Runge-Kutta discretizations of singularly perturbed gradient equations

Wolf–Jürgen Beyn*
Dept. of Mathematics
University of Bielefeld
P.O. Box 100131
D-33501 Bielefeld

Johannes Schropp †
Dept. of Mathematics and Computer Science
University of Konstanz
P.O. Box 5560
D-78434 Konstanz

Abstract

We analyze Runge-Kutta discretizations applied to singularly perturbed gradient systems. It is shown in which sense the discrete dynamics preserve the geometric properties and the longtime behavior of the underlying ordinary differential equation. If the continuous system has an attractive invariant manifold then numerical trajectories started in some neighbourhood (the size of which is independent of the step-size and the stiffness parameter) approach an equilibrium in a nearby manifold. The proof combines invariant manifold techniques developed by Nipp and Stoffer for singularly perturbed systems with some recent results of the second author on the global behavior of discretized gradient systems. The results support the favorable behavior of ODE methods for stiff minimization problems

1 Introduction

Dynamical systems with a gradient structure occur quite naturally when modelling systems that show a constant decay of some energy function. They also occur as auxiliary systems for minimization problems when searching for trajectories that approach a minimum. It is well known that the ω-limit sets of orbits in gradient systems consist of equilibria and the connecting orbits between them, see e.g. Hirsch, Smale (1974). In addition, these equilibria together with their unstable manifolds determine the general structure of the global attractor, see e.g. Hale and Raugel (1989).

Despite this seemingly simple situation gradient systems can show a rich transient and asymptotic behavior. Further analytic treatment of the systems is often impossible and

*supported by SFB 343 "Diskrete Strukturen in der Mathematik".
†The second author thanks the SFB 343 for supporting two visits to the University of Bielefeld.

numerical simulations become important in gaining a deeper understanding of the global behavior. Here the question arises which qualitative properties of the system are preserved by a numerical method. Some results about discretized gradient systems can be found for one-step methods in Stuart and Humphries (1994), (1996), Schropp (1995) and for linear multistep methods in Schropp (1997).

These results apply to nonstiff situations. In this paper we will treat the stiff gradient case. Such stiff systems occur quite often when the function f in the gradient system $\dot{x} = -\nabla f(x)$ arises from a highly nonlinear unconstrained minimization problem (see e.g. Schropp (1995), Schropp (1997)). In such a case the norm of the Jacobian typically becomes large, the allowed step sizes in the results of Humphries, Stuart (1994), Schropp (1995) tend to zero and the corresponding convergence theorems become useless for practical purposes.

In the present paper we analyze the behavior of Runge-Kutta discretizations applied to gradient equations of singular perturbation type. We prove global convergence of numerical trajectories for step sizes that are not restricted by the stiffness parameter. Moreover, the region of attraction will be a neighborhood of some invariant manifold that does not depend on both parameters.

We use the invariant manifold theorem of Nipp, Stoffer (1995) in order to reduce the discrete singularly perturbed problem to a parameter dependent regular problem. Our main result will then be obtained by a parameter dependent version of the global convergence properties for the nonstiff case, see Schropp (1995). Our results support the advantage of using stiff integrators on gradient systems for solving nonlinear least squares problems, see Schropp (1997). An illustrative example of this type will be given in section 2.

2 The main results

We consider a singularly perturbed autonomous system

$$\begin{aligned} \dot{x} &= f(x,y), \\ \epsilon\dot{y} &= g(x,y), \ 0 < \epsilon \ll 1 \end{aligned} \tag{2.1}$$

for $x \in \mathbb{R}^M$, $y \in \mathbb{R}^N$ and $\epsilon \in]0, \epsilon_0[$ of gradient type

$$f(x,y) := -\frac{\partial k}{\partial x}(x,y), \qquad g(x,y) := -\frac{\partial k}{\partial y}(x,y), \ \ k : \mathbb{R}^{M+N} \to \mathbb{R}. \tag{2.2}$$

Notice that (2.1) is in fact a gradient system as can be seen from the scaling

$$\xi = x, \quad \eta = \sqrt{\epsilon}y, \quad \kappa(\xi,\eta) = k(x, \frac{\eta}{\sqrt{\epsilon}}).$$

For the dynamical system (2.1) we assume the following

A1: The functions f, g are sufficiently smooth with globally bounded derivatives.

A2: There is a smooth function $s_0 : \mathbb{R}^M \to \mathbb{R}^N$ such that $g(x, s_0(x)) = 0$ for $x \in \mathbb{R}^M$.

A3: $\mu_2(\frac{\partial g}{\partial y}(x, s_0(x))) \leq -b_0$ holds for some $b_0 > 0$ and for all $x \in \mathbb{R}^M$.

In (A3) we used the logarithmic norm μ_2 of a matrix B defined by

$$\mu_2(B) \quad = \quad \lim_{\delta \to 0} \frac{1}{\delta}(\| I + \delta B \|_2 - 1).$$

Our final assumption is typical when dealing with the longtime behavior of gradient systems

A4: The equilibria of the reduced system

$$\dot{x} = f(x, s_0(x))$$

are hyperbolic, that is, at any stationary point of this system the Jacobian of the right hand side has no eigenvalues on the imaginary axis.

The Jacobian of the reduced system at an equilibrium is given by

$$\Phi(x) = \left(\frac{\partial f}{\partial x} - \frac{\partial f}{\partial y} \frac{\partial g}{\partial y}^{-1} \frac{\partial g}{\partial x} \right) (x, s_0(x)),$$

i.e. the Schur complement of $\frac{\partial f}{\partial x}$ with respect to the Jacobian of (f, g). Moreover, it is also easily seen that a similarity transformation of the Jacobian of $(f, \epsilon^{-1} g)$ leads to

$$\begin{pmatrix} \Phi(x) & 0 \\ \frac{1}{\epsilon}\frac{\partial g}{\partial x} & \frac{1}{\epsilon}\frac{\partial g}{\partial y} + \frac{\partial g}{\partial x}\frac{\partial f}{\partial y}\frac{\partial g}{\partial y}^{-1} \end{pmatrix} \; + \; O(\epsilon).$$

Thus, using (A3) we obtain that (A4) holds iff the stationary points of the system (2.1) on the manifold

$$M_0 = \{(x, y) \in \mathbb{R}^{M+N} \mid \; y = s_0(x)\}$$

are hyperbolic (compare Lubich, Nipp and Stoffer (1995)). Moreover the number of unstable eigenvalues is the same for the reduced and the full system.

Under the assumptions (A1)-(A3) Nipp (1985) (see also the general theory by Fenichel (1979)) has shown that for $\epsilon > 0$ small enough equation (2.1) admits an attractive, invariant manifold

$$M_\epsilon = \{(x, y) \in \mathbb{R}^{M+N} \mid \; y = s_\epsilon(x)\}$$

which is $O(\epsilon)$ close to M_0. In addition, the property of asymptotic phase holds (see Nipp, Stoffer (1995), Th. 1). The characterization of the longtime behaviour of the singularly perturbed gradient dynamics is the content of

2.1 Lemma: *Consider the model equation (2.1), (2.2) and assume (A1)-(A4) . Then there exist $\delta_0 > 0$ and $\epsilon_0 > 0$ such that any bounded solution of (2.1),(2.2) with $0 < \epsilon < \epsilon_0$ and* *initial value* (x_0, y_0) *satisfying* $\| y_0 - s_0(x_0) \| < \delta_0$ *converges towards a stationary point in* M_ϵ *as* $t \to \infty$.

As noted above, the system (2.1), (2.2) may be written in gradient form. Lemma 2.1 can then be proved by applying the general convergence result for gradient systems to the transformed system. The assumptions (A1)-(A4) guarantee that the equilibria of the transformed system are all hyperbolic in a neighborhood of the invariant manifold. Nevertheless we present an alternative direct proof in Section 3 because we will mimic this approach in the discrete case.

