1+9h\alpha) < 0$, which is implied by $h\alpha < \frac{1}{9}$.

Of course, the above inequalities remain true, if $\mathcal{N}_\varphi$ is replaced by $\mathcal{N}_\Phi$, also noticing that, by construction, $J^E(x_{n+1}) = \mathcal{N}_\Phi^E(J^E(x_n))$, so the induction is complete. ∎

Remark 4.2.1 The induction would fail, if, in estimate $\mathcal{N}_\varphi^E(x) < (1+h\alpha)x + \frac{6}{5}h\,x^2$, the constant $\frac{6}{5}$ was replaced by, say, $\frac{7}{5}$. (The explanation resides in the particular choice of the constant $\frac{3}{4}$ in the definition of a_n, since $\frac{3}{4}\cdot\frac{6}{5} < 1 < \frac{3}{4}\cdot\frac{7}{5}$.)

Remark 4.2.2 The upper estimate a_n in our first main lemma has been found by computer experiments with *Mathematica* based on the parametrized model function in [6].

In order to prove the boundedness of (25), the sum $\sum_{i=0}^{n}$ will be split into two. An appropriate index to split at is $\lceil \frac{const}{h\alpha} \rceil$, as established by the following lemma.

Lemma 4.4 *Suppose that $n > \lceil \frac{6}{h\alpha} \rceil$. Then*

$$\max\left(x_n, J^E(x_n)\right) < -\frac{2}{3}\alpha,$$

hence

$$1 + h\alpha + \frac{7}{4}h\max\left(x_n, J^E(x_n)\right) < 1 - \frac{h\alpha}{6}$$

holds for $n > \lceil \frac{6}{h\alpha} \rceil$.

Proof. By Lemma 4.3 it is sufficient to show that $n > \lceil \frac{6}{h\alpha} \rceil$ implies $a_n < -\frac{2}{3}\alpha$. This latter inequality is equivalent to $(1+h\alpha)^n(1+9h\alpha) > 16$. But if $n > \lceil \frac{6}{h\alpha} \rceil$, then

$$(1+h\alpha)^n > (1+h\alpha)^{\lceil \frac{6}{h\alpha} \rceil} = \left(1+\frac{1}{\frac{1}{h\alpha}}\right)^{(1+\frac{1}{h\alpha})\cdot\frac{h\alpha}{1+h\alpha}\cdot\lceil \frac{6}{h\alpha} \rceil}.$$

However, it is known that $\left(1+\frac{1}{A}\right)^{A+1} > e$, if $A \ge 1$, and it is easy to see that $\frac{B}{1+B}\cdot\lceil \frac{6}{B} \rceil > 3$, if $0 < B < 1$. Since $e^3 > 16$, the proof is complete. ■

Now we can turn to (25). Fix $h \in (0, h_0]$, $\alpha \in (0, \alpha_0]$ and $n \in \mathbb{N}^+$. (If $n \le \lceil \frac{6}{h\alpha} \rceil$, then the sums $\sum_{i=\lceil \frac{6}{h\alpha} \rceil+1}^{n}$ below are, of course, not present, making the proof even simpler.) Since now $\omega_{\varphi,-} < x_i < 0$, by Lemma 3.1 $|x_i| \le \frac{3}{2}\alpha$, and by monotonicity $\max\left(x_j, J^E(x_j)\right) \le x_0 \equiv J^E(x_0) \equiv -\frac{\alpha}{3}$, further, by using Lemma 4.4, assumption $h\alpha < 1$ from (23) and inequality $(1+\frac{1}{A})^A \le e$ (if $A \ge 1$), we get that

$$h\sum_{i=0}^{n} |x_i|^3 \prod_{j=i}^{n-1}\left(1+h\alpha+\frac{7}{4}h\max\left(x_j, J^E(x_j)\right)\right) \le$$

$$\frac{27h\alpha^3}{8}\sum_{i=0}^{\lceil \frac{6}{h\alpha} \rceil}\prod_{j=1}^{\lceil \frac{6}{h\alpha} \rceil-1}\left(1+h\alpha-\frac{7}{4}\cdot\frac{h\alpha}{3}\right) + \frac{27h\alpha^3}{8}\sum_{i=\lceil \frac{6}{h\alpha} \rceil+1}^{n}\prod_{j=i}^{n-1}\left(1-\frac{h\alpha}{6}\right) \le$$

