



Published in final edited form as:

Math Biosci. 2020 July ; 325: 108339. doi:10.1016/j.mbs.2020.108339.

The Quasi-Steady-State Approximations revisited: Timescales, small parameters, singularities, and normal forms in enzyme kinetics

Justin Eilertsen^a, Santiago Schnell^{a,b,1}

^aDepartment of Molecular & Integrative Physiology, University of Michigan Medical School, Ann Arbor, MI 48109, USA

^bDepartment of Computational Medicine & Bioinformatics, University of Michigan Medical School, Ann Arbor, MI 48109, USA

Abstract

In this work, we revisit the scaling analysis and commonly accepted conditions for the validity of the standard, reverse and total quasi-steady-state approximations through the lens of dimensional Tikhonov–Fenichel parameters and their respective critical manifolds. By combining Tikhonov–Fenichel parameters with scaling analysis and energy methods, we derive improved upper bounds on the approximation error for the standard, reverse and total quasi-steady-state approximations. Furthermore, previous analyses suggest that the reverse quasi-steady-state approximation is only valid when initial enzyme concentrations greatly exceed initial substrate concentrations. However, our results indicate that this approximation can be valid when initial enzyme and substrate concentrations are of equal magnitude. Using energy methods, we find that the condition for the validity of the reverse quasi-steady-state approximation is far less restrictive than was previously assumed, and we derive a new “small” parameter that determines the validity of this approximation. In doing so, we extend the established domain of validity for the reverse quasi-steady-state approximation. Consequently, this opens up the possibility of utilizing the reverse quasi-steady-state approximation to model enzyme catalyzed reactions and estimate kinetic parameters in enzymatic assays at much lower enzyme to substrate ratios than was previously thought. Moreover, we show for the first time that the critical manifold of the reverse quasi-steady-state approximation contains a singular point where normal hyperbolicity is lost. Associated with this singularity is a transcritical bifurcation, and the corresponding normal form of this bifurcation is recovered through scaling analysis.

¹Corresponding author schnells@umich.edu (Santiago Schnell).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Keywords

Enzyme Kinetics; Singular Perturbation; Quasi-Steady-State; Singularity; Normal Hyperbolicity; Transcritical Bifurcation

1. Introduction

Perhaps the most well-known reaction in biochemistry is the Michaelis-Menten (MM) reaction mechanism, (1), which describes the catalytic conversion of a substrate, S , into a product, P . Catalysis is achieved through means of an enzyme, E , that reversibly binds with the substrate, forming an intermediate complex, C . In turn, C irreversibly disassociates into enzyme and product molecules:



For theorists and applied mathematicians, whose aim is to mathematically describe the dynamics of chemical networks and metabolic pathways [1, 2], the MM reaction mechanism serves as a building block to these more complex systems. In general, there are two ways to mathematically model enzyme catalyzed reactions. At low concentrations of chemical species stochastic models are generally favorable since they describe the seemingly random collisions of reactant molecules in intracellular environments [3, 4]. In contrast, when concentrations are high and the chemical species are well-mixed, the MM mechanism can be appropriately modeled as a set of nonlinear ordinary differential equations. Although one can argue that deterministic models are in some sense more manageable than stochastic models, nonlinear deterministic models rarely admit closed form solutions, and therefore model reduction techniques must be employed in order to simplify the model equations so that approximate solutions can be obtained and analyzed. Typically, model reduction is synonymous with approximating the model dynamics on an invariant manifold, and the advent of powerful computer algebra systems aids in the systematic reduction of high-dimensional or even infinite-dimensional (i.e., partial differential equations and delay differential equations) dynamical systems. Not surprisingly, there is a large body of literature that clearly illustrates how model reduction can uncover, quantify, and explain the various nonlinear phenomena arising not only in chemistry, but also in biology [5, 6, 7] and physics [8, 9, 10, 11].

Historically, the most widely-utilized reduction technique in deterministic enzyme kinetics has been the singular perturbation method [12], in which a reduced model is constructed by approximating the flow of the model equations on a *slow invariant manifold* (SIM). Since the dimension of the SIM is less than the dimension of the phase-space, approximation of the dynamics on the SIM permits a reduction in the dimension of the problem. The singular perturbation method exploits the presence of disparate fast and slow timescales; the work of Tikhonov [13] and Gradshteyn [14] provides the necessary rigorous foundation for the construction of a reduced model. Briefly, when fast and slow timescales are present, the

differential equations that model the MM reaction with a timescale separation can be expressed in the form

$$\dot{x} = f(x, y; \epsilon), \quad (2a)$$

$$\epsilon \dot{y} = g(x, y; \epsilon), \quad (2b)$$

where $0 < \epsilon \ll 1$. Setting $\epsilon = 0$ yields a differential-algebraic-equation

$$\dot{x} = f(x, y; 0), \quad (3a)$$

$$0 = g(x, y; 0), \quad (3b)$$

and, according to Tikhonov's theorem, equations (3a)-(3b) provide a very good approximation to the dynamics of (2a)-(2b) when ϵ is sufficiently small. The reduced system (3a)-(3b) provides a simpler, and often times more tractable, reduced mathematical model that is commonly referred to as a *quasi-steady-state approximation* (QSSA).

The hope is that the condition that supports the validity of (3a)-(3b) (i.e., $\epsilon \ll 1$) is easy to implement in enzymatic assays, so that precise and accurate measurements of the kinetic parameters pertinent to the reaction can be made by fitting the QSSA model to experimental time course data [15, 16]. Due to the necessity of disparate timescales, slow manifold reduction may not be possible in every physical scenario. Therefore, the challenge for theorists is not only to derive a reduced model that has suitable utility, but also to determine the unique physical and chemical conditions that *permit* the validity of the associated reduction. Thus, an important task of the theoretician is determine the exact conditions for which the reduced model is valid [17]. Mathematically, this translates to determining “ ϵ ,” a (typically) dimensionless parameter that may not be unique. The non-uniqueness of “ ϵ ” adds complication, since some “epsilons” are better than others. The best-known example of the “non-uniqueness dilemma” resides the history of the derivation of the Michaelis–Menten (MM) equation

$$v = \frac{V_s}{K_M + s}, \quad (4)$$

which is obtained by applying the QSSA to the MM reaction mechanism (1). In (4), v is the velocity of product formation in the reaction, V is the limiting rate of the reaction, K_M is the Michaelis constant, and s is the free substrate concentration for the reaction. Alternatively, the MM equation is often referred to as the *standard* quasi-steady-state approximation (sQSSA). In 1967, Heineken et al. [18] formally applied, for the first time, the standard QSSA to the nonlinear differential equations governing the MM reaction mechanism (1) via singular perturbation analysis. Based on the findings of Laidler [19], a pioneer in chemical kinetics and authority on the physical chemistry of enzymes, Heineken et al. [18] introduced a consistent “ ϵ ” for the MM reaction mechanism through scaling analysis. A more rigorous analysis of the sQSSA was introduced by Reich and Sel'kov [20] and Schauer and Heinrich [21], from which other “epsilons” were determined by proposing conditions that minimized

the errors in the implementation of the sQSSA. In 1988, Lee A. Segel [22] derived the widely accepted criterion for the validity of the sQSSA and the derivation of the MM equation (4) by estimating the slow and fast timescales of the reaction. Segel [22] illustrated that prior physico-chemical knowledge about the reaction dynamic is instrumental in deriving the fast and slow timescales and uncovering the most general criteria for the validity of the sQSSA. As a direct result of Segel's scaling method, the conditions for the validity of the sQSSA were derived for suicide substrates [23, 24], alternative substrates [25], fully competitive enzyme reactions [26], zymogen activation [27], and coupled enzyme assays [28, 29]. Segel's scaling approach was also applied to the analysis of the MM reaction mechanism (1) to extend the validity of the QSSA in different regions of the initial enzyme and substrate concentration parameter space via the reverse QSSA (rQSSA) [30, 31], and the total QSSA (tQSSA) [32, 33, 34].

Despite the power of Segel's scaling and simplification analysis for the MM reaction mechanism (1), there is still a fundamental challenge with its implementation via the rQSSA: there has never been a small parameter (i.e., a specific " ϵ "), analogous to the one obtained by Segel [22] for the sQSSA, that is as effective at determining when the rQSSA is valid. Unfortunately, estimating the fast timescale associated with the rQSSA is difficult, and there have been fundamental disagreements in the reported estimates [31]. Thus, timescale estimation continues to be the "Achilles' heel" of the rQSSA. This raises the obvious question of whether or not timescale estimation is truly the best approach towards resolving this problem.

Recently, Walcher and his collaborators [35, 36, 37] demonstrated that identifying a Tikhonov–Fenichel parameter (TFP) is an effective way to determine a priori conditions that suggest the validity of the QSSA. Essentially, a TFP is a dimensional parameter – such as a rate constant or initial concentration of a species – that, when identically zero, ensures the existence of a manifold of equilibrium points. Such manifolds are central to *geometric singular perturbation theory* (GSPT) and, as a result of Fenichel [38], it is well-understood that when certain conditions hold, the existence of a critical manifold of equilibria ensures the existence of a SIM once the TFP is small but non-zero. In this sense the *origin* of the SIM can be linked to the vanishing of a specific dimensional parameter, and a sufficiently small TFP ensures the existence of a SIM and the corresponding validity of a QSSA. However, the identification of a TFP does not diminish the importance of the asymptotic small parameter ϵ , since it is essential to define what physically constitutes "small" when a parameter is non-zero. In other words, we must still answer the question: *how* small should Tikhonov–Fenichel parameters be in comparison to other dimensional parameters in order to yield an accurate reduced model?

In this paper, our primary objective is to determine a specific small parameter that determines the validity of the rQSSA to the MM reaction mechanism (1), but also to convey subject matter that can be quite technical in a language capable over reaching a wider audience that is not limited to applied mathematicians and physical chemists. In Section 2, we review the conditions for the validity of the various quasi-steady-state approximations that are commonly employed to approximate the long-time dynamics of the MM reaction mechanism (1): namely, the sQSSA, the rQSSA, and tQSSA. In Section 3, we assess the

validity of specific QSS reductions by employing geometric singular perturbation theory, and illustrate how assumptions about the validity of the sQSSA based on Segel's timescale separation can lead to erroneous conclusions. In Section 4, we introduce two methods that do not rely on scaling or timescale separation: Tikhonov–Fenichel parameters and energy methods, both of which can be employed to determine the validity of the QSSA. In Section 5, we analyze the validity of the rQSSA using the methods discussed in Section 4, and in Section 6, we discuss timescale separation and the *hierarchy* of small parameters that support the justification of the sQSSA, rQSSA, and the tQSSA. Finally, in our discussion (Section 7), we summarize our results and critique some of the conclusions drawn about the validity of the sQSSA and the rQSSA in the previous analyses of Segel and Slemrod [30], and Schnell and Maini [31], respectively. We also discuss the impact our results will have on experimental assays, and how the methods we utilize can be employed to analyze more complicated reactions and experimental assays.

2. Applying the Quasi-Steady-State Approximations to the Michaelis–Menten reaction mechanism: Scaling and simplification approaches

In this section we review the application of the different versions of the QSSA for the MM reaction mechanism (1), and the mathematical justification for the application of each approximation. We also present the timescales and criterion for the validity of the QSSA originally derived by Segel [22] and Segel and Slemrod [30].

2.1. Heuristic estimation of fast and slow timescales using the standard Quasi-Steady-State Approximation

The mathematical model that describes (deterministically) the reaction mechanism (1) is a set of nonlinear ordinary differential equations,

$$\dot{s} = -k_1(e_0 - c)s + k_{-1}c, \quad (5a)$$

$$\dot{c} = k_1(e_0 - c)s - (k_{-1} + k_2)c, \quad (5b)$$

$$\dot{p} = k_2c. \quad (5c)$$

The lowercase s , c and p denote concentrations of S , C and P , respectively, and “ $\dot{}$ ” denotes differentiation with respect to time. The equation that models the time evolution of the enzyme concentration, e , has been eliminated via the enzyme conservation law, $e+c = e_0$, and note that adding together (5a)-(5c) yields the substrate conservation law, $s+c+p = s_0$. From the model equations (5a)-(5c), we see that the MM reaction mechanism (1) is parametrically controlled by the initial substrate concentration, s_0 , the initial enzyme concentration, e_0 , and the magnitudes of the rate constants k_1 , k_{-1} and k_2 .

