Invariant manifolds and global error estimates of numerical integration schemes applied to stiff systems of singular perturbation type – Part I: RK-methods

[*Running head:* On numerical integration of stiff systems]

K. Nipp[1], D. Stoffer[1]

Abstract

For implicit RK-methods applied to singularly perturbed systems of ODEs it is shown that the resulting discrete systems preserve the geometric properties of the underlying ODE. This invariant manifold result is used to derive sharp bounds on the global error of RK-solutions.

Keywords: singular perturbation, attractive invariant manifold,
stiff systems, global error, implicit RK-method

Subject Classification: 65L, 34C

[1]Departement Mathematik, ETH Zürich, CH-8092 Zürich, Switzerland

1. Introduction

Numerical integration of ODEs in applications often leads to so-called stiff problems. Stiff problems are considerably more difficult to solve than nonstiff ones. Although efficient methods for stiff problems have been developed since the late sixties very little theoretical work has been done for nonlinear stiff problems. In 1988 Hairer, Lubich, Roche [1] derived bounds for the global error of RK-methods applied to stiff systems of singular perturbation type. Singularly perturbed systems of ODEs are a model class for stiff systems. The smaller the perturbation parameter ϵ the stiffer the system. Under usual assumptions a singularly perturbed system has an important geometric property: It admits a highly attractive smooth invariant manifold M_ϵ. In Section 2 we state the invariant manifold result for singularly perturbed ODEs proved in Nipp [3] and [4].

A numerical method applied to a system of ODEs defines a map in phase space. In Section 3 we deal with the question whether such a map preserves the geometric property of the underlying singularly perturbed ODE. In general, this geometric property carries over to the discrete system if the step size h is of the same order as the perturbation parameter ϵ. For such small step sizes h the numerical integration is very inefficient, however. For explicit methods the step size h has to be of order ϵ due to their poor stability properties. This has been shown in Kirchgraber, Nipp [2] for the explicit Euler method. We show that for implicit RK-methods the step size h may be chosen such that h is independent of ϵ. For h fixed and for all ϵ small enough such stiff discrete systems admit an attractive invariant manifold $M_{h,\epsilon}$ close to the invariant manifold M_ϵ of the ODE. This geometric result for stiff RK-methods is derived in Section 3.

In Section 4 we use the geometric concept and the results od Sections 2 and 3 to derive global error bounds of RK-methods applied to singularly perturbed ODEs. The general idea is to reduce a complicated system to a simpler one by restricting the complicated system to an attractive invariant manifold. In this paper this means reducing a stiff problem to a nonstiff one on M_ϵ. A solution of a singularly perturbed ODE quickly approaches a solution on the manifold M_ϵ. The discrete RK-solution quickly approaches a discrete solution on $M_{h,\epsilon}$. The global error of the RK-solution is estimated by combining the classical (nonstiff) estimate for a "RK-solution on M_ϵ", the properties of attractivity and asymptotic phase of M_ϵ and $M_{h,\epsilon}$, respectively, and the closeness of the two manifolds. Estimates for the global error of RK-solutions with special starting values on M_ϵ were first obtained by Hairer, Lubich, Roche [1] using a completely different approach not involving invariant manifold theory.

In Part II we shall investigate multistep methods applied to stiff systems of singular perturbation type.

2. An invariant manifold result for singularly perturbed ODEs

We consider the singularly perturbed autonomous system

$$
\begin{aligned}
\frac{dx}{dt} &= f(x,y) \\
\epsilon\,\frac{dy}{dt} &= g(x,y)
\end{aligned}
\tag{1}
$$

where $x \in \mathbb{R}^m, y \in \mathbb{R}^n$ and $\epsilon \in (0, \epsilon_0)$. By C_b^r we denote spaces of functions of class C^r with bounded derivatives.

We make the following

Hypothesis H_{DE}

1) f and g are bounded and there is r with $3 \le r < \infty$ such that $f \in C_b^r(\mathbb{R}^m \times \mathbb{R}^n, \mathbb{R}^m)$, $g \in C_b^r(\mathbb{R}^m \times \mathbb{R}^n, \mathbb{R}^n)$.

2) There is a function $s_0 \in C_b^r(\mathbb{R}^m, \mathbb{R}^n)$ such that $g(x, s_0(x)) = 0$ for $x \in \mathbb{R}^m$.

3) There is a positive constant b_0 such that all eigenvalues of the Jacobian $B_0(x) := g_y(x, s_0(x))$ have real parts smaller than $-b_0$ for all $x \in \mathbb{R}^m$.

Under the above assumptions it can be shown that for all $\epsilon > 0$ small enough Eq.(1) admits an attractive invariant manifold M_ϵ which is $O(\epsilon)$-close to the so-called *reduced manifold* $M_0 := \{(x,y) \mid x \in \mathbb{R}^m,\ y = s_0(x)\}$. This result was proved in Nipp [3] and the slightly improved version given below is proved in Nipp [4].

Theorem 1 *For every $\beta \in (0, b_0)$ there are positive constants ϵ^*, δ, K and a function $s \in C_b^r(\mathbb{R}^m \times (-\epsilon^*, \epsilon^*), \mathbb{R}^n)$ such that the following assertions hold for $\epsilon \in (0, \epsilon^*)$.*

i) Invariance. *The set $M_\epsilon = \{(x,y) \mid x \in \mathbb{R}^m,\ y = s(x,\epsilon)\} \subset \mathbb{R}^m \times \mathbb{R}^n$ is invariant under Eq.(1), i.e., $P_t(M_\epsilon) = M_\epsilon$, $t \in \mathbb{R}$, for the time-t map $P_t : (x,y) \longmapsto (x(t), y(t))$.*

ii) Attractivity. *Every solution $(x(t), y(t))$ of Eq.(1) with $|y(0) - s_0(x(0))| \le \delta$ satisfies*

$$|y(t) - s(x(t),\epsilon)| \le K\, e^{-\beta t/\epsilon}\, |y(0) - s(x(0),\epsilon)| \quad \textit{for all } t \ge 0.$$

iii) "Asymptotic phase". *For every solution $(x(t), y(t))$ of Eq.(1) with initial conditions (x_0, y_0) at $t = 0$ satisfying $|y_0 - s_0(x_0)| \le \delta$ there is $(\hat{x}_0, \hat{y}_0) \in M_\epsilon$ such that for $(\hat{x}(t), \hat{y}(t))$ being the solution of Eq.(1) with $(\hat{x}(0), \hat{y}(0)) = (\hat{x}_0, \hat{y}_0)$*

$$
\left.
\begin{aligned}
|\,\hat{x}(t) - x(t)\,| &\le K\,\epsilon\, e^{-\beta t/\epsilon}\, |\,y_0 - s(x_0,\epsilon)\,| \\
|\,\hat{y}(t) - y(t)\,| &\le K\, e^{-\beta t/\epsilon}\, |\,y_0 - s(x_0,\epsilon)\,|
\end{aligned}
\right\} \quad \textit{holds for } t \ge 0.
$$

iv) Closeness to M_0.

