Research article

On the quasi-steady-state approximation in an open Michaelis–Menten reaction mechanism

Justin Eilertsen[1], Marc R. Roussel[2], Santiago Schnell[1,3] and Sebastian Walcher[4,*]

[1] Department of Molecular & Integrative Physiology, University of Michigan Medical School, Ann Arbor, Michigan 49109, USA

[2] Alberta RNA Research and Training Institute, Department of Chemistry and Biochemistry, University of Lethbridge, Lethbridge, Alberta, Canada, T1K 3M4

[3] Department of Computational Medicine & Bioinformatics, University of Michigan Medical School, Ann Arbor, Michigan 49109, USA

[4] Mathematik A, RWTH Aachen, D-52056 Aachen, Germany

* **Correspondence:** Email: walcher@matha.rwth-aachen.de

Abstract: The conditions for the validity of the standard quasi-steady-state approximation in the Michaelis–Menten mechanism in a closed reaction vessel have been well studied, but much less so the conditions for the validity of this approximation for the system with substrate inflow. We analyze quasi-steady-state scenarios for the open system attributable to singular perturbations, as well as less restrictive conditions. For both settings we obtain distinguished invariant manifolds and time scale estimates, and we highlight the special role of singular perturbation parameters in higher order approximations of slow manifolds. We close the paper with a discussion of distinguished invariant manifolds in the global phase portrait.

1. Introduction

Cellular function involves a large network of transformations of substrates, denoted S, into products, P, which in turn may be further transformed, eliminated, or cycled back into a useful form. While the chemical conversion of S into P can occur spontaneously

$$S \xrightarrow{k} P, \tag{1.1}$$

the rate constant, k, that regulates the speed of the reaction (1.1) will often be very small, so that spontaneous conversion is too slow to sustain life. Moreover, spontaneous conversion allows only the crudest forms of control. Consequently, the reaction must be *catalyzed* or "sped up." Enzymes, denoted E, are biochemical catalysts that accelerate the conversion of S into P, and the chemical process by which the conversion of a substrate molecule into a product molecule is accelerated by an enzyme is called an *enzymatic reaction*.

The simplest description of an enzymatic reaction for a single-substrate, single-product reaction is the Michaelis–Menten mechanism [23, 29, 50],

$$S + E \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} C \overset{k_2}{\longrightarrow} E + P, \tag{1.2}$$

in which the conversion of S into P is achieved via two elementary reactions: the reversible formation of the enzyme-substrate complex, C, and the conversion of S to P in the complex C with (in this simple model) simultaneous dissociation into E and P. Enzymes lower the free-energy barrier separating reactants from products, with the result that (1.2) is generally faster than (1.1) by many orders of magnitude [30, Section 6.2].

The modeling and quantification of enzymatic reaction rates is of particular importance, especially since metabolic disease and dysfunction may arise when these reactions are too slow due, e.g., to a mutation in the corresponding gene. At or near the thermodynamic limit, enzymatic reactions are modeled by nonlinear ordinary differential equations (ODEs), known as rate equations, that obey the law of mass action. While the nonlinear terms in the model equations of enzymatic reactions make the mathematical treatment of the reaction mechanism challenging, avenues for simplification often exist. Specifically, if the rates of the elementary reactions that comprise the catalytic reaction are disproportionate, the ODE model will be *multiscale*, meaning the complete reaction will consist of disparate slow and fast time scales. Under the influence of distinct fast and slow time scales, the rate of change of c (using lower-case italic letters to represent the concentrations of the corresponding species) is very small relative to the rate of rate of change of s. The exploitation of this almost negligible rate of change warrants a simplification of the form

$$S \overset{k_{\text{eff}}}{\longrightarrow} P, \tag{1.3}$$

where k_{eff} is the *effective*—but non-elementary—rate function. In the case of the Michaelis–Menten mechanism, k_{eff} is a hyperbola in the variable s; in more complicated mechanisms it may adopt the form of, for instance, a Hill-type function. The advantage offered by (1.3) is that the entire reaction is describable in terms of the reactant concentration, s, since the explicit dependence on e and c has been eliminated. The most widely studied example of this kind of reduction is probably the Michaelis–Menten rate law, which can be obtained using the standard quasi-steady-state approximation (sQSSA). More generally, rate laws of the form (1.3) are referred to as quasi-steady-state (QSS) reductions or quasi-steady-state approximations (QSSA). The term QSS speaks to the fact that the concentration of at least one chemical species (typically an intermediate) changes very slowly for the majority of the reaction.* In fact, the rate of change is so small that it is *nearly zero* (steady-state) but not quite; hence the expression *quasi*-steady-state.

---

*In singular perturbation settings, these concentrations frequently correspond to fast variables that quickly equilibrate, and subsequently have a small net rate of change. Note the correspondence to the multiscale concept mentioned earlier.

The principal value of QSS approximations is that they yield a reduction of dimension [10]. In the biochemical arena, the related equilibrium approximation was initially justified via biochemical arguments by Henri [23] and by Michaelis and Menten [29]. Briggs and Haldane [4] later provided a mathematical justification of the sQSSA using an argument that hints at later singular perturbation treatments but lacked formal justification. Only the development of singular perturbation theory some decades later (with seminal contributions by Tikhonov [46], and later Fenichel [9]) laid a solid mathematical foundation, which was used by Heineken et al. [22] to develop criteria for the validity of the sQSSA for the closed Michaelis–Menten system. This history was paralleled in inorganic chemistry, with the initial development of the sQSSA based on ad hoc chemical reasoning [2, 6], followed eventually by more rigorous treatments based on singular perturbation theory [3].

Singular perturbation theory in this context applies to ODEs that depend on a small nonnegative parameter ε, and admit non-isolated stationary points at $\varepsilon = 0$. In practice, e.g. for systems with polynomial or rational right-hand side, the set of stationary points then contains a submanifold of positive dimension, which is called a critical manifold. Given appropriate conditions (see Appendix A for details), one obtains a reduction to a system of smaller dimension constrained to evolve on the critical manifold. The challenge in any application of Fenichel theory resides in finding a small parameter from a given parameter dependent system. Traditional analyses of enzymatic reactions rely heavily on scaling and non-dimensionalization in order to transform the model equations into a standard form, and the utility of scaling analysis is that the small parameter often emerges naturally from the dimensionless equations [41]. A different, more recent approach [16] starts with determining so-called Tikhonov–Fenichel parameter values (TFPV), by searching for parameter combinations at which the system admits non-isolated stationary points, and satisfies further technical conditions (see Section 3.1). From such (dimensional) TFPV one then obtains singular perturbation reductions via small perturbations along a curve in parameter space. In chemical applications, critical manifolds frequently emerge when specific system parameters (such as rate constants) vanish.

While singular perturbation theory provides a very satisfactory toolbox for reduction of chemical reaction networks, examples from the literature indicate that the approach may be too narrow for some applications. Thus in some scenarios, at a certain parameter value there exists a distinguished invariant manifold which is, however, not comprised of stationary points. Formally, this means that a QSS reduction which approximates the system when $0 < \varepsilon \ll 1$ is not attributable to Fenichel theory. Nevertheless the QSS reduction is still sometimes a good approximation to the full system when Fenichel theory is inapplicable, and this raises several important questions. First, given the lack of a critical manifold and a fixed reduction procedure, how does one justify a QSS reduction, and how does one go about quantifying its efficacy? Second, if Fenichel theory is not applicable but a QSS reduction still proves to be an accurate approximation, will there be a distinguished invariant manifold that attracts nearby trajectories? In other words, what phase-space structures make the QSS reduction possible in situations where Fenichel theory is extraneous? In the present paper we contribute, on the one hand, to answering these questions for an open Michaelis–Menten system with constant substrate influx. On the other hand, we provide sharper estimates for the accuracy of the sQSSA in singular perturbation scenarios. Finally, we consider distinguished invariant manifolds from a global perspective for the system on the Poincaré sphere.

2. An open Michaelis–Menten reaction mechanism

The open Michaelis–Menten reaction mechanism we consider here is the classical Michaelis–Menten reaction mechanism with a constant influx of substrate, S, at a rate k_0:

$$\emptyset \xrightarrow{k_0} S, \qquad S + E \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} C \xrightarrow{k_2} E + P, \tag{2.1}$$
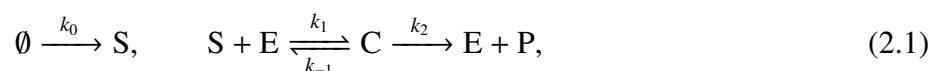
where k_0, k_1, k_{-1} and k_2 are rate constants.

Mathematical models for (2.1) come in both deterministic and stochastic forms. Here we consider only the deterministic ODE model that follows the law of mass action near the thermodynamic limit. For a thorough analysis of the chemical master equation corresponding to (2.1), we invite the reader to consult [1, 45].

The mass action model corresponding to (2.1) is given by the following set of nonlinear ODEs:

$$\dot{s} = k_0 - k_1 es + k_{-1}c, \tag{2.2a}$$
$$\dot{c} = k_1 es - (k_{-1} + k_2)c, \tag{2.2b}$$
$$\dot{e} = -k_1 es + (k_{-1} + k_2)c, \tag{2.2c}$$
$$\dot{p} = k_2 c, \tag{2.2d}$$

where a dot denotes differentiation with respect to time. Summing equations (2.2b) and (2.2c) reveals the conservation law

$$c + e = e_T, \tag{2.3}$$

where e_T denotes the total enzyme concentration. Employing (2.3) to eliminate (2.2c), and noting that (2.2d) is not coupled to (2.2a) or (2.2b), yields the simplified model

$$\begin{aligned} \dot{s} &= k_0 - k_1(e_T - c)s + k_{-1}c, \\ \dot{c} &= k_1(e_T - c)s - (k_{-1} + k_2)c, \end{aligned} \tag{2.4}$$

from which the time dependence of p and e are readily obtained from (2.2d) and (2.3) once the solution to (2.4) is procured.

In contrast, the mass-action system for the *closed* Michaelis–Menten reaction mechanism is recovered by setting $k_0 = 0$:

$$\dot{s} = -k_1(e_T - c)s + k_{-1}c, \tag{2.5a}$$
$$\dot{c} = k_1(e_T - c)s - (k_{-1} + k_2)c. \tag{2.5b}$$

One distinguishing difference between the open and closed systems is that the total substrate concentration, s_T, is a conserved quantity when the reaction is closed. Therefore, (2.2) with $k_0 = 0$ is equipped with the additional conservation law $s_T = s + c + p$, whereas with $k_0 > 0$ one has only one conservation law, (2.3).

It is well known that further simplification of (2.5) is possible via a QSS reduction. The most common reduction is the sQSSA, in which (2.5) is approximated with a differential-algebraic equation consisting of the algebraic equation obtained by setting the right-hand side of equation (2.5b) equal

to zero ("$\dot{c} = 0$") along with the differential equation (2.5a). This reduces to the single differential equation

$$\dot{s} = -\frac{k_2 e_T s}{K_M + s}, \quad K_M := \frac{k_{-1} + k_2}{k_1}, \tag{2.6a}$$

$$c = \frac{e_T s}{K_M + s}, \tag{2.6b}$$

where K_M is the Michaelis constant.

The legitimacy of the sQSSA (2.6) for the closed Michaelis–Menten reaction mechanism (2.5) is well-understood. Following an early effort by Briggs and Haldane [4], Heineken, Tsuchiya, and Aris [22] were perhaps the first to prove with some degree of rigor that (2.6) is valid provided $e_T \ll s_0$. The qualifier, $e_T \ll s_0$, was justified via singular perturbation analysis. Defining $\bar{s} := s/s_0$, $\bar{c} := c/e_T$, and $T := k_1 e_T t$ generates the singularly perturbed dimensionless form of (2.5)

$$\bar{s}' = -\bar{s} + \bar{c}(\bar{s} + \kappa - \lambda), \tag{2.7a}$$

$$\mu \bar{c}' = \bar{s} - \bar{c}(\bar{s} + \kappa), \tag{2.7b}$$

where prime denotes differentiation with respect to T, $\lambda := k_2/k_1 s_0$, $\kappa := K_M/s_0$, and $\mu := e_T/s_0$. Consequently, the sQSSA (2.6) is justified via Tikhonov's theorem [46]. Throughout the years, refinements and variations of the condition $\mu \ll 1$ have been made. Perhaps most famously, Segel [42] and Segel and Slemrod [41] extended the results of Heineken et al. [22] and demonstrated that (2.6) is valid whenever $e_T \ll K_M + s_0$. Embedded in Segel's estimate is the more restrictive condition, $e_T \ll K_M$, which is independent of the initial substrate concentration, and is nowadays the almost universally accepted qualifier that justifies (2.6) [7].