$$\frac{27h\alpha^3}{8}\left(1+\frac{5}{12}h\alpha\right)^{\frac{6}{h\alpha}}\left(\left\lceil \frac{6}{h\alpha} \right\rceil + 1\right) + \frac{27h\alpha^3}{8}\sum_{i=\lceil \frac{6}{h\alpha} \rceil+1}^{n}\left(1-\frac{h\alpha}{6}\right)^{n-i} \le$$

$$\frac{27h\alpha^3}{8}\left(1+\frac{5}{12}h\alpha\right)^{\frac{12}{5h\alpha}\cdot\frac{5h\alpha}{12}\cdot\frac{6}{h\alpha}}\left(\frac{6+2h\alpha}{h\alpha}\right) + \frac{27h\alpha^3}{8}\sum_{i=0}^{\infty}\left(1-\frac{h\alpha}{6}\right)^{i} \le$$

$$\frac{27h\alpha^3}{8}\cdot e^{\frac{30}{12}}\cdot\frac{8}{h\alpha} + \frac{27h\alpha^3}{8}\cdot\frac{6}{h\alpha} \le 350\,\alpha^2.$$

Therefore, $\sup_{[x_{n+1},x_n]} |\, id - J^E| \le 350c \cdot h^p\alpha^2$ for any $h \in (0, h_0]$, $\alpha \in (0, \alpha_0]$ and $n \ge 1$, further, as we have seen, $\sup_{[x_1,x_0]} |\, id - J^E| \le \frac{c}{27}h^{p+1}\alpha^3$, which yield the following lemma.

Lemma 4.5 *Under assumption (23)*

$$\sup_{(\omega_{\varphi,-},x_0]} |\, id - J^E| \le 350c \cdot h^p\alpha^2.$$

Now the closeness estimate is proved in the interval $(y_0, \omega_{\varphi,0})$. Recall that $y_0 = x_0 = J^E(x_0) \equiv -\frac{\alpha}{3}$ and $\omega_{\varphi,0} = \omega_{\Phi,0} \equiv 0$.

Suppose that $n \ge 1$. (The case $n = 0$ will be examined later.) Then

$$\sup_{[y_n,y_{n+1}]} |\, id - J^E| = \sup_{[y_n,y_{n+1}]} \left|\left(\mathcal{N}_\Phi^E\right)^{[-1]} \circ \mathcal{N}_\Phi^E - \left(\mathcal{N}_\Phi^E\right)^{[-1]} \circ J^E \circ \mathcal{N}_\varphi^E\right| \le$$

$$\sup_{x\in[y_n,y_{n+1}]}\left[\left(\sup_{[\{\mathcal{N}_\Phi^E(x), J^E\circ\mathcal{N}_\varphi^E(x)\}]}\left(\left(\mathcal{N}_\Phi^E\right)^{[-1]}\right)'\right)\left(\left|\mathcal{N}_\Phi^E - \mathcal{N}_\varphi^E\right|(x) + \left|\mathcal{N}_\varphi^E - J^E\circ\mathcal{N}_\varphi^E\right|(x)\right)\right]$$

$$\leq \left[\sup_{x\in[y_n,y_{n+1}]} \sup_{[\{\mathcal{N}_\Phi^E(x),J^E\circ\mathcal{N}_\varphi^E(x)\}]} \left((\mathcal{N}_\Phi^E)^{[-1]} \right)' \right] \left[c\cdot h^{p+1}|y_n|^3 + \sup_{[y_{n-1},y_n]} |\, id - J^E| \right],$$

provided that $\sup_{[\{\mathcal{N}_\Phi^E(x),J^E\circ\mathcal{N}_\varphi^E(x)\}]} \left((\mathcal{N}_\Phi^E)^{[-1]} \right)'$ is nonnegative.