$$s(x,0) = s_0(x)\,, \qquad x \in \mathbb{R}^m\,.$$

3. The invariant manifold result for the RK-map

In this section we investigate the geometric behaviour of RK-methods applied to Eq.(1). Since Eq.(1) is stiff for small ϵ, stiff RK-methods are needed to integrate such a system. For small ϵ the step size h of the integration method usually is much larger than the perturbation parameter ϵ.

For our approach it is essential that a RK-method is considered as a map in phase space. A RK-method applied to the differential equation $\dot{w} = F(w)$, $w \in \mathbb{R}^\ell$, is the map which takes $w \in \mathbb{R}^\ell$ to

$$\overline{w} \;=\; w + h \sum_{i=1}^{s} b_i\, F(W_i)$$

where the W_i are defined by

$$W_i \;=\; w + h \sum_{j=1}^{s} a_{ij}\, F(W_j), \quad i = 1, ..., s\,.$$

For implicit methods, i.e., $a_{ij} \neq 0$ for some $i \leq j$, one has to make sure that the stages W_i are well defined for the step size h and the vector field F considered. It is convenient to introduce the following vectors in $\mathbb{R}^{s\ell}$:

$$W := \begin{pmatrix} W_1 \\ \vdots \\ W_s \end{pmatrix}, \quad F(W) := \begin{pmatrix} F(W_1) \\ \vdots \\ F(W_s) \end{pmatrix}, \quad \mathbf{w} = \begin{pmatrix} w \\ \vdots \\ w \end{pmatrix}.$$

In this notation the RK-map may be written as

$$\begin{aligned} \overline{w} &= w + h\,(b^T \otimes I_\ell)\, F(W) \\ W &= \mathbf{w} + h\,(A \otimes I_\ell)\, F(W)\,. \end{aligned}$$

We make the following assumptions on the RK-method which are appropriate to integrate stiff systems.

Hypothesis H_{RK}

1) The RK-method has order p and stage order $1 \leq q < p$.

2) The RK-matrix A is invertible.

3) The stability function $R(z) := 1 + z\, b^T (I_s - zA)^{-1} 1\!\!1$, $z \in \mathbb{C}$, where $1\!\!1 = (1, ..., 1)^T \in \mathbb{R}^s$, satisfies $|R(\infty)| < 1$.

Remarks:

1) If $q = p$ then for $p > 1$ one may redefine $q = p - 1$ for Hypothesis H_{RK} 1) to hold. In this way, the results are not weakened. For $p = 1$ see Remark 7) in Section 4.

2) Since A is invertible $R(\infty)$ may be written as $R(\infty) = 1 - b^T A^{-1} 1\!\!1$.

3) Our Hypothesis H_{RK} is sufficient to show the results below for $\epsilon << h$ which is the important case for approximating solutions of Eq.(1) near the invariant manifold M_ϵ. The same results also hold for $\epsilon \le c\,h$, $c > 0$, under the following additional assumptions (used in Hairer, Lubich, Roche [1]):

   i) The system (1) satisfies $\mu(g_y(x, s_0(x))) \le -b_0$ for all $x \in \mathbb{R}^m$ where $\mu(\cdot)$ denotes the logarithmic norm with respect to some inner product.

   ii) The RK-method is A-stable and the eigenvalues of the RK-matrix A have positive real part. ⊣

We now apply a RK-method satisfying Hypothesis H_{RK} with $p < r$ to Eq.(1). This yields

$$\begin{aligned} \overline{x} &= x + h(b^T \otimes I_m)\, f(X, Y) \\ \overline{y} &= y + \frac{h}{\epsilon}(b^T \otimes I_n)\, g(X, Y) \end{aligned} \tag{2}$$

where X and Y are given by

$$\begin{aligned} X &= \mathbf{x} + h(A \otimes I_m)\, f(X, Y) \\ Y &= \mathbf{y} + \frac{h}{\epsilon}(A \otimes I_n)\, g(X, Y)\ . \end{aligned}$$

It will be shown that these equations define a smooth map

$$P: \ \mathbb{R}^m \times \mathbb{R}^n \ni \begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} \overline{x} \\ \overline{y} \end{pmatrix} \in \mathbb{R}^m \times \mathbb{R}^n\ . \tag{2}$$

We introduce the new variables z, Z measuring the difference to the manifold M_ϵ of the differential equation (1). Writing $s(x)$ instead of $s(x, \epsilon)$, for short, we define

$$y = s(x) + z\,, \quad Y = s(X) + Z\ .$$

We expand the function g about $z = 0$ and have

$$g(x, s(x) + z) = g(x, s(x)) + (B(x) + \hat{B}(x, z))\, z$$

with $B(x) := g_y(x, s(x)) = B_0(x) + O(\epsilon)$ and $\hat{B}(x, z) = O(|z|)$. In the new variables the y-component of the RK-map P takes the form

$$s(\overline{x}) + \overline{z} = s(x) + z + \frac{h}{\epsilon}\,(b^T \otimes I_n)\Big\{g(X, s(X)) \;+\; \mathrm{diag}\,\Big[B(X) + \hat{B}(X,Z)\Big]\,Z\Big\} \tag{3}$$

$$s(X) + Z = s(\mathbf{x}) + \mathbf{z} + \frac{h}{\epsilon}\,(A \otimes I_n)\Big\{g(X, s(X)) \;+\; \mathrm{diag}\,\Big[B(X) + \hat{B}(X,Z)\Big]\,Z\Big\} \tag{4}$$