While the QSS reductions of the closed Michaelis–Menten reaction are well-studied, analyses pertaining to the validity of the QSSA in open reaction environments are somewhat sparse [1, 18, 44, 45]. The question we address is therefore: when is further reduction of (2.4) possible? The trajectories illustrated in Figure 1 show that there are certainly conditions under which the QSSA estimate of the enzyme-substrate complex, given by Eq (2.6b), which applies equally to the open system, is close to an invariant manifold (here, a trajectory) that attracts nearby trajectories and along which the equilibrium point is eventually approached from almost all initial conditions [11, 20, 39]. We thus ask under what condition is the open sQSSA

$$\dot{s} = k_0 - \frac{k_2 e_T s}{K_M + s}, \tag{2.8}$$

permissible? At first glance, it seems rather intuitive to postulate that the open sQSSA (2.8) is valid under the same condition that legitimizes the closed sQSSA: $e_T \ll K_M$. In fact, following the earlier work of Segel and Slemrod [41], Stoleriu et al. [44] introduce the parameter $\alpha := k_0/(k_2 e_T)$, and suggest that (2.6b) and (2.8) are applicable whenever

$$e_T(1 - \alpha) \ll s_0 + \frac{K_M}{1 - \alpha} \tag{2.9}$$

holds. The inequality (2.9) is less restrictive than the Segel and Slemrod condition, since (2.9) is satisfied as long as k_0 is sufficiently close to $k_2 e_T$ [Implicitly, Stoleriu et al. assume that $\alpha < 1$ in Eq (2.9).]
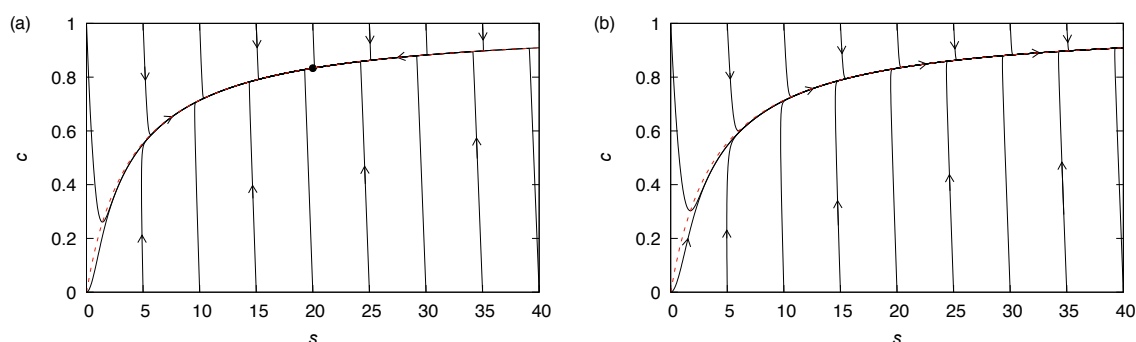
Figure 1. Trajectories of the open Michaelis–Menten equations (2.4) for (a) $k_1 = 1$, $e_T = 1$, $k_{-1} = 1$ $k_2 = 3$ and $k_0 = 2.5$ (in arbitrary units), i.e. under conditions where there is an equilibrium point in the first quadrant, marked by a dot; and (b) with parameters as in (a), except $k_0 = 3.5$, under which conditions there is not an equilibrium point in the first quadrant, and the s component of the solution grows without bound. The arrows show the direction of the flow. The dashed curve in both figures is defined by the QSSA equation (2.6b).

The approach used to derive (2.9) was based on the traditional method of comparing time scales: a singular perturbation parameter was recovered through scaling analysis of the mass action equations (2.4). However, it is possible to derive erroneous conclusions regarding the validity of the QSSA, even when great care is taken in scaling and non-dimensionalization methodology (see, for example [13], Section 4). It thus seems prudent to reexamine the basis for the sQSSA in the open Michaelis–Menten mechanism using tools of singular perturbation theory that go beyond scaling arguments.

3. The Quasi-Steady-State Approximation: Justification from singular perturbation theory

In this section we derive the QSSA directly from Fenichel theory. Details covering projection onto the critical manifold can be found in Appendix A.

3.1. The critical manifolds: Tikhonov–Fenichel parameter values

To apply Fenichel theory to the open Michaelis–Menten reaction mechanism, we need a curve of non-isolated equilibrium solutions to form in the first quadrant of $\mathbb{R}^2$; see [16]. The following Lemma addresses the conditions that ensure the existence of a critical manifold, and records some general qualitative features.

Lemma 1. *(a) System* (2.4) *admits an infinite number of stationary points if and only if one of the following conditions holds.*

- $k_0 = k_1 = 0$;
- $k_0 = e_T = 0$;
- $k_0 = k_2 = 0$.

(b) If the number of stationary points in the plane is finite then it is equal to zero or one. There exists one stationary point if and only if the genericity conditions

$$k_1 \neq 0, \quad k_2 \neq 0 \text{ and } k_2 e_T - k_0 \neq 0 \tag{3.1}$$

are satisfied. In that case the stationary point is equal to

$$P_0 := (\widehat{s}, \widehat{c}) = \left(\frac{(k_{-1} + k_2)k_0}{k_1(k_2 e_T - k_0)}, \frac{k_0}{k_2} \right). \tag{3.2}$$

This point lies in the first quadrant if and only if

$$k_2 e_T - k_0 > 0, \tag{3.3}$$

in which case it is an attracting node. The stationary point lies in the second quadrant if and only if $k_2 e_T - k_0 < 0$, in which case it is a saddle point.

(c) *The first quadrant is positively invariant for system (2.4), and solutions starting in the first quadrant exist for all $t \geq 0$. When $k_{-1} + k_2 > 0$ then every solution that starts in the first quadrant enters the (positively invariant) subset defined by $c \leq e_T$ at some positive time.*

(d) *System (2.4) admits no non-constant closed trajectory.*

Sketch of proof. Parts (a) and (b) are straightforward, as is the first statement in part (c). For the second statement note $\dot{s} + \dot{c} \leq k_0$, hence solutions starting in the first quadrant remain in a compact set for all finite $t > 0$. Finally, when $c \geq e_T$ then (2.4) shows that $\dot{c} \leq -(k_{-1} + k_2)e_T$, hence the second statement of part (c) holds. We turn to the proof of part (d): If there exists a non-constant closed trajectory then its interior contains a stationary point. Given a degenerate situation from part (a), the variety of stationary points is unbounded, hence would intersect a closed trajectory if it intersects its interior; a contradiction. This leaves the setting with an isolated stationary point, necessarily of index one, which is only possible when the stationary point (3.2) lies in the first quadrant. By part (c) the closed trajectory must be contained in the strip defined by $c \leq e_T$. But in this strip the divergence of the vector field equals $-(k_1(e_T - c) + k_1 s + k_{-1} + k_2) < 0$, and no closed non-constant trajectory can exist by Bendixson's criterion. $\qquad \square$

Remark 1. The case $k_0 > k_2 e_T$, in which the inflow exceeds the enzyme's clearance capacity, is not physiologically irrelevant since the gene coding for a particular enzyme may suffer a mutation that results in an enzyme with reduced catalytic activity, for example. As a rule, the accumulation of a metabolite will eventually become toxic (or possibly oncogenic) to the cell, and the rate at which S accumulates is therefore of interest. Other situations, e.g. the existence of an alternative but less efficient pathway for eliminating S, or the permeation of S through the cell membrane, would require more elaborate models for their study. Nevertheless, the model under study here would yield useful initial insights into the cellular effects of a mutation to an enzyme.

Lemma 1 ensures the existence of a critical manifold comprised of equilibrium points whenever k_0 vanishes along with either e_T, k_1 or k_2. We note that in the context of the closed reaction (2.5), parameters with $e_T = 0$ (with all remaining parameters > 0), resp. $k_1 = 0$ (remaining parameters > 0), resp. $k_2 = 0$ (remaining parameters > 0), are TFPV. Generally, a TFPV $[\widehat{k_0} \ \widehat{e_T} \ \widehat{k_1} \ \widehat{k_2} \ \widehat{k_{-1}}]$ is characterized by the property that the variety of stationary points has positive dimension, and a generic small perturbation of the parameters results in the formation of a normally hyperbolic invariant manifold, called a slow manifold [17]. Choosing a curve in parameter space which is parameterized by ε and starts at a TFPV, one obtains a system which admits a singular perturbation reduction.

Let $\pi \in \mathbb{R}_+^5$ denote the parameter vector: $\pi := [k_0 \ \ e_T \ \ k_1 \ \ k_2 \ \ k_{-1}]^T$. By Lemma 1, and upon checking normal hyperbolicity (see Appendix A below), the TFPVs and the critical manifolds, M, are as follows:

$$\pi = [0 \ 0 \ k_1 \ k_2 \ k_{-1}] \implies M := \{(s, c) \in \mathbb{R}^2 : c = 0\}, \tag{3.4a}$$

$$\pi = [0 \ e_T \ 0 \ k_2 \ k_{-1}] \implies M := \{(s, c) \in \mathbb{R}^2 : c = 0\}, \tag{3.4b}$$

$$\pi = [0 \ e_T \ k_1 \ 0 \ k_{-1}] \implies M := \{(s, c) \in \mathbb{R}^2 : c = k_1 e_T s / (k_{-1} + k_1 s)\}. \tag{3.4c}$$

Fenichel theory ensures that perturbing π in (3.4) along a curve in parameter space through the TFPV results in the formation of an invariant slow manifold that attracts nearby trajectories at an exponential rate. Formally, the QSSA may be seen as an approximation of the dynamics on the slow manifold, perturbing from a TFPV.

3.2. Singular perturbations and the geometry of parameter space

The justification of the QSSA from singular perturbation theory requires us to implicitly equip parameter space with some additional geometric structure. For example, consider the case where both e_T and k_0 vanish in the singular limit. In order to formally apply singular perturbation theory, it must hold that $k_0 \sim O(e_T)$ (in a sense discussed below). Generally speaking, this means that we can apply singular perturbation theory along a parametric curve, Γ, in (e_T, k_0) parameter space, $\Gamma := (e_T, z(e_T))$, provided $z(0) = 0$ and

$$\lim_{e_T \to 0^+} \frac{z(e_T)}{e_T} < \infty. \tag{3.5}$$

However, a *small* perturbation suggests that the parameter values will be *close* to the parameter plane origin located at $(e_T, k_0) = (0, 0)$. In this case $z(e_T)$ is well-approximated by its tangent line at $e_T = 0$, thus it is enough to only consider rays of the form $k_0 = \gamma e_T$, where γ is a positive constant with dimension t^{-1}. To eliminate the need for a dimensional slope γ, define a ray in parameter space by

$$e_T \mapsto \varepsilon e_T^* \quad \text{and} \quad k_0 \mapsto \varepsilon k_0^*, \tag{3.6}$$

where the parameters k_0^* and e_T^* are nominal values of k_0 and e_T, respectively.

The additional constraint of sampling parameter space along a ray [or in a more general way along a curve satisfying (3.5)] must be imposed in order to justify the open sQSSA from singular perturbation theory. In their analysis of the open Michaelis–Menten reaction (2.4), Stoleriu et al. [44] implicitly performed their analysis along a ray defined by

$$e_T = e(0) + k_0 / k_2, \tag{3.7}$$

where $e(0) > 0$ is the initial free enzyme concentration. This ray in parameter space is encoded in their initial conditions, which allow for an arbitrary positive value of $e(0)$, but which specify $c(0) = \widehat{c} = k_0/k_2$. The advantage of working along the ray defined by (3.7) is that there is no possibility that the inflow can exceed the clearance capacity of the enzyme, i.e. inequality (3.3) is automatically satisfied.

In order to apply singular perturbation theory, we need to start from a critical manifold, i.e. from one of the cases in the set (3.4). Note that the ray through the (e_T, k_0) parameter plane chosen by Stoleriu et

al. [44], Equation (3.7), does not satisfy (3.5) unless $e(0) = 0$. This leads to difficulties. For example, the condition (2.9) *along* the ray defined by (3.7) translates to

$$k_1 e(0) \ll k_1 s_0 + (k_{-1} + k_2)\left(\frac{1}{1-\alpha}\right). \tag{3.8}$$

The inequality (3.8) is satisfied by taking $k_1 \to 0$, but this limit alone does not produce a critical manifold. Hence, the singular perturbation machinery is not applicable to legitimizing the open sQSSA (2.8) by this route.

Another issue with the constrained set of initial conditions imposed by (3.7) is that it excludes many initial conditions that are physiologically relevant. For example, a natural initial condition is $(s, e, c, p)(0) = (0, e_T, 0, 0)$, corresponding to the substrate flow being turned on at time zero (e.g. because the cell is placed in a new environment, or because it has turned on a previously dormant metabolic pathway that produces S), but this initial point is *inaccessible* if the parametric constraint (3.7) has been imposed. Consequently, it remains an open question whether the results of the analysis apply at arbitrary points in parameter space and for arbitrary initial conditions. In particular, there is no guarantee that the analysis of Stoleriu et al. [44] applies when the inflow exceeds the clearance capacity of the enzyme which, as argued previously, is not an irrelevant case. By contrast, a transformation informed by the basic requirements of singular perturbation theory such as (3.6) allows us to make rigorous statements about the manifold structure of the problem, and imposes no constraints on the initial conditions.