Since the model equations (5a)-(5c) are nonlinear, closed form solutions are intractable. However, it is well-established that whenever the quantity, ε_{SS} , is very small,

$$\varepsilon_{SS} = \frac{e_0}{K_M + s_0} \ll 1, \quad (6)$$

equations (5a)-(5c) can be *approximated* by the system of differential-algebraic equations:

$$\dot{s} \sim -\frac{Vs}{K_M + s} \sim -\dot{p}, \quad c \sim \frac{e_0 s}{K_M + s}. \quad (7)$$

Here the Michaelis constant is defined as $K_M \equiv (k_{-1} + k_2)/k_1$, and the limiting rate is $V \equiv k_2 e_0$. Formally, the equations (7) are collectively referred to as the sQSSA, and condition (6) is known as the *reactant-stationary-assumption* (RSA) [39]. When the RSA holds, and the sQSSA (7) is valid, the intermediate complex reaches its maximum value very quickly. By comparison, the conversion of S to P is slow, and the MM reaction mechanism (1) is characterized (temporally) by the two disparate timescales, t_C and $t_{\mathcal{D}}$, that respectively quantify the approximate amount of time it takes c to become maximal, and for s to significantly deplete. Thus, when the RSA holds, the reaction consists of a fast timescale, t_C , and a slow timescale, $t_{\mathcal{D}}$. Both timescales depend on the rate constants, as well as the initial substrate and enzyme concentrations:

$$t_C = \frac{1}{k_1(s_0 + K_M)}, \quad (8a)$$

$$t_{\mathcal{D}} = \frac{K_M + s_0}{V}. \quad (8b)$$

Segel [22] first derived t_C and $t_{\mathcal{D}}$ using heuristic methods. To estimate t_C , Segel noted, no doubt from the earlier work of Heineken, Tsuchiya, and Aris [18], that if $e_0 \ll s_0$, then c should reach its maximum value very quickly, and there should be correspondingly very little substrate depletion during the initial accumulation of c . Consequently, one can assume $s \approx s_0$ during the initial accumulation of c . To obtain an equation that models the initial increase in c , we set $s = s_0$ in (5b), which gives rise to a linear ordinary differential equation for c :

$$\dot{c} \sim k_1(e_0 - c)s_0 - (k_{-1} + k_2)c. \quad (9)$$

Since (9) is linear, its solution is easily attainable, and is given by:

$$c(t) \sim \varepsilon_{SS} s_0 (1 - e^{-t/t_C}). \quad (10)$$

Thus, t_C is the *characteristic* timescale of the initial fast transient, and it is valid as long as s is approximately constant during the initial accumulation of c . The reader should be aware of the fact that (9) is only valid for timescales on the order of t_C . Hence, the solution (10) is

only an initial, temporary approximation to the kinetics of (1) when the RSA is valid. As such, it is often referred to as an *initial layer*.

Since t_C was derived under the presupposition of a negligible depletion of s during the initial accumulation of c , how then, can we determine a dimensionless parameter that must be small in order to ensure the depletion of s is in fact negligible over t_C ? Segel [22] reasoned that since the maximum rate of depletion of s is identically $-k_1 e_0 s_0$, it suffices to demand that

$$\{\max |\dot{s}| \cdot t_C \ll s_0\} \equiv \{\epsilon_{SS} \ll 1\} \quad (11)$$

hold in order to justify the use of (9). Thus, the RSA (6) ensures that the approximation (9) is valid.

After c reaches its maximum value, the QSS phase of the reaction ensues, and c starts to slowly deplete. During the depletion phase \dot{c} is very small but non-zero, and thus we say that c evolves in a QSS or is “slaved” by s since, according to the system (7), the approximate concentration of c is parametrically determined by the concentration of s . To estimate the depletion timescale, Segel [22] assumed that the sQSSA (7) was valid once $t_C \lesssim t$, and that the rate of depletion of substrate was approximately

$$\dot{s} \sim -\frac{V_s}{K_M + s} \leq 0. \quad (12)$$

Since $s \approx s_0$ for $t \lesssim t_C$, it stands to reason that the sQSSA (12) can be supplied with the boundary condition $s(t_C) = s_0$, and thus the maximum rate of depletion in the QSS phase of the reaction is approximately

$$\max |\dot{s}| \approx \frac{V_s s_0}{K_M + s_0}. \quad (13)$$

Finally, to calculate the depletion timescale, Segel [22] divided the total change in substrate over the course of the reaction, $s = s_0$, by the maximum rate of depletion in the QSS phase (13):

$$t_{\mathcal{D}} = \frac{|\Delta s|}{\max |\dot{s}|} = \frac{s_0}{\frac{V_s s_0}{K_M + s_0}} = \frac{K_M + s_0}{V}. \quad (14)$$

As noted in the introduction, the effective use of the singular perturbation method relies on separation of fast and slow timescales, and the question that naturally arises from the heuristic line of reasoning arises is: Does the RSA (6) ensure separation of timescales? That is, does it hold that $t_C \ll t_{\mathcal{D}}$ whenever $\epsilon_{SS} \ll 1$? The answer is yes: the RSA not only ensures that there is negligible formation of product (or depletion of substrate) for $t \lesssim t_C$, it also ensures that the timescales, t_C and $t_{\mathcal{D}}$, are widely separated:

$$\frac{t_C}{t_{\mathcal{D}}} \equiv \epsilon = \frac{\epsilon_{SS}}{\left(1 + \frac{k_{-1}}{k_2}\right)\left(1 + \frac{s_0}{K_M}\right)} \leq \epsilon_{SS}. \quad (15)$$

Thus, the RSA (i.e., the condition that $\epsilon_{SS} \ll 1$) induces the separation of fast and slow timescales. Consequently, the RSA is understood to be a sufficient condition for the validity of the sQSSA (7), and the separation of fast and slow timescales (i.e., $\epsilon \ll 1$) is a necessary (but not sufficient) condition for the sQSSA.

When the RSA is valid, the trajectories in the (s, c) -plane have a very recognizable form: during fast transient phase, the trajectories move almost vertically towards the c -nullcline, given by the curve $\{(s, c) \in \mathbb{R}^2: c - e_0 s / (K_M + s) = 0\}$, on which $\dot{c} = 0$. The intermediate complex c reaches its maximum value once the trajectory reaches the c -nullcline, at which point the QSS phase begins and the sQSSA is valid. After c reaches its maximum value, the trajectory closely follows the c -nullcline towards the equilibrium point $(s, c) = (0, 0)$ (see FIGURE 1).

The fact that the phase-plane trajectory follows the c -nullcline on the slow timescale, and approaches the c -nullcline in vertical fashion over the fast timescale is the hallmark of the sQSSA. While the trajectory in FIGURE 1 appears as though it is *on* the c -nullcline, it is important to note that it is not. In fact, the trajectory approaches an invariant manifold, \mathcal{M}^ϵ , that lies just above the c -nullcline. However, since \mathcal{M}^ϵ and the c -nullcline are very close together when (6) holds, the c -nullcline is used to *approximate* \mathcal{M}^ϵ (again, see figure 1 for a detailed explanation and illustration). The accuracy of using the c -nullcline as an approximation is assessed asymptotically through scaling and non-dimensionalization methods, and we review these methods in the subsection that follows.

2.2. Asymptotic justification for the sQSSA: Scaling and non-dimensionalization

The identification of the small parameter (6) as an appropriate condition for the validity of the standard QSSA is justified through scaling and non-dimensionalization. Introducing the dimensionless variables $\tau \equiv t / t_C$, $\bar{s} \equiv s / s_0$ and $\bar{c} \equiv c / \epsilon_{SS} s_0$ yields

$$\frac{d\bar{s}}{d\tau} = \epsilon_{SS}[\mu\bar{c}\bar{s} - \bar{s} + \alpha\beta\bar{c}], \quad (16a)$$

$$\frac{d\bar{c}}{d\tau} = \bar{s} - \mu\bar{c}\bar{s} - \beta\bar{c}, \quad (16b)$$

where $\beta \equiv 1/(1 + \sigma)$, $\sigma \equiv s_0/K_M$, $\kappa \equiv k_{-1}/k_2$, $\mu \equiv 1 - \beta$ and $\alpha \equiv \kappa/(1 + \kappa)$. Thus, we see from (16a) and (16b) that if $\epsilon_{SS} \ll 1$, then $s \approx s_0$ during the transient phase. Over the slow timescale, $T \equiv \epsilon_{SS}\tau$, we obtain

$$\frac{d\bar{s}}{dT} = \mu\bar{c}\bar{s} - \bar{s} + \alpha\beta\bar{c}, \quad (17a)$$

$$\epsilon_{SS} \frac{d\bar{c}}{dT} = \bar{s} - \mu\bar{c}\bar{s} - \beta\bar{c}. \quad (17b)$$

Formally, the standard QSSA is an *asymptotic* approximation of the dynamics on the slow timescale obtained by setting $\epsilon_{SS} = 0$ in (17b) and solving for c in terms of s . Thus, the standard QSSA is the zeroth-order approximation to (17a)-(17b) and, when $\epsilon_{SS} \ll 1$, we refer to s as a *slow* variable, and c as a *fast* variable.

In summary, we have two takeaways from this subsection:

- i. The RSA is synonymous with $\epsilon_{SS} \ll 1$, and guarantees that there is a negligible loss of substrate during the transient (fast) phase of the reaction. When the RSA is valid, the QSSA system (7) is an appropriate asymptotic approximation to (5a)-(5c).
- ii. Timescale separation is synonymous with $\epsilon \equiv t_C / t_{\mathcal{D}} \ll 1$.

While (i) certainly implies (ii), the converse does not hold, which begs the question: What happens when $\epsilon \ll 1$ but $\epsilon_{SS} \sim 1$? We discuss this special case in the subsection that follows.

2.3. The extended Quasi-Steady-State Approximation

The RSA and timescale separation provide two different small parameters: ϵ_{SS} and $\epsilon \equiv t_C / t_{\mathcal{D}}$. While $\epsilon_{SS} \ll 1$ ensures that $\epsilon \ll 1$, the converse is not true. The obvious question is: Do QSS dynamics still prevail when $\epsilon_{SS} \sim 1$ but $t_C / t_{\mathcal{D}} \ll 1$? Segel and Slemrod [30] proposed that the sQSSA (7) should still be valid at some point in the time course of the reaction as long as $t_C \ll t_{\mathcal{D}}$, even if $\epsilon_{SS} \sim 1$. Rescaling the mass action equations (5a)-(5b) with respect to $\bar{T} = t / t_{\mathcal{D}}$ yields

$$\nu\beta \frac{d\bar{s}}{d\bar{T}} = \mu\bar{c}\bar{s} - \bar{s} + \alpha\beta\bar{c}, \quad (18a)$$

$$\nu\beta\epsilon_{SS} \frac{d\bar{c}}{d\bar{T}} = \bar{s} - \mu\bar{c}\bar{s} - \beta\bar{c}, \quad (18b)$$

where again, $\nu = k_2 / (k_{-1} + k_2)$, and $\beta = K_M / (K_M + s_0)$. If $\epsilon_{SS} = 1$, then setting $\nu = 0$ (which implies $\alpha = 1$) in (18a)-(18b) yields a (dimensional) manifold:

$$c = \frac{e_0 s}{K_S + s}, \quad (19)$$

where $K_S = k_{-1} / k_1$ is the enzyme-substrate dissociation constant. Interestingly, the manifold (19) is identical to the s -nullcline. However, since the c -nullcline and the s -nullcline are indistinguishable in the limiting case that corresponds to $k_2 = 0$, it stands to reason that trajectories should closely follow both the s - and c -nullcline when k_2 is incredibly small, but non-zero. Numerical simulations confirm that phase-plane trajectories closely follow the c -nullcline after a brief transient (see, Figure 2).