We are interested in the behaviour of s-stage Runge-Kutta methods with Butcher tableau

$$\begin{array}{c|c} c & A \\ \hline & b^T \end{array}, \quad A = (a_{ij})_{1 \le i,j \le s} \in \mathbb{R}^{s,s}, \; b, c \in \mathbb{R}^s$$

as applied to (2.1) with step size $h \ge \epsilon$. The discrete iteration has the form

$$\begin{aligned} x_{n+1} &= x_n + h(b^T \otimes I)\bar{f}(X^n, Y^n), \\ y_{n+1} &= y_n + \frac{h}{\epsilon}(b^T \otimes I)\bar{g}(X^n, Y^n) \end{aligned} \tag{2.3}$$

where $X^n = (X_1^n, \ldots, X_s^n) \in \mathbb{R}^{Ms}$, $Y^n = (Y_1^n, \ldots, Y_s^n) \in \mathbb{R}^{Ns}$ denote the solution of the algebraic system

$$\begin{aligned} U - (1\!\!1 \otimes x_n) &= h(A \otimes I)\bar{f}(U, V), \\ V - (1\!\!1 \otimes y_n) &= \frac{h}{\epsilon}(A \otimes I)\bar{g}(U, V). \end{aligned} \tag{2.4}$$

Here, $\bar{f}, \bar{g}$ stand for $\bar{f}(X^n, Y^n) = (f(X_1^n, Y_1^n), \ldots, f(X_s^n, Y_s^n))$, $\bar{g}(X^n, Y^n) = (g(X_1^n, Y_1^n), \ldots, g(X_s^n, Y_s^n))$ and $1\!\!1$ denotes $(1, \ldots, 1)^T \in \mathbb{R}^s$. The Runge-Kutta method has stage order q, if

$$\sum_{j=1}^{s} a_{ij} c_j^{k-1} = \frac{c_i^k}{k}, \; k = 1, \ldots, q, \; i = 1, \ldots, s$$

holds. As in Hairer, Lubich and Roche (1988) we impose the following condition on the Runge-Kutta method (2.3):

- B1: The Runge-Kutta method is A-stable, that is, the stability function $R(z) := 1 + zb^T(I - zA)^{-1}1\!\!1$ satisfies $| \mathrm{R}(z) | \le 1$ for $\mathrm{Re}(z) \le 0$.
- B2: $\mathrm{Re}(\lambda) > 0$ holds for $\lambda \in \sigma(A)$.
- B3: $R(\infty) = 1 - b^T A^{-1} 1\!\!1$ satisfies $| R(\infty) | < 1$.
- B4: The method has order p in the classical sense and stage order $q \ge 1$. In addition, let $p \ge q + 1$ for $p \ge 2$ or $p = q = 1$ hold.

In the case $\epsilon \ll h$, it is sufficient to require the invertibility of A, (B3) and (B4).
When applying the Runge-Kutta method (2.3) to equation (2.1) one is interested in the geometric properties of the discrete dynamical system. In a remarkable paper Nipp and Stoffer (1995) showed under the assumptions above that the resulting discrete dynamics inherits the crucial properties from the smooth analog. To be more precise, the one-step mapping possesses an attractive, invariant manifold

$$M_{\epsilon,h} = \{(x,y) \in \mathbb{R}^{M+N} \mid y = s_{\epsilon,h}(x)\}$$

which is $O(h^{q+1})$ close to M_ϵ and for which the property of asymptotic phase holds.

The main result of this paper is the following discrete analog of Lemma 2.1 for singularly perturbed gradient systems.

2.2 Theorem: *Let the conditions (A1)-(A4) hold for the system (2.1), (2.2) and suppose that the Runge-Kutta method satisfies (B1)-(B4). By (x_n, y_n) we denote the sequences generated by the Runge-Kutta method (2.3) for the system (2.1), (2.2). Then $d_0 > 0$ exists such that the Runge-Kutta iteration is well defined for $0 < \epsilon < \epsilon_0$, $0 < h < h_0$ and initial values (x_0, y_0) satisfying $\| y_0 - s_0(x_0) \| < d_0$. Consider a compact set $K = K_1 \times K_2 \subset \mathbb{R}^M \times \mathbb{R}^N$ and parameter values $0 < \epsilon \leq h < h_0$. Then all Runge Kutta iterates that stay in K for $n \in \mathbb{N}$ satisfy*

$$\lim_{n\to\infty} (x_n, y_n) = (\bar{x}, \bar{y})$$

for some stationary point $(\bar{x}, \bar{y})$ of the system (2.1), (2.2).

Let us discuss the consequences of Theorem 2.2. The ϵ independent convergence result for our model equation (2.1), (2.2) confirmes that stiff gradient equations should be attacked with implicit A-stable Runge-Kutta methods. In particular, one should use Runge-Kutta methods satisfying $R(\infty) = 0$, because otherwise the attractivity of M_ϵ is only poorly reproduced by the attractivity of $M_{\epsilon,h}$ (see Remark 5 to Th. 2 in Nipp, Stoffer (1995)). For such methods, after a short transient phase of approximating $M_{\epsilon,h}$, the discrete dynamics is governed by the gradient dynamics restricted to $M_{\epsilon,h}$. This process continues till the convergence towards a fixed point takes place.
Let us now illustrate this scenario by an example from nonlinear least squares which is taken from Schropp (1995). With the function f from formula (4.2) in Schropp (1995) playing the role of k we set up our model equation (2.1), (2.2). Then, the corresponding phase portraits for various values of ϵ are displayed in Figures 1 and 2.

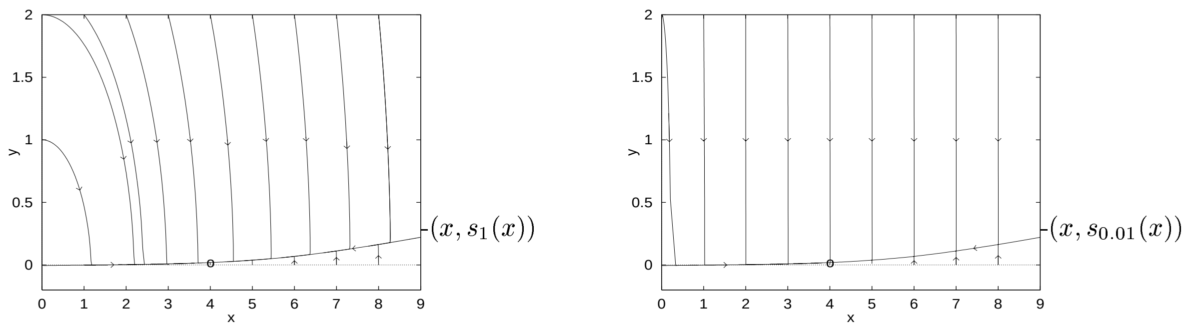

Fig. 1, 2: Phase portrait for $\epsilon = 1$ (left) and $\epsilon = 0.01$ (right)

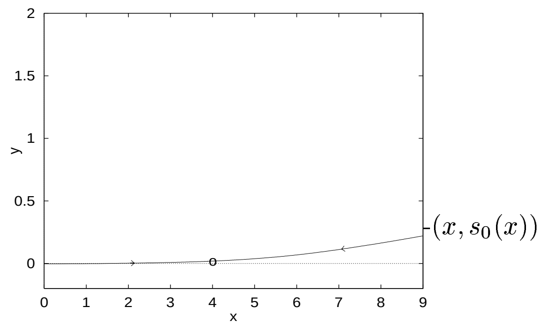

Fig. 3: Phase portrait of the differential algebraic equation ($\epsilon = 0$)

In the above figures the label "$\circ$" indicates the stationary point which is the local solution of the underlying minimization problem. Figure 1 shows the stiffness of the original problem with $\epsilon = 1$. This means that $\epsilon = 1$ is already "small" in our example and $(x, s_1(x))$ is the attractive, invariant manifold. For $\epsilon = 10^{-2}$ the dynamics in y-direction is nearly vertical. Finally, Figure 3 shows the graph of s_0 which is quite close to the graph of s_1.

The rest of the paper is organized as follows. In Section 3 we present a proof of Lemma 2.1 and in Section 4 we show that the Runge-Kutta iteration is well defined. In Section 5 a proof of Theorem 2.2 will be given. Here we use some technical properties of stable manifolds which will finally be proved in Section 6.