3.3. Quasi-steady-state reductions: Projecting onto the critical manifold

Let us now consider the first scenario in which e_T and k_0 vanish in the singular limit. The perturbation of the singular vector field is

$$\begin{aligned}
\dot{s} &= \varepsilon k_0^* - k_1(\varepsilon e_T^* - c)s + k_{-1}c, \\
\dot{c} &= k_1(\varepsilon e_T^* - c)s - (k_{-1} + k_2)c.
\end{aligned} \tag{3.9}$$

The singular limit obtained by setting $\varepsilon = 0$ in (3.9) yields a critical manifold, M, that is identically the s axis:

$$M := \{(s, c) \in \mathbb{R}^2 : c = 0\}. \tag{3.10}$$

To compute the corresponding singular perturbation reduction (see Appendix A for specific details), we rewrite the right hand side of (3.9) as $h(s, c) + \varepsilon G(s, c, \varepsilon)$. Furthermore $h(s, c) = P(s, c)f(s, c)$, with

$$P(s, c) := \begin{bmatrix} k_1 s + k_{-1} \\ -k_1 s - (k_{-1} + k_2) \end{bmatrix}, \quad f(s, c) := c, \quad G(s, c, \varepsilon) := \begin{bmatrix} k_0 - k_1 e_T s \\ k_1 e_T s \end{bmatrix}. \tag{3.11}$$

Since $DfP = -(k_1 s + k_{-1} + k_2)$ is negative on M, M satisfies the attracting hyperbolicity condition, and Tikhonov-Fenichel reduction is applicable (see [15] and also Appendix A). The singular perturbation reduction is then obtained by projecting $G(s, c, 0)$ onto the tangent space of M at $x = (s, c)$, $T_x M$, via the linear operator Π^M which projects onto the kernel along the image N_x of the Jacobian $Dh(x)$.

$$\Pi^M|_{c=0} G(s, 0, 0). \tag{3.12}$$

Note that N_x, which is equal to the range of $P(x)$, is a complementary subspace to T_xM. For our specific problem (3.9), Π^M is given by

$$\Pi^M := \begin{bmatrix} 1 & u(s) \\ 0 & 0 \end{bmatrix}, \quad u(s) := \frac{(s + K_S)}{(s + K_M)}, \quad K_S := k_{-1}/k_1, \tag{3.13}$$

and the corresponding reduction, which agrees with the QSS reduction, is

$$\dot{s} = k_0 - \frac{k_2 e_T s}{K_M + s}. \tag{3.14}$$

Equation (3.14) is, of course, the open sQSSA. A similar calculation is easily carried out for the case of small k_1 and small k_0, as well as small k_0 and k_2, and we refer the reader to Appendix A for details. The specific QSS reduction that accompanies the perturbation defined by $k_1 \mapsto \varepsilon k_1^*$ and $k_0 \mapsto \varepsilon k_0^*$ is

$$\dot{s} = k_0 - \frac{k_2 e_T}{K_M} s, \tag{3.15}$$

which is the linear limiting law obtained in the small-s limit of (3.14).

Accordingly, we have confirmation that the open sQSSA (2.8) is valid under any condition that invokes a scaling of the form $k_0 \mapsto \varepsilon k_0^*$ and $e_T \mapsto \varepsilon e_T^*$. We further note that a QSS reduction based on Fenichel theory is also possible in case $k_0 \mapsto \varepsilon k_0^*$ and $k_2 \mapsto \varepsilon k_2^*$ so that both k_0 and k_2 vanish in the singular limit. This reduction yields the classical equilibrium approximation, (See Section 5.3 and Appendix A for details.)

Several questions remain. First, what is ε? We have shown that the open sQSSA is valid provided k_0 and e_T are sufficiently small, but *what is small* when k_0 and e_T are nonzero? Second, from the work of Goeke et al. [17], the QSS may still hold in certain regions of the phase-plane even if Fenichel theory is not applicable. The analysis of Stoleriu et al. [44] is also indirectly suggestive of the idea that the validity of the open sQSSA may not necessarily stem from singular perturbation theory. These observations raise the deeper question: is a scaling of the form $k_0 \mapsto \varepsilon k_0^*, e_T \mapsto \varepsilon e_T^*$ necessary for the validity of the QSSA, or merely sufficient? We address these questions directly in the sections that follow.

4. Revisiting quasi-steady state for the complex species

4.1. The notion of QSS

Singular perturbation theory provides a natural setting for developing conditions under which QSSA holds, but the literature (notably Stoleriu et al. [44] for the open Michaelis–Menten mechanism) suggests that one should consider less restrictive notions as well. In the following we will sketch one such notion. This goes back to Schauer and Heinrich [40], who were the first to note that the minimal requirement for the validity of QSS reduction for some species should be the *near-invariance* of a corresponding variety, which we call the QSS variety. This variety is defined as the zero set of the rate of change for the species under consideration. The idea of near-invariance was expounded upon by Noethen et al. [32], and further analyzed by Goeke et al. [17]:

- A fundamental feature of QSS is that the rate of change of certain sets of species should be close to zero for an extended period of time. (In the Michaelis–Menten reaction, QSS for complex thus means that $\dot{c} \approx 0$ for an extended period of time.) In the phase space interpretation, a sizable part of the trajectory should thus be close to the *QSS variety* which is defined by evaluating the condition $\dot{c} = 0$. (In the Michaelis–Menten mechanism the QSS variety for complex is thus defined by $k_1(e_T - c)s - (k_{-1} + k_2)c = 0$; see Eq (2.6b).) The validity of such a condition will depend on the parameters.

- According to [17], Section 3.3, the minimal requirement for QSS should therefore be near-invariance of the QSS variety, in the sense that the system parameters are small perturbations of QSS parameter values. By definition, at a QSS parameter value the QSS variety is an invariant set for system (2.4). (In the Michaelis–Menten mechanism one thus requires invariance of the variety defined by Eq (2.6b) for system (2.4) at a QSS parameter value.) The arguments in [17] show that this condition is necessary if one requires arbitrary accuracy of the QSS approximation for suitable parameters. By standard continuous dependence theorems for initial values and parameters (see e.g. Perko [34], Section 2.3), small perturbations of a QSS parameter value yield trajectories that remain close to the QSS variety on compact time intervals; thus the condition is also sufficient. One practical advantage of this notion is that QSS parameter values, similarly to TFPV, are algorithmically accessible for polynomial or rational systems.

- The near-invariance condition alone may not be sufficiently strong to satisfy expectations about QSS. One may also require that solutions quickly approach the QSS variety in an initial transient phase. Since the combination of these two features is automatically satisfied in singular perturbation settings, singular perturbations naturally enter the picture. But the singular perturbation scenario is both broader and narrower than QSS for chemical species: It is broader since it also is applicable to settings with slow and fast reactions. On the other hand, we will see below that it is, in a sense, too narrow for sQSS in the open Michaelis–Menten reaction mechanism.

4.2. Open Michaelis–Menten: QSS parameter values for complex

The QSS variety for (2.4) is given by

$$c = w(s) := \frac{k_1 e_T s}{k_1 s + k_{-1} + k_2}. \tag{4.1}$$

We prefer this to the usual notation $w(s) = e_T s / (K_M + s)$, which may obscure the role of k_1. We first determine all QSS parameter values.

Lemma 2. *The QSS parameters of system* (2.4) *are as follows:*

(i) $e_T = 0$ *with the other parameters arbitrary;*

(ii) $k_1 = 0$ *with the other parameters arbitrary;*

(iii) $k_0 = k_2 = 0$*;*

(iv) $k_{-1} = k_2 = 0$.

Proof. We proceed along the lines of [17], Section 3.4, using an invariance criterion that employs the Lie derivative, $L[\cdot]$, corresponding to (2.4). The Lie derivative is defined by

$$
\begin{aligned}
L[\varphi](s,c) &= \dot{s}\frac{\partial \varphi}{\partial s} + \dot{c}\frac{\partial \varphi}{\partial c} \\
&= (k_0 - k_1(e_T - c)s + k_{-1}c)\frac{\partial \varphi}{\partial s} + (k_1(e_T - c)s - (k_{-1} + k_2)c)\frac{\partial \varphi}{\partial c}
\end{aligned}
$$

for any polynomial (more generally, smooth) function φ. For the variety defined by $\varphi = 0$ to be invariant it is necessary that

$$
L[\varphi](s,c) = 0 \text{ whenever } \varphi(s,c) = 0.
$$

Moreover, the condition is sufficient when φ is irreducible, and it is applicable to the irreducible factors of φ; for details see [17] and the references therein.

Now let $\psi(s,c) = 0$ define the QSS manifold, thus

$$
\psi(s,c) := k_1(e_T - c)s - (k_{-1} + k_2)c. \tag{4.2}
$$

The invariance condition for the curve $\psi(s,c) = 0$ is

$$
L[\psi](s,c) = -k_1(e_T - c)(-\psi(s,c) + k_0 - k_2c) - (k_1s + k_{-1} + k_2)\psi(s,c) = 0 \tag{4.3}
$$

whenever $\psi(s,c) = 0$, thus

$$
k_1(e_T - c)(k_0 - k_2c) = 0 \text{ whenever } \psi(s,c) = 0. \tag{4.4}
$$

This product yields three conditions which can be evaluated. Clearly $k_1 = 0$ works and yields (ii). The second condition, $e_T - c = 0$, holds on $\psi = 0$ if and only if $(k_{-1} + k_2)e_T = 0$, which yields respectively (iv) and (i). The third condition yields $k_0 = k_2 = 0$ when $k_2 = 0$, i.e. (iii). In case $k_2 \neq 0$ one obtains $c = k_0/k_2$, and

$$
k_1(e_T - k_0/k_2)s - (k_{-1} + k_2)k_0/k_2 = 0 \text{ for all } s;
$$

here the coefficient of s and the constant must vanish. This again leads to conditions already discussed. □

Remark 2. (a) In cases (i) and (ii), the QSS variety is given by $c = 0$, provided that the other parameters are positive, and the QSS parameter conditions are less restrictive than for singular perturbations, which also require $k_0 = 0$. This is a notable difference to the closed Michaelis–Menten scenario, for which all complex-QSS parameter values are also TFPV. Case (iii) corresponds to a singular perturbation scenario. The dynamics in case (iv) is of some interest in the Michaelis–Menten reaction mechanism without inflow; see [7].

(b) Classical QSS reduction is tantamount to exploiting the fact that if $\psi(s,c) = 0$ defines a nearly invariant curve, then $c \approx w(s)$, from which the open sQSSA (2.8) presumably follows. However, a word of caution is in order. When a QSS parameter value is also consistent with a singular perturbation and gives rise to a critical manifold, the classical QSS reduction may differ from the

reduction obtained from Fenichel theory (see [17], Section 3.5). For example, $\psi(s, c) = 0$ is nearly invariant if k_0 and k_2 are small, but the classical QSS reduction, given by

$$\dot{s} = k_0 - \frac{k_2 e_T s}{k_1 s + k_{-1}}, \tag{4.5}$$

does not agree with the reduction obtained from singular perturbation theory, which is given by (A.18). Convergence to the singular perturbation reduction is guaranteed by Fenichel theory, hence the QSS reduction (4.5) cannot correctly describe the dynamics at lowest order.

For the QSS parameters which do not correspond to singular perturbations, there remains to investigate whether solutions approach this variety, and if so, how fast and how close the approach is. Furthermore, even in the singular perturbation scenario one needs estimates on the initial (boundary layer) behavior, since Fenichel's theory applies directly only to a neighborhood of the critical variety.

These problems will be addressed via direct estimates, which will also be of help in answering a quantitative question, i.e. how small should e_T be in order to justify (2.8)? Ultimately, the term *small* is relative in nature. Therefore, the appropriate question to ask is: For (2.8) to be approximately accurate, e_T and k_0 must be much smaller than *what*? Before we start this investigation we establish an auxiliary result about the phase plane geometry of (2.4).