The c -nullcline “following” that occurs when $\nu \ll 1$ and $\epsilon_{SS}, \beta \sim 1$ is defined by Segel and Slemrod [30] as the *extended* QSSA. One obvious difference between the sQSSA and the extended QSSA is that there is a noticeable amount of substrate depletion when $\nu \ll 1$ but $\epsilon_{SS} \sim 1$. Segel and Slemrod [30] proposed that the QSSA system (7) should still hold after the transient, but that the primary difference between the sQSSA and the extended QSSA was that one could not take $s = s_0$ at the onset of the QSS phase in the case of the extended QSSA. Thus, Segel and Slemrod [30] proposed that, when the extended QSSA is valid but the RSA is invalid, the approximation

$$\dot{s} \sim -\frac{Vs}{K_M + s}, \quad (20)$$

must be supplied with a boundary condition “ s_0^e ” that is less than s_0 . They employed a graphical method to estimate s_0^e in their original manuscript [30], and found that when $\beta = \epsilon = 1$, $s_0^e \approx (\sqrt{2} - 1)s_0$. Thus, they concluded that the QSS phase of the reaction could be approximated with the initial value problem:

$$\dot{s} \sim -\frac{Vs}{K_M + s}, \quad s(0) = (\sqrt{2} - 1)s_0. \quad (21)$$

Moreover, they found that $s_0^e \rightarrow 0$ as β increases, and determined that the parameter regime where $\nu\beta \ll 1$ but $\epsilon_{SS} \sim 1$ results in a rather “uninteresting” domain of applicability for extended QSSA². Hence, the extended QSSA is formally defined to be the case when $\nu \ll 1$, but $\epsilon_{SS}, \beta \sim 1$. Up to this point, we have two dynamical regimes of interest:

- i. The RSA, in which $\epsilon_{SS} \ll 1$ and (7) holds.
- ii. The extended QSSA, in which $\nu \ll 1$, but $\epsilon_{SS}, \beta \sim 1$ and *possibly* (21) holds.

As a final remark of this subsection, we note that while the extended QSSA seems reasonable, there are a few technical deficiencies that engulf the validity of the approximation. First, setting $\nu = 0$ results in the coalescence of the s - and c -nullclines, which implies that s and c are in some sense both in a QSS after the fast transient. Second, there is no scaling justification for (21). Consequently, other than the observation that the phase-plane trajectory follows the c -nullcline after the transient phase, there is really no rigorous justification for (21). We will return to these observations and discuss them in more detail in a later section.

2.4. The reverse Quasi-Steady-State Approximation

In contrast to the intermediate complex evolving in a QSS after the fast transient, the rQSSA occurs when the substrate evolves in QSS after the initial fast transient. Nguyen and Fraser

²Segel and Slemrod [30] did not take into account product formation, and assumed that the complete depletion of substrate was synonymous with the completion of the reaction. Unfortunately, this assumption is false. For more details, we invite the reader to consult [29].

[40] observed that the rQSSA corresponds to trajectories closely following the s -nullcline instead of the c -nullcline in the s - c phase-plane (see Figure 3).

To formulate an approximation to (5a)-(5c), it is assumed that virtually all of the initial substrate will deplete during the initial transient phase, in which case

$$\frac{dc}{dt} + \frac{dp}{dt} = 0 \quad (22)$$

approximately holds during the QSS phase. It is also assumed that there is negligible formation of product during the transient phase. Thus, in the rQSSA, the underlying assumption is that there is a preliminary transient phase in which nearly all of the substrate is depleted, but a negligible amount of product is generated. As a consequence of this assumption, (22) admits an approximate conservation law, $s_0 \sim c + p$, that implies

$$\frac{dp}{dt} = k_2(s_0 - p) \quad (23)$$

is the leading order approximation to \dot{p} in the QSS phase of the reaction. The above expression governs the production formation under the rQSSA or *rapid equilibrium approximation* [41, 42]. Equation (23) is linear and admits a closed-form solution

$$p(t) = s_0(1 - e^{-k_2 t}), \quad (24)$$

as well as a corresponding slow timescale: $\tilde{T} = k_2 t$.

Is there a small parameter homologous to ϵ_{SS} for the rQSSA that justifies the validity of (23)? Typically, less-than-rigorous heuristic arguments are employed to determine conditions for the validity of (23). By utilizing timescale and geometrical arguments, Segel and Slemrod [30] *indirectly* proposed³

$$\epsilon^* \equiv \frac{K_M}{e_0} = \eta^{-1} \ll 1 \quad (25)$$

as a necessary condition for the validity of the rQSSA. However, there are several problems with labeling (25) as a *sufficient* condition. First, as Schnell and Maini [31] point out, $\epsilon^* \ll 1$ does not ensure that $d s/d t$ vanishes after the initial transient, since $K_M \ll e_0 \ll s_0$ would seem to imply the validity of the QSSA, and Segel and Slemrod [30] only considered cases where $e_0 \gg s_0$ in their analysis of the rQSSA. Schnell and Maini [31] suggested that there should be both a negligible loss of enzyme concentration and a maximal depletion of free substrate over the fast transient; they proposed

$$\epsilon_{SM} \equiv \eta^{-1} + \frac{s_0}{e_0} = \frac{1}{\epsilon_{SS}} \ll 1 \quad (26)$$

³They directly proposed that $K_S/e_0 \ll 1$, but noted that $k_2 \approx k_{-1}$, which implies $K_M/e_0 \ll 1$.

as an appropriate qualifier for the validity of the rQSSA. The condition (26) suggests that since e is in excess with respect to s , the majority of the substrate should vanish over the fast transient. Second, from a more technical perspective, Goeke et al. [43] remark that $K_s/e_0 \ll 1$ is not sufficient for the validity of the rQSSA; we will say more about Goeke's observation in the section that follows. For now we note that, heuristically, if $1/k_2$ is a slow timescale, then it stands to reason that k_2 should also be small in some sense.

In short, most of the conditions in the literature *basically* imply that $\epsilon_{SS} \gg 1$ in order for the rQSSA to be valid. In this work, the question we want to address through our analysis is: How small should k_{-1} and k_2 be in comparison to other variables to ensure the validity of the rQSSA (i.e., what is the equivalent of ϵ_{SS} in this case)? Clearly, it must hold that $\epsilon^* \ll 1$, but how does the initial concentration of substrate, s_0 , impact the accuracy of the approximation (23)? This is still an open problem and, using non-scaling methods, we will close it in a later section.

2.5. The total Quasi-Steady-State Approximation

The tQSSA provides an approximation that is valid over a broader parameter domain. Utilizing the conservation law, $s_0 = s + c + p$, and replacing $s = s_0 - (c + p)$ in (5b) yields

$$\dot{c} = k_1(e_0 - c)(s_0 - c - p) - k_1 K_M c, \quad (27a)$$

$$\dot{p} = k_2 c. \quad (27b)$$

Setting the left hand side of (27a) to zero and solving for c gives

$$c = \frac{1}{2}(e_0 + K_M + s_0 - p) \pm \frac{1}{2}\sqrt{(e_0 + K_M + s_0 - p)^2 - 4e_0(s_0 - p)}. \quad (28)$$

Of the two possible roots given in (28), only the “−” root is *physical*. Taking

$$c \approx \frac{1}{2}(e_0 + K_M + s_0 - p) - \frac{1}{2}\sqrt{(e_0 + K_M + s_0 - p)^2 - 4e_0(s_0 - p)} \equiv c_0(p), \quad (29)$$

and inserting it into (29) yields the tQSSA:

$$\dot{p} \approx k_2 c_0(p). \quad (30)$$

Thus, the complex is taken to be the fast variable that evolves in a QSS after a brief transient, and p , rather than s , is the corresponding slow variable. Due to the fact that p is the designated slow variable, the tQSSA is valid over much larger domain than either the sQSSA or the rQSSA. As a result of the more recent work of Borghans et al. [32], the term $s_0 - p$ is usually replaced by $s_T = s + c$ in the literature (see, for example, [33, 44, 45, 46, 47, 48, 16]), where s_T is a lumped variable that denotes the “total” substrate. Consequently, the equations

$$\dot{c} = k_1(e_0 - c)(s_T - c) - k_1 K_M c, \quad (31a)$$

$$\dot{s}_T = -k_2 c, \quad (31b)$$

are more commonly employed than (27a)-(27b). Either way, as Pedersen [49] points out, the coordinate systems (s_T, c) and (p, c) are merely “two sides of the same coin”. We will employ the (p, c) coordinate system, and avoid the lumped variable “ s_T ” in this paper as it is not a uniquely distinguishable physical variable in experimental assays.

The validity of the tQSSA was first analyzed in the (p, c) coordinate system by Laidler [19], and later by Borghans et al. [32]. However, the accepted condition for the validity of the tQSSA was derived by Tzafirri [34]. Briefly, Tzafirri noted that if there is zero formation of product during the initial transient, then the formation of c during the transient phase can be approximated by the Riccati equation

$$\dot{c} = k_1(e_0 - c)(s_0 - c) - k_1 K_M c. \quad (32)$$

The solution to (32) admits a natural *fast* timescale:

$$t_C^* = \frac{1}{k_1 \sqrt{(e_0 + K_M + s_0)^2 - 4e_0 s_0}}. \quad (33)$$

The corresponding *slow* timescale, t_P is calculated from the method of Segel [22]

$$t_P = \frac{|\Delta p|}{\max |k_2 c_0(p)|} = \frac{2s_0}{(e_0 + K_M + s_0) - \sqrt{(e_0 + K_M + s_0)^2 - 4e_0 s_0}}, \quad (34)$$

and the separation of t_C^* and t_P is the accepted condition for validity of the tQSSA:

$$\varepsilon_T \equiv \frac{t_C^*}{t_P} = \frac{(e_0 + K_M + s_0) - \sqrt{(e_0 + K_M + s_0)^2 - 4e_0 s_0}}{2s_0 k_1 \sqrt{(e_0 + K_M + s_0)^2 - 4e_0 s_0}} \ll 1. \quad (35)$$

Several scaling analyses have been carried out that justify the asymptotic validity of the tQSSA, and we refer the reader to Schnell and Maini [33] and Dell’Acqua and Bersani [47] for detailed perturbation analyses, and Bersani et al. [50] for a thorough review of the topic. Since the tQSSA is valid whenever there is negligible formation of product during the fast transient, the validity of the sQSSA, as well as the rQSSA and extended QSSA, implies the validity of the tQSSA. In this sense, the tQSSA *contains* both the standard QSSA and the rQSSA. Thus, $\varepsilon_{SS}, \nu \ll 1 \Rightarrow \varepsilon_T \ll 1$, but $1 \ll \varepsilon_{SS}$ also implies $\varepsilon_T \ll 1$ (again, see [32]).

It is worth noting that the fast and slow timescales derived by Tzafirri are valid whenever there is negligible product formation during the fast transient. Thus, they are valid in both the sQSSA and rQSSA regimes. Additionally, the upper bound on c is easily obtained:

$$\sup c \equiv \lambda = c_0(0) = \frac{1}{2}(e_0 + K_M + s_0) - \frac{1}{2}\sqrt{(e_0 + K_M + s_0)^2 - 4e_0s_0}. \quad (36)$$

Consequently, we obtain universal timescales as well an upper bound on c from the tQSSA formulation. At first glance the timescales t_C^* and t_B as well as the upper bound λ , appear quite complicated, and their utility in the context of scaling analysis appears limited. However, when we introduce non-scaling methods, we will show that in fact these “ingredients” are quite useful.

The tQSSA wraps up our introduction to each “QSSA” (i.e., the sQSSA, rQSSA, and tQSSA) that is employed as a reduced model in enzyme kinetics. In summary, we have:

- i. The RSA is valid when $\varepsilon_{SS} \ll 1$, and the corresponding QSSA is given by (7).
- ii. The extended QSSA is valid when $\nu \ll 1$, but ($\varepsilon_{SS} \sim 1$ and $\beta \sim 1$). The corresponding QSSA is *possibly* given by equation (21).
- iii. The rQSSA is at least valid when $1 \ll \varepsilon_{SS}$, and the corresponding QSSA is given by (23).
- iv. The tQSSA is valid whenever $\varepsilon_T \ll 1$, and the corresponding QSSA is given by (30).

3. Applying the Quasi-Steady State Approximations with geometric singular perturbation theory

The work of Fenichel [38] consists of a group of theorems that warrant the existence of a slow invariant manifold, \mathcal{M} , in fast/slow systems of the form of (2a)-(2b). In Section 3.1, we give a *basic* introduction to the results obtained by Fenichel and introduce relevant terminology for our paper. For a more detailed introduction, we refer the reader to [51, 52] and [53]. In Section 3, we apply Geometric Singular Perturbation Theory (GSPT) to (5a)-(5c) under the assumption that the RSA is valid. In Section 3.3, we apply GSPT to (5a)-(5c) in the parameter-space region where the extended QSSA is valid, and we show how GSPT can be used to prove that Segel and Slemrod’s conclusion that the approximation given (21) is valid after the initial transient is incorrect.