3 Singularly perturbed gradient equations

Here we analyze the qualitative behavior of the solutions of the singularly perturbed initial value problem

$$\begin{aligned} \dot{x} &= f(x,y) = -\frac{\partial k}{\partial x}(x,y),\ x(0)=x_0, \\ \epsilon\dot{y} &= g(x,y) = -\frac{\partial k}{\partial y}(x,y),\ y(0)=y_0. \end{aligned} \tag{3.1}$$

For $0<\epsilon<\epsilon_0$, $\epsilon_0>0$ sufficiently small, due to Nipp (1985), equation (3.1) possesses an attractive invariant manifold

$$M_\epsilon = \{(x,y)\in \mathrm{I\!R}^{M+N} \mid y=s_\epsilon(x)\}.$$

Differentiating the invariance relation yields $g(x,s_\epsilon(x))=\epsilon Ds_\epsilon(x)f(x,s_\epsilon(x))$. Thus, equation (3.1) restricted to M_ϵ is equivalent to the smooth parameter dependent autonomous system

$$\dot{x} = f(x,s_\epsilon(x)) = -\frac{\partial k}{\partial x}(x,s_\epsilon(x)),\ x(0)=x_0. \tag{3.2}$$

Now let $\psi(t,\epsilon,x_0)$ denote the solution of the initial value problem (3.2), and let $\phi(t,\epsilon,x_0):=(\psi(t,\epsilon,x_0),s_\epsilon(\psi(t,\epsilon,x_0)))$. In what follows, we examine the Liapunov properties of $r_\epsilon(t):=k(\phi(t,\epsilon,x_0))$ for the system (3.2) in such a way that the analogy with the discretized system becomes transparent. By a straightforward calculation we have

$$\begin{aligned} \dot{r}_\epsilon(t) &= -\frac{\partial k}{\partial x}(\phi(t,\epsilon,x_0))^T[I+\epsilon Ds_\epsilon(\psi(t,\epsilon,x_0))^T Ds_\epsilon(\psi(t,\epsilon,x_0))]\frac{\partial k}{\partial x}(\phi(t,\epsilon,x_0)), \\ \ddot{r}_\epsilon(t) &= 2[\frac{\partial^2 k}{\partial x^2}(\phi(t,\epsilon,x_0))\frac{\partial k}{\partial x}(\phi(t,\epsilon,x_0)) \\ &\quad +\frac{\partial^2 k}{\partial x\partial y}(\phi(t,\epsilon,x_0))Ds_\epsilon(\psi(t,\epsilon,x_0))\frac{\partial k}{\partial x}(\phi(t,\epsilon,x_0))]^T \\ &\quad \cdot[I+\epsilon Ds_\epsilon(\psi(t,\epsilon,x_0))^T Ds_\epsilon(\psi(t,\epsilon,x_0))]\frac{\partial k}{\partial x}(\phi(t,\epsilon,x_0)) \\ &\quad +\frac{\partial k}{\partial x}(\phi(t,\epsilon,x_0))^T[2\epsilon[D^2 s_\epsilon(\psi(t,\epsilon,x_0))\frac{\partial k}{\partial x}(\phi(t,\epsilon,x_0))]^T \\ &\quad \cdot Ds_\epsilon(\psi(t,\epsilon,x_0))]\frac{\partial k}{\partial x}(\phi(t,\epsilon,x_0)). \end{aligned}$$

This yields the inequalities

$$\begin{aligned} \dot{r}_\epsilon(0) &\leq -\parallel \tfrac{\partial k}{\partial x}(x_0,s_\epsilon(x_0))\parallel_2^2, \\ \mid \ddot{r}_\epsilon(t)\mid &\leq (C_1+\epsilon C_2)\parallel \tfrac{\partial k}{\partial x}(\phi(t,\epsilon,x_0))\parallel_2^2 \end{aligned}$$

for suitable constants C_1, $C_2 > 0$. Using Taylors formula we can establish the estimate

$$\begin{aligned} k(\phi(h,\epsilon,x_0)) - k(x_0, s_\epsilon(x_0)) \;\; &\leq \;\; - \parallel \tfrac{\partial k}{\partial x}(x_0, s_\epsilon(x_0)) \parallel_2^2 h \\ &\quad + \int_0^h (h-\sigma)(C_1 + \epsilon C_2) \parallel \tfrac{\partial k}{\partial x}(\phi(\sigma,\epsilon,x_0)) \parallel_2^2 d\sigma. \end{aligned}$$

Since $\frac{\partial k}{\partial x}(\phi(t,\epsilon,x_0))$ solves the variational equation associated to (3.2), we obtain from the Gronwall lemma the inequality

$$\parallel \tfrac{\partial k}{\partial x}(\phi(t,\epsilon,x_0)) \parallel_2 \;\; \leq \;\; \parallel \tfrac{\partial k}{\partial x}(x_0, s_\epsilon(x_0)) \parallel_2 \; \exp(\gamma t)$$

where

$$\gamma := \;\; \sup\{ \parallel \tfrac{\partial^2 k}{\partial x^2}(x, s_\epsilon(x)) - \tfrac{\partial^2 k}{\partial x \partial y}(x, s_\epsilon(x)) D s_\epsilon(x) \parallel_2 | \;\; x \in \mathbb{R}^M \}.$$

Thus, we can estimate

$$k(\phi(h,\epsilon,x_0)) - k(x_0, s_\epsilon(x_0)) \;\; \leq \;\; h l_0(h,\epsilon) \parallel \tfrac{\partial k}{\partial x}(x_0, s_\epsilon(x_0)) \parallel_2^2 \tag{3.3}$$

with $l_0(h,\epsilon) := -1 + \frac{h}{2}(C_1 + \epsilon C_2) \exp(2\gamma h)$.
Inequality (3.3) states that $k(\cdot, s_\epsilon(\cdot))$ is a Liapunov function for equation (3.2) for $0 \leq \epsilon \leq \epsilon_0$ and $\epsilon_0 > 0$ sufficiently small. In fact, an inequality similar to (3.3) will be obtained with the one-step mapping replacing the h-flow. Since the right hand side of equation (3.3) can only vanish at equilibria of equation (3.2), the invariance principle of La Salle (see, e.g., La Salle (1976), Ch. 2, Theorem 6.4) assures that the solutions of equation (3.2) converge to a stationary point, provided the positive halforbit has compact closure. Here we have used the fact that (A4) implies that the equilibria of the reduced equation (3.2) are hyperbolic for ϵ sufficiently small. Finally, by the asymptotic phase result of Nipp and Stoffer (see Nipp, Stoffer (1995), Theorem 1) convergence also holds for initial values (x_0, y_0) satisfying $\parallel y_0 - s_0(x_0) \parallel < \delta_0$ and Lemma 2.1 is proved.

4 Existence and uniqueness of the Runge-Kutta solution

Our proof of Theorem 2.2 makes essential use of the results of Nipp and Stoffer (1995) in the case $\epsilon \leq h$. By the remark at the end of Section 4 in Lubich, Nipp and Stoffer (1995), this is possible under the stronger conditions we have imposed. A proof can be obtained by applying the techniques in Section 5 of Hairer, Lubich and Roche (1988). However, we notice that Lemmata 5 and 6 in that paper use the assumption $g(x_0, y_0) = O(h)$. This means that the domain of the Runge-Kutta iteration will be h dependent and will shrink to zero if h tends to zero. An existence and uniqueness proof in an h independent domain is the content of the following Lemma.

4.1 Lemma: *Let the assumptions (A1)-(A3) be fulfilled for equation (2.1) and let the Runge-Kutta method satisfy (B1)-(B3). Then there exists some $d_0 > 0$ (independent of h) such that for $0 < h < h_0$, $0 < \epsilon < \epsilon_0$ and initial values (x_0, y_0) with $\| y_0 - s_0(x_0) \| < d_0$ the Runge-Kutta iteration (2.3) is well defined.*

To prove Lemma 4.1 we employ the concept of vector norms. A functional $| \cdot |: V \to \mathbb{R}^k$ on a vector space V is called a generalized norm, if

$$| v | \geq 0, \ | v | = 0 \iff v = 0,$$
$$| v_1 + v_2 | \leq | v_1 | + | v_2 |$$

holds with the natural ordering "$\leq$" on $\mathbb{R}^k$. Every norm $\| \cdot \|_*$ in $\mathbb{R}^k$ defines a norm $\| \cdot \|$ in V via

$$\| v \| \ = \ \| \, | v | \, \|_* \, .$$

Furthermore, any two norms of this type are equivalent.