where, e.g., $\mathrm{diag}[B(X)]$ denotes the $sn \times sn$ block diagonal matrix with $n \times n$ blocks $B(X_1), ..., B(X_s)$. We denote the RK-map in the new variables by $\widetilde{P}$ and we consider this map in the space $\mathbb{R}^m \times D_d$, $D_d := \{z \mid z \in \mathbb{R}^n,\ |z| \leq d\}$, where d will be determined later. We here suppose that $Z \in [D_d]^s$. We will see that Eq.(4) indeed may be uniquely solved for Z in $[D_d]^s$. Collecting the linear terms in Z in Eq.(4) yields

$$Z = \frac{\epsilon}{h}\,C(X,Z)^{-1}(\mathbf{z} - E(x,X))$$

with

$$C(X,Z) := -(A \otimes I_n)\ \mathrm{diag}\,\Big[B(X) + \hat{B}(X,Z)\Big] + \frac{\epsilon}{h}\,(I_s \otimes I_n)\,, \tag{5}$$

$$E(x,X) := s(X) - s(\mathbf{x}) - \frac{h}{\epsilon}\,(A \otimes I_n)\,g(X, s(X))\,. \tag{6}$$

Note that due to H_{DE} 3) and H_{RK} 2) the matrix C is invertible for d and ϵ/h small enough. Inserting the expression obtained for Z into Eq.(3) we find

$$\overline{z} = z - e(x, \overline{x}, X) + (b^T \otimes I_n)\ \mathrm{diag}\,\Big[B(X) + \hat{B}(X,Z)\Big]\,C(X,Z)^{-1}\,(\mathbf{z} - E(x,X))$$

where

$$e(x, \overline{x}, X) := s(\overline{x}) - s(x) - \frac{h}{\epsilon}\,(b^T \otimes I_n)\,g(X, s(X))\,. \tag{7}$$

In what follows we mostly suppress the arguments of E and e for simplicity of notation. Note that E and e are of order $O(h)$. This is due to the fact that $g(X, s_0(X)) = 0$ and $s(X) - s_0(X) = O(\epsilon)$ and that $X - \mathbf{x}$ and $\overline{x} - x$ are $O(h)$. By means of Eq.(5) we may replace $\mathrm{diag}\Big[B(X) + \hat{B}(X,Z)\Big]$ by $(A^{-1} \otimes I_n)\Big(\frac{\epsilon}{h}\,(I_s \otimes I_n) - C(X,Z)\Big)$. Hence, we find that the map $\widetilde{P}$ may be written in the form

$$\begin{aligned}
\overline{x} &= x + h(b^T \otimes I_m) f(X, s(X) + Z)\\
\overline{z} &= \Big(R(\infty)\,I_n + \frac{\epsilon}{h}\,\Delta(\mathbb{1} \otimes I_n)\Big)\,z + \Big((b^T A^{-1} \otimes I_n) - \frac{\epsilon}{h}\,\Delta\Big)\,E - e
\end{aligned} \tag{8}$$

where $\Delta := (b^T A^{-1} \otimes I_n)\ C(X,Z)^{-1}$ and where X, Z are defined by

$$\begin{aligned}
X &= \mathbf{x} + h(A \otimes I_m) f(X, s(X) + Z)\\
Z &= \frac{\epsilon}{h}\,C(X,Z)^{-1}\,(\mathbf{z} - E)\,.
\end{aligned} \tag{9}$$

By means of the contraction principle it can be shown that for h and ϵ/h sufficiently small Eq.(9) has a unique solution $(X(x,z,h,\epsilon), Z(x,z,h,\epsilon))$ which is $O(h+\epsilon/h)$-close to $(\mathbf{x},0)$. From the implicit function theorem it follows that for small h and ϵ/h this solution is smooth with bounded derivatives. Therefore, the map $\widetilde{P}$ given in Eq.(8) is well defined and is of the form

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

where $\hat{F}$ and G are of class C_b^r with respect to x and z.

We now apply the invariant manifold result of Nipp, Stoffer [5]. The cylinder $\mathbb{R}^m \times D_d$ is mapped into itself by the map $\widetilde{P}$ if h and ϵ/h are sufficiently small. The functions $\hat{F}$ and G have the following Lipschitz constants with respect to x and z:

$$\begin{aligned} L_{11} &= O(h)\,, & L_{12} &= O(\epsilon)\,, \\ L_{21} &= O(h) + O\Big(\frac{\epsilon}{h}\Big)\,, & L_{22} &= |R(\infty)| + O\Big(\frac{\epsilon}{h}\Big)\,. \end{aligned}$$

The estimates for L_{12} and L_{22} might not be obvious. They may be obtained as follows. Differentiating Eq. (9) with respect to z one gets $dX/dz = O(h)\,dZ/dz$ and $dZ/dz = O(\epsilon/h)$. Differentiating Eq. (8) with respect to z and using these estimates yields the above estimates for L_{12} and L_{22}. We have to verify the two conditions

$$\begin{aligned} 2\sqrt{L_{12}\,L_{21}} &< 1 - L_{11} - L_{22} \\ L_{22} + L_{12}\,\lambda &< (1 - L_{11} - L_{12}\,\lambda)^r \end{aligned}$$

where

$$\lambda := \frac{2L_{21}}{1 - L_{11} - L_{22} + \sqrt{(1 - L_{11} - L_{22})^2 - 4L_{12}\,L_{21}}}\,.$$

These two conditions are satisfied if h and ϵ/h are sufficiently small. (Note that the larger r the smaller h and ϵ/h have to be taken. In the case $|R(\infty)| = 0$, however, h may be chosen independent of r.) Now, Theorem 5 of [5] implies the existence of a smooth attractive invariant manifold $\widetilde{M}_{h,\epsilon}$ of the map $\widetilde{P}$. More precisely: Let Ω_{h^*,δ^*} be the domain $\{(x,h,\epsilon)|\; x \in \mathbb{R}^m,\; h \in (0,h^*),\; \epsilon \in (0,h\delta^*)\}$.