3.1. The invariance equation

The sQSSA, rQSSA, and the tQSSA are, formally, zeroth-order approximations to the reaction dynamics on the corresponding slow timescale. However, there is much more to the story. Whenever the tQSSA is valid, trajectories are attracted to a *slow* invariant manifold (SIM). The rigorous study of SIMs is referred to as geometric singular perturbation theory (GSPT). Briefly, given a system of the form

$$\dot{x} = f(x, y; \varepsilon), \quad (37a)$$

$$\varepsilon \dot{y} = g(x, y; \varepsilon), \quad (37b)$$

where x is a slow variable and y is a fast variable, one can approximate the slow manifold by first assuming that the fast variable is expressible in terms of the slow variable as $y = h(x)$ ⁴, where h is a function that is to be determined. Typically, the curve $y = h(x)$ is denoted as \mathcal{M}^ε , and we will use this notation from here on. Since the slow manifold is invariant, it must satisfy the differential equation (37a)-(37b). Hence, $y = h(x)$ must satisfy

$$\frac{dh}{dx} \cdot f(x, h(x); \varepsilon) = \frac{1}{\varepsilon} g(x, h(x); \varepsilon), \quad (38)$$

which is known as the *invariance equation*. The power of the invariance equation is that, given a scaled equation and an appropriate small parameter, it allows us to systematically determine the higher-order asymptotic approximations to the SIM. The function $h(x)$ admits a unique asymptotic expansion in terms of ε :

$$h(x) = h_0(x) + \varepsilon h_1(x) + \varepsilon^2 h_2(x) + \mathcal{O}(\varepsilon^3), \quad (39)$$

where $h_0(x)$ satisfies $g(y, h_0(x); 0) = 0$. Thus, when $g(x, y)$ is not explicitly dependent on ε , the nullcline associated with the fast variable commonly serves as a zeroth-order approximation to the slow manifold, $\mathcal{M}_0^\varepsilon = \{(x, y) \in \mathbb{R}^2 : y = h_0(x)\}$, and formally the asymptotic approximation to the slow manifold can be expressed as:

$$\mathcal{M}^\varepsilon = \mathcal{M}_0^\varepsilon + \varepsilon \mathcal{M}_1^\varepsilon + \varepsilon^2 \mathcal{M}_2^\varepsilon + \dots + \mathcal{O}(\varepsilon^N) + \dots \quad (40)$$

The terms $\varepsilon^0, \varepsilon^1, \varepsilon^2, \dots$ in (40) are called *gauge functions*, and the terms $\mathcal{M}_0^\varepsilon, \mathcal{M}_1^\varepsilon, \dots$ are the coefficients that accompany the gauge functions. Generally speaking, the slow manifold is not necessarily *unique*, meaning there can be slow manifolds. However, they are separated by a distance that is $\mathcal{O}(e^{-K/\varepsilon})$ where K is $\mathcal{O}(1)$ [53].

In addition to the slow manifold, the work of Fenichel also guarantees the existence of *fast fibers*, $\mathcal{F}^{\varepsilon, p}$, which can be utilized to approximate the dynamics during the transient stage as the trajectory rapidly approaches the slow manifold.⁵ The zeroth-order fiber, $\mathcal{F}_0^{\varepsilon, p}$, which is the zeroth-order approximation to the fibers $\mathcal{F}^{\varepsilon, p}$, is the solution to the fast subsystem:

$$x' = 0, \quad (41a)$$

$$y' = f(x, y; 0). \quad (41b)$$

⁴The underlying assumption here is that the slow manifold can be expressed in terms of a global coordinate chart. More complicated manifolds, such as an n -dimensional torus, may require an atlas and a collection of transition maps.

⁵In contrast to the critical manifold, the individual zeroth-order fast fibers do not perturb to individual invariant fibers when ε is non-zero. However, the collective family of fast fibers, $\bigcup_p \mathcal{F}^{\varepsilon, p}$, is invariant in a certain way. For more details, we invite the reader to consult [54].

The fast subsystem contains a continuous branch of equilibrium points given by the curve $y = h_0(x)$, which is referred to in this context as the *critical manifold*. Thus, each zeroth-order fiber links an initial condition (x_0, y_0) to a unique *base point* (equilibrium point) $p \in \mathcal{M}_0^\varepsilon$, and the superscript p in $\mathcal{F}^{\varepsilon, p}$ is a reference to the unique base point, p . Since we are only interested in the zeroth-order approximation to $\mathcal{F}^{\varepsilon, p}$, we will henceforth drop the superscript p for convenience, and use “ $\mathcal{F}_0^\varepsilon$ ” to denote an *individual* zeroth-order fiber. We will say more about fast fibers and the perspective of the slow manifold from the context of the fast subsystem when we introduce the idea of Tikhonov-Fenichel parameters in Section 4.2. However, for now we will just associate the zeroth-order fast fiber as being the solution to (41a)-(41b).

3.2. Standard Quasi-Steady-State dynamics: Geometric singular perturbation theory when $\varepsilon_{SS} \ll 1$.

The RSA ensures negligible depletion of substrate during the fast transient, and *suggests*⁶ that the zeroth-order fast fiber, $\mathcal{F}_0^{\varepsilon_{SS}}$, is a straight line that connects the initial condition $(\bar{s}, \bar{c}) = (1, 0)$ to the base point $(\bar{s}, \bar{c}) = (1, 1)$:

$$\mathcal{F}_0^{\varepsilon_{SS}} = \{(\bar{s}, \bar{c}) \in \mathbb{R}^2 : \bar{s} = 1, 0 \leq \bar{c} \leq 1\}. \quad (42)$$

If $0 < \varepsilon_{SS} \ll 1$, then the time it takes the phase-plane trajectory to reach a $\mathcal{O}(\varepsilon_{SS})$ -neighborhood of the slow manifold is $\tau_{\varepsilon_{SS}} \sim \ln \varepsilon_{SS}$ or, in dimensional time $t \sim t_C \ln \varepsilon_{SS}$. Moreover, there exists an invariant manifold, $\mathcal{M}^{\varepsilon_{SS}}$, that is $\mathcal{O}(\varepsilon_{SS})$ from the critical manifold, and admits a unique expansion in powers of ε_{SS} in the form of (40), where $\mathcal{M}_0^{\varepsilon_{SS}}$ is identically the \bar{c} -nullcline. These terms are straightforward to compute by constructing an asymptotic solution to the invariance equation:

$$\frac{1}{\varepsilon_{SS}} g(\mathcal{M}^{\varepsilon_{SS}}, \bar{s}) = D_{\bar{s}} \mathcal{M}^{\varepsilon_{SS}} f(\mathcal{M}^{\varepsilon_{SS}}, \bar{s}), \quad (43)$$

and by inserting (40) into (43), it is easy to recover

$$\mathcal{M}_0^{\varepsilon_{SS}} = \{(\bar{s}, \bar{c}) \in \mathbb{R}_+^2 : \bar{c} - \bar{s}(1 + \sigma) / (1 + \bar{s}\sigma) = 0\}. \quad (44)$$

Finally, inserting the expression for $\mathcal{M}_0^{\varepsilon_{SS}}$ into (17a) yields, in dimensional variables, the MM equation

$$\frac{ds}{dt} \sim -\frac{Vs}{K_M + s}, \quad (45)$$

⁶This is, of course, a violation of the conservation law for substrate in the MM reaction mechanism (1), but we will proceed as if it were permissible, and discuss this observation in more detail in section 4.

which approximates the long-time dynamics of the mass action equations when the RSA is valid.

3.3. Extended Quasi-Steady-State dynamics: Geometric singular perturbation theory when $\nu \ll 1$, $\beta \sim 1$ and $\epsilon \sim 1$.

Now let us consider the case of the extended QSSA. The problem can be reformulated in the (p, c) coordinate system so that it is of the form (37a)-(37b) (see, for example, [18, 55]). However, we will work out the solution in the (s, c) coordinate system. In this case, we can guess the asymptotic expansion of the slow manifold in this coordinate system, and assume it takes the form

$$c = \mathcal{M}^{\tilde{\nu}} = \frac{\bar{s}}{\alpha\beta + \mu\bar{s}} + \nu \mathcal{M}_1^{\tilde{\nu}} + \mathcal{O}(\tilde{\nu}^2), \quad \tilde{\nu} \equiv 1 / \kappa = k_2 / k_{-1}, \quad (46)$$

where the curve

$$\bar{c} = \frac{\bar{s}}{\alpha\beta + \mu\bar{s}}, \quad (47)$$

is the dimensionless form of the s -nullcline (19). Taking just the zeroth-order expansion of (46) (i.e., (47)) and inserting it into equation (18a) while zeroing all terms of order ν or higher yields

$$\frac{d\bar{s}}{dT} = 0. \quad (48)$$

Since the \bar{c} - and \bar{s} -nullclines coalesce when $k_2 = 0$, the zeroth-order approximation to the dynamics on the slow manifold results in the flow being infinitely slow (i.e., an equilibrium solution). Consequently, we must go to first-order in $\tilde{\nu}$ to determine the QSS approximation to the dynamics on the slow manifold. Solving for $\mathcal{M}_1^{\tilde{\nu}}$ by approximating the solution to the invariance equation⁷

$$g(\mathcal{M}^{\tilde{\nu}}, \bar{s}) = D_{\bar{s}} \mathcal{M}^{\tilde{\nu}} f(\mathcal{M}^{\tilde{\nu}}, \bar{s}) \quad (49)$$

in powers of $\tilde{\nu}$, we recover

$$\frac{d\bar{s}}{dT} = \tilde{\nu} \frac{((\bar{s} - 1)\theta + 1)(\theta - 1)\bar{s}}{(\bar{s} - 1)^2\theta^2 + (2\bar{s} - 2\epsilon - 1)\theta + \epsilon + 1} + \mathcal{O}(\tilde{\nu}^2), \quad \theta = \frac{\sigma}{\alpha + \sigma} < 1, \quad (50)$$

by inserting $\mathcal{M}_1^{\tilde{\nu}}$ into (18a), which is the leading-order non-trivial solution to the dynamics on the slow manifold. Thus, the QSS approximation for s (50) is different when $\nu \ll 1$ versus when $\epsilon_{SS} \ll 1$. In dimensional variables, (50) translates to,

⁷ $D_{\bar{s}}$ denotes differentiation with respect to \bar{s} .

$$\frac{ds}{dt} \sim - \frac{V_S(s + K_S)}{e_0 K_S + (s + K_S)^2}, \quad K_S \equiv k_{-1} / k_1. \quad (51)$$

The result (51) is not new, and the approximation that can be found in several papers [56, 57, 58, 43, 55]. The point is that by utilizing the invariance equation, we have shown that in fact Segel and Slemrod's extended QSS approximation (21) is *invalid*. The difference between the QSS approximations for s when $\nu \ll 1$ versus when $\epsilon_{SS} \ll 1$ has a nice history, and Roussel [55] recently published a thorough explanation of the different QSS approximations through the lens of Tikhonov's Theorem and geometric singular perturbation theory.