Lemma 4.6 *Suppose that $n \geq 1$, then under assumption (23) we have that*

$$\sup_{x\in[y_n,y_{n+1}]} \sup_{[\{\mathcal{N}_\Phi^E(x),J^E\circ\mathcal{N}_\varphi^E(x)\}]} \left((\mathcal{N}_\Phi^E)^{[-1]} \right)' \leq 1 - \frac{h\alpha}{8}.$$

Proof.

$$\sup_{x\in[y_n,y_{n+1}]} \sup_{[\{\mathcal{N}_\Phi^E(x),J^E\circ\mathcal{N}_\varphi^E(x)\}]} \left((\mathcal{N}_\Phi^E)^{[-1]} \right)' = \sup_{x\in[y_n,y_{n+1}]} \sup_{[\{\mathcal{N}_\Phi^E(x),J^E\circ\mathcal{N}_\varphi^E(x)\}]} \frac{1}{(\mathcal{N}_\Phi^E)'\circ(\mathcal{N}_\Phi^E)^{[-1]}}$$

$$= \sup_{x\in[y_n,y_{n+1}]} \sup_{[\{x,(\mathcal{N}_\Phi^E)^{[-1]}\circ J^E\circ\mathcal{N}_\varphi^E(x)\}]} \frac{1}{(\mathcal{N}_\Phi^E)'} = \dots$$

But, by definition, $(\mathcal{N}_\Phi^E)^{[-1]}\circ J^E\circ\mathcal{N}_\varphi^E(x) = J^E(x)$, if $x\in[y_n,y_{n+1}]$, and $[\{x, J^E(x)\}] = [\min(x, J^E(x)), \max(x, J^E(x))]$, further, by the monotonicity of id and J^E we obtain that

$$\dots = \sup_{[\min(y_n,J^E(y_n)),\max(y_{n+1},J^E(y_{n+1}))]} \frac{1}{(\mathcal{N}_\Phi^E)'} \leq \dots$$

By construction, however, $[\min(y_n, J^E(y_n)), \max(y_{n+1}, J^E(y_{n+1}))] \subset (y_0, 0) = (-\frac{\alpha}{3}, 0)$ and $(\mathcal{N}_\Phi^E)'$ is nonnegative here by assumption (23), justifying the computations just above the lemma. We now continue the proof of the lemma.

$$\dots \leq \sup_{(-\frac{\alpha}{3},0)} \frac{1}{(\mathcal{N}_\Phi^E)'} \leq \dots$$

It is easy to see that assumption (23) together with $x<0$ imply that $(\mathcal{N}_\Phi^E)'(x) \geq 1 + h\alpha + \frac{9}{4}hx \geq 0$. So

$$\dots \leq \sup_{x\in(-\frac{\alpha}{3},0)} \frac{1}{1+h\alpha+\frac{9}{4}hx} \leq \frac{1}{1+h\alpha+\frac{9}{4}h\left(-\frac{\alpha}{3}\right)} = \frac{1}{1+\frac{1}{4}h\alpha} \leq 1 - \frac{h\alpha}{8},$$

since $\frac{1}{1+A} \leq 1 - \frac{A}{2}$, if $A\in[0,1]$. ∎

We have thus proved (also using $|y_n| \leq \frac{\alpha}{3}$) that for $n \geq 1$

$$\sup_{[y_n,y_{n+1}]} |\, id - J^E| \leq \left(1 - \frac{h\alpha}{8}\right) \left[\frac{c}{27}\cdot h^{p+1}\alpha^3 + \sup_{[y_{n-1},y_n]} |\, id - J^E| \right] \qquad (26)$$

For $n=0$, similarly as before, we get that

$$\sup_{[y_0,y_1]} |\, id - J^E| \leq \left[\sup_{x\in[y_0,y_1]} \sup_{[\{\mathcal{N}_\Phi^E(x),J^E\circ\mathcal{N}_\varphi^E(x)\}]} \left((\mathcal{N}_\Phi^E)^{[-1]} \right)' \right] \left[c\cdot h^{p+1}|y_0|^3 + \sup_{[y_{-1},y_0]} |\, id - J^E| \right].$$

But $[y_{-1}, y_0] \equiv [x_1, x_0]$, so the second factor $[\dots]$ is bounded by $2\cdot\frac{c}{27}h^{p+1}\alpha^3$. As for the first factor $[\dots]$, we notice that $y_0 < (\mathcal{N}_\Phi^E)^{[-1]}(y_0)$ (since this is equivalent to $x_1 < x_0$), which implies that

$$\sup_{x\in[y_0,y_1]} \sup_{[\{\mathcal{N}_\Phi^E(x),J^E\circ\mathcal{N}_\varphi^E(x)\}]} \left((\mathcal{N}_\Phi^E)^{[-1]} \right)' = \sup_{x\in[y_0,y_1]} \sup_{[\{x,(\mathcal{N}_\Phi^E)^{[-1]}\circ J^E\circ\mathcal{N}_\varphi^E(x)\}]} \frac{1}{(\mathcal{N}_\Phi^E)'} =$$

$$\sup_{[y_0,y_1]\cup[y_0,(\mathcal{N}_\Phi^E)^{[-1]}(y_0)]} \frac{1}{(\mathcal{N}_\Phi^E)'} \le \sup_{[y_0,0)} \frac{1}{(\mathcal{N}_\Phi^E)'} \le 1,$$

therefore

$$\sup_{[y_0,y_1]} |\,id - J^E| \le 2\cdot \frac{c}{27} h^{p+1}\alpha^3. \tag{27}$$