4.2 Lemma: *Let $(V, | \cdot |)$ be a Banach space with generalized norm $| \cdot |$ and let $B := \{ v \in V \mid | v - v_0 | \leq r \}$ for $r > 0$. Let the map $F : B \mapsto V$ be continuously differentiable with invertible $DF(v_0)$. Moreover, for some nonnegative matrices $P, K \in \mathbb{R}^{k,k}$ we assume*

$$\begin{aligned} | DF(v_0)^{-1} z | \ &\leq \ P \, | z |, \ z \in V, \\ | (DF(v_0) - DF(v)) z | \ &\leq \ K \, | z |, \ z \in V, \ v \in B, \\ P \, | F(v_0) | \ &< \ (I - PK) r. \end{aligned}$$

Then, the equation $F(v) = 0$ has a unique solution in B. In addition, the matrix $I - PK$ is nonsingular and we have the stability inequality

$$| v - w | \ \leq \ (I - PK)^{-1} P \, | F(v) - F(w) | \quad \forall v, w \in B.$$

Lemma 4.2 is a version of Banachs fixed point theorem on a ball in terms of generalized norms. It can be proved for example by applying Theorem B1, Appendix B, Beyn (1994) to the fixed point equation $T(v) = v$ where

$$T(v) \ := \ v - DF(v_0)^{-1} F(v), \ T : B \to V.$$

Proof of Lemma 4.1: We prove Lemma 4.1 by applying Lemma 4.2 to the equation

$$F(U, V, \tau, h) \ = \ \begin{pmatrix} U - (1\!\!1 \otimes x) - h(A \otimes I) \bar{f}(U, V) \\ \tau (V - (1\!\!1 \otimes y)) - (A \otimes I) \bar{g}(U, V) \end{pmatrix} \ = \ 0, \ \tau = \frac{\epsilon}{h}. \tag{4.1}$$

We recall that (4.1) is equivalent to the algebraic system (2.4). To apply Lemma 4.2 we use the generalized norm $| (U, V) | = (\| U \|, \| V \|) \in \mathbb{R}^2$ and the central point $v_0 = (1\!\!1 \otimes x, 1\!\!1 \otimes y)$.

For the derivative of F we find

$$\begin{aligned}
\frac{\partial}{\partial(U,V)} F(1\!\!1 \otimes x, 1\!\!1 \otimes y, \tau, h) &= \begin{pmatrix} I - h(A \otimes \frac{\partial f}{\partial x}(x,y)) & -h(A \otimes \frac{\partial f}{\partial y}(x,y)) \\ -(A \otimes \frac{\partial g}{\partial x}(x,y)) & \tau I - (A \otimes \frac{\partial g}{\partial y}(x,y)) \end{pmatrix} \\
&= \begin{pmatrix} I + O(h) & O(h) \\ O(1) & \tau I - (A \otimes \frac{\partial g}{\partial y}(x,y)) \end{pmatrix}, \ \tau \geq 0.
\end{aligned}$$

Now condition (B2) and the spectral relation

$$\mathrm{Re}(\lambda) < 0 \ \ \forall \lambda \in \sigma(\frac{\partial g}{\partial y}(x,y)), \ \| \, y - s_0(x) \, \| =: d > 0 \text{ sufficiently small}$$

imply that $\tau I - (A \otimes \frac{\partial g}{\partial y}(x,y))$ is nonsingular for $\tau \geq 0$.
In addition, using the techniques of Lemma 5, Section 5 in Hairer, Lubich and Roche (1988) we can show that $L_\tau := (\tau I - (A \otimes \frac{\partial g}{\partial y}(x,y)))^{-1}$ satisfies an estimate $\| \, L_\tau \, \| \leq C$ for all $\tau \geq 0$. From

$$\frac{\partial}{\partial(U,V)} F(1\!\!1 \otimes x, 1\!\!1 \otimes y, \tau, h) = \begin{pmatrix} I & 0 \\ O(1) & L_\tau^{-1} \end{pmatrix} \cdot (I + O(h))$$

we can deduce that $\frac{\partial}{\partial(U,V)} F(1\!\!1 \otimes x, 1\!\!1 \otimes y, \tau, h)$ is invertible for $\tau \geq 0$, $0 < h \leq h_0$, $h_0 > 0$ sufficiently small. The inverse is of the form

$$\frac{\partial}{\partial(U,V)} F(1\!\!1 \otimes x, 1\!\!1 \otimes y, \tau, h)^{-1} = \begin{pmatrix} I + O(h) & O(h) \\ O(1) & L_\tau + O(h) \end{pmatrix}.$$

This leads to

$$\begin{aligned}
\left| \frac{\partial}{\partial(U,V)} F(1\!\!1 \otimes x, 1\!\!1 \otimes y, \tau, h)^{-1}(X,Y) \right| &\leq \begin{pmatrix} 1 + O(h) & O(h) \\ O(1) & O(1) \end{pmatrix} \cdot \begin{pmatrix} \| \, X \, \| \\ \| \, Y \, \| \end{pmatrix} \\
&=: P_h \cdot \, | \, (X,Y) \, | \, .
\end{aligned}$$

Moreover, we obtain

$$\begin{aligned}
&\left| \left(\frac{\partial}{\partial(U,V)} F(1\!\!1 \otimes x, 1\!\!1 \otimes y, \tau, h) - \frac{\partial}{\partial(U,V)} F(U,V,\tau,h)\right)(X,Y) \right| \\
&\leq [O(\| \, U - (1\!\!1 \otimes x) \, \|) + O(\| \, V - (1\!\!1 \otimes y) \, \|)] \cdot \begin{pmatrix} O(h) & O(h) \\ O(1) & O(1) \end{pmatrix} \cdot \begin{pmatrix} \| \, X \, \| \\ \| \, Y \, \| \end{pmatrix} \\
&=: K_h \cdot \, | \, (X,Y) \, | \ \text{ for } (X,Y) \in \mathbb{R}^{Ms+Ns}.
\end{aligned}$$

Thus, for $(U,V) \in B_r := \{(X,Y) \in \mathbb{R}^{Ms+Ns} \mid \, | \, (U,V) - (1\!\!1 \otimes x, 1\!\!1 \otimes y) \, | \leq r\}$, $r = (r_1, r_2) > 0$ we can show the estimate

$$K_h = O(r_1 + r_2) \cdot \begin{pmatrix} O(h) & O(h) \\ O(1) & O(1) \end{pmatrix}.$$

Next we calculate

$$| F(1\!\!1 \otimes x, 1\!\!1 \otimes y, \tau, h) | \;=\; \begin{pmatrix} O(h) \parallel f(x,y) \parallel \\ O(1) \parallel g(x,y) \parallel \end{pmatrix}.$$

Using Taylor expansion and $g(x, s_0(x)) = 0$, we find

$$\parallel g(x,y) \parallel \;\leq\; C \parallel y - s_0(x) \parallel = \; O(d)$$

with $d := \parallel y - s_0(x) \parallel$. Hence, $| F(1\!\!1 \otimes x, 1\!\!1 \otimes y, \tau, h) | = (O(h), O(d))^T$ and

$$P_h \mid F(1\!\!1 \otimes x, 1\!\!1 \otimes y, \tau, h) \mid \;=\; \begin{pmatrix} O(h) + O(hd) \\ O(h) + O(d) \end{pmatrix}$$

follow. Now, Lemma 4.2 is applicable provided $r > 0$ is chosen in such a way that

$$P_h \mid F(1\!\!1 \otimes x, 1\!\!1 \otimes y, \tau, h) \mid \;<\; (I - P_h K_h) r \tag{4.2}$$

holds. To assure (4.2), we fix $h_0, r_1, r_2 > 0$ sufficiently small, such that

$$(I - P_h K_h) r \;\geq\; \frac{1}{2} \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} \tag{4.3}$$

is satisfied for $0 < h < h_0$. By choosing additionally $d_0 > 0$ sufficiently small we obtain

$$P_h \mid F(1\!\!1 \otimes x, 1\!\!1 \otimes y, \tau, h) \mid \;=\; \begin{pmatrix} O(h) + O(hd) \\ O(h) + O(d) \end{pmatrix} \;<\; \frac{1}{2} \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} \tag{4.4}$$

for $0 < h < h_0$, $0 < d < d_0$ possibly after diminishing $h_0 > 0$ once more. Now (4.2) is an immediate consequence of (4.3), (4.4) and Lemma 4.2 guarantees the unique solvability of equation (4.1) for $r = (r_1, r_2) > 0$ sufficiently small in the set

$$B_r \;=\; \{(X,Y) \in \mathbb{R}^{Ms+Ns} \mid \, |(X,Y) - (1\!\!1 \otimes x, 1\!\!1 \otimes y)| \leq r\}.$$

Remark: The stability inequality of Lemma 4.2 yields exactly the statement of Lemma 6, Section 5 in Hairer, Lubich and Roche (1988) without the assumption $g(x_0, y_0) = O(h)$.