In addition to the change in leading-order dynamics on the slow manifold when $\nu \ll 1$ instead of $\epsilon_{SS} \ll 1$, the zeroth-order approximation to the fast fiber, \mathcal{F}_0^ν , will also change. One major difference between the transient dynamics of the sQSSA and the extended QSSA is that we can no longer take $s \approx s_0$ during the trajectory's initial approach to the slow manifold. Hence, we must determine an appropriate starting point in order to "match" the inner and outer solutions. Locating an appropriate starting position on a slow manifold is not necessarily a trivial exercise. However, mathematicians have developed methods for such a situation (in [59], A. J. Roberts develops a clever method for estimating starting positions on centre manifolds). Luckily, in our case, the conservation of total substrate, $s_0 = s + c + p$, can be utilized to estimate a starting point. Over the τ -timescale, the total system is given by:

$$\frac{d\bar{s}}{d\tau} = \epsilon_{SS}[\mu\bar{c}\bar{s} - \bar{s} + \alpha\beta\bar{c}], \quad (52a)$$

$$\frac{d\bar{c}}{d\tau} = \bar{s} - \mu\bar{c}\bar{s} - \beta\bar{c}, \quad (52b)$$

$$\frac{d\bar{p}}{d\tau} = \epsilon\bar{c}. \quad (52c)$$

Since $\nu \rightarrow 0$ implies $\epsilon \rightarrow 0$, we set $\epsilon = 0$ and $\alpha = 1$ in (52c) and (52a), respectively. Setting the right hand side of (52c) to zero implies that the *total* substrate, " $\bar{s}_T = \bar{s} + \bar{c}$," is conserved over the transient timescale when $\nu = 0$,

$$\bar{s} + \bar{c} = 1, \quad (53)$$

and insertion of (53) into the scaled equations yields a Riccati equation for \bar{c} :

$$\frac{d\bar{c}}{d\tau} = 1 - 2\bar{c} + \mu\bar{c}^2. \quad (54)$$

Two observations can be made. First, the correction to the duration of the fast transient when $\nu \ll 1$ (as opposed to ϵ_{SS}) can be found by solving the Riccati equation (54). In dimensional

form, the solution to (54) yields (33). Again, this timescale was originally introduced by Tzafiriri [34], and can be taken to be characteristic of the fast transient whenever there is minimal product formation during the time it takes c to reach its maximum value (i.e., even when $\varepsilon_{SS} \ll 1$).⁸ Second, the base point (\bar{c}^*, \bar{s}^*) of the zeroth-order fast fiber is given by⁹

$$\bar{c}^* = \frac{2 - \sqrt{4(1 - \mu)}}{2\mu}, \quad \bar{s}^* = 1 - \bar{c}^*. \quad (55)$$

Consequently, if $\varepsilon_{SS} = \sigma = 1$ then, after the fast transient, we expect $\bar{c} \approx 2 - \sqrt{2}$ and $\bar{s} \approx \sqrt{2} - 1$ (see figure 4). In dimensional variables, this corresponds to $c \approx (2 - \sqrt{2}) s_0$ and $s = (2 - \sqrt{2}) s_0$, and we recover the original estimate given by Segel and Slemrod [30] (see figure 4).

GSPT, or really just the invariance equation, provides a powerful framework from which QSS approximations can be carefully derived. Moreover, it can help to rule out erroneous conclusions about specific QSS approximations such as (21). Note that the utility of the invariance equation still relies on direct knowledge of a suitable small parameter in order to construct an asymptotic expansion to the slow manifold. However, the appropriate small parameter that warrants the validity of the rQSSA is still an open problem. This raises the question: are there suitable non-scaling methods that can be employed to determine the validity of the rQSSA? We introduce two such methods in the section that follows.

4. Applying the Quasi-Steady-State Approximations to the MichaelisMenten reaction mechanism: Non-scaled approaches

In this section, we introduce non-scaling approaches to finding parameters or combinations of parameters that, when made very small, justify the validity of the sQSSA, the extended QSSA or the rQSSA. We also remark on the *origins* of slow manifolds, as this will be critical to uncovering sufficient conditions for the validity of the rQSSA in Section 5.

4.1. Geometric singular perturbation theory: The fast subsystem

Let us consider once more a fast/slow system of the form

$$x' = \varepsilon f(x, y; \varepsilon), \quad (56a)$$

$$y' = g(x, y; \varepsilon), \quad (56b)$$

which we will refer to as being in *Tikhonov standard form*. The associated fast subsystem,

$$x' = 0, \quad (57a)$$

⁸The usage of “characteristic” is slightly abused in this context since the Riccati equation is nonlinear.

⁹The zeroth-order fast fiber is not a vertical line in the (s, c) phase-plane. However, it is a vertical line in the (p, c) phase-plane, since p is *the* slow variable in this case and not s .

$$y' = g(x, y; 0), \quad (57b)$$

contains a branch of fixed points $y = h_0(x)$, where $g(x, h_0(x); 0) = 0$. Again, each fixed point in the set $y = h_0(x)$ is called a base point, and the complete set $y = h_0(x)$, denoted hereby as \mathcal{M} , is said to be normally-hyperbolic provided

$$\Re\left(\left.\frac{\partial g(x, y; 0)}{\partial y}\right|_{y=h_0(x)}\right) \neq 0. \quad (58)$$

In this context, normal hyperbolicity is partially¹⁰ a statement about the linear stability along directions that are transverse to the critical manifold. As previously stated, the zeroth-order fast fibers are solutions to (56a)-(56b), but they are also individual manifolds themselves. To see this, note that if (58) is less than zero, then \mathcal{M} is attracting, and trajectories that start *close* to \mathcal{M} move closer to \mathcal{M} as $t \rightarrow \infty$ (see Figure 5). In contrast, if (58) is positive, then \mathcal{M} is repelling, and trajectories that start close to \mathcal{M} move away from \mathcal{M} as $t \rightarrow \infty$. Thus, when the critical manifold is normally hyperbolic and attracting, the zeroth-order fast fiber is the stable manifold corresponding to the base point “ p ” of \mathcal{M} , and the stable manifold of \mathcal{M} is comprised of (foliated by) the entire union of zeroth-order fast fibers. Why are we interested in manifolds that are comprised of fixed points? Fenichel’s Theorems describe what happens to normally hyperbolic critical manifolds of fixed points once the vector field is perturbed. Formally, we introduce the following theorem due to Fenichel [38], which can be found in [53]:

Theorem 1: Let F be a C^r vector field on \mathbb{R}^n with $r \geq 1$. Let \mathcal{M} be a compact and connected C^1 manifold embedded in \mathbb{R}^n . Suppose that \mathcal{M} is normally hyperbolic and invariant under the flow of F . Then, given any C^1 vector field F^ε that is sufficiently C^1 -close to F , there exists a C^1 -manifold \mathcal{M}^ε that is invariant under the flow of F^ε and diffeomorphic to \mathcal{M} .

The notation “ C^r ” means the vector field has at least r derivatives, all of which are continuous. The terms “compact” and “connected” are referring to specific topological properties that \mathcal{M} must be equipped with. Connected means that \mathcal{M} cannot be expressed as the union of two disjoint and non-empty sets, and compact means every open cover has a finite sub-cover. The term C^1 -close is a statement about the distance between F and F^ε :

Definition: Let F and F^ε be two C^1 vectors fields on \mathbb{R}^n , and Λ be a compact set. F is said to be ϑ -close to F^ε on Λ if:

$$\sup_{x \in \Lambda} \|F - F^\varepsilon\| \leq \vartheta, \quad (59a)$$

$$\sup_{x \in \Lambda} \|DF - DF^\varepsilon\| \leq \vartheta, \quad (59b)$$

¹⁰Normally hyperbolicity provides information about the dominant directions of the flow near the critical manifold.

where D denotes differentiation. **Theorem 1** is one of several theorems that provides the foundation for GSPT (see [54] for a more technical survey). In a nutshell, one starts with a vector field that contains a critical set that consists of fixed points. Fenichel's theorems tell us that if a compact subset of the critical set is normally hyperbolic and attracts nearby trajectories, then perturbing the vector field (i.e., allowing ε in (57a)-(57b) to be greater than zero) in an appropriate way results in an attracting slow manifold that is at least locally invariant.

Let us now discuss a specific point, " $p \in \mathcal{M}$ ", where (58) fails and

$$\left. \frac{\partial g(x, y; 0)}{\partial y} \right|_{p \in \mathcal{M}} = 0. \quad (60)$$

A point that satisfies (60) is called a *singular point*. Singular points can indicate where a change in the stability of the critical manifold \mathcal{M} may occur, and can correspond to dynamic bifurcation points. In particular, a transcritical bifurcation occurs at a singular point where two normally hyperbolic submanifolds cross and undergo an exchange of stability (see Figure 6). This specific bifurcation will be of interest when we analyze the rQSSA.

4.2. Tikhonov-Fenichel parameters

We now want to take the idea of a critical manifold that is composed of fixed points one step further. Revisiting the mass action equations (5a)-(5c), we can ask the following question: Are there specific dimensional parameters that, when identically zero, result in the formation of a normally hyperbolic invariant manifold that is composed of fixed points? If so, then it seems reasonable to assume that these parameters should in some sense be small whenever a specific QSS reduction is valid. For example, let us first consider the rQSSA in the (s, p) phase-plane, for which the mass action equations are:

$$\dot{s} = -k_1(e_0 - (s_0 - s - p))s + k_{-1}(s_0 - s - p), \quad (61a)$$

$$\dot{p} = k_2(s_0 - s - p). \quad (61b)$$

We know that the rQSSA implies that $s \approx 0$ during the QSS phase of the reaction. Define $k_2 = \varepsilon_2 k_2^*$ and $k_{-1} = \varepsilon_1 k_{-1}^*$, where ε_1 and ε_2 are dimensionless and small, $0 < \varepsilon_1, \varepsilon_2 \ll 1$, and k_{-1}^*, k_2^* are identically one with the same units as k_{-1} and k_2 respectively. Expressing (61a)-(61b) as

$$\dot{s} = -k_1(e_0 - (s_0 - s - p))s + \varepsilon_1 k_{-1}^*(s_0 - s - p), \quad (62a)$$

$$\dot{p} = \varepsilon_2 k_2^*(s_0 - s - p), \quad (62b)$$

it is clear that the curve $s = 0$ is filled with equilibrium points when $\epsilon_1 = \epsilon_2 = 0$ and $e_0 > s_0$. Furthermore, it is straightforward to show that the curve $s = 0$ is normally hyperbolic and attracting when $s_0 < e_0$, since

$$\frac{\partial}{\partial s} [-k_1(e_0 - (s_0 - s - p))s] \Big|_{s=0} < 0. \quad (63)$$

Consequently, we anticipate

$$\dot{p} \sim k_2(s_0 - p) \quad (64)$$

during the slow QSS phase of the reaction when $0 < \epsilon_1, \epsilon_2 \ll 1$. The takeaway message from this example is two-fold. First, without scaling and non-dimensionalization, we have predicted the validity of the rQSSA by choosing dimensional parameters (i.e., k_{-1} and k_2) that, upon vanishing, give rise to an invariant manifold of equilibrium points. This is consistent with the notion of a normally hyperbolic critical manifold from fast/slow analysis and GSPT. Second, any dimensionless parameter homologous to ϵ_{SS} that acts as a qualifier for the validity of the rQSSA should vanish whenever k_{-1} and k_2 vanish. In planar systems, a parameter that, upon vanishing, gives rise to a manifold of fixed points that is normally hyperbolic, is called a *Tikhonov-Fenichel parameter* (TFP); Walcher and collaborators were the first to conceptualize TFPs [37, 58, 35, 43]. A manifold of equilibrium points is a special case of a normally hyperbolic manifold; the formal definition of a normally hyperbolic invariant manifold can be found in [53]. The informal definition of a normally hyperbolic invariant manifold is that the linearized flow that is tangent to the manifold is *dominated* by the flow that is normal (transversal) to the manifold. This is more or less intuitive when it comes to manifolds comprised of equilibria, since the flow on such a manifold is trivial. Nevertheless, manifolds of equilibria can lose normal hyperbolicity and, as we will see, the loss can play a significant role in determining when QSS approximations are valid.

The challenging part in the direct utilization of Tikhonov-Fenichel parameters is that often times the perturbed mass action equations obtained by regulating a Tikhonov-Fenichel parameter do not necessarily result in a set of equations that are in Tikhonov standard form. This can make computing a QSS reduction (without resorting to scaling) less than obvious. For example, the parameters e_0 and k_1 are Tikhonov-Fenichel parameters that warrant the validity of the sQSSA [35]. Treating k_1 as a small parameter, we can express (5a)-(5b) as

$$\dot{s} = -\epsilon k_1^*(e_0 - c)s - k_{-1}c, \quad (65a)$$

$$\dot{c} = \epsilon k_1^*(e_0 - c)s + (k_{-1} + k_2)c. \quad (65b)$$

The (s, c) phase-plane is equipped with an invariant manifold of equilibria, $c = 0$, when $k_1 = 0$. If $0 < k_1 \ll 1$, then one would expect that $s_0 \sim s + p$ on the slow timescale, provided the curve $c = 0$ is normally hyperbolic and attracts nearby trajectories. However, determining a QSS approximation for s or c from (65a)-(65b) is not obvious without scaling, since the

system is not in Tikhonov standard form. Nonetheless, using algebraic methods, Walcher and collaborators [43, 36, 57] developed a surprisingly straightforward method for obtaining QSS approximations from non-standard systems such as (65a)-(65b). In so doing, they have circumnavigated the need for scaling analysis to justify the QSSA.