4.3. Phase plane arguments

From here on we restrict attention to system (2.4) on the positively invariant strip W defined by $s \geq 0$ and $0 \leq c \leq e_T$. A priori we impose no requirements on the parameters. We look at isoclines, noting that

$$\dot{c} = 0 \Leftrightarrow c = \mathcal{N}_c(s) \equiv \frac{k_1 e_T s}{k_1 s + k_{-1} + k_2}, \qquad \dot{c} \geq 0 \Leftrightarrow c \leq \mathcal{N}_c(s); \tag{4.6}$$

and

$$\dot{s} = 0 \Leftrightarrow c = \mathcal{N}_s(s) \equiv \frac{k_1 e_T s - k_0}{k_1 s + k_{-1}}, \qquad \dot{s} \geq 0 \Leftrightarrow c \geq \mathcal{N}_s(s), \tag{4.7}$$

where $\mathcal{N}_x$ denotes the x nullcline. These nullclines define positively invariant sets:

Lemma 3. *Consider the "wedge"*

$$W_1 := \max\left\{0, \frac{k_1 e_T s - k_0}{k_1 s + k_{-1}}\right\} \leq c \leq \frac{k_1 e_T s}{k_1 s + k_{-1} + k_2}, \quad s \geq 0.$$

Then the following hold:

(a) *If the system admits no positive stationary point, thus $k_0 > k_2 e_T$, then the c-isocline lies above the s-isocline for all $s \geq 0$, and W_1 extends to $s \to \infty$. If the system admits the positive stationary point $(\widehat{s}, \widehat{c})$ then the isoclines meet at this point, and $s \leq \widehat{s}$, $c \leq \widehat{c}$ for all points of W_1.*

(b) *W_1 is positively invariant for system (2.4), and on W_1 one has $\dot{s} \geq 0$.*

Proof. Part (a) is straightforward. As for part (b), from (2.4) one sees that $\dot{s}+\dot{c} = k_0 - k_2 c \geq k_0 - k_2 \widehat{c} = 0$ on W_1, thus

$$\dot{s} = 0 \Rightarrow \dot{c} \geq 0, \quad \dot{c} = 0 \Rightarrow \dot{s} \geq 0.$$

This implies the positive invariance of W_1, since the vector field points to the interior of W_1 at the boundary (Figure 2). Clearly $\dot{s} \geq 0$ on W_1. □
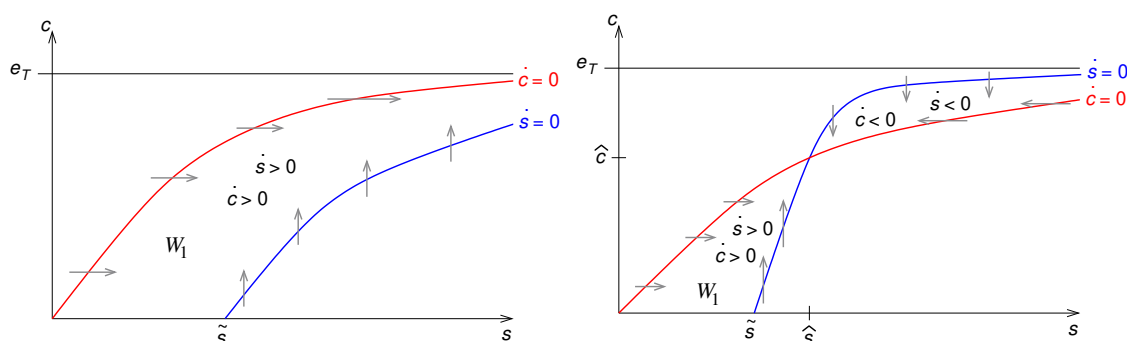
Figure 2. Sketches of the positively invariant sets W_1 in the phase plane for the open Michaelis–Menten reaction mechanism (2.1). The curves are the nullclines, and the arrows show the direction of motion of trajectories as they cross the nullclines. Both nullclines tend asymptotically to $c = e_T$ as $s \to \infty$. $\tilde{s}$ is the s intercept of the s nullcline. Left: $k_0 > k_2 e_T$ and the two nullclines never meet. Right: $k_2 e_T > k_0$ and the nullclines cross at the stationary point $(\widehat{s}, \widehat{c})$. The flow points into the region delimited by the two nullclines, making this region a funnel [24].

Remark 3. Smallness of e_T and existence of a positive stationary point imply smallness of k_0; this leads automatically to the singular perturbation setting. Matters are different when k_1 is small.

Everywhere inside the wedge, $\dot{c} > 0$ and $\dot{s} > 0$. Thus, all trajectories inside the wedge have positive slope. Since the flow points into the wedge, the slow manifold must also lie inside the wedge. Thus, the slow manifold has a positive slope for $s \leq \widehat{s}$ in the first quadrant. Moreover, the slow manifold must enter the first quadrant by crossing through the s axis in the interval $(0, \tilde{s})$, where $\tilde{s} = k_0/k_1 e_T$ is the s intercept of the s nullcline (Figure 2).

In the case that there is a positive equilibrium point, for $s > \widehat{s}$, $\dot{c} < 0$ and $\dot{s} < 0$ between the two nullclines so that trajectories in this region still have positive slope. The flow is, again, into the region between the two nullclines (Figure 2), so the slow manifold must lie within this region. The slow manifold therefore has positive slope here as well. Moreover, $\lim_{s \to \infty} \mathcal{N}_c(s) = \lim_{s \to \infty} \mathcal{N}_s(s) = e_T$. Thus, the two nullclines pinch together asymptotically. Although we do not pursue this idea here, this property would allow the anti-funnel theorem to be used to prove the existence of a unique slow manifold to the right of the equilibrium point [5, 24]. (See Section 6 for correspondence to the global behavior.)

4.4. *How small is small: A direct estimate*

Given that we are interested in obtaining a condition that ensures phase plane trajectories closely follow the QSS variety corresponding to the c nullcline, we compute an upper bound on the limit

supremum (lim sup) of

$$L := |c - w(s)|$$

for a solution of (2.4), where $w(s)$ is given by (4.1). To determine such an upper bound, we calculate

$$\frac{1}{2}\frac{d}{dt}L^2 = (c - w(s))(\dot{c} - w'(s)\dot{s}). \tag{4.8}$$

The derivative $\dot{c}$ given in (2.4) factors:

$$\dot{c} = -k_1(s + K_M)(c - w(s)) =: -\tau(s)(c - w(s)), \tag{4.9}$$

and substitution of (4.9) into (4.8) yields

$$\frac{1}{2}\frac{d}{dt}L^2 = -\tau(s)L^2 - (c - w(s))(w'(s)\dot{s}) \tag{4.10a}$$

$$\leq -\tau_0 L^2 + |L| \max |w'(s)| \max |\dot{s}|, \quad \tau_0 := \tau(0). \tag{4.10b}$$

Differentiating $w(s)$ with respect to s reveals $\max |w'(s)| = k_1 e_T/(k_{-1} + k_2)$. Denote $\max |\dot{s}|$ by v and note that $v \leq k_0$ on W_1, due to $\dot{s} \geq 0$.

With

$$\varepsilon_c := \frac{k_1 e_T}{k_{-1} + k_2},$$

Cauchy's inequality

$$ab \leq \sigma a^2 + \frac{b^2}{4\sigma}, \quad \forall \sigma > 0 \tag{4.11}$$

implies

$$\varepsilon_c v|L| \leq \sigma L^2 + \frac{(\varepsilon_c v)^2}{4\sigma} \quad \forall \sigma > 0, \tag{4.12}$$

which yields

$$\frac{1}{2}\frac{d}{dt}L^2 \leq (\sigma - \tau_0)L^2 + \frac{(\varepsilon_c v)^2}{4\sigma} \quad \forall \sigma > 0. \tag{4.13}$$

A natural choice for σ is $\sigma := \tau_0/2$ leading to the inequality

$$\frac{d}{dt}L^2 \leq -\tau_0 L^2 + \frac{(\varepsilon_c v)^2}{\tau_0}. \tag{4.14}$$

Applying Gronwall's lemma to (4.14) generates an upper estimate for L^2:

Proposition 1. *(a) For every solution of* (2.4) *with initial value in W_1 one has the estimates*

$$L^2 \leq L^2(0)e^{-\tau} + \frac{(\varepsilon_c v)^2}{\tau_0^2}(1 - e^{-\tau}); \tag{4.15a}$$

$$L^2 \leq L^2(0)e^{-\tau} + \frac{(\varepsilon_c k_0)^2}{(k_{-1} + k_2)^2} \tag{4.15b}$$

with $\tau := \tau_0 t = (k_{-1} + k_2)t$.

(b) Thus with

$$\varepsilon^* := \frac{k_0 k_1 e_T}{(k_{-1} + k_2)^2}, \tag{4.16}$$

the solution approaches the QSS variety up to an error of ε^{*2}, with time constant $\lambda_\tau := (k_{-1} + k_2)^{-1}$.

Note that the estimates from the proposition explain the rapid approach of the trajectories in Figure 1 to the QSS variety.

From our analysis of the mathematical energy, L^2, we have both a time constant, $\lambda_\tau = \tau_0^{-1}$, as well as a parameter, ε^*. The time constant yields a natural *dimensional* fast time scale, τ, that is equivalent to the fast time scale obtained by Segel [42] for the closed Michaelis–Menten reaction mechanism. Moreover, ε^* should in some sense be small for the open sQSSA to be accurate. The difficulty here is that ε^* has dimension, and we must scale ε^* appropriately to recover a dimensionless parameter. To scale, note that if $k_0 < k_2 e_T$, then

$$\frac{e_T k_0}{K_M(k_{-1} + k_2)} < \frac{k_2 e_T^2}{K_M(k_{-1} + k_2)}. \tag{4.17}$$

Since $c \le e_T$, we divide the (4.17) through by e_T, and take the inequality,

$$\varepsilon_o := \frac{k_2 e_T}{K_M(k_{-1} + k_2)} \ll 1, \tag{4.18}$$

to be the general qualifier for the validity of open sQSSA (2.8) in W_1, when a finite stationary point is located in the first quadrant.

Note that ε_o vanishes if either k_1, e_T or k_2 vanish. However, the use of Fenichel theory also requires k_0 to vanish in the singular limit, otherwise the perturbation is non-singular and the accuracy of a specific QSS reduction is attributable *only* to the near-invariance of the QSS manifold [hence the difference in the justification of the open sQSSA that occurs from the mapping $(k_0, e_T) \mapsto \varepsilon(k_0^*, e_T^*)$ versus the mapping $(k_0, e_T) \mapsto (k_0, \varepsilon e_T^*)$]. This observation is a definitive difference between our work and that of Stoleriu et al. [44].

5. Additional insights from solutions of the invariance equation

The sQSSA can be thought of as an attempt to approximate the slow invariant manifold [11]. There are many other methods for approximating the slow manifold, ranging from the method of intrinsic low-dimensional manifolds [28], which is accurate to $O(\varepsilon)$ [25], to methods that can be improved order-by-order such as singular-perturbation theory [3, 22, 41], computational singular perturbation theory [27], and Fraser's iterative method [11, 31]. Here, we study solutions of the invariance equation, the equation that the exact slow manifold satisfies, in order to gain further insights into the role of the TFPV in determining the validity of the sQSSA. The Fraser iterative method will be a major tool, but we will also consider various small-parameter expansions of the iterates.

5.1. The invariance equation

Assume that, in accordance with the arguments in Section 4.3, we can represent the slow manifold (at least locally) as the graph of a function $c = C(s)$. If $\dot{s} = \dot{s}(s, c)$ and $\dot{c} = \dot{c}(s, c)$, then differentiating

the assumed representation of the slow manifold with respect to time, we get

$$\dot{c}(s, C) = \frac{dC(s)}{ds} \dot{s}(s, C), \tag{5.1}$$

the invariance equation [11, 19, 21, 26, 35].

The invariance equation could be solved using a perturbation method. A strategy suggested by the work of the previous sections is to perturb from a TFPV along a curve in parameter space with the TFPV as its endpoint, e.g. the ray (3.6). The scaling parameter ε can then serve as a perturbation parameter, and a perturbation problem of the typical form results, i.e. to compute the i'th term in the perturbation series, we solve an *algebraic* equation that only depends on the previous terms. However, suppose that we did not know about TFPVs. Then we might try to use the same small parameter as in the closed system, viz. some scaled version of e_T [4, 22, 42, 44]. In the current framework, we would write $e_T \mapsto \varepsilon e_T^*$, and expand $C(s) = \chi_1(s)\varepsilon + \chi_2(s)\varepsilon^2 + \dots$ If we implement this program, we find that $\chi_1(s)$ satisfies the *differential* equation

$$\frac{d\chi_1}{ds} = \frac{1}{k_0} \left[k_1 e_T^* s - \chi_1(k_1 s + k_{-1} + k_2) \right]. \tag{5.2}$$

Higher-order terms also satisfy differential rather than algebraic equations. These difficulties are linked to the fact that $e_T = 0$, of itself, is not a TFPV for the open system. For the TFPVs (3.4a) and (3.4b), since the leading-order term in $C(s)$ is $O(\varepsilon)$, rescaling the TFPVs balances the terms in the invariance equation such that, to leading order, $\dot{c}$, $\dot{s}$ and dC/ds are all $O(\varepsilon)$. As a result, (with slight abuse of notation) $d\chi_i/ds$ first appears to $O(\varepsilon^{i+1})$, and we obtain an algebraic equation for χ_i. The case of TFPV (3.4c) is slightly different. If we rescale $(k_0, k_2) \mapsto \varepsilon(k_0^*, k_2^*)$ and take $C(s) = \zeta_0(s) + \zeta_1(s)\varepsilon + \zeta_2(s)\varepsilon + \dots$, the ε^0 terms of the invariance equation can be rearranged to

$$\left[k_1 s(e_T - \zeta_0) - k_{-1}\zeta_0 \right] \left(1 + \frac{d\zeta_0}{ds} \right) = 0. \tag{5.3}$$

The term in square brackets gives us the critical manifold (3.4c) for ζ_0. (The other solution, $d\zeta_0/ds = -1$, gives the fast foliations of the manifold in the limit $\varepsilon \to 0$.) At higher orders, ζ_i first appears with the $O(\varepsilon^i)$ terms. However, because in the limit $\varepsilon \to 0$ for this TFPV, the k_0 term in $\dot{s}$ vanishes, the coefficient of $d\zeta_i/ds$ at $O(\varepsilon^i)$ is the term in square brackets in Eq (5.3), which vanishes. Thus, $d\zeta_i/ds$ first appears with a non-vanishing coefficient at $O(\varepsilon^{i+1})$, and we again have a perturbation problem involving only algebraic equations.