For h^* and δ^* sufficiently small there is a λ-Lipschitz function $\widetilde{\sigma} : \Omega_{h^*,\delta^*} \longrightarrow D_d \subset \mathbb{R}^n$ of class C_b^r with respect to x, $\lambda = O(h) + O(\epsilon/h)$, such that the following assertions hold.

i) $\widetilde{M}_{h,\epsilon} = \{(x,z)\,|\; x \in \mathbb{R}^m,\; z = \widetilde{\sigma}(x,h,\epsilon)\}$ *is an invariant set of the map* $\widetilde{P}$, *i.e.,* $\widetilde{P}(\widetilde{M}_{h,\epsilon}) = \widetilde{M}_{h,\epsilon}$.

ii) $\widetilde{M}_{h,\epsilon}$ *is uniformly attractive for* $\widetilde{P}$ *with attractivity constant* $\chi(\lambda) = L_{22} + L_{12}\,\lambda = |R(\infty)| + O(\epsilon/h) < 1$.

iii) The "property of asymptotic phase" holds, i.e., for every $(x_0, z_0) \in \mathbb{R}^m \times D_d$ *there is* $(\tilde{x}_0, \tilde{z}_0) \in \widetilde{M}_{h,\epsilon}$ *such that for* $(x_j, z_j) := \widetilde{P}^j(x_0, z_0)$ *and* $(\tilde{x}_j, \tilde{z}_j) := \widetilde{P}^j(\tilde{x}_0, \tilde{z}_0) \in \widetilde{M}_{h,\epsilon}$, $j \in \mathbb{N}_0$,

$$\begin{aligned} |\tilde{x}_j - x_j| &\leq c\,\chi(\lambda)^j\, |z_0 - \tilde{\sigma}(x_0, h, \epsilon)| \\ |\tilde{z}_j - z_j| &\leq (1 + \lambda c)\, \chi(\lambda)^j\, |z_0 - \tilde{\sigma}(x_0, h, \epsilon)| \end{aligned}$$

holds with

$$c = \frac{L_{12}}{\sqrt{(1 - L_{11} - L_{22})^2 - 4 L_{12} L_{21}}} = O(\epsilon)\ .$$

iv) If the function $G(x, z, h, \epsilon)$ *has the form* $G(x, z, h, \epsilon) = H(x, z, h, \epsilon) z + \hat{G}(x, z, h, \epsilon)$ *with* $|H(x, z, h, \epsilon)| \leq \rho < 1$ *then* $\tilde{\sigma}$ *may be estimated by*

$$|\tilde{\sigma}(x, h, \epsilon)| \leq \frac{1}{1-\rho} \sup_{x \in \mathbb{R}^m,\, |z| \leq d} |\hat{G}(x, z, h, \epsilon)|\ .$$

According to Eq.(8) the function G may be split as in iv) with $|H| = |R(\infty)| + O(\epsilon/h)$ and

$$\hat{G} = \Big((b^T A^{-1} \otimes I_n) - \frac{\epsilon}{h}\,\Delta\Big)\, E(x, X) - e(x, \overline{x}, X)\,. \tag{10}$$

Since E and e are of order $O(h)$ as noted above the function $\hat{G}$ is of order $O(h)$. It therefore follows from iv) that $\tilde{\sigma}$ is of order $O(h)$ as well. This estimate may be improved, however.

Claim *The following estimates hold:*

$$Z \;=\; O\Big(\frac{\epsilon}{h}\Big)\, \mathbf{z} \;+\; O(\epsilon\, h^q)\ .$$

$$\tilde{\sigma} \;=\; \begin{cases} O(\epsilon\, h^q) & \textit{if } a_{sj} = b_j \textit{ for } j = 1, \ldots, s\ , \\ O(h^{q+1}) & \textit{else}\ . \end{cases}$$

To prove this claim we have to derive better estimates for $e(x, \overline{x}, X)$ and $E(x, X)$. We need some preparations. We consider solutions $(u(t), v(t))$ of Eq.(1) on the manifold M_ϵ established in Theorem 1. These solutions satisfy the differential equation

$$\begin{aligned} \dot{u} &= f(u, s(u)) \\ \dot{v} &= \frac{1}{\epsilon}\, g(u, s(u)) = s'(u) f(u, s(u))\ . \end{aligned} \tag{11}$$

Here, $s(u)$ is the function defining the manifold M_ϵ (for simplicity we again drop the parameter ϵ in $s(u, \epsilon)$). The identity $g(u, s(u)) = \epsilon s'(u) f(u, s(u))$ follows from $v(t) = s(u(t))$. s' denotes the derivative of $s(u)$ with respect to u. Applying a RK-method to Eq.(11) we obtain

$$\begin{aligned} \overline{u} &= u + h(b^T \otimes I_m)\, f(U, s(U)) \\ \overline{v} &= v + h(b^T \otimes I_n)\, \mathrm{diag}\, [s'(U)]\, f(U, s(U)) \end{aligned} \tag{12}$$

with

$$\begin{aligned} U &= \mathbf{u} + h(A \otimes I_m)\, f(U, s(U)) \\ V &= \mathbf{v} + h(A \otimes I_n)\, \mathrm{diag}\, [s'(U)]\, f(U, s(U)) \,. \end{aligned} \tag{13}$$

For $u = x$ we have (cf. Eqs.(6) and (7))

$$\begin{aligned} e(x, \overline{x}, X) &= e(u, \overline{u}, U) + O(1)(\overline{x} - \overline{u}) + O(h)(X - U) \\ E(x, X) &= E(u, U) + O(1)(X - U) \,. \end{aligned} \tag{14}$$

Since the method is of order p and has stage order q and since f, g are of class C_b^r with $r > p$ we conclude that

$$\begin{aligned} O(h^{p+1}) &= \overline{v} - v(h) &&= \overline{v} - s(u(h)) &&= \overline{v} - s(\overline{u}) + O(h^{p+1}) \\ O(h^{q+1}) &= V - \mathbf{v}(c_i h) &&= V - s(\mathbf{u}(c_i h)) &&= V - s(U) + O(h^{q+1}) \end{aligned}$$

where $c_i = \sum_{j=1}^{s} a_{ij}$ and $\mathbf{v}(c_i h) := \Big(v(c_1 h)^T, ..., v(c_s h)^T\Big)^T$. Using Eqs.(6), (7) and (12), (13) it follows that

$$\begin{aligned} e(u, \overline{u}, U) &= s(\overline{u}) - \overline{v} &&= O(h^{p+1}) \\ E(u, U) &= s(U) - V &&= O(h^{q+1}) \,. \end{aligned} \tag{15}$$