For simple planar systems, Tikhonov-Fenichel parameters provide direct insight into the conditions that contribute to QSS dynamics, as well the *origin* (i.e., the critical manifold) of the corresponding slow manifold, and in this sense they are invaluable. However, in order to determine the validity of a particular QSS reduction so that experiments can be prepared with the intention of determining enzyme activity, the necessary “smallness” of Tikhonov-Fenichel parameters must be defined. For example, if $s_0 = e_0$, then the critical set obtained by setting $\varepsilon_1 = \varepsilon_2 = 0$ in (62a)-(62b) yields

$$e_0 s - (s_0 - s - p)s = 0, \quad (66)$$

since the term $e_0 - (s_0 - s - p)$ can vanish if $e_0 = s_0$. In this situation, it is not entirely clear how to go about properly defining the critical manifold when $s_0 = e_0$. This raises the question as to whether or not the rQSSA is still valid when $e_0 \lesssim s_0$ and, if so, then how small must the parameters k_{-1} and k_2 be in order for the approximation (64) to hold over the slow timescale? Obviously we expect the standard QSSA to be valid at extremely large values of s_0 , but what happens in parameter regions where s_0 and e_0 are the same order of magnitude? This question is central not only to the validity of the rQSSA, but is also central to understanding the validity of the sQSSA. Clearly, it seems that something more than scaling analysis and the identification of Tikhonov-Fenichel parameters is needed in order to address the validity of the rQSSA. We will introduce a method that permits a possible solution to the validity of the rQSSA in the subsection that follows.

4.3. Energy methods: Revisiting the sQSSA and Michaelis–Menten equation

To obtain a fundamental small parameter that ensures the validity of the sQSSA, we need to understand *where* the corresponding slow manifold comes from and, to do this, we must determine the associated critical manifold. One *possible* caveat with the traditional analysis of the sQSSA in the (s, c) phase-plane is that, if we are only interested in bounded trajectories (again, this rules out the particular case of taking $s_0 \rightarrow \infty$), then taking $\varepsilon \rightarrow 0$ implies that either k_1 or e_0 must vanish, as Walcher and his collaborators have carefully pointed out [43, 58, 35, 37]. In both of these cases c must vanish, but the dimensionless variable $\bar{c} = c / \varepsilon s_0$ contains a zero denominator in the singular limit. According to Walcher and Lax [60] this is not a problem. However, in order to avoid having to deal with this, it is useful to write down the dimensionless equations in the (\bar{s}, \bar{p}) coordinate system:

$$\frac{d\bar{s}}{d\tau} = -\varepsilon_S S \bar{s} + \mu \bar{s}(1 - \bar{s} - \bar{p}) + \alpha \beta (1 - \bar{s} - \bar{p}) \quad (67a)$$

$$\frac{d\bar{p}}{d\tau} = \beta(1 - \alpha)(1 - \bar{s} - \bar{p}). \quad (67b)$$

Setting $\varepsilon_{SS} = 0$ in (67a)-(67b) reveals the critical manifold to be $\bar{p} = 1 - \bar{s}$, which corresponds to the critical manifold of the dimensional system (i.e., $c = 0$) obtained by setting k_1 or e_0 to zero in (65a)-(65b). Since the curve $s = s_0 - p$ implies $c = 0$, we avoid the *other* caveat encountered in the analysis in the (s, c) coordinate system in section 3: there is no violation of the substrate conservation law $s_0 = s + c + p$ along a zeroth-order fast fiber, since initial conditions lie *on* the critical manifold, $c = 0$. If a *non-trivial* c -nullcline is defined to be the critical manifold, then there is zero loss of substrate along the one-dimensional zeroth-order fibers when $\varepsilon_{SS} = 0$, which is a violation of the substrate conservation law, since $s + c > s_0$ as the trajectory approaches a base point that lies above the s -axis.

It is expected that trajectories will closely follow the curve $p = s_0 - s$ once the perturbation is “turned on” and $0 < \varepsilon_{SS} \ll 1$. To get an idea of how “good” the approximation $p \sim s_0 - s$ is for all time, we need to compute an upper bound “ M ” such that¹¹

$$(1 - \bar{s} - \bar{p}) \leq M. \quad (68)$$

We will start with a differential equation for the energy¹² $\mathcal{E}^2 \equiv (s_0 - s - p)^2$. Multiplying both sides of the dimensional form of (67a) by “ $(s_0 - s - p) \equiv \mathcal{E}$ ” yields:

$$\frac{1}{2} \frac{d\mathcal{E}^2}{dt} = k_1 e_0 s \mathcal{E} - k_1 (s + K_M) \mathcal{E}^2, \quad (69a)$$

$$\leq k_1 e_0 s_0 |\mathcal{E}| - k_1 K_M \mathcal{E}^2. \quad (69b)$$

Now we can use Cauchy’s inequality with “ δ ,”

$$ab \leq \delta a^2 + \frac{b^2}{4\delta}, \quad (70)$$

and expand the term $k_1 e_0 s_0 \mathcal{E}$ in (69b):

$$k_1 e_0 s_0 |\mathcal{E}| \leq \frac{(k_1 e_0 s_0)^2}{4\delta} + \delta \mathcal{E}^2. \quad (71)$$

Choosing $\delta = k_1 K_M / 2$ yields

$$k_1 e_0 s_0 |\mathcal{E}| \leq \frac{(k_1 e_0 s_0)^2}{2k_1 K_M} + \frac{k_1 K_M}{2} \mathcal{E}^2. \quad (72)$$

¹¹We do not need to write $|\mathcal{E}|$ since $0 = s_0 - s - p$, but later on we will consider cases where the quantity of interest can vary in sign, so it helps to keep the notation consistent through each procedure.

¹²This is the *mathematical* energy of a function, and should not be confused with the *physical* energy of a particle.

After combining (72) with (69b) we have

$$\frac{d\mathcal{E}^2}{dt} \leq \frac{(k_1 e_0 s_0)^2}{k_1 K_M} - k_1 K_M \mathcal{E}^2, \quad (73)$$

and integrating both sides of (73) yields

$$\mathcal{E}^2(t) \leq \mathcal{E}^2(0)e^{-(k_1 K_M)t} + s_0^2 \eta^2 (1 - e^{-k_1 K_M t}). \quad (74)$$

Replacing “ $(1 - e^{-k_1 K_M t})$ ” with “1” and dividing both sides of (74) by s_0^2 reveals

$$\limsup_{t \rightarrow \infty} \bar{\mathcal{E}}^2 \leq \eta^2, \quad \bar{\mathcal{E}} \equiv \mathcal{E} / s_0, \quad (75)$$

where the limit supremum or “limsup” of a bounded function $w(t)$ is given by:

$$\limsup_{t \rightarrow \infty} w(t) = \lim_{T \rightarrow \infty} \left(\sup_{t > T} w(t) \right). \quad (76)$$

It is clear from (75) that the long-time error in the approximation is controlled by η . We can simplify things further: utilizing the fact that $\mathcal{E}(0) = 0$, and taking the square root of both sides yields

$$(1 - \bar{s} - \bar{p}) \leq \eta. \quad (77)$$

The bound (77) obtained from the energy analysis is promising, but what we are ultimately after is an expression for $p(t)$ that can be utilized for parameter estimation. In practice, the approximation¹³

$$\frac{dp}{dt} \sim \frac{k_2 e_0 (s_0 - p)}{K_M + (s_0 - p)}, \quad (78)$$

is often used to estimate K_M and k_2 ; the assumption in (78) is that c is negligible and is well approximated by

$$c \approx \frac{e_0 (s_0 - p)}{K_M + s_0 - p}. \quad (79)$$

The question that immediately follows is: how *good* is the approximation (79)? In order to assess the validity of (79), we will derive an upper bound on $|\mathcal{E}_c|$,

¹³This approximation is easily justified through scaling analysis. See [16] for details.

$$\mathcal{E}_c \equiv \left(c - \frac{e_0(s_0 - p)}{K_M + s_0 - p} \right). \quad (80)$$

To begin,

$$\frac{1}{2} \frac{d\mathcal{E}_c^2}{dt} = \left(c - \frac{e_0(s_0 - p)}{K_M + s_0 - p} \right) \left(\frac{dc}{dt} + \frac{e_0 K_M}{(K_M + s_0 - p)^2} \frac{dp}{dt} \right), \quad (81a)$$

$$= \mathcal{E}_c \left(k_1 e_0 (s_0 - p) - k_1 c (e_0 + K_M + s_0 - p) + k_1 c^2 + \frac{e_0 K_M k_2 c}{(K_M + s_0 - p)^2} \right), \quad (81b)$$

$$= \mathcal{E}_c \left(-k_1 (K_M + s_0 - p) \mathcal{E}_c + k_1 c (c - e_0) + \frac{e_0 K_M k_2 c}{(K_M + s_0 - p)^2} \right), \quad (81c)$$

$$= -k_1 (K_M + s_0 - p) \mathcal{E}_c^2 \left(k_1 c (c - e_0) + \frac{e_0 K_M k_2 c}{(K_M + s_0 - p)^2} \right). \quad (81d)$$

Next, we need to find an upper bound on the right and side of (81d):

$$\frac{1}{2} \frac{d\mathcal{E}_c^2}{dt} \leq -k_1 K_M \mathcal{E}_c^2 + |\mathcal{E}_c| \max \left| k_1 c (c - e_0) + \frac{e_0 K_M k_2 c}{(K_M + s_0 - p)^2} \right|, \quad (82a)$$

$$\leq -k_1 K_M \mathcal{E}_c^2 + |\mathcal{E}_c| \max |k_1 c (c - e_0)| + |\mathcal{E}_c| \max \left| \frac{e_0 K_M k_2 c}{(K_M + s_0 - p)^2} \right|, \quad (82b)$$

$$\leq -k_1 K_M \mathcal{E}_c^2 + |\mathcal{E}_c| \left(\frac{k_1 e_0^2}{4} + \frac{e_0 k_2 \lambda}{K_M} \right). \quad (82c)$$

Using Cauchy's inequality with $\delta = k_1 K_M / 2$, it follows that

$$\frac{d\mathcal{E}_c^2}{dt} \leq -k_1 K_M \mathcal{E}_c^2 + \frac{1}{k_1 K_M} \left(\frac{k_1 e_0^2}{4} + \frac{e_0 k_2 \lambda}{K_M} \right)^2. \quad (83)$$

Applying Gronwall's lemma and integrating both sides of (83) yields

$$\mathcal{E}_c^2 \leq \mathcal{E}_c^2(0) e^{-k_1 K_M t} + \frac{1}{k_1^2 K_M^2} \left(\frac{k_1 e_0^2}{4} + \frac{e_0 k_2 \lambda}{K_M} \right)^2. \quad (84)$$

Although the exponential term (84) is in terms of the dimensional time, t , we expect that if (84) is written in terms of the slow time, $\bar{T} = tk_2\varepsilon_{SS}$, then it should decay relatively quickly in \bar{T} . To do this we need to solve

$$k_1 K_M t = C_{\bar{T}} k_2 \varepsilon_{SS} t \quad (85)$$

for the constant, $C_{\bar{T}}$. Solving (85) reveals $C_{\bar{T}} = 1 / \nu\varepsilon$, and it clear that the exponential term in (84) will vanish rapidly in \bar{T} whenever $\varepsilon\nu \ll 1$. Finally, taking the square root of both sides while noting that $\sqrt{x+y} \leq \sqrt{x} + \sqrt{y} \forall x, y \geq 0$, and dividing through by ε_0 produces the following upper bound:

$$|\tilde{c} - h(\bar{p})| \leq \mu e^{-\bar{T}/2\varepsilon_{SS}\nu} + \frac{\eta}{4} + \nu \frac{\lambda}{K_M}, \quad h(\bar{p}) \equiv \frac{\sigma(1-\bar{p})}{1+\sigma(1-\bar{p})}, \quad (86)$$

where $\tilde{c} \equiv c / \varepsilon_0$. The bound (86) is actually quite revealing. First, we immediately see that taking $\nu \rightarrow 0$ does not support the validity of the sQSSA (i.e., the MM equation), even though the exponential term decays rapidly in the slow time whenever $\varepsilon_{SS}\nu \ll 1$, and “nullcline following” in (s, c) phase-plane is observed when $k_2 \ll k_{-1}$. Thus, we do not encounter any sort of dilemma between the sQSSA and the extended QSSA when energy methods are appropriately employed. Second, if $s_0 < \varepsilon_0$, then $\lambda < s_0$, and we have

$$|\tilde{c} - h(\bar{p})| \leq \mu e^{-\bar{T}/2\varepsilon_{SS}\nu} + \frac{\eta}{4} + \nu\sigma. \quad (87)$$