5 Proof of the main result

For the given compact set K and $0 < \epsilon \leq h < h_0$, let B_ϵ denote the set of initial values (x_0, y_0) that satisfy $\parallel y_0 - s_0(x_0) \parallel < d_0$ and that generate a Runge-Kutta sequence $(x_n, y_n)_{n \in \mathbb{N}}$ that stays in K for all $n \in \mathbb{N}$. By Theorem 2 of Nipp and Stoffer (1995), the Runge-Kutta mapping admits an invariant manifold

$$M_{\epsilon,h} \;=\; \{(x,y) \in \mathbb{R}^{M+N} \mid \; y = s_{\epsilon,h}(x)\}$$

that attracts all initial values in a neighborhood $\| y_0 - s_0(x_0) \| < d_0$. Moreover, if $d_0 > 0$ is chosen sufficiently small then the property of asymptotic phase holds. That is, for every (x_0, y_0) there exists some $(\tilde{x}_0, \tilde{y}_0) \in M_{\epsilon,h}$ such that the corresponding Runge-Kutta evolutions (x_n, y_n) and $(\tilde{x}_n, \tilde{y}_n)$ satisfy

$$\lim_{n\to\infty}((x_n, y_n) - (\tilde{x}_n, \tilde{y}_n)) = 0.$$

Thus, it is sufficient to show our convergence result for initial values $(x_0, y_0) \in M_{\epsilon,h}$. Restricted to that manifold, the Runge-Kutta iteration has the form

$$x_{n+1} = x_n + h(b^T \otimes I)\bar{f}(X^n, Y^n) =: G_{h,\epsilon}(x_n),$$

where $U = X^n, V = Y^n$ solve the system (see (2.4))

$$\begin{aligned} U - (1\!\!1 \otimes x_n) &= h(A \otimes I)\bar{f}(U, V), \\ \epsilon(V - (1\!\!1 \otimes s_{\epsilon,h}(x_n))) &= h(A \otimes I)\bar{g}(U, V). \end{aligned}$$

Now in complete analogy to our Liapunov estimate for the continuous system in Section 3 (see formula (3.3)), we can show with $L_\epsilon(x) = k(x, s_\epsilon(x))$ and the solution flow ψ of (3.2) the following inequality

$$L_\epsilon(G_{h,\epsilon}(x)) - L_\epsilon(x) \le C_0 \| G_{h,\epsilon}(x) - \psi(h, \epsilon, x) \| + hl_0(h, \epsilon) \| \tfrac{\partial k}{\partial x}(x, s_\epsilon(x)) \|_2^2 .$$

Let $\tilde{G}_{h,\epsilon}$ denote the Runge-Kutta map with step-size h when applied to (3.2). Then we obtain

$$\begin{aligned} \| G_{h,\epsilon}(x) - \psi(h, \epsilon, x) \| &\le \| G_{h,\epsilon}(x) - \tilde{G}_{h,\epsilon}(x) \| + \| \tilde{G}_{h,\epsilon}(x) - \psi(h, \epsilon, x) \| \\ &\le C_1 \epsilon h^{q+1} + C_2 h^{p+1} \le C_3 h^{q+1}. \end{aligned}$$

The first part of the second inequality is valid due to Lemma 3 of Lubich, Nipp and Stoffer (1995). Lemma 3 of that paper applies under our assumptions to an arbitrary x_0, since (A1) holds as well as

$$\epsilon\, Ds_\epsilon(x) f(x, s_\epsilon(x)) = g(x, s_\epsilon(x)) \text{ for } x \in \mathbb{R}^M.$$

Moreover, part two of the second inequality follows from (B4) by the standard local error estimate for Runge-Kutta methods. Combining the last two estimates we find for the discrete dynamics

$$L_\epsilon(x_{n+1}) - L_\epsilon(x_n) \le h[l_0(h, \epsilon) \| \tfrac{\partial k}{\partial x}(x_n, s_\epsilon(x_n)) \|_2^2 + Ch^q],\ n \in \mathbb{N}. \tag{5.1}$$

Let S denote the set of equilibria of equation (3.2) in K_1. By assumption (A4) and the closeness of s_0 and s_ϵ all equilibria in S are hyperbolic. Thus we can deduce that either $S = \emptyset$ or $S = \{\bar{x}_1, \ldots, \bar{x}_l\}$ for some $l \ge 1$.
Suppose first that $S = \emptyset$ holds. Setting

$$\gamma_\epsilon := \min\{\| \tfrac{\partial k}{\partial x}(x, s_\epsilon(x)) \|_2 |\ x \in K_1\} > 0$$

and taking $h > 0$ sufficiently small we then find the inequality

$$L_\epsilon(x_{n+1}) - L_\epsilon(x_n) \quad \leq \quad h\frac{l_0(0,\epsilon)}{2}\gamma_\epsilon^2 \text{ for } n \in \mathbb{N},\ (x_0, s_{\epsilon,h}(x_0)) \in B_\epsilon.$$

This implies $\lim_{n\to\infty} L_\epsilon(x_n) = -\infty$ in contrast to the boundedness of L_ϵ on K_1.
In the case of a nonempty finite set S we analyze the asymptotic behavior of the iterates starting in $M_{\epsilon,h}$ by using the discrete ω- limit set

$$\omega_{h,\epsilon}(x_0) \quad = \quad \{\hat{x} \in \mathbb{R}^M \mid \lim_{k\to\infty} x_{n_k} = \hat{x} \text{ for some subsequence } n_k \to \infty\}.$$

The general properties of discrete limit sets are described in Theorem 5.2, Ch.1 in La Salle (1976). In our situation, we have $\omega_{h,\epsilon}(x_0) \neq \emptyset$. Since the invariant manifold Theorem of Nipp and Stoffer (1995) reduces the singular perturbed problem to a regular parameter dependent problem on $M_{\epsilon,h}$, one can extend in a straightforward manner the techniques from the Appendix of Schropp (1998) to show

$$\begin{aligned} \bar{r}(h) &:= \{\| \tfrac{\partial k}{\partial x}(\hat{x}_{h,\epsilon}, s_\epsilon(\hat{x}_{h,\epsilon})) \|_2^2 \mid \ \hat{x}_{h,\epsilon} \in \omega_{h,\epsilon}(x_0),\ (x_0, s_{\epsilon,h}(x_0)) \in B_\epsilon,\ 0 < \epsilon \leq h\} \\ &= o(1). \end{aligned} \tag{5.2}$$