In order to estimate $e(x, \overline{x}, X)$ and $E(x, X)$ it remains to estimate $\overline{x} - \overline{u}$ and $X - U$ where $\overline{x}$ and X are defined by Eqs.(8), (9) and $\overline{u}$, U by Eqs.(12), (13). For $u = x$ we have

$$\begin{aligned} \overline{x} - \overline{u} &= h(b^T \otimes I_m)\Big(f(X, s(X) + Z) - (f(U, s(U))\Big) \\ X - U &= h(A \otimes I_m)\, \Big(f(X, s(X) + Z) - (f(U, s(U))\Big) \end{aligned}$$

or

$$\begin{aligned} \overline{x} - \overline{u} &= O(h)\, Z + O(h)\, (X - U) \\ \Big((I_s \otimes I_m) + O(h)\Big)(X - U) &= O(h)\, Z \end{aligned}$$

implying

$$\begin{aligned} \overline{x} - \overline{u} &= O(h)\, Z \\ X - U &= O(h)\, Z \,. \end{aligned} \tag{16}$$

Using Eqs.(14), (15) and (16) we find

$$
\begin{aligned}
e(x,\overline{x},X) &= O(h^{p+1}) + O(h)\, Z \\
E(x,X) &= O(h^{q+1}) + O(h)\, Z\ .
\end{aligned}
$$

Inserting the second estimate into Eq.(9) gives

$$
Z = O\Big(\frac{\epsilon}{h}\Big)\ \mathbf{z}\ +\ O(\epsilon\, h^q)\ . \tag{17}
$$

Hence, we have

$$
\begin{aligned}
e(x,\overline{x},X) &= O(\epsilon)\ \mathbf{z}\ +\ O(\epsilon\, h^{q+1}) + O(h^{p+1}) \\
E(x,X) &= O(\epsilon)\ \mathbf{z}\ +\ O(h^{q+1})\ .
\end{aligned} \tag{18}
$$

We insert these estimates into the expression for $\hat{G}$ in Eq.(10) and obtain a new splitting of the function G:

$$
G(x,z,h,\epsilon) = (H(x,z,h,\epsilon) + O(\epsilon))\ z + O(h^{q+1})\ .
$$

Now, property iv) implies $\widetilde{\sigma} = O(h^{q+1})$. In the case $b_i = a_{si}$ one has $\overline{z} = Z_s$ and therefore for $(x,z) \in \widetilde{M}_{h,\epsilon}$ Eq.(17) implies $Z = O(\epsilon\, h^q)$ and hence $\widetilde{\sigma} = O(\epsilon\, h^q)$. This completes the proof of the claim above.

Expressing the results above in the original variables x, y and defining $\sigma(x,h,\epsilon) := s(x,\epsilon) + \widetilde{\sigma}(x,h,\epsilon)$ we have shown

Theorem 2 *Let the differential equation (1) satisfy Hypothesis H_{DE}. Apply a RK-method with Hypothesis H_{RK} to Eq.(1) and assume $p < r$.*

Then there are constants h_0, δ_0, d, c, K and a function $\sigma : \Omega_{h_0,\delta_0} \longrightarrow \mathbb{R}^n$, $\Omega_{h_0,\delta_0} := \{(x,h,\epsilon)|\ x \in \mathbb{R}^m,\ h \in (0,h_0),\ \epsilon \in (0, h\delta_0)\}$, σ of class C_b^r with respect to x, such that for all h, ϵ with $h \le h_0$, $\epsilon/h \le \delta_0$ the following assertions hold.

i) Invariance. *The set $M_{h,\epsilon} = \{(x,y)\,|\ x \in \mathbb{R}^m,\ y = \sigma(x,h,\epsilon)\}$ is an invariant set of the map P given in Eq.(2), i.e., $P(M_{h,\epsilon}) = M_{h,\epsilon}$.*

ii) Attractivity. *The manifold $M_{h,\epsilon}$ is uniformly attractive for P with attractivity constant $\chi(h,\epsilon) = |R(\infty)| + c\,\epsilon/h < 1$, i.e., for all (x,y) with $|y - s(x,\epsilon)| \le d$ the inequality*

$$
|\overline{y} - \sigma(\overline{x},h,\epsilon)|\ \le\ \chi(h,\epsilon)\,|y - \sigma(x,h,\epsilon)|
$$

holds.

iii) "Asymptotic phase". *For every* (x_0, y_0) *with* $|y_0 - s(x_0,\epsilon)| \le d$ *there is* $(\tilde{x}_0, \tilde{y}_0) \in M_{h,\epsilon}$ *such that for* $(x_j, y_j) := P^j(x_0, y_0)$ *and* $(\tilde{x}_j, \tilde{y}_j) := P^j(\tilde{x}_0, \tilde{y}_0) \in M_{h,\epsilon}$, $j \in \mathbb{N}_0$,

$$\begin{aligned} |\tilde{x}_j - x_j| &\le K\epsilon\,\chi(h,\epsilon)^j\,|\,y_0 - \sigma(x_0,h,\epsilon)| \\ |\tilde{y}_j - y_j| &\le (1 + K\epsilon)\,\chi(h,\epsilon)^j\,|\,y_0 - \sigma(x_0,h,\epsilon)|\,. \end{aligned}$$

iv) Closeness to M_ϵ. *For all* $x \in \mathbb{R}^m$ *the following estimate holds:*

$$|\sigma(x,h,\epsilon) - s(x,\epsilon)| \le \begin{cases} K\,\epsilon\,h^q & \text{if } b_i = a_{si},\ i = 1,\ldots,s\,, \\ K\,h^{q+1} & \text{else}\,. \end{cases}$$

Moreover, the stages of the RK-method (2) satisfy

$$|Y - s(X,\epsilon)| \le K\,\frac{\epsilon}{h}\left(|y - s(x,\epsilon)| + h^{q+1}\right).$$

Remarks:

4) Under Hypothesis H_{DE} the singularly perturbed differential equation (1) admits an attractive invariant manifold. Theorem 2 states that under Hypothesis H_{RK} the RK-map (2) inherits this geometric property.