Repeated application of (26), further (27) yield for $n \ge 1$ that

$$\sup_{[y_n,y_{n+1}]} |\,id - J^E| \le \left(1-\frac{h\alpha}{8}\right)^n \sup_{[y_0,y_1]} |\,id - J^E| + \frac{c}{27}h^{p+1}\alpha^3 \sum_{i=1}^{n}\left(1-\frac{h\alpha}{8}\right)^i \le$$

$$1\cdot 2\cdot \frac{c}{27}h^{p+1}\alpha^3 + \frac{c}{27}h^{p+1}\alpha^3\cdot\frac{8}{h\alpha} \le \frac{c}{3}h^p\alpha^2,$$

due to $h\alpha \le \frac{1}{2}$ by (23). The same upper estimate is valid for $n = 0$, so we have proved the following result.

Lemma 4.7 *Under assumption (23)*

$$\sup_{(x_0,0)} |\,id - J^E| \le \frac{c}{3}h^p\alpha^2.$$

5 The outer region

In this section, we first prove an $\mathcal{O}(h^p)$ closeness-estimate in the interval $[z_0, \omega_{\varphi,-})$ for $\alpha > 0$. Then, in the second part, the closeness is proved on $[z_0, \omega_{\Phi,0}) \equiv [z_0, 0)$ for $\alpha \le 0$.

The derivation of the following formulae is similar to their counterparts in the inner region, with the difference that—since this time the sequence z_n is increasing—an extra term and an index-shift occur.

For $n \ge 1$ (also using (23)) we have that

$$\sup_{[z_n,z_{n+1}]} |\,id - J^E| \le c\cdot h^{p+1}|z_0|^3 \prod_{j=1}^{n}\left(1+h\alpha+\frac{7}{4}h\max\left(z_j, J^E(z_j)\right)\right) +$$

$$c\cdot h^{p+1}\sum_{i=0}^{n-1}|z_i|^3 \prod_{j=i+2}^{n}\left(1+h\alpha+\frac{7}{4}h\max\left(z_j, J^E(z_j)\right)\right), \tag{28}$$

where, again $\prod_{j=n+1}^{n}$ above is 1, and

$$\sup_{[z_0,z_1]} |\,id - J^E| \le c\cdot h^{p+1}|z_0|^3.$$

The following main lemma, as a counterpart of Lemma 4.3, gives a lower estimate of the sequence z_n, if $\alpha > 0$.

Lemma 5.1 *For $n \ge 0$, set*

$$b_n(h,\alpha) := -2\alpha\cdot\frac{(1+h\alpha)^{n+1}}{-1+\alpha+(1+h\alpha)^n},$$

then $b_n \le \min\left(z_n, J^E(z_n)\right)$.

Proof. $b_0 = 2 - 2h\alpha < -2 \le -1 \le -\varepsilon_0 \le z_0 = J^E(z_0)$ holds due to assumption (23). Suppose that the statement is true for some $n \ge 0$. Since $\mathcal{N}_\varphi^E(x) \ge (1 + h\alpha)x + \frac{3}{5}h\,x^2$ follows from $|x| \le \varepsilon_0 < \frac{2}{5K}$, further $(1+h\alpha)id + \frac{3}{5}h\,id^2$ is monotone increasing (which is implied by, *e.g.*, $|x| \le \frac{5}{6h}$, but it is easy to see that $h \le \frac{5}{18}$ and $-3 < b_n < 0$ follows from (23), hence $|b_n| \le \frac{5}{6h}$), so we obtain that

$$z_{n+1} = \mathcal{N}_\varphi^E(z_n) \ge (1+h\alpha)z_n + \frac{3}{5}h\,z_n^2 \ge (1+h\alpha)b_n + \frac{3}{5}h\,b_n^2,$$

thus it is sufficient to show that

$$(1+h\alpha)b_n + \frac{3}{5}h\,b_n^2 \ge b_{n+1}.$$

However, this is equivalent to

$$0 \le \frac{2h\alpha^2(1+h\alpha)^{2+2n}}{5\,(-1+\alpha+(1+h\alpha)^n)^2} \cdot \frac{-1+\alpha+(1+h\alpha)^n(1+6h\alpha)}{-1+\alpha+(1+h\alpha)^{n+1}},$$

which is true since $\alpha > 0$ and $h > 0$.