To recapitulate, rescaling the TFPVs yields a tractable perturbation problem precisely because the TFPVs define critical manifolds. Choosing any path through parameter space that does not reduce to a TFPV as $\varepsilon \to 0$ will, by contrast, necessarily yield a troublesome perturbation problem.

Each TFPV yields a different perturbation problem. Rather than studying the perturbation expansions of the slow manifold directly, we turn to Fraser's method [11, 31], which will allow us to compute a sequence of approximations in a TFPV-agnostic manner. Series expansions of the approximations can then be obtained for any desired TFPV scaling parameter.

In Fraser's iterative method, we think of the invariance equation as an equation to be solved for C in terms of dC/ds. In this case, we can explicitly rearrange the invariance equation to the functional

equation [11]

$$C = \frac{k_1 e_T s \left(1 + \dfrac{dC}{ds}\right) - k_0 \dfrac{dC}{ds}}{(k_1 s + k_{-1}) \left(1 + \dfrac{dC}{ds}\right) + k_2}. \tag{5.4}$$

Observe that if we rescale k_0 and e_T as in (3.6) and let $\varepsilon \to 0$ in the functional equation, we recover the critical manifold (3.4a). Similar comments can be made for (k_0, k_1) and (k_0, k_2) and the corresponding critical manifolds (3.4b) and (3.4c), respectively. Thus, the critical manifolds are recovered in suitable limits of the functional equation.

We now want to solve Eq (5.4) for the slow manifold. If we knew the derivative of C with respect to s along the slow manifold, we could immediately compute $C(s)$ from (5.4). Since we do not, we solve the invariance equation by iteration: From some initial guess $C_0(s)$, we compute the derivative, substitute it into (5.4) to obtain $C_1(s)$, and iterate. In practice, iterative solution of a functional equation such as (5.4) tends to converge specifically to the slow manifold [11, 36] if it converges at all [37], even though every trajectory is a solution of the invariance equation.

The critical manifolds associated with the TFPVs suggest potential initial functions for iteration. Suppose then that we start iteration from the critical manifold [under either TFPV (3.4a) or (3.4b)] $C_0(s) = 0$. Then $C_1(s)$ is the sQSSA (2.6b). Figure 3 shows a sequence of iterates calculated from this initial function. Convergence is rapid, although much more so away from the s axis.
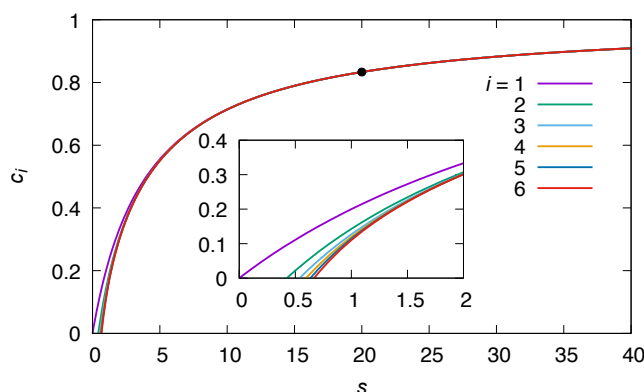
Figure 3. Iterates of Eq (5.4) for the open Michaelis–Menten reaction mechanisms starting from the initial function $C_0(s) = 0$ for the parameters of Figure 1(a). The solid dot marks the location of the equilibrium point. The inset shows an expanded view of the behavior of the iterates near the origin.

As a side note, consider using a vertical initial function, i.e. one for which $dC/ds = \infty$. The first iterate from such an initial function is the s nullcline, which intercepts the s axis at $s = k_0/k_1 e_T$, i.e. at the extreme right end of the possible range of s intercepts of the slow manifold. The sQSSA, on the other hand, is the c nullcline, obtained in one iterative step from the initial function $C_0(s) = 0$, and it intercepts the s axis at $s = 0$. The two nullclines thus arise naturally as approximations of the slow manifold by iteration from coordinate axes, and serve as upper and lower bounds for the slow manifold. Similar comments about the relationship of the functional equation to the nullclines have previously been made about closed systems [11, 12, 31].

5.2. The TFPVs (k_0, e_T) and the small parameters rediscovered

If we obtain higher iterates using a symbolic algebra system, then make the substitution (3.6), and finally expand in powers of ε, we find that the i'th iterate is consistent with the previous iterate to order ε^{i-1}. In other words, the iterative method builds the perturbation series term-by-term, as was previously observed for various perturbative solutions of the closed system [25, 36]. However, this property does *not* hold if we, for instance, expand in powers of e_T, since $e_T = 0$ is not, of itself, a TFPV for the open system. These properties parallel those of the direct perturbation calculations (not shown).

The first two non-zero terms of the perturbation series computed along the ray (3.6) can be written as follows:

$$\frac{C(s)}{e_T^*} = \frac{s}{s + K_M}\varepsilon + \frac{K_M\left[s(k_2 e_T^* - k_0^*) - k_0^* K_M\right]}{k_1(s + K_M)^4}\varepsilon^2 + O(\varepsilon^3). \tag{5.5}$$

Division by e_T^*, the nominal value of the enzyme concentration, has made this expression dimensionless. Thus, the ε^2 term represents an error term for the sQSSA. Specifically, the absolute value of the coefficient of ε^2,

$$\delta(s) = \left|\frac{K_M\left[s(k_2 e_T^* - k_0^*) - k_0^* K_M\right]}{k_1(s + K_M)^4}\right|, \tag{5.6}$$

is a dimensionless error parameter such that the error in the sQSSA is small provided this coefficient is small. An elementary calculation shows that $\delta(s)$ has a local maximum of

$$\delta_m = \frac{27 k_2 e_T^*\left(1 - \frac{k_0^*}{k_2 e_T^*}\right)^4}{256 k_1 K_M^2} \tag{5.7}$$

in $s \in (0, \infty)$ provided $k_2 e_T^* > k_0^*$. The global maximum of $\delta(s)$ for $s \geq 0$ is either this local maximum or

$$\delta(0) = \varepsilon^*/e_T = \frac{k_0^*}{k_1 K_M^2}, \tag{5.8}$$

where the dimensional parameter ε^* is defined in Eq (4.16). When the inflow exceeds the enzyme's clearance capacity, the situation is straightforward, and $\delta(0)$ is the correct small parameter. Otherwise, δ_m will be larger than $\delta(0)$ when

$$\frac{27}{256}\left(1 - \frac{k_0}{k_2 e_T}\right)^4 > \frac{k_0}{k_2 e_T}. \tag{5.9}$$

We dropped the asterisks here because $k_0^*/k_2 e_T^* = k_0/k_2 e_T$. This inequality can be solved numerically. It yields $k_0/k_2 e_T < 0.0767$. Putting it all together, we have the following:

- The sQSSA is a good approximation to the slow manifold globally if $k_0/k_2 e_T < 0.0767$ and $\delta_m \ll 1$. Comparing Eq (5.7) to (4.18), and noting that in this parameter range, $\delta_m < \frac{27}{256}\varepsilon_o$, we conclude that $\varepsilon_o \ll 10$ is sufficient for the validity of the sQSSA. This is a somewhat more permissive bound than (4.18).

- If $k_0/k_2 e_T > 0.0767$, then $\delta(0) \ll 1$ is the appropriate condition for the validity of the sQSSA in the open system.

Note that this analysis has recovered both of the small parameters identified in Section 4.4, but has also established a sharp boundary for switching from one small parameter to the other. We thus have two complementary methods to obtain small parameters. In any given problem, one or the other method might be unworkable, thus our presentation of both methods here.

5.3. The TFPVs (k_0, k_2) and the equilibrium approximation

We can also expand the iterates using the small parameter implied by (3.4c). If we take $(k_0, k_2) \mapsto \varepsilon(k_0^*, k_2^*)$, and then expand the second (or higher) iterate in powers of ε, we get

$$\frac{C(s)}{e_T} = \frac{s}{s + K_E} - \frac{K_E \left(k_2^* s + k_0^*\right) + k_2^* s^2}{k_1 (s + K_E) \left[(s + K_E)^2 + K_E e_T\right]} \varepsilon + O(\varepsilon^2), \tag{5.10}$$

where $K_E = k_{-1}/k_1$. Note that the $O(\varepsilon^0)$ term is the classical quasi-equilibrium approximation (QEA) for the Michaelis–Menten reaction mechanism. Contrast equations (5.5) and (5.10): The QEA for the open system is only accurate to order ε^0, unlike the sQSSA which is accurate to order ε. (This fact is also reflected in Remark 2(b) and in the second example in Appendix A.) This is easily understood given that the QEA lies above the sQSSA at any $s > 0$, and that the slow manifold, which enters the first quadrant by passing through the positive s semi-axis, lies below the sQSSA for $s < \widehat{s}$. In the interval $s \in [0, \widehat{s}]$, the sQSSA will therefore always be closer to the slow manifold than the QEA. This is unlike the situation in the closed system, where the slow manifold lies between the QEA and sQSSA, and where it is possible to choose parameters such that one or the other approximation is more accurate near the origin. The difference is that the QEA is a nullcline in the closed system, but not in the open system. One implication of this result is that the TFPV (3.4a) is the most natural one to use as a basis for a geometric singular perturbation treatment of the slow manifold. (See also Appendix A for further notes on the expansion from the TFPV (3.4c).)

5.4. The TFPVs (k_0, k_1) and the linear regime

Finally, turning to the TFPV (3.4b), we define a perturbation parameter ε by

$$(k_0, k_1) \mapsto \varepsilon(k_0^*, k_1^*). \tag{5.11}$$

A perturbation series based on this small parameter is a polynomial in s due to the appearance of k_1 and s together in the rate equations. The first nonzero terms of this series are

$$\frac{C(s)}{e_T} = \frac{k_1^* s}{k_{-1} + k_2} \varepsilon - \frac{k_1^* \left\{k_1^* s \left[s(k_{-1} + k_2) - k_2 e_T\right] + k_0^*(k_{-1} + k_2)\right\}}{(k_{-1} + k_2)^3} \varepsilon^2 + O(\varepsilon^3). \tag{5.12}$$

Substituting this series along with the parameter definitions (5.11) into $\dot{s}$ from (2.4), we get, to lowest order in ε,

$$\dot{s} \approx \left(k_0^* - \frac{v_{\max} s}{K_M^*}\right) \varepsilon, \tag{5.13}$$

where $v_{\max} = k_2 e_T$ and $K_M^* = (k_{-1} + k_2)/k_1^*$ or, restoring the small parameters from (5.11),

$$\dot{s} \approx k_0 - \frac{v_{\max} s}{K_M}. \tag{5.14}$$

This is of course the small-s linear limit of the sQSSA, the previously seen Eq (3.15). An alternative route to this equation is presented in Appendix A.

6. The open Michaelis–Menten reaction on the Poincaré sphere

It is worthwhile to consider the global behavior of system (2.4) and its distinguished invariant sets to illuminate the role of QSS varieties in a broader context, particularly for systems which do not admit a stationary point in the first quadrant.

It is a standard technique to extend planar polynomial ODE systems to the Poincaré sphere. A good description of the procedure is given in Perko [34], Section 3.10: Given a sphere in $\mathbb{R}^3$, let the phase plane be tangent to its north pole, and consider the bijective central projection from the upper half sphere to the phase plane. Then points on the equator of the sphere may be viewed as points at infinity for the planar system, with each line through the origin corresponding to a pair of antipodal points on the equator.[†] Finally, for the purpose of visualization one applies a parallel projection in the north-south direction from the upper hemisphere to the equatorial plane.

A discussion of the system on the Poincaré sphere thus allows us to understand the behavior of the planar system at infinity. Note that all solutions of the system on the Poincaré sphere, which is compact, exist for all $t \in \mathbb{R}$, while this is not necessarily the case for solutions of (2.4) when $t \leq 0$ or outside the first quadrant. Any reference to limit sets in the following arguments is to be understood for the system on the sphere. In our analysis we will mostly be interested in the first quadrant.