Formula (5.1) states that the longtime behaviour of the discrete trajectories takes place near the zeroes of $\frac{\partial k}{\partial x}(x, s_\epsilon(x))$. Now let $\bar{x}$ be a stationary point of (3.2). Then $(\bar{x}, s_\epsilon(\bar{x}))$ is a stationary point of (3.1). An easy calculation shows that Runge-Kutta methods retain equilibria as fixed points. Then the attractivity of $M_{\epsilon,h}$ (see Theorem 2, ii) in Nipp, Stoffer (1995)) yields $s_\epsilon(\bar{x}) = s_{\epsilon,h}(\bar{x})$ for $h > 0$ sufficiently small. Hence, $\bar{x}$ is a fixed point of $x_{n+1} = G_{h,\epsilon}(x_n)$ and we need more information about the behavior of this iteration in the neighbourhood of fixed points.
Let U denote a neighbourhood of a fixed point $\bar{x}$. The fixed point $\bar{x}$ is called hyperbolic, if the linearization at $\bar{x}$ has no eigenvalues of modulus one. We consider the set of initial values that have bounded trajectories in U

$$B^U_{s,h,\epsilon}(\bar{x}) \quad = \quad \{x_0 \in \mathbb{R}^M \mid \ x_n \in U \text{ for } n \in \mathbb{N}\}$$

as well as the stable set of $\bar{x}$ in $B^U_{s,h,\epsilon}(\bar{x})$ given by

$$W^U_{s,h,\epsilon}(\bar{x}) \quad = \quad \{x_0 \in B^U_{s,h,\epsilon}(\bar{x}) \mid \ x_n \to \bar{x} \text{ as } n \to \infty\}.$$

It is well known that for any fixed values of the parameters these two sets coincide and form the local stable manifold of the fixed point. However, we need that the neighborhood can be chosen independently of ϵ and h. Such a result will be proved in Section 6 by an extension of Theorem 4.1 in Schropp (1997) to the parameter dependent case. More precisely, we will show that $\bar{x}_i$, $i = 1, \ldots, l$ are hyperbolic fixed points of $G_{h,\epsilon}$ and that for some $\delta > 0$ independent of h, ϵ and i we have

$$W^{K_\delta(\bar{x}_i)}_{s,h,\epsilon}(\bar{x}_i) \quad = \quad B^{K_\delta(\bar{x}_i)}_{s,h,\epsilon}(\bar{x}_i),\ 0 < \epsilon \leq h < h_0,\ \bar{x}_i \in S,\ i = 1, \ldots, l \tag{5.3}$$

where $K_\delta(\bar{x}_i) = \{w \in \mathbb{R}^M \mid \| w - \bar{x}_i \| < \delta\}$, $i = 1, \ldots, l$.
Now we define

$$\rho \;:=\; \begin{cases} \min\{\| \bar{x}_i - \bar{x}_j \| \mid \bar{x}_i, \bar{x}_j \in S,\ i \neq j,\ i, j \in \{1, \ldots, l\}\} \text{ if } l \geq 2, \\ \infty, \text{ else} \end{cases}$$

and consider the neighborhood $V_b := (\cup_{i=1}^{l} K_b(\bar{x}_i))$ with $b := \frac{1}{2} \min\{\delta, \frac{\rho}{2}\}$. In addition we define the quantity

$$\gamma_b \;:=\; \min\{\| \tfrac{\partial k}{\partial x}(x, s_\epsilon(x)) \|_2^2 \mid x \in K_1 \backslash V_b,\ 0 \leq \epsilon \leq h_0\} \;>\; 0.$$

Next we diminish $h_0 > 0$ such that $\bar{r}(h) < \gamma_b$ holds for $0 < \epsilon \leq h < h_0$. Since

$$\| \tfrac{\partial k}{\partial x}(\hat{x}_{h,\epsilon}, s_\epsilon(\hat{x}_{h,\epsilon})) \|_2^2 \leq \bar{r}(h) \;<\; \gamma_b$$

is valid for $\hat{x}_{h,\epsilon} \in \omega_{h,\epsilon}(x_0)$, $0 < \epsilon \leq h < h_0$ we obtain the inclusion $\omega_{h,\epsilon}(x_0) \subset V_b \cap K_1$. Now we use that the ω-limit set $\omega_{h,\epsilon}(x_0)$ is invariant and invariantly connected (see La Salle (1976), Ch. 1, Section 5 for definition), which leads to the conclusion

$$\omega_{h,\epsilon}(x_0) \subset K_b(\bar{x}_\eta) \cap K_1 \text{ for some } \eta \in \{1, \ldots, l\},\ 0 < \epsilon \leq h < h_0.$$

Since $\operatorname{dist}(x_n, \omega_{h,\epsilon}(x_0)) \to 0$ as $n \to \infty$, $0 < \epsilon \leq h < h_0$ there exists a number $N_1 = N_1(h, \epsilon, b, x_0) \in \mathbb{N}$ such that

$$x_n \;\subset\; K_{2b}(\bar{x}_\eta) \;\subset\; K_\delta(\bar{x}_\eta) \text{ for } n \geq N_1,\ (x_0, s_{\epsilon,h}(x_0)) \in B_\epsilon,\ 0 < \epsilon \leq h < h_0.$$

This implies $x_{N_1} \in B_{s,h,\epsilon}^{K_\delta(\bar{x}_\eta)}(\bar{x}_\eta) = W_{s,h,\epsilon}^{K_\delta(\bar{x}_\eta)}(\bar{x}_\eta)$, and our convergence result is proved.

6 Stable and bounded sets of hyperbolic fixed points

It is well known that the local stable manifold of a hyperbolic fixed point may be equivalently defined as the set of initial data that lead to bounded sequences in the neighborhood. In this Section we show that such a statement holds for a fixed neighborhood uniformly in h and ϵ, see equation (5.3).

6.1 Lemma: *Let $(\bar{x}, \bar{y} = s_0(\bar{x}))$ be a hyperbolic stationary point of the equation (2.1) and let the assumptions (A1)-(A3) hold. Consider a Runge-Kutta method (2.3) satisfying (B1)-(B4) with invariant manifolds $M_{\epsilon,h}, 0 < \epsilon \leq h$ close to M_0. Define the map $G_{h,\epsilon}$ in a neighborhood of $\bar{x}$ such that the x-component of the Runge-Kutta iteration on $M_{\epsilon,h}$ reads $x_{n+1} = G_{h,\epsilon}(x_n)$. Then $\bar{x}$ is a hyperbolic fixed point of the map $G_{\epsilon,h}$ and there exists a $\delta > 0$ such that*

$$W_{s,h,\epsilon}^{K_\delta(\bar{x})}(\bar{x}) \;=\; B_{s,h,\epsilon}^{K_\delta(\bar{x})}(\bar{x})$$

holds for all $0 < \epsilon \leq h < h_0$.

Proof: Following Lubich, Nipp and Stoffer (1995), Section 2 the Runge-Kutta method on $M_{\epsilon,h}$ can be written as

$$(x_{n+1}, y_{n+1}) = (G_{\epsilon,h}(x_n), s_{\epsilon,h}(x_{n+1}))$$

where

$$G_{h,\epsilon}(x) := x + h(b^T \otimes I)\bar{f}(X,Y), \tag{6.1}$$

and $(X,Y) \in \mathbb{R}^{Ms} \times \mathbb{R}^{Ns}$ is the solution of

$$\begin{aligned} X - (\mathbb{1} \otimes I)x &= h(A \otimes I)\bar{f}(X,Y), \\ \epsilon(Y - (\mathbb{1} \otimes I)s_{\epsilon,h}(x)) &= h(A \otimes I)\bar{g}(X,Y). \end{aligned}$$

In what follows, we derive the representation

$$DG_{h,\epsilon}(\bar{x}) \;=\; I + h[\Phi(\bar{x}) + O(\epsilon) + O(h)]. \tag{6.2}$$

where

$$\bar{y} = s_0(\bar{x}), \quad \Phi(\bar{x}) = (\frac{\partial f}{\partial x} - \frac{\partial f}{\partial y}\frac{\partial g}{\partial y}^{-1}\frac{\partial g}{\partial x})(\bar{x},\bar{y}).$$

Using this we indicate how the proof of the stable manifold theorem can be repeated in such a way that our assertion follows.