The proof remains valid if $\mathcal{N}_\varphi$ is replaced by $\mathcal{N}_\Phi$ (and $J^E(z_n)$ is written instead of z_n), hence $b_n \le J^E(z_n)$ also holds. ■

Now, since $z_j < \omega_{\varphi,-}$ and $J^E(z_j) < \omega_{\Phi,-}$, by Lemma 3.1 we get that the right-hand side of (28) is at most

$$c \cdot h^{p+1}|z_0|^3 \prod_{j=1}^{n}\left(1 - \frac{h\alpha}{2}\right) + c \cdot h^{p+1}\sum_{i=0}^{n-1}|z_i|^3 \prod_{j=i+2}^{n}\left(1 - \frac{h\alpha}{2}\right) \le$$

$$c \cdot h^{p+1}|z_0|^3 + c \cdot h^{p+1}\sum_{i=0}^{n-1}|z_i|^3\left(1 - \frac{h\alpha}{2}\right)^{n-1-i}.$$

We will verify that $h\sum_{i=0}^{n}|z_i|^3\left(1 - \frac{h\alpha}{2}\right)^{n-i}$ is uniformly bounded for any $n \ge 0$, $0 < h \le h_0$ and $0 < \alpha \le \alpha_0$.

If $n \ge \lceil \frac{1}{h\alpha} \rceil$, then by Lemma 5.1 (also using that $h\alpha \le \frac{1}{9}$ and $z_j < 0$)

$$h\sum_{i=\lceil \frac{1}{h\alpha}\rceil}^{n}|z_i|^3\left(1 - \frac{h\alpha}{2}\right)^{n-i} \le h\sum_{i=\lceil \frac{1}{h\alpha}\rceil}^{n}|b_i|^3\left(1 - \frac{h\alpha}{2}\right)^{n-i} \le$$

$$11h\alpha^3\sum_{i=\lceil \frac{1}{h\alpha}\rceil}^{n}\left(\frac{(1+h\alpha)^i}{-1+\alpha+(1+h\alpha)^i}\right)^3\left(1 - \frac{h\alpha}{2}\right)^{n-i} \le \ldots$$

for these i indices however $\frac{(1+h\alpha)^i}{-1+\alpha+(1+h\alpha)^i} \le 3$ holds (since this is implied by $\frac{3}{2} \le (1+h\alpha)^i$, being true by $(1+h\alpha)^i \ge (1+h\alpha)^{\frac{1}{h\alpha}} \ge 1 + \frac{1}{h\alpha}\cdot h\alpha > \frac{3}{2}$), thus

$$\ldots \le 27 \cdot 11\alpha^2 h\alpha \sum_{i=0}^{\infty}\left(1 - \frac{h\alpha}{2}\right)^i = 594\alpha^2.$$

On the other hand, if $n < \lceil \frac{1}{h\alpha} \rceil$, then (using that $|z_i| \le 1$ and $h\alpha \le \frac{1}{9}$ again)

$$h\sum_{i=0}^{n}|z_i|^3\left(1 - \frac{h\alpha}{2}\right)^{n-i} \le h\sum_{i=0}^{n}|z_i|^2\left(1 - \frac{h\alpha}{2}\right)^{n-i} \le \tag{29}$$

$$5h\sum_{i=0}^{n}\left(\frac{\alpha(1+h\alpha)^i}{-1+\alpha+(1+h\alpha)^i}\right)^2\left(1-\frac{h\alpha}{2}\right)^{n-i}\leq\ldots$$

now using inequalities $e^{\frac{x}{2}}\leq 1+x$ ($x\in[0,1]$) and $1+x\leq e^x$ ($x\in\mathbb{R}$) we get that $(1+h\alpha)^{2i}\leq e^{h\alpha 2i}\leq e^{h\alpha 2n}\leq e^2<8$, further, $\left(1-\frac{h\alpha}{2}\right)^{n-i}\leq e^{-\frac{h\alpha}{2}(n-i)}$ and $e^{\frac{h\alpha}{2}i}\leq(1+h\alpha)^i$, therefore

$$\ldots\leq 40h\sum_{i=0}^{n}\left(\frac{\alpha e^{-\frac{h\alpha}{4}(n-i)}}{-1+\alpha+e^{\frac{h\alpha}{2}i}}\right)^2.$$