5) The invariant manifold of the differential equation is highly attractive, i.e., the manifold M_ϵ of the time-h map of Eq.(1) has attractivity $O(e^{-\beta h/\epsilon})$. The invariant manifold $M_{h,\epsilon}$ of the RK-map (2) has attractivity $|R(\infty)| + O(\epsilon/h)$. Hence, the attractivity property of the manifold M_ϵ is poorly reproduced by the attractivity of $M_{h,\epsilon}$ unless $R(\infty) = 0$.

6) In general, the distance between the manifolds M_ϵ and $M_{h,\epsilon}$ is $O(h^{q+1})$. If $|y - s(x,\epsilon)| = O(h^{q+1})$ then Property iv) shows that $Y - s(X,\epsilon) = O(\epsilon\,h^q)$. This means that starting $O(h^{q+1})$-close to M_ϵ or $M_{h,\epsilon}$ all the information to step forward the numerical method is taken $O(\epsilon\,h^q)$-close to M_ϵ. $\dashv$

4. Global error bounds for stiff RK-methods

Our main result in Section 3 was obtained by investigating one step of the RK-method, only. The approximation error after one step is the local error. However, integrating an ODE one is interested in the global error, i.e., the error after N steps, at time $T = Nh$. Since the invariant manifold result in fact is a global one this result is very helpful to derive bounds on the global error of the RK-method. According to Theorem 1 iii) the flow of

Eq.(1) close to the manifold M_ϵ is essentially described by the m-dimensional system $\dot{x} = f(x, s(x,\epsilon))$ which is no longer stiff with $\epsilon \to 0$. Similarly, according to Theorem 2 iii) the dynamics of the RK-map (2) near the manifold $M_{h,\epsilon}$ is essentially described by an m-dimensional map (the x-part of (2) on $M_{h,\epsilon}$). Thus, our geometric approach reduces the stiff problem to a nonstiff one.

Theorem 3 *Let the differential equation (1) satisfy Hypothesis H_{DE} and consider a solution $(x(t), y(t))$. Let (x_k, y_k) be a RK-approximation by a method satisfying Hypothesis H_{RK} with initial values $(x_0, y_0) = (x(0), y(0))$ and assume $p < r$.*

Then there are constants h_0, δ_0, d, c such that for all $(x(0), y(0))$ with $|y(0) - s(x(0),\epsilon)| \leq d$ and for all h, ϵ with $h \leq h_0$, $\epsilon/h \leq \delta_0$ the following error bounds hold for $kh \leq Nh = T$, T fixed:

$$\begin{aligned} x_k - x(kh) &= O(h^p) + O(\epsilon\, h^{q+1}) + O\Big(\epsilon\, |y_0 - s(x_0,\epsilon)|\Big) \\ y_k - y(kh) &= O(h^{q+1}) + O\Big((\epsilon + \chi(h,\epsilon)^k)\, |y_0 - s(x_0,\epsilon)|\Big) \end{aligned}$$

where $\chi(h,\epsilon) = |R(\infty)| + c\,\epsilon/h < 1$.

Moreover, if $b_i = a_{si}$, $i = 1, ..., s$, the estimate

$$y_k - y(kh) = O(h^p) + O(\epsilon\, h^q) + O\Big((\epsilon + \chi(h,\epsilon)^k)\, |y_0 - s(x_0,\epsilon)|\Big)$$

holds.

Remarks:

7) If $q = p = 1$ then $x_k - x(kh) = O(h) + O\Big(\epsilon\, |y_0 - s(x_0,\epsilon)|\Big)$, $y_k - y(kh) = O(h) + O\Big((\epsilon + \chi(h,\epsilon)^k)\, |y_0 - s(x_0,\epsilon)|\Big)$. See also Remark 1).

8) As stated in Remark 5) the attractivity of $M_{h,\epsilon}$ is weak compared to the attractivity of M_ϵ unless $R(\infty) = 0$. For the global error, however, the strong attractivity of $M_{h,\epsilon}$ is not crucial. In any case, after a few steps, $\chi(h,\epsilon)^k\, |y_0 - s(x_0,\epsilon)|$ becomes small compared to the other terms in the global error estimate.

9) The result for solutions starting on M_ϵ, i.e., $(x_0, y_0) = (x(0), y(0))$ with $y_0 = s(x_0,\epsilon)$, was first proved in Hairer, Lubich, Roche [1] using ϵ-expansions of the solutions of Eqs.(1) and (2). $\dashv$

Proof: We prove Theorem 3 in the following way. In a first part we compare a numerical solution on $M_{h,\epsilon}$ with a "numerical solution on M_ϵ". In the second part of the proof it is shown that the assertions of Theorem 3 then follow from the estimate of the first part, from the attractivity and the property of asymptotic phase of the manifolds $M_{h,\epsilon}$ and M_ϵ and from the closeness of the two manifolds.

Part I: We consider a RK-solution (x_k, y_k) of Eq.(1) with starting values on $M_{h,\epsilon}$, i.e., $y_0 = \sigma(x_0, h, \epsilon)$, and as in Section 3 we compare x_k with the RK-solution u_k of the differential equation $\dot{u} = f(u, s(u))$. (We again write $s(u)$ instead of $s(u,\epsilon)$.) We set $u_0 = x_0$. The x-part of the RK-solution of the full system satisfies

$$
\begin{aligned}
x_{k+1} &= x_k + h(b^T \otimes I_m)\, f(X_k, s(X_k) + Z_k) \\
X_k &= \mathbf{x}_k + h(A \otimes I_m)\, f(X_k, s(X_k) + Z_k)
\end{aligned} \tag{19}
$$

(where $s(X_k) + Z_k = Y_k$) and the RK-solution of the simplified system satisfies

$$
\begin{aligned}
u_{k+1} &= u_k + h(b^T \otimes I_m)\, f(U_k, s(U_k)) \\
U_k &= \mathbf{u}_k + h(A \otimes I_m)\, f(U_k, s(U_k))\, .
\end{aligned} \tag{20}
$$

In this section $X_k, U_k \in \mathbb{R}^{sm}$, $Z_k \in \mathbb{R}^{sn}$ denote the vectors of all stages in the $(k+1)$-st step (and not the k-th stages).

We want to estimate $x_k - u_k$. We first derive a weaker bound stated in the claim below. This preliminary bound then simplifies the proof of the desired estimate (29).