With $(A^{-1} \otimes I)(X - (\mathbb{1} \otimes I)x) = h\bar{f}(X,Y)$ we can rewrite $G_{h,\epsilon}$ in the form

$$G_{h,\epsilon}(x) \;=\; R(\infty)x + (b^T A^{-1} \otimes I)X(h,\epsilon,x,s_{\epsilon,h}(x)).$$

In particular, the linearization of $G_{h,\epsilon}$ has the structure

$$\begin{aligned} DG_{h,\epsilon}(x) \;=\; & R(\infty)I + (b^T A^{-1} \otimes I)(\frac{\partial}{\partial x}X(h,\epsilon,x,s_{\epsilon,h}(x)) \\ & + \frac{\partial}{\partial y}X(h,\epsilon,x,s_{\epsilon,h}(x))\, Ds_{\epsilon,h}(x)). \end{aligned} \tag{6.3}$$

The dependence on the arguments $(h,\epsilon,x,s_{\epsilon,h}(x))$ will be suppressed in the following. By implicit differentiation we find from (4.1) the derivatives

$$\frac{\partial X}{\partial x} = C_{h,\tau}^{-1}(\mathbb{1} \otimes I) \qquad \frac{\partial X}{\partial y} = C_{h,\tau}^{-1}[h\tau(A \otimes \frac{\partial f}{\partial y})L_\tau(\mathbb{1} \otimes I)], \tag{6.4}$$

where

$$C_{h,\tau} \;:=\; I - h[(A \otimes \frac{\partial f}{\partial x}) + (A \otimes \frac{\partial f}{\partial y})L_\tau(A \otimes \frac{\partial g}{\partial x})]$$

and $L_\tau = (\tau I - (A \otimes \frac{\partial g}{\partial y}))^{-1}$ as in Section 4. Recall that our assumptions imply $\| L_\tau \| \leq C$ for $\tau \geq 0$ so that $C_{h,\tau}$ is nonsingular for small h and

$$C_{h,\tau}^{-1} \;=\; I + h[(A \otimes \frac{\partial f}{\partial x}) + (A \otimes \frac{\partial f}{\partial y})L_\tau(A \otimes \frac{\partial g}{\partial x})] + O(h^2). \tag{6.5}$$

Inserting (6.4) into (6.3) leads to

$$DG_{h,\epsilon}(\bar{x}) = R(\infty)I + (b^T A^{-1} \otimes I)C_{h,\tau}^{-1}[(\mathbb{1} \otimes I) + h\tau(A \otimes \frac{\partial f}{\partial y})L_\tau(\mathbb{1} \otimes I)\, Ds_{\epsilon,h}].$$

In this expression we want to replace $Ds_{\epsilon,h}$ by the derivative $Ds_0(x)$ through

$$\begin{aligned} Ds_{\epsilon,h}(x) &= Ds_0(x) + O(h) + O(\epsilon) \\ &= -\left(\frac{\partial g}{\partial y}^{-1}\frac{\partial g}{\partial x}\right)(x, s_0(x)) + O(h) + O(\epsilon). \end{aligned} \tag{6.6}$$

For this relation we need some smoothness of $s_{\epsilon,h}(x)$ with respect to (x,h,ϵ). In order to establish such a result one has to modify the proof of Nipp and Stoffer (1995) (see the arguments following equation (9) there). Instead of their operator $\tilde{P}$ we consider $\hat{P}$ defined by

$$\hat{P}(x,h,\epsilon,z) \;:=\; \begin{pmatrix} x + \hat{F}(x,z,h,\epsilon) \\ h \\ \epsilon \\ G(x,z,h,\epsilon) \end{pmatrix}$$

and apply their invariant manifold result (see Theorem 5, Nipp, Stoffer (1992)) to $\hat{P}$. This yields an invariant and smooth manifold

$$\hat{M} \;:=\; \{(x,h,\epsilon,z) \mid \; (x,h,) \in \mathbb{R}^M \times]0,h_0[, 0 \le \epsilon \le \delta_0 h, \; z = \hat{\sigma}(x,h,\epsilon)\}$$

for the map $\hat{P}$. Using the structure of $\hat{P}$, we can deduce

$$\hat{M} \;=\; \cup_{0<h<h_0} \cup_{0<\epsilon<h\delta_0} M_{h,\epsilon}.$$

In particular, $s_0(x) = \hat{\sigma}(0,0,x)$ and $s_{\epsilon,h}(x) = \hat{\sigma}(h,\epsilon,x)$ is smooth in (h,ϵ,x). The estimate (6.6) then follows from a Taylor expansion.

Using (6.6) we find

$$\begin{aligned} DG_{h,\epsilon}(\bar{x}) &= R(\infty)I + (b^T A^{-1} \otimes I)C_{h,\tau}^{-1} \\ &\quad \cdot[(\mathbb{1} \otimes I) - h\tau(A \otimes \frac{\partial f}{\partial y})L_\tau(\mathbb{1} \otimes I)\frac{\partial g}{\partial y}^{-1}\frac{\partial g}{\partial x}] + O(h\epsilon) + O(\epsilon^2) \end{aligned}$$

and combining this with (6.4) some algebraic manipulations lead to

$$\begin{aligned} DG_{h,\epsilon}(\bar{x}) &= I + h(b^T A^{-1} \otimes I)[(A \otimes \frac{\partial f}{\partial x})(\mathbb{1} \otimes I) \\ &\quad +(A \otimes \frac{\partial f}{\partial y})L_\tau[(A \otimes \frac{\partial g}{\partial x})(\mathbb{1} \otimes I) - \tau(\mathbb{1} \otimes I)\frac{\partial g}{\partial y}^{-1}\frac{\partial g}{\partial x}]] \\ &\quad +O(\epsilon h) + O(\epsilon^2) + O(h^2). \end{aligned}$$

Finally, with the help of $b^T 1\!\!1 = 1$ and the equations

$$\begin{aligned}(1\!\!1 \otimes I)\frac{\partial g}{\partial y}^{-1}\frac{\partial g}{\partial x} &= (A^{-1} \otimes \frac{\partial g}{\partial y}^{-1})(A \otimes \frac{\partial g}{\partial x})(1\!\!1 \otimes I),\\ L_\tau(I - \tau(A^{-1} \otimes \frac{\partial g}{\partial y}^{-1})) &= -(A^{-1} \otimes \frac{\partial g}{\partial y}^{-1}),\end{aligned}$$

we obtain the representation (6.2).

By (A4) the matrix $\Phi(\bar{x})$ is hyperbolic. Let M_s and $M_u = M - M_s$ denote the number of its stable and unstable dimensions respectively. Then $\bar{x}$ is a hyberbolic fixed point of $DG_{h,\epsilon}(\bar{x})$ with the same number of stable and unstable eigenvalues. The corresponding invariant subspaces $Z_s^{h,\epsilon}$, $Z_u^{h,\epsilon}$ can be obtained by a simple perturbation argument (compare Lemma 3.4 in Beyn (1987)) as follows. For $0 \leq \epsilon \leq h \leq h_0$ and h_0 sufficiently small there exist $\mu > 0$, matrices $T(h,\epsilon)$ and norms $\|\cdot\|_s$, $\|\cdot\|_u$ on $\mathbb{R}^{M_s}$, $\mathbb{R}^{M_u}$ such that

$$T(h,\epsilon)^{-1} DG_{h,\epsilon}(\bar{x}) T(h,\epsilon) = \begin{pmatrix} Q_u(h,\epsilon) & 0 \\ 0 & Q_s(h,\epsilon) \end{pmatrix}$$

and

$$\begin{aligned}\max\{\| Q_u(h,\epsilon)^{-1} \|_u, \| Q_s(h,\epsilon) \|_s\} &\leq 1 - \mu h,\\ \| T(\epsilon,h) \| + \| T(\epsilon,h)^{-1} \| &\leq C.\end{aligned} \tag{6.7}$$

The first M_u columns of $T(h,\epsilon)$ form a basis for $Z_u^{h,\epsilon}$ and the last M_s columns for $Z_s^{\epsilon,h}$. In the following we will use the norms $\|\cdot\|_{\epsilon,h}$ defined by

$$\| T(\epsilon,h)(x_u,x_s) \|_{\epsilon,h} = \max(\| x_u \|_u, \| x_s \|_s) \quad \text{for } x_u \in \mathbb{R}^{M_u}, x_s \in \mathbb{R}^{M_s}.$$

By (6.7) these norms are uniformly equivalent to the given norm and henceforth we will drop the indices ϵ, h. By construction we have the relations

$$\begin{aligned}\| x \| = \ &\max\{\| [x]_s \|, \| [x]_u \|\}, \ x = [x]_s + [x]_u,\\ \max\{\| DG_{h,\epsilon}(\bar{x})_s \|, \| (DG_{h,\epsilon}(\bar{x})_u)^{-1} \|\} &\leq 1 - \mu h.\end{aligned}$$

Here $[x]_u$, $[x]_s$ denote the projections of x onto $Z_u^{h,\epsilon}$, $Z_s^{h,\epsilon}$ and $DG_{h,\epsilon}(\bar{x})_s$, $DG_{h,\epsilon}(\bar{x})_u$ stand for the restrictions of $DG_{h,\epsilon}(\bar{x})$ to $Z_s^{h,\epsilon}$, $Z_u^{h,\epsilon}$.