Set $g_{h,\alpha,n}(x)\equiv g(x):=\left(\frac{\alpha\exp\left(-\frac{1}{4}h\alpha(n-x)\right)}{-1+\alpha+\exp\left(\frac{1}{2}h\alpha x\right)}\right)^2$, if $x\in[0,\infty)$. Notice that g is bounded at $x=0$. For this function we have that

$$g'(x)=-\frac{1}{2}h\alpha^3\,e^{-\frac{1}{2}h\alpha(n-x)}\cdot\frac{1-\alpha+e^{\frac{1}{2}hx\alpha}}{\left(-1+\alpha+e^{\frac{1}{2}hx\alpha}\right)^3},$$

meaning that g is strictly monotone decreasing, if $\alpha<1$. Hence

$$40h\sum_{i=0}^{n}\left(\frac{\alpha e^{-\frac{h\alpha}{4}(n-i)}}{-1+\alpha+e^{\frac{h\alpha}{2}i}}\right)^2=40h+40h\sum_{i=1}^{n}g_{h,\alpha,n}(i)\leq$$

$$40h+40h\int_0^n g_{h,\alpha,n}(x)\mathrm{d}x=40h+40h\left[-2\alpha\frac{\exp\left(-\frac{1}{2}h\alpha n\right)}{h\left(-1+\alpha+\exp\left(\frac{1}{2}h\alpha x\right)\right)}\right]_{x=0}^{n}=$$

$$40h+40h\left(\frac{2\left(1-\exp\left(-\frac{1}{2}h\alpha n\right)\right)}{h\left(\exp\left(\frac{1}{2}h\alpha n\right)-1+\alpha\right)}\right)\leq 40h+80\left(\frac{1-\exp\left(-\frac{1}{2}h\alpha n\right)}{\exp\left(\frac{1}{2}h\alpha n\right)-1}\right)=$$

$$40h+80e^{-\frac{1}{2}h\alpha n}\leq 120,$$

since $h\leq 1$.

Now combining all the estimates so far in the section, under assumption (23) we get that if $\alpha>0$, then

$$\sup_{[z_0,\omega_{\varphi,-})}|id-J^E|=\sup_{n\in\mathbb{N}}\sup_{[z_n,z_{n+1}]}|id-J^E|\leq$$

$$\sup_{n\in\mathbb{N}}\max\left(c\cdot h^{p+1}|z_0|^3,\ c\cdot h^{p+1}|z_0|^3+c\cdot h^{p+1}\sum_{i=0}^{n}|z_i|^3\left(1-\frac{h\alpha}{2}\right)^{n-i}\right)\leq$$

$$c\cdot h^{p+1}|z_0|^3+c\cdot h^p\cdot(120+594\alpha^2)\leq 130c\cdot h^p.$$

Remark 5.1 If, in (29), the exponent of $|z_i|$ had not been changed to 2, then the integral of g would have been significantly more complicated. (Interestingly, similar complication occurs, if one considers simply $|z_i|$ instead of $|z_i|^2$.) The rational pair $\frac{1}{4}$ and $\frac{1}{2}$ in the definition of g has also been a fortunate choice: when working with the numbers $\frac{1}{5}$ and $\frac{1}{2}$ instead, for example, *Mathematica* produced so complicated integrals that were practically useless from the viewpoint of further analysis.

Finally, we prove a closeness estimate on $[z_0, 0)$ for $\alpha \le 0$. We begin with a simple observation on monotonicity of the sequence $z_n \equiv z_n(\alpha)$. (As before, for brevity, the dependence on h is still suppressed.)