Stationary points at infinity (i.e. on the equator) for a polynomial planar system are of particular interest. Antipodal pairs of stationary points generally correspond to invariant lines for the homogeneous part of highest degree; see e.g. [48]. For system (2.4) with $k_1 \neq 0$ we thus need to consider the homogeneous quadratic part

$$\begin{aligned} \dot{s} &= k_1 c s, \\ \dot{c} &= -k_1 c s. \end{aligned}$$

This homogeneous vector field admits three invariant lines, viz.

$$\mathbb{R} \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \mathbb{R} \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \mathbb{R} \cdot \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

The stationary points at infinity which are relevant for the first quadrant correspond to the rays

$$\mathbb{R}_+ \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ and } \mathbb{R}_+ \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

and we call the corresponding stationary points at infinity P_1, resp. P_2. Moreover we denote by P_3 the stationary point at infinity which corresponds to $\mathbb{R}_+ \cdot \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

We present the pertinent results for system (2.4) on the Poincaré sphere (see Appendix B for computations and proofs).

Lemma 4. *Assume that the genericity conditions* (3.1) *are satisfied. Then the following hold for the system on the Poincaré sphere.*

[†]The central projection also yields a bijection from the lower hemisphere to the phase plane, and one thus obtains a vector field on the sphere which is mirror symmetric relative to the equatorial plane, and has the equator as an invariant set. One could furthermore pass to a direction field on the projective plane, but we will not do so.

(a) *The stationary point P_1 at infinity is a degenerate saddle when $k_2 e_T > k_0$, with the stable manifold contained in the equator. In case $k_2 e_T < k_0$ this point is a degenerate attracting node.*

(b) *The stationary point P_2 at infinity is a saddle-node, with a repelling node part on the upper hemisphere.*

(c) *The stationary point P_3 at infinity is a repelling node.*

We first describe the behavior of system (2.4) on the relevant part of the Poincaré sphere when there is an isolated stationary point in the first quadrant as illustrated in Figure 4.
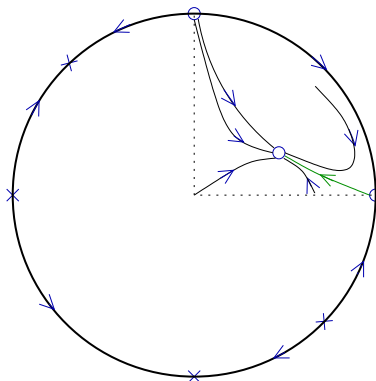
Figure 4. The system on the Poincaré Sphere for the open Michaelis–Menten reaction mechanism in case $k_2 e_T > k_0$. The distinguished trajectory is colored green.

Proposition 2. *Assume that the genericity conditions* (3.1) *hold, and let $k_2 e_T > k_0$. Then every solution starting in the first quadrant converges toward $P_0 = (\widehat{s}, \widehat{c})$ as $t \to \infty$. There is a unique distinguished trajectory that connects the saddle P_1 at infinity to P_0. Moreover this trajectory is asymptotic in the phase plane to the line $c = e_T$ as $t \to -\infty$.*

We turn to the case when P_0 lies in the second quadrant; see Figure 5. Here, considering the system on the Poincaré sphere is necessary to understand the global dynamics, and moreover a proper understanding requires us to look beyond the first quadrant.
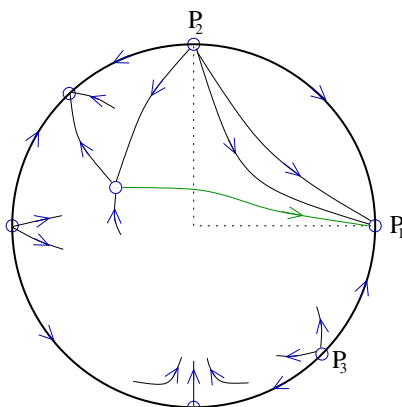
Figure 5. The system on the Poincaré Sphere for the open Michaelis–Menten reaction mechanism in case $k_2 e_T < k_0$. The distinguished trajectory is colored green.

Proposition 3. *Assume that the genericity conditions* (3.1) *hold, and let $k_2 e_T < k_0$. Then every solution that starts in the first quadrant converges to P_1 as $t \to \infty$, and the corresponding trajectory in the phase plane is asymptotic to the line $c = e_T$. There is a unique distinguished trajectory that connects the saddle P_0 to P_1.*

Remark 4. In view of Proposition 3, the mathematically distinguished trajectory connecting P_0 and P_1 may be seen as a natural candidate for a "global" slow manifold in appropriate parameter regimes. We provide a few more details here. From the proof (see also Figure 5) one finds that the two components of the unstable manifold of P_0 connect respectively to P_1 and to the antipode of P_3. Solutions in the open upper hemisphere, unless they start on the stable manifold of P_0, converge either to P_1 or to the antipode of P_3 as $t \to \infty$. Moreover one component of the stable manifold of P_0 connects to P_2 (which is the only available alpha limit point), and the other may connect either to the antipode of P_1, or to the antipode of P_2, or to P_3. (Topological arguments do not yield more precise information, and we will not delve further into this matter.) The stable manifold of P_0 separates the regions of attraction for P_1 and the antipode of P_3 in the open upper hemisphere. In turn, the region of attraction for P_1 is separated by the distinguished trajectory into two subregions. For one of these subregions, the alpha limit set of all points is equal to $\{P_2\}$, thus one may briefly say that all trajectories in this region come down from $c = \infty$. For the other subregion, a similarly concise statement does not seem possible: The set of alpha limit points certainly includes P_3, but it may also include the antipode of P_2 or of P_1.

7. Discussion

The open Michaelis–Menten reaction mechanism, although of definitive relevance in biochemistry, has attracted less attention than the classical closed mechanism without influx. We investigated the sQSSA for this system from two perspectives:

1. We considered QSS as a singular perturbation phenomenon. We determined all parameter combinations (TFPVs) from which singular perturbation reductions emanate via a small perturbation, and in particular we identified the relevant parameter values for sQSSA, which are given by $k_0 = e_T = 0$, with all other parameters positive. By singular perturbation theory one obtains the familiar QSS reduction.

2. Motivated by a more general notion of QSS (proposed by Schauer and Heinrich [40]) and by the results of Stoleriu et al. [44], we obtained quasi-steady state reduction by direct estimates from QSS parameter values given by $e_T = 0$, all other parameters positive. Thus sQSS reduction for the open MM reaction is applicable to a wider range of parameters than for singular perturbation reduction. Note that such a phenomenon does not appear in the closed Michaelis–Menten system. By these estimates we obtained a justification of central results in [44], and could also determine their range.

However, considering the fine structure of slow manifolds by analysis of higher order approximations revealed the special role (and higher accuracy of approximation) of Tihkonov-Fenichel parameter values in contrast to QSS parameter values.

Finally, we took a global mathematical perspective to investigate scenarios with no positive equilibrium.

Acknowledgments

Conflict of interest

The authors declare no conflict of interest.

References

1. A. V. P. Bobadilla, B. J. Bartmanski, R. Grima, H. G. Othmer, The status of the QSSA approximation in stochastic simulations of reaction networks, in *2018 MATRIX Annals*, Springer International Publishing, 2020, 137–147.

2. M. Bodenstein, Eine Theorie der photochemischen Reaktionsgeschwindigkeiten, *Z. Phys. Chem.*, **85** (1913), 329–397.

3. J. R. Bowen, A. Acrivos, A. K. Oppenheim, Singular perturbation refinement to quasi-steady state approximation in chemical kinetics, *Chem. Eng. Sci.*, **18** (1963), 177–187.

4. G. E. Briggs, J. B. S. Haldane, A note on the kinetics of enzyme action, *Biochem. J.*, **19** (1925), 338–339.

5. M. S. Calder, D. Siegel, Properties of the Michaelis-Menten mechanism in phase space, *J. Math. Anal. Appl.*, **339** (2008), 1044–1064.

6. D. L. Chapman, L. K. Underhill, LV.—The interaction of chlorine and hydrogen. The influence of mass, *J. Chem. Soc., Trans.*, **103** (1913), 496–508.

7. J. Eilertsen, S. Schnell, The quasi-steady-state approximations revisited: Timescales, small parameters, singularities, and normal forms in enzyme kinetics, *Math. Biosci.*, **325** (2020), 108339.

8. N. Fenichel, Persistence and smoothness of invariant manifolds for flows, *Indiana U. Math. J.*, **21** (1971), 193–226.

9. N. Fenichel, Geometric singular perturbation theory for ordinary differential equations, *J. Differ. Equations*, **31** (1979), 53–98.

10. E. H. Flach, S. Schnell, Use and abuse of the quasi-steady-state approximation, *IEEE Proc. Syst. Biol.*, **153** (2006), 187–191.

11. S. J. Fraser, The steady state and equilibrium approximations: A geometrical picture, *J. Chem. Phys.*, **88** (1988), 4732–4738.

12. S. J. Fraser, M. R. Roussel, Phase-plane geometries in enzyme kinetics, *Can. J. Chem.*, **72** (1994), 800–812.

13. A. Goeke, C. Schilli, S. Walcher, E. Zerz, Computing quasi-steady state reductions, *J. Math. Chem.*, **50** (2012), 1495–1513.

14. A. Goeke, S. Walcher, Quasi-steady state: Searching for and utilizing small parameters, in *Recent Trends in Dynamical Systems* (eds. A. Johann, H.-P. Kruse, F. Rupp and S. Schmitz), Springer Basel, Basel, 2013, 153–178.

15. A. Goeke, S. Walcher, A constructive approach to quasi-steady state reductions, *J. Math. Chem.*, **52** (2014), 2596–2626.

16. A. Goeke, S. Walcher, E. Zerz, Determining "small parameters" for quasi-steady state, *J. Differ. Equations*, **259** (2015), 1149–1180.

17. A. Goeke, S. Walcher, E. Zerz, Classical quasi-steady state reduction – A mathematical characterization, *Phys. D*, **345** (2017), 11–26.

18. D. Gonze, W. Abou-Jaoudé, D. A. Ouattara, J. Halloy, How molecular should your molecular model be? On the level of molecular detail required to simulate biological networks in systems and synthetic biology, *Meth. Enzymol.*, **487** (2011), 171–215.

19. A. N. Gorban, Model reduction in chemical dynamics: Slow invariant manifolds, singular perturbations, thermodynamic estimates, and analysis of reaction graph, *Curr. Opin. Chem. Eng.*, **21** (2018), 48–59.

20. A. N. Gorban, I. V. Karlin, Method of invariant manifold for chemical kinetics, *Chem. Eng. Sci.*, **58** (2003), 4751–4768.

21. A. N. Gorban, I. V. Karlin, A. Yu. Zinovyev, Constructive methods of invariant manifolds for kinetic problems, *Phys. Rep.*, **396** (2004), 197–403.

22. F. G. Heineken, H. M. Tsuchiya, R. Aris, On the mathematical status of the pseudo-steady hypothesis of biochemical kinetics, *Math. Biosci.*, **1** (1967), 95–113.

23. V. Henri, Théorie générale de l'action de quelques diastases, *C. R. Acad. Sci.*, **135** (1902), 916–919.

24. J. H. Hubbard, B. H. West, *Differential Equations: A Dynamical Systems Approach*, vol. 5 of Texts in Applied Mathematics, Springer, New York, 1991.

25. H. G. Kaper, T. J. Kaper, Asymptotic analysis of two reduction methods for systems of chemical reactions, *Phys. D*, **165** (2002), 66–93.

26. C. Kuehn, *Multiple Time Scale Dynamics*, vol. 191 of Applied Mathematical Sciences, Springer, 2015.

27. S. H. Lam, Using CSP to understand complex chemical kinetics, *Combust. Sci. Technol.*, **89** (1993), 375–404.

28. U. Maas, S. Pope, Simplifying chemical kinetics: Intrinsic low-dimensional manifolds in composition space, *Combust. Flame*, **88** (1992), 239–264.

29. L. Michaelis, M. L. Menten, Die Kinetik der Invertinwirkung, *Biochem. Z.*, **49** (1913), 333–369.

30. D. L. Nelson, M. M. Cox, *Lehninger Principles of Biochemistry*, 5th edition, Freeman, New York, 2008.

31. A. H. Nguyen, S. J. Fraser, Geometrical picture of reaction in enzyme kinetics, *J. Chem. Phys.*, **91** (1989), 186–193.

32. L. Noethen, S. Walcher, Quasi-steady state and nearly invariant sets, *SIAM J. Appl. Math.*, **70** (2009), 1341–1363.

33. L. Noethen, S. Walcher, Tikhonov's theorem and quasi-steady state, *Discrete Contin. Dyn. Syst. Ser. B*, **16** (2011), 945–961.

34. L. Perko, *Differential equations and dynamical systems*, 3rd edition, no. 7 in Texts in Applied Mathematics, Springer, New York, 2001.

35. A. J. Roberts, The utility of an invariant manifold description of the evolution of a dynamical system, *SIAM J. Math. Anal.*, **20** (1989), 1447–1458.