For the remainder we follow the proof of Theorem 4.1 in Schropp (1997).

To simplify the notation we assume without loss of generality $\bar{x} = 0$. Let S_0, S_b denote the Banach spaces of zero convergent, respectively, bounded $\mathbb{R}^M$-valued sequences $(x_n)_{n \in \mathbb{N}}$ with the norm

$$\| (x_n)_{n\in\mathbb{N}} \|_\infty := \sup\{\| x_n \| \mid n \in \mathbb{N}\}.$$

Consider the scaled and cut-off vector field

$$\hat{G}_{h,\epsilon}(x) := DG_{h,\epsilon}(0)x + R_{h,\epsilon}(x), \ R_{h,\epsilon}(x) := \chi(x)\frac{1}{\delta}(G_{h,\epsilon}(\delta x) - DG_{h,\epsilon}(0)\delta x)$$

for $0 < \epsilon \leq h \leq h_0$ and scaling factor $\delta > 0$. Here $\chi \in C_b^\infty(\mathbb{R}^M, [0,1])$ is a cut-off function satisfying

$$\chi(x) = 1, \text{ if } \| x \| \leq 1, \qquad \chi(x) = 0, \text{ if } \| x \| \geq 2.$$

For the transformed map $\hat{G}_{h,\epsilon}$, we define the operators

$$\begin{aligned} \Gamma_i^{h,\epsilon} : S_i &\rightarrow Z_s^{h,\epsilon} \times S_i =: Y_i \\ (x_n)_{n\in\mathbb{N}} &\rightarrow ([x_0]_s, (x_{n+1} - \hat{G}_{h,\epsilon}(x_n))_{n\in\mathbb{N}}) \end{aligned}$$

for $i \in \{0, b\}$. The aim is to apply the Lipschitz inverse mapping Theorem (see the Appendix of Irwin (1990) or Lemma 4.2) to the equations

$$\Gamma_i^{h,\epsilon}((x_n)_{n\in\mathbb{N}}) = (v, (0)_{n\in\mathbb{N}}), \; v \in Z_s^{h,\epsilon}$$

simultaneously and with the same data for $i = 0$ and $i = b$. Since $S_b \subset S_0$ our conclusion then follows from the uniqueness of the solutions. The main ingredient for the application of the Lipschitz inverse mapping theorem is the estimate

$$\| DR_{h,\epsilon}(x) \| \leq \frac{\mu h}{2} \text{ for } 0 < \epsilon \leq h < h_0, \; x \in \mathbb{R}^M. \tag{6.8}$$

Using $\chi(x) = 0$ für $\| x \| \geq 2$, we obtain

$$\begin{aligned} \| DR_{h,\epsilon}(x) \| \leq\; & \sup\{\| D\chi(x) \| \mid \| x \| \leq 2\} \\ & * \sup\{\| \frac{1}{\delta}(G_{h,\epsilon}(\delta x) - G_{h,\epsilon}(0) - DG_{h,\epsilon}(0)\delta x) \| \mid \| x \| \leq 2\} \\ & + \sup\{\| DG_{h,\epsilon}(\delta x) - DG_{h,\epsilon}(0) \| \mid \| x \| \leq 2\} \quad \forall x \in \mathbb{R}^M. \end{aligned}$$

A simple calculation yields

$$\frac{1}{\delta}(G_{h,\epsilon}(\delta x) - G_{h,\epsilon}(0) - DG_{h,\epsilon}(0)\delta x) = \frac{h}{\delta}(\Delta_{h,\epsilon}(\delta x) - \Delta_{h,\epsilon}(0) - D\Delta_{h,\epsilon}(0)\delta x)$$

where $\Delta_{h,\epsilon}(x) := (b^T \otimes I)\bar{f}(X(h, \epsilon, x, s_{\epsilon,h}(x)), Y(h, \epsilon, x, s_{\epsilon,h}(x)))$. Furthermore, we find for the difference of the derivatives

$$DG_{h,\epsilon}(\delta x) - DG_{h,\epsilon}(0) = h(D\Delta_{h,\epsilon}(\delta x) - D\Delta_{h,\epsilon}(0)).$$

Combining this with the smoothness properties of $\Delta_{h,\epsilon}$ the estimate (6.8) follows for h and δ sufficiently small.

The further steps are (cf. Lemma 4.2 with $k = 1$ and $r \geq 1$) a bound on the inverse

$$\| D\Gamma_i^{h,\epsilon}((0)_{n\in\mathbb{N}})^{-1} \|_\infty \leq \frac{1}{\mu h}$$

and a small residual

$$\| \Gamma_i^{\epsilon,h}((0)_{n\in\mathbb{N}}) \|_\infty < \frac{\mu h}{2}.$$

These can be proved along the lines of Theorem 4.1 in Schropp (1997) and our Lemma is shown.

References

Beyn, W.-J. (1987): On the Numerical Approximation of Phase Portraits near Stationary Points, SIAM J. Numer. Anal. **24**, p.1095-1113.

Beyn, W.-J. (1994): Numerical analysis of homoclinic orbits emanating from a Takens-Bogdanov point, IMA J. of Numer. Anal. **14**, p.381-410.

Fenichel, N. (1979): Geometric singular perturbation theory for ordinary differential equations, J. Differential Equations **31**, 53-98.

Hairer, E., Lubich, Ch., Roche, M. (1988): Error of Runge-Kutta methods for stiff problems studied via differential algebraic equations, BIT **28**, p.678-700.

Hale, J. K., Raugel, G. (1989): Lower semicontinuity of attractors of gradient systems and applications. Annali di Mat. pura ed appl. Vol. CLIV, p.281-326.

Hirsch, M.W., Smale, S. (1974): Differential equations, dynamical systems and linear algebra, Academic Press, London.

Humphries, A.R., Stuart, A.M. (1994): Runge-Kutta methods for dissipative and gradient systems, SIAM J. Numer. Anal. **31**, p.1452-1485.

Irwin, M.C. (1990): Smooth dynamical systems, Academic Press, London and New York.

La Salle, J.P. (1976): The stability of dynamical systems, Regional Conference Series in Applied Mathematics **25**, SIAM.

Lubich, Ch., Nipp, K., Stoffer, D. (1995): Runge-Kutta solutions of stiff differential equations near stationary points, SIAM J. of Numer. Anal. **32**, p.1296-1307.

Nipp, K. (1985): Invariant manifolds of singularly perturbed ordinary differential equations, J. of Appl. Math. and Phys. **36**, p.309-320.

Nipp, K., Stoffer, D. (1992): Attractive invariant manifolds for maps, SAM Research report No. 92-11, ETH Zürich.

Nipp, K., Stoffer, D. (1995): Invariant manifolds and global error estimates of numerical integration schemes applied to stiff systems of singular perturbation type - Part I: RK-methods, Numer. Math. **70**, p.245-257.

Schropp, J. (1995): Using dynamical systems methods to solve minimization problems, Applied Numerical Mathematics **18**, p.321-335.

Schropp, J. (1997): A note on minimzation problems and multistep methods, Numer. Math. **78**, p.87-101.

Schropp, J. (1998): One- and Multistep procedures for constrained minimization problems, to appear in IMA Journal of Numerical Analysis.

Stuart, A. M., Humphries, A. R. (1996): Dynamical Systems and Numerical Analysis. Cambridge University Press.