Lemma 5.2 *Suppose that $\alpha \le 0$ and assumption (23) hold. Then for any $0 < h \le h_0$, $-\alpha_0 \le \alpha \le \beta \le 0$ and $n \in \mathbb{N}$ we have that*

$$0 > z_n(\alpha) \ge z_n(\beta).$$

Proof. By definition, we have that $z_0(\alpha) = z_0(\beta) = z_0$, so suppose that for some n we already know that $z_n(\alpha) \ge z_n(\beta)$. Then, by the definition of the sequence z_n, further by the facts that the function $z \mapsto \mathcal{N}_\varphi(h, z, \alpha)$ is monotone *increasing* and the function $\alpha \mapsto \mathcal{N}_\varphi(h, z, \alpha)$ is monotone *decreasing*, we get that

$$z_{n+1}(\alpha) = \mathcal{N}_\varphi(h, z_n(\alpha), \alpha) \ge \mathcal{N}_\varphi(h, z_n(\beta), \alpha) \ge \mathcal{N}_\varphi(h, z_n(\beta), \beta) = z_{n+1}(\beta),$$

which completes the induction. ■

This means that $0 > z_n(\alpha) \ge z_n(0)$ holds for $\alpha \le 0$, hence it is enough to give a lower estimate for $z_n(0)$. But such an estimate has been constructed in Lemma 3.3 [1], namely we recall the following.

Lemma 5.3 *Under assumption (23), we have for $n \in \mathbb{N}$ that*

$$z_n(0) \ge z_0$$

and for $n \ge \lfloor \frac{1}{h} \rfloor + 1$

$$z_n(0) \ge -\frac{2}{nh}.$$

Then we can simply estimate (28) for $\alpha \le 0$ as follows. Supposing that $n \ge 1$ we get that

$$\sup_{[z_n, z_{n+1}]} |\, id - J^E| \le c \cdot h^{p+1} |z_0|^3 \prod_{j=1}^{n} \left(1 + h\alpha + \frac{7}{4} h \max\left(z_j, J^E(z_j)\right)\right) +$$

$$c \cdot h^{p+1} \sum_{i=0}^{n-1} |z_i|^3 \prod_{j=i+2}^{n} \left(1 + h\alpha + \frac{7}{4} h \max\left(z_j, J^E(z_j)\right)\right) \le$$

$$c \cdot h^{p+1} |z_0|^3 \cdot 1^{n-1} + c \cdot h^p \cdot h \sum_{i=0}^{n} |z_i(0)|^3 \cdot 1^{n-i-1} \le$$

$$c \cdot h^p \left(h|z_0|^3 + h \sum_{i=0}^{\lfloor \frac{1}{h} \rfloor} |z_i(0)|^2 + h \sum_{i=\lfloor \frac{1}{h} \rfloor + 1}^{n} |z_i(0)|^2 \right),$$

where, of course, for $n \le \lfloor \frac{1}{h} \rfloor$, the sum above $\sum_{i=\lfloor \frac{1}{h} \rfloor + 1}^{n}$ should be omitted. But

$$h \sum_{i=0}^{\lfloor \frac{1}{h} \rfloor} |z_i(0)|^2 \le h \cdot \frac{1}{h} \cdot z_0^2 = z_0^2,$$

and

$$h \sum_{i=\lfloor \frac{1}{h} \rfloor +1}^{n} |z_i(0)|^2 \leq h \sum_{i=\lfloor \frac{1}{h} \rfloor +1}^{n} \frac{4}{i^2 h^2} \leq \frac{4}{h} \int_{\frac{1}{h}-1}^{\infty} \frac{1}{i^2} = \frac{4}{1-h} \leq 8.$$

We have thus proved that

$$\sup_{[z_0,0)} |\, id - J^E| \leq 10c \cdot h^p.$$

References

[1] L. Lóczi *Construction of a conjugacy and closeness estimates in the discretized fold bifurcation*, Preprint 04/022, DFG research group "Spectral analysis, asymptotic distributions and stochastic dynamics", University of Bielefeld 2004, available via
`www.mathematik.uni-bielefeld.de/fgweb/Preprints/index04.html`

[2] S. Wiggins, *Introduction to Applied Nonlinear Dynamical Systems and Chaos*, Springer-Verlag, New York, 1990.

[3] P. Glendinning, *Stability, Instability and Chaos*, Cambridge Univ. Press, Cambridge, 1994.

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

[5] L. Lóczi *A normal form for the fold bifurcation and its discretization*, Preprint 03/006, DFG research group "Spectral analysis, asymptotic distributions and stochastic dynamics", University of Bielefeld 2003, available via
`www.mathematik.uni-bielefeld.de/fgweb/Preprints/index03.html`

[6] T. Hüls, *A Model Function for Polynomial Rates in Discrete Dynamical Systems*, Appl. Math. Letters **17** (2004) 1–5.