Since $\sigma < \eta$ when $s_0 < \varepsilon_0$, it follows that $\eta \rightarrow 0$ in order for the sQSSA to be valid. Third, note that making s_0 arbitrarily large only makes the exponential term decay faster, since $\varepsilon_{SS} \rightarrow 0$ as $s_0 \rightarrow \infty$. However, large s_0 has no influence on the long-time bound: only K_M and ε_0 influence the long-time bound. Finally, if $\varepsilon_0 < s_0$, then $\lambda = \varepsilon_0$ and it holds that

$$|\tilde{c} - h(\bar{p})| \leq \mu e^{-\bar{T}/2\varepsilon_{SS}\nu} + \frac{5}{4}\eta. \quad (88)$$

In all cases the natural “small parameter” that arises from the energy analysis is η . What can happen if $\varepsilon_{SS} \ll 1$ but $\eta \gg 1$? The answer is that a curious situation arises: the (s, c) phase-plane trajectory will rapidly approach the c -nullcline, and proceed to follow it closely for a finite amount of time. However, as the reaction progresses, the phase-plane trajectory, which follows the invariant manifold \mathcal{M}^ε , will start to move away from the c -nullcline, and eventually follow the s -nullcline (see figure 7).

In conclusion of this section, our analysis suggests that η is the fundamental small parameter that justifies the validity of the sQSSA. The condition $\eta \ll 1$ can be viewed as a special case of $\varepsilon_{SS} \ll 1$. However, if $\varepsilon_{SS} \ll 1$ and $1 \ll \eta$, then we observe a scenario in the (s, c) phase-plane in which the sQSSA is initially valid after the fast transient, but steadily loses validity until the rQSSA becomes valid for the remainder of the reaction. It seems as though the situation is quite unique, since the sQSSA appears to valid over the slow timescale, but the

rQSSA is valid over a super-slow timescale. If we can understand the origin of the slow manifold in this case (i.e., its associated critical manifold), then perhaps it will be possible to explain why this scenario occurs. Of course, this raises the question: “Can a unique critical manifold be determined in this case?” The answer to this question resides in the analysis of the rQSSA, and we will analyze this particular approximation in great detail in section 5.

5. The reverse Quasi-Steady-State Approximation

In this section, we derive the validity of the rQSSA using a combination of scaling and energy methods.

5.1. Justification of the reverse Quasi-Steady-State Approximation in the (p, c) phase-plane: Scaling and bifurcation analysis

It is straightforward to verify the validity of the rQSSA through scaling analysis. Assuming that $s_0 < e_0$, it is obvious that c is bounded by s_0 , which suggests that $\hat{c} = c / s_0$ is the appropriate scaled variable when $s_0 < e_0$. Rescaling the mass action equations (27a)-(27b) in the (p, c) phase-plane yields

$$\varepsilon^* v \frac{d\hat{c}}{dT} = (1 - \ell\hat{c})(1 - \hat{c} - \bar{p}) - \varepsilon^* \hat{c}, \quad (89a)$$

$$\frac{d\bar{p}}{dT} = \hat{c}, \quad (89b)$$

and the corresponding critical manifold, \mathcal{M}_0 , is obtained by setting $\varepsilon^* = 0$,

$$\mathcal{M}_0 = \{(\bar{p}, \hat{c}) \in \mathbb{R}_+^2 \mid (1 - \ell\hat{c})(1 - \hat{c} - \bar{p}) = 0\}. \quad (90)$$

If $s_0 < e_0$, then $(1 - \ell\hat{c}) > 0$, and we can take the critical manifold to be $1 - \hat{c} - \bar{p} = 0$. Furthermore, the critical manifold is normally hyperbolic and attracting, since

$$\frac{\partial}{\partial \hat{c}} (1 - \ell\hat{c})(1 - \hat{c} - \bar{p}) < 0. \quad (91)$$

Thus, due to the fact that the critical manifold is both normally hyperbolic and attracting, we expect trajectories to approach the curve $c = s_0 - p$ when $s_0 < e_0$ and $0 < \varepsilon^* \ll 1$. But what happens when $e_0 = s_0$? Let us start by setting $e_0 = s_0$, in which case $\ell = 1$. The dimensionless equations are given by

$$\varepsilon^* v \frac{d\hat{c}}{dT} = (1 - \hat{c})(1 - \hat{c} - \bar{p}) - \varepsilon^* \hat{c}, \quad (92a)$$

$$\frac{d\bar{p}}{dT} = \hat{c}, \quad (92b)$$

and the associated critical set is

$$\mathcal{M}_0 = \{(\bar{p}, \hat{c}) \in \mathbb{R}_+^2 \mid (1 - \hat{c})(1 - \hat{c} - \bar{p}) = 0\}. \quad (93)$$

The set (93) fails to be a manifold at the point where the curves $\hat{c} = 1$ and $\hat{c} = 1 - \bar{p}$ intersect, which is precisely the point $(\bar{p}, \hat{c}) = (0, 1)$. Moreover, there is a loss of normal hyperbolicity at this point, since

$$\frac{\partial}{\partial \hat{c}} [(1 - \hat{c})(1 - \hat{c} - \bar{p})] \Big|_{(\bar{p}, \hat{c}) = (0, 1)} = 0. \quad (94)$$

What is happening at the point where normal hyperbolicity is lost? The sub-manifolds $\hat{c} = 1$ and $\bar{p} = 1 - \hat{c}$ intersect and exchange stability at the point $(\bar{p}, \hat{c}) = (0, 1)$ when $K_M = 0$. To see this, rescale time and define $\tau^* = \tilde{T} / \varepsilon^* \nu$:

$$\frac{d\hat{c}}{d\tau^*} = (1 - \hat{c})(1 - \hat{c} - \bar{p}) - \varepsilon^* \hat{c}, \quad (95a)$$

$$\frac{d\bar{p}}{d\tau^*} = \varepsilon^* \nu \hat{c}. \quad (95b)$$

Since $e_0 = s_0$, setting $\varepsilon^* = 0$ implies that $K_M = 0$, and we recover

$$\frac{d\hat{c}}{d\tau^*} = (1 - \hat{c})(1 - \hat{c} - \bar{p}), \quad (96a)$$

$$\frac{d\bar{p}}{d\tau^*} = 0. \quad (96b)$$

Treating \bar{p} as a slowly-varying parameter in the fast subsystem given by (96a)-(96b), the normal form of the transcritical bifurcation is recovered by making the change of variables $u = 1 - \hat{c}$

$$\frac{du}{d\tau^*} = \bar{p}u - u^2. \quad (97)$$

The normal form equation (97) indicates that the critical submanifolds¹⁴ undergo an exchange of stability at the transcritical singularity. Clearly, this transcritical singularity influences the dynamics whenever $K_M = 0$ and $e_0 = s_0$ (see Figure 8).

¹⁴The transcritical bifurcation is also easily recoverable in (s, p) coordinates.

5.2. A heuristic analysis based on timescale separation

The bifurcation analysis from the previous section gave us a global picture of what the phase-plane dynamics look like when $K_M = 0$. Moreover, we can use the bifurcation structure to our advantage and attempt to determine a suitable “small parameter” that warrants the validity of the rQSSA. To begin, if $e_0 < s_0$, then there are essentially three stages to the reaction (see figure 9).

The first stage is the transient stage, and the phase-plane trajectory moves almost vertically towards the curve $c = e_0$. The approximate duration of this stage is given by Tzafirri’s fast timescale, t_C^* . After the initial fast transient the differential equation for product formation is roughly

$$\dot{p} \sim k_2 e_0. \quad (98)$$

Integrating the above expression yields $p \sim k_2 e_0 t$. Since $c \approx e_0$ in this regime, the trajectory will reach the vicinity of the c -axis once $p = s_0 - e_0$, which, based on the substrate conservation law, would imply that $s \approx 0$. Thus, the time it takes the trajectory to reach the curve $s_0 = c + p$ is given by t_ℓ

$$t_\ell = \frac{s_0 - e_0}{k_2 e_0}. \quad (99)$$

Next, the trajectory can be assumed to be very close to the curve $c = s_0 - p$ and, at this point, the time it takes for the reaction to effectively complete is $t_{\text{slow}} = 1/k_2$. If the ratio of t_ℓ to t_{slow} is small, i.e.,

$$0 < \frac{t_\ell}{t_{\text{slow}}} = \frac{s_0 - e_0}{e_0} \ll 1, \quad (100)$$

then, from a *practical* point of view (as opposed to a rigorous point of view), the rQSSA can be considered valid if $s_0 > e_0$ and K_M is incredibly small.

The heuristic analysis at least partially supports the idea that parameter domain over which the rQSSA is valid may in fact be larger than previously assumed. However, the combined conditions, $e^* \ll 1$ and $(s_0 - e_0)/e_0 \ll 1$, still do not explain how increasing s_0 impacts the validity of the approximation. This is a reasonable question that has plagued theorists for some time now (see, for example [30] and [31]). The usual methods of scaling analysis and timescale estimation have not provided very satisfying answers. Consequently, this suggests that a new method needs to be employed to tackle this problem.

5.3. The reverse quasi-steady-state approximation: energy methods

In order to determine a small parameter that ensures the validity of the rQSSA, it helps to first consider what we expect to happen *chemically* when such a parameter is made arbitrarily small. Historically, thanks to the careful work of Nguyen and Fraser [40], the geometric interpretation of the rQSSA’s validity has been associated with a phase-plane

trajectory closely following the s -nullcline in the (s, c) phase-plane. However, instead of thinking about the validity of the rQSSA in terms of a trajectory's propensity for following the s -nullcline, it is perhaps better to think of it in purely chemical terms: when the rQSSA is valid, the initial substrate concentration vanishes rapidly with respect to the timescale $\tilde{T} \equiv k_2 t$. Translated mathematically, this means that s must dissipate rapidly with respect to \tilde{T} . Thus, we want an equation that describes the dissipation of $s^2/2$:

$$\frac{1}{2} \frac{ds^2}{dt} = -k_1(e_0 - c)s^2 + k_{-1}cs, \quad (101a)$$

$$\leq -k_1(e_0 - \lambda)s^2 + k_{-1}s\lambda. \quad (101b)$$

The term $k_{-1}s\lambda$ can be expanded thanks to Cauchy's inequality

$$k_{-1}s\lambda \leq \delta s^2 + \frac{k_{-1}^2\lambda^2}{4\delta}, \quad (102)$$

and choosing $\delta = k_1(e_0 - \lambda)/2$ yields

$$\frac{1}{2} \frac{ds^2}{dt} \leq -\frac{1}{2}k_1(e_0 - c)s^2 + \frac{k_{-1}^2\lambda^2}{2k_1(e_0 - \lambda)}. \quad (103)$$

From Gronwall's lemma we have¹⁵

$$s^2(t) \leq s_0^2 e^{-k_1(e_0 - \lambda)t} + \frac{k_{-1}^2\lambda^2}{k_1^2(e_0 - \lambda)^2} (1 - e^{-k_1(e_0 - \lambda)t}), \quad (104a)$$

$$\leq s_0^2 e^{-k_1(e_0 - \lambda)t} + \frac{k_{-1}^2\lambda^2}{k_1^2(e_0 - \lambda)^2}. \quad (104b)$$

Diving both sides by s_0^2 , taking the square root of both sides, and invoking the triangle inequality yields:

$$s(t) \leq s_0 e^{-k_1(e_0 - \lambda)t/2} + \frac{k_{-1}\lambda}{k_1(e_0 - \lambda)}. \quad (105)$$