36. M. R. Roussel, S. J. Fraser, Geometry of the steady-state approximation: Perturbation and accelerated convergence methods, *J. Chem. Phys.*, **93** (1990), 1072–1081.

37. M. R. Roussel, Forced-convergence iterative schemes for the approximation of invariant manifolds, *J. Math. Chem.*, **21** (1997), 385–393.

38. M. R. Roussel, Heineken, Tsushiya and Aris on the mathematical status of the pseudo-steady state hypothesis: A classic from volume 1 of *Mathematical Biosciences*, *Math. Biosci.*, **318** (2019), 108274.

39. M. R. Roussel, S. J. Fraser, On the geometry of transient relaxation, *J. Chem. Phys.*, **94** (1991), 7106–7113.

40. M. Schauer, R. Heinrich, Analysis of the quasi-steady-state approximation for an enzymatic one-substrate reaction, *J. Theor. Biol.*, **79** (1979), 425–442.

41. L. A. Segel, M. Slemrod, The quasi-steady-state assumption: A case study in perturbation, *SIAM Rev.*, **31** (1989), 446–477.

42. L. A. Segel, On the validity of the steady state assumption of enzyme kinetics, *Bull. Math. Biol.*, **50** (1988), 579–593.

43. H. L. Smith, P. Waltman, *The Theory of the Chemostat. Dynamics of Microbial Competition*, no. 13 in Cambridge Studies in Mathematical Biology, Cambridge University Press, Cambridge, 1995.

44. I. Stoleriu, F. A. Davidson, J. L. Liu, Quasi-steady state assumptions for non-isolated enzyme-catalysed reactions, *J. Math. Biol.*, **48** (2004), 82–104.

45. P. Thomas, A. V. Straube, R. Grima, Limitations of the stochastic quasi-steady-state approximation in open biochemical reaction networks, *J. Chem. Phys.*, **135** (2011), 181103.

46. A. Tikhonov, Systems of differential equations containing small parameters in their derivatives, *Mat. Sb. (N.S.)*, **31** (1952), 575–586.

47. S. Walcher, On transformations into normal form, *J. Math. Anal. Appl.*, **180** (1993), 617–632.

48. S. Walcher, On the Poincaré problem, *J. Differ. Equations*, **166** (2000), 51–78.

49. M. Wechselberger, *Geometric Singular Perturbation Theory Beyond the Standard Form*, no. 6 in Frontiers in Applied dynamical systems: Tutorials and Reviews, Springer, 2020.

50. A. Wurtz, Sur la papaïne. Nouvelle contribution à l'histoire des ferments solubles, *C. R. Acad. Sci.*, **91** (1880), 787–791.

A. Projecting onto a critical manifold

This appendix is intended to give the reader a short and user-friendly overview of pertinent methods in geometric singular perturbation theory. For a more elaborate presentation, we encourage the reader to consult [8, 9, 15, 26, 33, 49]. The references [15, 33] focus specifically on the QSSA.

Singular perturbation reductions are straightforward for sufficiently smooth systems in standard form

$$\begin{aligned}
\dot{x}_1 &= \varepsilon f_1(x_1, x_2, \varepsilon), \\
\dot{x}_2 &= f_2(x_1, x_2, \varepsilon),
\end{aligned}$$

which depend on a "small parameter" ε. We assume that the reader is familiar with this procedure, which we sketch without giving details: In slow time $\tau = \varepsilon t$ the system may be rewritten as

$$\begin{aligned}
x_1' &= f_1(x_1, x_2, \varepsilon), \\
\varepsilon x_2' &= f_2(x_1, x_2, \varepsilon).
\end{aligned}$$

Given suitable hyperbolicity conditions, solutions of the latter system converge toward solutions of the reduced (differential-algebraic) system

$$x_1' = f_1(x_1, x_2, 0), \quad f_2(x_1, x_2, 0) = 0.$$

The procedure just sketched requires an a priori separation of slow and fast variables, which is not necessarily given. This difficulty is overcome by a coordinate-free version of singular perturbation reductions as developed by Fenichel [9]. (Almost all the relevant information is contained in pp. 65–66 of this reference, but in rather condensed form.) Here we present the basic theory and computation-relevant facts, loosely following Wechselberger [49], Chapter 3, as well as [14, 15] specifically for the QSS reduction procedure.

For systems not in standard form, one must first *define* the notion of a singular perturbation, according to Fenichel. Given a differential equation of the form

$$\dot{x} = F(x, \varepsilon), \quad x \in \mathbb{R}^n, \quad F : \mathbb{R}^n \times \mathbb{R} \mapsto \mathbb{R}^n \tag{A.1}$$

with sufficiently smooth F, one is interested in the dynamics in the asymptotic limit $\varepsilon \to 0+$. The singular points of $F(x, 0)$ determine the nature of the perturbation. It will be convenient to express F in the form

$$F(x, \varepsilon) =: h(x) + \varepsilon G(x, \varepsilon), \tag{A.2}$$

so that there is a clear distinction between the vector field, $h(x) = F(x, 0)$, and the perturbation, $\varepsilon G(x, \varepsilon)$. Let S denote the set of singular points of $F(x, 0)$:

$$S := \{x \in \mathbb{R}^n : h(x) = 0\}. \tag{A.3}$$

Note that S is an algebraic variety for polynomial or rational systems, which are quite common in reaction equations. If S is the empty set, or contains only isolated singularities, then the perturbation is called *regular*. In contrast, if $M \subseteq S$ is a differentiable manifold comprised of non-isolated singularities, then the perturbation is singular, and M is called a *critical manifold*.

We now outline the reduction procedure in the coordinate-free setting:

1. Given a singularly perturbed problem of the form (A.1), one can establish necessary and sufficient conditions for the existence of a local transformation to standard form, as follows: For every $x \in M$ the required conditions are

   (i) $T_x M = \ker Dh(x) = \{v \in \mathbb{R}^n : Dh(x) \cdot v = 0\}$.

   (ii) For the eigenvalue 0 of $Dh(x)$ the algebraic and the geometric multiplicities are equal.

   (iii) Normal hyperbolicity: All nonzero eigenvalues of $Dh(x)$ have nonzero real parts. In applications one often requires the stronger attracting hyperbolicity condition, viz. all nonzero eigenvalues of $Dh(x)$ have negative real parts.

2. Given the above conditions, one has a direct sum decomposition

$$\mathbb{R}^n = T_x M \oplus N_x, \quad \forall x \in M, \tag{A.4}$$

where the $Dh(x)$-invariant complementary subspace $N_x = \mathcal{R}(Dh(x))$ is the range of $Dh(x)$. The conditions in item 1 are necessary and sufficient for such a decomposition to exist.

Given this decomposition one can define the projection operator, Π^M, which for every $x \in M$ maps $\mathbb{R}^n$ onto the tangent space of M at x and has kernel N_x. (Recall that a projection is uniquely determined by its kernel and image.)

Once Π^M is known, the leading order singular perturbation reduction is computed by projecting the perturbation term onto the tangent space of M of x:

$$x' = \Pi^M G(x, 0), \quad G(x, 0) \in \mathbb{R}^n, \quad x \in M. \tag{A.5}$$

3. To compute Π^M explicitly, it is useful to employ a decomposition

$$h(x) = P(x)f(x), \tag{A.6}$$

where $P(x)$ is a rectangular matrix-valued function, M locally coincides with the zero level set of the vector-valued function $f(x)$, and $Df(x)$ has full rank when $x \in M$. The existence of such a decomposition is guaranteed by the implicit function theorem, and for polynomial or rational systems it can be obtained in an algorithmic manner. One obtains the operator Π^M as

$$\Pi^M := I - P(DfP)^{-1}Df. \tag{A.7}$$

4. In addition to the reduced equation, the initial value on the critical manifold M is also relevant. In the attracting hyperbolic case the fast equation $\dot{x} = h(x)$ admits $n - \dim M$ independent first integrals in a neighborhood of M, and to a given initial value $z \in \mathbb{R}^n$ corresponds (up to a correction of order ε) the point where M intersects the common level set of the first integrals containing z; see Fenichel [9], Lemma 5.3, and also [15], Proposition 2.

Formally, the procedure resulting in (A.5) is referred to as a *critical manifold projection*; see Figure 6 for a geometric illustration. The flow on M at $\varepsilon = 0$ is trivial, but Fenichel theory ensures that the perturbed vector field has an invariant *slow manifold* close to M, on which the flow is slow but non-trivial. The long-time evolution of x is given (approximately) by the projected dynamical system (A.5).
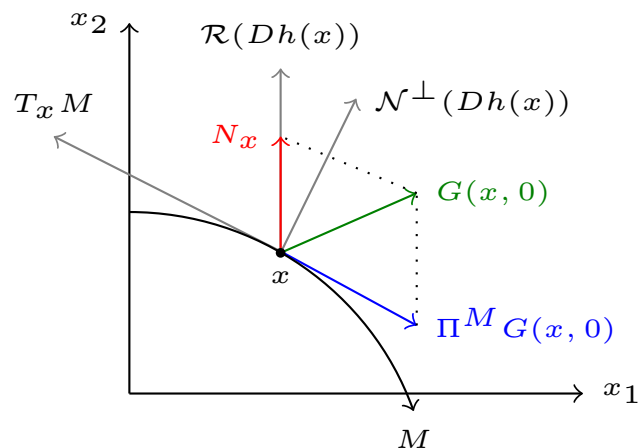
Figure 6. Projecting onto the slow manifold. In this figure, "$\mathcal{R}$" denotes range and "$\mathcal{N}$" denotes nullspace. The complementary subspaces $T_x M$ and N_x are invariant with respect to the linearization $Dh(x)$, and the components of $G(x,0) \in \mathbb{R}^n$ can be uniquely expressed as $G(x,0) = u + v$, with $u \in T_x M$ and $v \in N_x$. Π^M is constructed in the form of an oblique projection onto $T_x M$; note that $T_x M$ and N_x are not necessarily orthogonal. The perturbed dynamical system that is influenced by the presence of $G(x,0)$ is approximated by (A.5). Note that the critical manifold M is in fact filled with non-isolated equilibria.

As a first illustrating example we formally compute the singular perturbation reduction for the case of small $k_0 = \varepsilon k_0^*$ and small $k_1 = \varepsilon k_1^*$, thus we have the perturbation problem

$$
\begin{aligned}
\dot{s} &= \varepsilon k_0^* - \varepsilon k_1^*(e_T - c)s + k_{-1}c, \\
\dot{c} &= \varepsilon k_1^*(e_T - c)s - (k_{-1} + k_2)c.
\end{aligned}
\tag{A.8}
$$

The critical manifold, M, attained by setting $\varepsilon = 0$ in (A.8), corresponds to the s axis:

$$
M := \{(s,c) \in \mathbb{R}^2 : c = 0\}.
\tag{A.9}
$$

Furthermore, $h(s,c) = P(s,c)f(s,c)$ and $G(s,c,\varepsilon)$ are given by

$$
P(s,c) := \begin{bmatrix} k_{-1} \\ -(k_{-1} + k_2) \end{bmatrix}, \quad f(s,c) = c, \quad G(s,c,\varepsilon) := \begin{bmatrix} k_0^* - k_1^*(e_T - c)s \\ k_1^*(e_T - c)s \end{bmatrix},
\tag{A.10}
$$

and thus $Df = [0 \ \ 1]$. Next, DfP is the scalar $-(k_{-1} + k_2)$. This scalar is negative throughout M, and therefore (e.g. according to [15]) conditions (ii) and (iii) above are satisfied. The product of P and Df is

$$
PDf := \begin{bmatrix} 0 & k_{-1} \\ 0 & -(k_{-1} + k_2) \end{bmatrix}.
\tag{A.11}
$$

Combining the above results yields

$$
\Pi^M = I_{2\times 2} + \frac{1}{(k_{-1} + k_2)}PDf = \begin{bmatrix} 1 & \dfrac{k_{-1}}{k_{-1} + k_2} \\ 0 & 0 \end{bmatrix},
\tag{A.12}
$$

and the reduced equation is

$$s' := \Pi^M|_{c=0}G(s,0,0) = k_0^* - \frac{k_1^* k_2 e_T}{k_{-1} + k_2}s. \tag{A.13}$$

It is worth pointing out that the fast system here admits the first integral: $(k_{-1} + k_2)s + k_{-1}c$. Hence, given an initial value (s_0, c_0) for (A.8), the corresponding initial value for the reduced equation is just

$$\widetilde{s_0} = s_0 + \frac{k_{-1}}{k_{-1} + k_2}c_0.$$

In this example the singular perturbation reduction (A.13) coincides with the "classical" QSS reduction with respect to c in the linear regime where $s \ll K_M$. This is not accidental, but due to the special form of the critical manifold; see [17], Proposition 5.