Claim $\quad x_k - u_k = O(\epsilon\, h^q)\,, \qquad X_k - U_k = O(\epsilon\, h^q) \qquad \textit{for } kh \le T.$

We introduce the differences

$$
\Delta x_k = x_k - u_k\,, \quad \Delta X_k = X_k - U_k\,.
$$

From Eqs.(19) and (20) we find that

$$
\begin{aligned}
\Delta x_{k+1} &= \Delta x_k \;+\; O(h)\, Z_k \;+\; O(h)\, \Delta X_k \\
\Delta X_k &= \mathbf{\Delta x}_k \;+\; O(h)\, Z_k \;+\; O(h)\, \Delta X_k\, .
\end{aligned} \tag{21}
$$

From Theorem 2 we know that for $z_k = y_k - s(x_k) = \sigma(x_k, h, \epsilon) - s(x_k) = O(h^{q+1})$ the estimate $Z_k = Y_k - s(X_k) = O(\epsilon\, h^q)$ holds. We conclude

$$
\Delta X_k = O(1)\, \mathbf{\Delta x}_k \;+\; O(\epsilon\, h^{q+1})\,, \tag{22}
$$

$$
\Delta x_{k+1} = \big(I_m + O(h)\big)\, \Delta x_k + O(\epsilon\, h^{q+1})\,. \tag{23}
$$

A simple induction argument involving a geometric series implies $|\Delta x_k| \le C_0 \epsilon h^q$ for $k \le N$. Inserting this estimate into Eq.(22) terminates the proof of the claim.

The bounds in the claim may be improved, however. We investigate the first equation of Eq.(21) more precisely. From Eqs. (19) and (20) we have

$$\Delta x_{k+1} = \Delta x_k + h(b^T \otimes I_m)\,\mathrm{diag}[f_x^k + f_y^k s'^k]\,\Delta X_k + h(b^T \otimes I_m)\,\mathrm{diag}[f_y^k]\,Z_k + O(\epsilon\, h^{q+2}) \tag{24}$$

where $\mathrm{diag}[f_y^k]$, e.g., is the $sm \times sn$ block diagonal matrix with blocks $f_y(X_{k,i}, s(X_{k,i}))$, $X_{k,i}$ being the i-th stage of the k-th step. Subtracting Δx_k on both sides of Eq.(24), taking the sum from $k = 0$ to $k = j$, taking into account Eq.(22) and then taking norms yields

$$|\Delta x_{j+1}| \leq hC_0 \sum_{k=0}^{j} |\Delta x_k| + \left| h(b^T \otimes I_m) \sum_{k=0}^{j} \mathrm{diag}[f_y^k]\,Z_k + O(\epsilon\, h^{q+1}) \right| . \tag{25}$$

Eq.(25) is of the form $\delta_{j+1} \leq c_0 \sum_{k=0}^{j} \delta_k + c_1$, $c_0 > 0$, for $0 \leq j < N$. Using a discrete *Gronwall* argument leads to $\delta_{j+1} \leq c_1(1 + c_0)^j \leq c_1 e^{j\, c_0}$, $j < N$. Hence, we have

$$|\Delta x_{j+1}| \leq C_1(C_2(h, \epsilon) + \epsilon\, h^{q+1}) . \tag{26}$$

where $C_2(h, \epsilon)$ is an upper bound for

$$S_j := h(b^T \otimes I_m) \sum_{k=0}^{j} \mathrm{diag}[f_y^k]\,Z_k , \quad 0 \leq j < N . \tag{27}$$

We show that $C_2(h, \epsilon) = O(\epsilon\, h^{q+1})$. Using Eqs.(3), (7) and the fact that $Z_k = O(\epsilon\, h^q)$ we find that

$$z_{k+1} = z_k - e(x_k, x_{k+1}, X_k) + \frac{h}{\epsilon}(b^T \otimes I_n)\,\mathrm{diag}[g_y^k]\,Z_k + O(\epsilon\, h^{q+2}) .$$

Multiplying this equation by ϵ and using the estimate for e given in Eq. (18) we obtain

$$h(b^T \otimes I_n)\,\mathrm{diag}[g_y^k]\,Z_k = \epsilon\,(z_{k+1} - z_k) + T_k \tag{28}$$

with $T_k = O(\epsilon\, h^{q+2})$. Note that, if in Eq.(27) $\mathrm{diag}[f_y^k]$ were replaced by $\mathrm{diag}[g_y^k]$ we directly had $S_j = \epsilon\,(z_{j+1} - z_0) + O(\epsilon h^{q+1}) = O(\epsilon h^{q+1})$. We show that indeed a similar expression for S_j may be obtained with appropriate factors, however. We define for $0 \leq k < N$

$$H(x_k) := f_y(x_k, s(x_k))\; g_y(x_k, s(x_k))^{-1}$$

and

$$\begin{aligned} Q_k &:= h(b^T \otimes I_m)\,\mathrm{diag}[f_y^k]\,Z_k - H(x_k)\,h(b^T \otimes I_n)\,\mathrm{diag}[g_y^k]\,Z_k \\ R_k &:= \epsilon\,[H(x_k) - H(x_{k+1})]\,z_{k+1} . \end{aligned}$$

With these definitions we get from Eq.(28)

$$
\begin{aligned}
h(b^T \otimes I_m)\,\mathrm{diag}[f_y^k]\,Z_k &= H(x_k)\,h(b^T \otimes I_n)\,\mathrm{diag}[g_y^k]\,Z_k + Q_k \\
&= \epsilon\,[H(x_{k+1})z_{k+1} - H(x_k)z_k] + R_k + Q_k + H(x_k)\,T_k
\end{aligned}
$$

and hence

$$
S_j = \epsilon\,H(x_{j+1})\,z_{j+1} - \epsilon\,H(x_0)\,z_0 + \sum_{k=0}^{j} \Big(Q_k + R_k + H(x_k)\,T_k\Big)
$$

holds. We estimate R_k and Q_k. Since $H(x_k) - H(x_{k+1}) = O(|x_{k+1} - x_k|) = O(h)$ we conclude that $R_k = O(\epsilon\,h^{q+2})$. Similarly, from

$$
\mathrm{diag}[f_y^k] - (I_s \otimes H(x_k))\,\mathrm{diag}[g_y^k] = O(h)
$$

we have $Q_k = O(\epsilon\,h^{q+2})$. We conclude

$$
\sum_{k=0}^{j} \Big(Q_k + R_k + H(x_k)\,T_k\Big) = \sum_{k=0}^{j} O(\epsilon\,h^{q+2}) = O(\epsilon\,h^{q+1})
$$

implying $S_j = O(\epsilon\,h^{q+1})$ and thus $c_1(h,\epsilon) = O(\epsilon\,h^{q+1})$. From Eq.(26) we obtain

$$
\Delta x_k = x_k - u_k = O(\epsilon\,h^{q+1}). \tag{29}
$$

Part II: We first estimate $x_k - x(kh),\ kh \le Nh = T$.

- According to the property of asymptotic phase of $M_{h,\epsilon}$ (cf. Theorem 2 iii)) there is a RK-solution $(\tilde{x}_k, \tilde{y}_k = \sigma(\tilde{x}_k, h, \epsilon))$ in $M_{h,\epsilon}$ with

  $$
  \tilde{x}_k - x_k = O(\epsilon\,\chi(h,\epsilon)^k |y_0 - \sigma(x_0, h, \epsilon)|)\ . \tag{30}
  $$

- We consider the RK-solution $(\tilde{u}_k)$ of the reduced equation

  $$
  \dot{u} = f(u, s(u)) \tag{31}
  $$

  with $\tilde{u}_0 = \tilde{x}_0$. By the estimate given in Eq.(29) we have

  $$
  \tilde{x}_k - \tilde{u}_k = O(\epsilon\,h^{q+1})\ .
  $$

- According to the property of asymptotic phase of M_ϵ (cf. Theorem 1 iii)) there is a solution $(\hat{u}(t), s(\hat{u}(t)))$ of Eq.(1) in M_ϵ (note that $\hat{u}(t)$ satisfies Eq.(31)) with

  $$
  \hat{u}(kh) - x(kh) = O\Big(\epsilon\,e^{-\beta kh/\epsilon}\,|y(0) - s(x(0))|\Big)\ . \tag{32}
  $$

- We consider the RK-solution $(\hat{u}_k)$ of Eq.(31) with $\hat{u}_0 = \hat{u}(0)$. Since Eq.(31) is nonstiff for $\epsilon \to 0$ we have the global error estimate

$$\hat{u}_k - \hat{u}(kh) = O(h^p)\ .$$

- By the Lipschitz continuity of the RK-map we have $|\tilde{u}_{k+1} - \hat{u}_{k+1}| \leq (1+c_0 h)\,|\tilde{u}_k - \hat{u}_k|$ and hence $\tilde{u}_k - \hat{u}_k = O(|\tilde{u}_0 - \hat{u}_0|)$. Using Eqs.(30) and (32) for $k = 0$ we estimate

$$\begin{aligned}\tilde{u}_0 - \hat{u}_0 &= \tilde{x}_0 - \hat{u}(0) = [\tilde{x}_0 - x_0] + [x(0) - \hat{u}(0)] \\ &= O(\epsilon\,|y_0 - \sigma(x_0, h, \epsilon)|) + O(\epsilon\,|y_0 - s(x_0)|)\ .\end{aligned}$$

  Since $M_{h,\epsilon}$ is $O(h^{q+1})$-close to M_ϵ (cf. Theorem 2 iv)) we get

$$\tilde{u}_0 - \hat{u}_0 = O(\epsilon\,|y_0 - s(x_0)|) + O(\epsilon\, h^{q+1})\ .$$

- Combining the above estimates we find

$$\begin{aligned}x_k - x(kh) &= [x_k - \tilde{x}_k] + [\tilde{x}_k - \tilde{u}_k] + [\tilde{u}_k - \hat{u}_k] + [\hat{u}_k - \hat{u}(kh)] + [\hat{u}(kh) - x(kh)] \\ &= O(h^p) + O(\epsilon\, h^{q+1}) + O(\epsilon\,|y_0 - s(x_0)|)\ .\end{aligned}$$

It remains to estimate $y_k - y(kh)$, $kh \leq Nh = T$. Using the attractivity of the manifolds $M_{h,\epsilon}$ and M_ϵ, their closeness and the above estimate of the x-component we obtain

$$\begin{aligned}y_k - y(kh) &= [y_k - \sigma(x_k, h, \epsilon)] + [\sigma(x_k, h, \epsilon) - s(x_k)] \\ &\quad + [s(x_k) - s(x(kh))] + \ [s(x(kh)) - y(kh)] \\ &= [O(\chi(h,\epsilon)^k\,|y_0 - \sigma(x_0, h, \epsilon)|)] + [O(h^{q+1})] \\ &\quad + [O(h^p) + O(\epsilon\, h^{q+1}) + \ O(\epsilon\,|y_0 - s(x_0)|)] + [O(e^{-\beta kh/\epsilon}|y_0 - s(x_0)|)] \\ &= O(h^{q+1}) + O\Big((\epsilon + \chi(h,\epsilon)^k)\,|y_0 - s(x_0)|\Big)\ .\end{aligned}$$

If $b_i = a_{si}$ the $O(h^{q+1})$-term is replaced by $O(h^p) + O(\epsilon\, h^q)$. $\perp$

Acknowledgement

The starting point of this paper was a talk given by E. Hairer at ETH on the global error results of RK-methods applied to singularly perturbed ODEs. At that occasion U. Kirchgraber asked about the geometric properties of such discrete systems.

The authors learned a lot on the RK-formalism in a summer school on numerical integration methods for ODEs held by Ch. Lubich.

References

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

[2] U. Kirchgraber and K. Nipp, *Geometric properties of RK-methods applied to stiff ODE's of singular perturbation type*, Research Report No. 89-02, Seminar für Angewandte Mathematik, ETH Zürich (1989).

[3] K. Nipp, *Invariant manifolds of singularly perturbed ordinary differential equations*, ZAMP 36 (1985), 309-320.

[4] K. Nipp, *Smooth attractive invariant manifolds of singularly perturbed ODE's*, Research Report No. 92-13, Seminar für Angewandte Mathematik, ETH Zürich (1992).

[5] K. Nipp and D. Stoffer, *Attractive invariant manifolds for maps: Existence, smoothness and continuous dependence on the map*, Research Report No. 92-11, Seminar für Angewandte Mathematik, ETH Zürich (1992).