As a second example consider the singular perturbation reduction for the case of small $k_0 = \varepsilon k_0^*$ and small $k_2 = \varepsilon k_2^*$, thus yielding the perturbation problem

$$\begin{aligned} \dot{s} &= \varepsilon k_0^* - k_1(e_T - c)s + k_{-1}c, \\ \dot{c} &= k_1(e_T - c)s - (k_{-1} + \varepsilon k_2^*)c. \end{aligned} \tag{A.14}$$

Here the "classical" QSS reduction is significantly different from the singular perturbation reduction. The critical manifold is defined by

$$M := \{(s,c) \in \mathbb{R}^2 : f(s,c) = 0\} \tag{A.15}$$

with $f(s,c) := k_1(e_T - c)s - k_{-1}c$, and furthermore, $h(s,c) = P(s,c)f(s,c)$ with

$$P(s,c) := \begin{bmatrix} -1 \\ 1 \end{bmatrix}; \quad \text{moreover } G(s,c,\varepsilon) := \begin{bmatrix} k_0^* \\ -k_2^* c \end{bmatrix}. \tag{A.16}$$

A routine calculation yields $DfP = -(k_1(e_T - c) + k_1 s + k_{-1})$, which is < 0 on M, so again conditions (ii) and (iii) are satisfied. One obtains the reduced system

$$\begin{bmatrix} s' \\ c' \end{bmatrix} = \frac{k_0^* - k_2^* c}{k_1(e_T - c) + k_1 s + k_{-1}} \begin{bmatrix} k_1 s + k_{-1} \\ k_1(e_T - c) \end{bmatrix} \tag{A.17}$$

which is relevant only on the invariant manifold M. Using the parameterization $c = k_1 e_T s/(k_1 s + k_{-1})$ of M, one arrives at an equation for s alone:

$$s' = (k_1 s + k_{-1}) \cdot \frac{k_0^*(k_1 s + k_{-1}) - k_2^* k_1 e_T s}{k_1 k_{-1} e_T + (k_1 s + k_{-1})^2}. \tag{A.18}$$

The reduced system without inflow is known from the literature, see e.g. [14, Example 8.6], [38, Section 5] or [49, Section 3.4]. Note that $s + c$ is a first integral of the fast system; this may be employed to determine the initial value on M.

B. Computations and proofs for the Poincaré sphere

In this Appendix, we record the necessary computations, and give proofs for Lemma 4 as well as Propositions 2 and 3.

From a computational perspective it is convenient to project the system from the sphere to another tangent plane. The procedure was streamlined (for different purposes) in [48], and we are using it here, noting that the final result is the same as in [34].

To accommodate various transformations we rename the variables in system (2.4), thus obtaining

$$
\begin{aligned}
\dot{x}_1 &= k_0 - k_1(e_T - x_2)x_1 + k_{-1}x_2, \\
\dot{x}_2 &= k_1(e_T - x_2)x_1 - (k_{-1} + k_2)x_2.
\end{aligned}
\tag{B.1}
$$

We compute the Poincaré transform of this system with respect to x_1 (terminology from [48]), which corresponds to the transformed system on the tangent plane to the "east pole" (compare Perko [34], Section 3.10, Theorem 2).

- In a first step introduce a further variable x_3 and homogenize, to obtain

$$
\begin{aligned}
\dot{x}_1 &= k_0 x_3^2 - k_1 e_T x_1 x_3 + k_1 x_1 x_2 + k_{-1} x_2 x_3 &=: g_1, \\
\dot{x}_2 &= k_1 e_T x_1 x_3 - k_1 x_1 x_2 - (k_{-1} + k_2)x_2 x_3 &=: g_2.
\end{aligned}
$$

- In step 2, compute the projected system

$$
\begin{aligned}
\dot{x}_2 &= -x_2 g_1 + x_1 g_2, \\
\dot{x}_3 &= -x_3 g_1.
\end{aligned}
$$

- In step 3, dehomogenize by setting $x_1 = 1$, to obtain

$$
\begin{aligned}
\dot{x}_2 &= -k_1 x_2 + k_1 e_T x_3 - k_1 x_2^2 + (k_1 e_T - k_{-1} - k_2)x_2 x_3 - k_{-1} x_2^2 x_3 - k_0 x_2 x_3^2, \\
\dot{x}_3 &= -x_3 \left(k_1 x_2 - k_1 e_T x_3 + k_{-1} x_2 x_3 + k_0 x_3^2 \right).
\end{aligned}
\tag{B.2}
$$

with the equator corresponding to $x_3 = 0$.

Lemma 5. *Assume that the genericity conditions (3.1) are satisfied. Then the following hold:*

(a) System (B.2) admits two stationary points on the line $x_3 = 0$. These are $(0, 0)$, corresponding to P_1 for the homogeneous quadratic part of system (B.1), and $(-1, 0)$ which corresponds to P_3.

(b) At the stationary point $(0, 0)$ the Jacobian is

$$
\begin{bmatrix}
-k_1 & k_1 e_T \\
0 & 0
\end{bmatrix}.
$$

In case $k_2 e_T > k_0$ this point is a degenerate saddle with the equator as local stable manifold and a center-unstable manifold tangent to the line $x_2 - e_T x_3 = 0$. In case $k_2 e_T < k_0$ this point is a degenerate attracting node with all trajectories but the two on the equator approaching it tangent to the line $x_2 - e_T x_3 = 0$.

(c) At the stationary point $(-1, 0)$ (which is irrelevant for the dynamics on the first quadrant) the Jacobian is

$$\begin{bmatrix} k_1 & k_2 \\ 0 & k_1 \end{bmatrix},$$

hence this point is a repelling node.

Proof. (i) The stationary points with $x_3 = 0$ are determined from the equation $-k_1(x_2 + x_2^2) = 0$; thus part (a) follows. Computing the Jacobians is straightforward, and part (c) as well as the first statement of (b) follows. To prepare for proving the remaining statements, introduce new coordinates $y_2 = x_2 - e_T x_3$ and $y_3 = x_3$ to obtain

$$\begin{aligned} \dot{y}_2 &= -k_1 y_2 & - & (k_{-1} + k_2)e_T y_3^2 + \cdots \\ \dot{y}_3 &= & - & k_1 y_2 y_3 - (k_{-1}e_T + k_0)y_3^3 + \cdots \end{aligned}$$

with diagonalized linear part.

(ii) The following auxiliary result is a special case from the last section of [47]: Consider a system

$$\begin{aligned} \dot{z}_1 &= & \alpha_{111}z_1^2 + 2\alpha_{112}z_1 z_2 + \alpha_{122}z_2^2 & + & \gamma_{1111}z_1^3 + \cdots \\ \dot{z}_2 &= \lambda z_2 + & \alpha_{211}z_1^2 + 2\alpha_{212}z_1 z_2 + \alpha_{222}z_2^2 & + & \cdots \end{aligned} \tag{B.3}$$

with real parameters, and $\lambda \neq 0$. Then the normal form on an invariant manifold (NFIM) tangent to $z_2 = 0$, up to degree three, is given by

$$\dot{z}_1 = \alpha_{111}z_1^2 + (\gamma_{1111} - 2\alpha_{112}\alpha_{211}/\lambda)\, z_1^3 + \cdots \tag{B.4}$$

In case $\alpha_{111} \neq 0$ the stationary point 0 of system (B.3) is a saddle-node. In case $\alpha_{111} = 0$ and $\lambda < 0$ the stationary point 0 is a degenerate saddle when $\gamma_{1111} - 2\alpha_{112}\alpha_{211}/\lambda > 0$ and a degenerate attracting node when $\gamma_{1111} - 2\alpha_{112}\alpha_{211}/\lambda < 0$. In the latter case all but two trajectories are tangent to $z_2 = 0$.

(iii) Applying this result with the identifications $z_1 = y_3$, $z_2 = y_2$, $\alpha_{111} = 0$, $2\alpha_{112} = -k_1$, $\gamma_{1111} = -(k_{-1}e_T + k_0)$ and $\alpha_{211} = -(k_{-1} + k_2)e_T$ one obtains the NFIM up to degree three as

$$\dot{y}_3 = (k_2 e_T - k_0)y_3^3 + \cdots ,$$

and all assertions follow.

$\square$

There remains to discuss the stationary point P_2 at infinity, for which we use the Poincaré transform of (B.1) with respect to x_2. The first step is unchanged but the projection in the second step is now given as

$$\begin{aligned} \dot{x}_1 &= -x_1 g_2 + x_2 g_1, \\ \dot{x}_3 &= -x_3 g_2 \end{aligned}$$

and dehomogenization with $x_2 = 1$ yields

$$\begin{aligned} \dot{x}_1 &= -k_1(x_1 + x_1^2) + k_{-1}x_3 + (k_{-1} + k_2 + k_1 e_T)x_1 x_3 + k_0 x_3^2 - k_1 e_T x_1^2 x_3, \\ \dot{x}_3 &= -x_3 (k_1 e_T x_1 x_3 - k_1 x_1 - (k_{-1} + k_2)x_3) . \end{aligned} \tag{B.5}$$

There are two stationary points with $x_3 = 0$. The point $(-1, 0)$ corresponds to P_3, which has been taken care of. There remains $(0, 0)$, corresponding to P_2.

Lemma 6. *Assume that the genericity conditions* (3.1) *hold. Then the stationary point P_2 at infinity is a saddle-node, with a repelling node part on the upper hemisphere. Except for the trajectories on the equator, all trajectories of* (B.5) *with positive x_3 that emanate from this stationary point are tangent to the line given by $k_1 x_1 + k_{-1} x_3 = 0$.*

Proof. To diagonalize the Jacobian at $(0, 0)$, introduce new coordinates $y_1 = x_1 + k_{-1}/k_1 \cdot x_3$, $y_3 = x_3$ to obtain the system

$$\begin{aligned}
\dot{y}_1 &= k_1 y_1 + \cdots \\
\dot{y}_3 &= k_2 y_3^2 + k_1 y_1 y_3 + \cdots
\end{aligned}$$

with the dots denoting terms of higher order. By the result quoted in the proof of Lemma 5, the NFIM up to degree two on $y_1 = 0$ is given by $\dot{y}_3 = k_2 y_3^2 + \cdots$, and all assertions follow. $\qquad\square$

We note that Lemma 4 is thus proven. We turn to the Propositions.

Proof of Proposition 2. By Poincaré-Bendixson and Lemma 1, the omega limit set of every solution starting in the first quadrant must contain a stationary point. By Lemma 6, P_2 is a saddle-node with a repelling node part in the upper hemisphere, so P_2 cannot be an omega limit point. By Lemma 5 the point P_1 is a saddle with stable manifold on the equator, therefore an omega limit set containing P_1 cannot consist of P_1 alone. By the Butler-McGehee theorem (see e.g. Smith and Waltman [43]), the omega limit set has nonempty intersection with the stable manifold of the saddle, and by invariance and closedness it must contain P_2 or P_3; a contradiction since both these points are repelling. So, only P_0 remains, and by attractivity of P_0 the solution converges toward this point. The last two assertions are concerned with the center-unstable manifold of P_1, and are a consequence of Lemma 5(b), since dehomogenizing $x_2 - e_T x_3$ with respect to x_3 yields $x_2 - e_T$. $\qquad\square$

Proof of Proposition 3. By Lemmas 5 and 6, and by properties of antipodal points for systems of even degree, the stationary points at infinity are P_1 (attracting node), its antipode (repelling node), P_2 (repelling node part for upper hemisphere), its antipode (saddle part for the upper hemisphere, with stable manifold on the equator), P_3 (repelling node) and its antipode (attracting node).

The first two statements follow from Lemma 5 and Poincaré-Bendixson theory similarly to the previous case. For the last statement, consider the two local components of the unstable manifold of P_0. Such a component cannot connect to a component of the stable manifold, since the existence of a homoclinic orbit would imply the existence of a further stationary point. Therefore the omega limit set of a point on the unstable manifold must contain a different stationary point. This point cannot be a repelling node, which excludes P_3 or the antipode of P_1, and it cannot be P_2, which is repelling for the upper hemisphere. Finally, it cannot be the antipode of P_2 with saddle part in the upper hemisphere, by arguments analogous to those in the previous proof, invoking the Butler-McGehee theorem. Therefore, the omega limit set of a point in the unstable manifold of P_0 contains a single point, which is either P_1 or the antipode of P_3. Finally, not both local components of the unstable manifold can have the same point as omega limit point; this again would imply the existence of a further stationary point.

$\qquad\square$

Remark 5. In case $k_2 e_T = k_0$, system (2.4) admits no finite stationary point, and for the sake of completeness we record the pertinent result about this setting. By routine (albeit lengthy) computations

one finds that in case $k_2 e_T = k_0$, $k_1 \neq 0$ and $k_2 \neq 0$ the NFIM of (B.2) at 0 up to degree four is given by

$$\dot{y}_3 = -\frac{(k_{-1} + k_2)k_2 e_T}{k_1}y_3^4 + \cdots,$$

hence the stationary point P_1 is a degenerate attracting node. In the global picture, P_1 thus attracts all solutions starting in the first quadrant.