Next, re-writing (105) in terms of \tilde{T} ,

$$\bar{s}(\tilde{T}) \leq e^{-\tilde{T}/2\varepsilon\nu} + \tilde{\varepsilon} \frac{\lambda}{s_0}, \quad (106a)$$

$$\leq e^{-\tilde{T}/2\varepsilon\nu} + \tilde{\varepsilon}, \quad (106b)$$

reveals two small parameters:

$$\underline{\varepsilon} \equiv \frac{K_M}{e_0 - \lambda}, \quad \tilde{\varepsilon} \equiv \frac{K_S}{e_0 - \lambda}. \quad (107)$$

Now we have a choice to make: which small parameter corresponds to the potential validity of the rQSSA? The answer to this question resides in Fenichel theory: Setting $\underline{\varepsilon}$ to zero results in a critical manifold of fixed points, whereas making $\tilde{\varepsilon} = 0$ by setting $k_{-1} = 0$ results in the manifold $s_0 = c + p$ being invariant, but it is not a manifold of equilibrium points, since k_{-1} is not a TFP (see [58] for details). Moreover, it holds that

$$\bar{s}(\tilde{T}) \leq e^{-\tilde{T}/2\underline{\varepsilon}} + \underline{\varepsilon}, \quad (108)$$

since $\nu > 1$ and $\tilde{\varepsilon} \leq \underline{\varepsilon}$. Consequently, we expect s to vanish rapidly in \tilde{T} whenever $\underline{\varepsilon} \ll 1$. Furthermore, note that

$$\lim_{K_M \rightarrow 0} \underline{\varepsilon} = 0 \text{ if } s_0 < e_0, \quad \lim_{K_M \rightarrow 0} \underline{\varepsilon} = \frac{s_0 - e_0}{e_0} \text{ if } s_0 > e_0, \quad (109)$$

and we recover the exact small parameter (100) that was obtained heuristically in Section 5.1.

As a final check, we observe the influence of $\underline{\varepsilon}$ on the dynamics in the (p, c) phase-plane when $0 < \underline{\varepsilon} \ll 1$. First note that the c -nullcline can be approximated by the curve $s_0 = c + p$ as $K_M \rightarrow 0$ in any region where $s_0 < e_0$. Thus, if trajectories closely follow the c -nullcline when $K_M \ll 1$ and $s_0 < e_0$, then the rQSSA should be valid. Second, one can show that

$$|c - c_0(p)| \leq \lambda e^{-t\zeta_T/2} + \frac{\eta\lambda}{(1+\eta)(1+\kappa)} \left(\frac{K_M}{e_0 - \lambda + K_M} \right), \quad (110)$$

where $\zeta_T \equiv k_1(e_0 - \lambda + K_M)$ (we will derive this in the Section 6 using energy methods, but also in Section Appendix A using a method pioneered by Gradshteyn [14]). Dividing both sides by λ and expressing (110) in terms of \tilde{T} yields

$$|\bar{c} - \bar{c}_0(\bar{p})| \leq e^{-\tilde{T}/2\varepsilon\nu} + \frac{\eta\nu\underline{\varepsilon}}{(1+\eta)(1+\underline{\varepsilon})}, \quad (111)$$

from which it follows that

$$|\bar{c} - \bar{c}_0(\bar{p})| \leq e^{-\tilde{T}/2\underline{\varepsilon}} + \underline{\varepsilon}. \quad (112)$$

The upper bound (112) reveals that the phase-plane trajectory will closely follow the c -nullcline after a brief transient. Since the c -nullcline is approximately $s_0 = c + p$ whenever s_0

$< e_0$ and $\underline{\varepsilon} \ll 1$, it holds that the rQSSA is valid under these conditions, and we take $\underline{\varepsilon}$ to be *the* small parameter that ensures its validity.

In summary of this section, we have shown that, while scaling methods or Tikhonov-Fenichel parameters can be used to predict the presence of a singular point, more analysis is needed to assess the validity of the rQSSA in neighborhoods impacted by the critical singularity. Hence, the presence of the singular point in the critical set is what makes the analysis of the rQSSA challenging. As we have shown, energy methods can be easily employed to assess the validity of the rQSSA when initial substrate and enzyme concentrations are of similar magnitude. Moreover, the bound

$$|c - c_0(p)| \leq \lambda e^{-t\zeta_T/2} + \frac{\eta\lambda}{(1+\eta)(1+\kappa)} \left(\frac{K_M}{e_0 - \lambda + K_M} \right), \quad (113)$$

provides direct insight into the origin of each critical manifold, and suggests that there are three parameters, η , $\underline{\varepsilon}$ and ν , that, when adequately small, guarantee the long-time validity of the QSSA, the rQSSA and the tQSSA, respectively. We will discuss this observation in more detail in section 6.

6. Final remarks on timescales and small parameters

In this section we explore the relationship between the fast and slow timescales of the tQSSA, Tikhonov-Fenichel parameters, and the small parameters obtained from the energy analysis.

6.1. Timescale separation and small parameters: local versus global conditions for the accuracy of QSS approximations

As mentioned earlier, because the tQSSA encompasses the rQSSA and the sQSSA, the accepted criterion for the validity of a QSSA is separation of Tzafriri's timescales, $\varepsilon_T \ll 1$ (see 35). We can express the bound on the limsup of \underline{c} and $\bar{c}_0(\bar{p})$ in terms of $T_z = t/t_P$,

$$|\underline{c} - \bar{c}_0(\bar{p})| \leq e^{-T_z/2\varepsilon_D} + \varepsilon_L. \quad (114)$$

To compute (114), we start by computing,

$$\limsup_{t \rightarrow \infty} (c - h^-(p))^2, \quad (115)$$

where we will use the notation

$$h^-(p) = \frac{1}{2}(e_0 + s_0 + K_M - p) - \frac{1}{2}\sqrt{(e_0 + s_0 + K_M + p)^2 - 4e_0(s_0 - p)}, \quad (116a)$$

$$h^+(p) = \frac{1}{2}(e_0 + s_0 + K_M - p) + \frac{1}{2}\sqrt{(e_0 + s_0 + K_M + p)^2 - 4e_0(s_0 - p)}, \quad (116b)$$

for convenience. Thus, the reader should identify $\lambda = h^-(0)$ and note that $c_0(p) \equiv h^-(p)$. The differential equation for the energy, $\mathcal{E}^2 \equiv (c - h^-(p))^2$, is

$$\frac{1}{2} \frac{d\mathcal{E}^2}{dt} = \left[\frac{dc}{dt} - \frac{dh^-}{dp} \frac{dp}{dt} \right] \mathcal{E}. \quad (117)$$

Carefully note that the derivative for c can be factored, $\dot{c} = k_1(c - h^-(p))(c - h^+(p))$, and thus

$$\frac{1}{2} \frac{d\mathcal{E}^2}{dt} = \left[k_1(c - h^-(p))(c - h^+(p)) - \frac{dh^-}{dp} \frac{dp}{dt} \right] \mathcal{E} \quad (118a)$$

$$= \left[k_1 \mathcal{E}(c - h^+(p)) - \frac{dh^-}{dp} \frac{dp}{dt} \right] \mathcal{E} \quad (118b)$$

$$\leq \mathcal{E}^2 k_1(c - h^+(p)) + \max \left| \frac{dh^-}{dp} \right| \sup \left| \frac{dp}{dt} \right| |\mathcal{E}|. \quad (118c)$$

Next, the term “ $c - h^+(p)$ ” can be bounded above:

$$\max c - h^+(p) \leq \lambda - \min h^+(p) = -k_1(e_0 + K_M - \lambda). \quad (119)$$

Using the Cauchy- δ inequality with $\delta = k_1(e_0 + K_M - \lambda)/2$, we obtain

$$\begin{aligned} \max \left| \frac{dh^-}{dp} \right| \sup \left| \frac{dp}{dt} \right| |\mathcal{E}| &\leq \frac{k_1(e_0 + K_M - \lambda)}{2} \mathcal{E}^2 \\ &+ \frac{\left(\max \left| \frac{dh^-}{dp} \right| \right)^2 \left(\sup \left| \frac{dp}{dt} \right| \right)^2}{2k_1(e_0 + K_M - \lambda)}. \end{aligned} \quad (120)$$

Inserting (120) into (118c) yields

$$\frac{d\mathcal{E}^2}{dt} \leq -\mathcal{E}^2 k_1(e_0 + K_M - \lambda) + \frac{\left(\max \left| \frac{dh^-}{dp} \right| \right)^2 \left(\sup \left| \frac{dp}{dt} \right| \right)^2}{k_1(e_0 + K_M - \lambda)}, \quad (121)$$

and from Gronwall’s Lemma and the triangle inequality we recover:

$$|\mathcal{E}(t)| \leq |\mathcal{E}(0)| e^{-k_1(e_0 + K_M - \lambda)t} + \frac{\max \left| \frac{dh^-}{dp} \right| \sup \left| \frac{dp}{dt} \right|}{k_1(e_0 + K_M - \lambda)}. \quad (122)$$

Calculating the “max” and “sup” on the right hand side of (122), dividing both sides by λ , and expressing the exponential in terms of $T_z = t/t_P$ yields (114), where ε_D denotes the “exponential ε ,” and ε_L denotes the “long-time ε .”

$$\varepsilon_D \equiv \left(\frac{\lambda}{s_0}\right) \left(\frac{1}{1+\kappa}\right) \left(\frac{\underline{\varepsilon}}{1+\underline{\varepsilon}}\right), \quad \varepsilon_L \equiv \left(\frac{\eta}{1+\eta}\right) \left(\frac{1}{1+\kappa}\right) \left(\frac{\underline{\varepsilon}}{1+\underline{\varepsilon}}\right). \quad (123)$$

As expected, ε_D is dependent on initial data, meaning its magnitude is influenced by the starting location of the trajectory within the vector field. Moreover, ε_D can be “factored” into three small parameters:

- i. λ/s_0 : vanishes as a result of zero enzyme to substrate ratio.
- ii. $1/(1+\kappa) \equiv \nu$: vanishes in the limit of infinitely slow product formation.
- iii. $\underline{\varepsilon}/(1+\underline{\varepsilon})$: vanishes at infinitely high enzyme concentration or, zero K_M whenever $s_0 = e_0$.

What is the relationship between Tzafirri’s timescales and ε_D and ε_L ? Tzafirri’s fast timescale, t_C^* , can be written as

$$t_C^* = \frac{1}{k_1(K_M + e_0 - \lambda + s_0 - \lambda)}. \quad (124)$$

Using the identity (124), a simple calculation reveals $\varepsilon_T \sim \varepsilon_D$. In fact, $\varepsilon_T \sim \varepsilon_D$ at very high enzyme concentration, and $\varepsilon_T \rightarrow 0$ if $\varepsilon_D \rightarrow 0$. Thus, we can control timescale separation by controlling ε_D . However, $\varepsilon_D \ll 1$ only implies that the phase-plane trajectory moves rapidly towards the c -nullcline; it does not ensure that the trajectory stays close to the c -nullcline as $T_z \rightarrow \infty$. The long-time validity of the zeroth-order approximation is regulated by ε_L , and $\varepsilon_D \sim \varepsilon_L$. Hence, ε_L is *the* fundamental small parameter, since it is independent of a particular choice of initial data and provides an upper bound on the lim sup:

$$\limsup_{t \rightarrow \infty} (c - c_0(p))^2 \leq \lambda^2 \varepsilon_L^2. \quad (125)$$

For example, ε_D can be made arbitrarily small by taking s_0 to be arbitrarily large, but the accuracy of the approximation in regimes where s_0 is large may only be temporary at best (see figure 10). Finally, the magnitude of ε_L is regulated by the TFPs. The parameter ε_L determines “how small” a TFP must be in comparison to other parameters in order to ensure the validity of the tQSSA. Thus, there is a fundamental “hierarchy” or “ordering” of parameters:

$$\text{TFP} < \varepsilon_L < \varepsilon_D < \varepsilon_T. \quad (126)$$

From the perspective of the ordering given in (126), timescale separation should be seen as an effect, not a cause. Segel and Slemrod [30] understood this; they were very clear in issuing their warning of sufficient versus necessary conditions for the validity of QSS reductions, and stressed the fact that timescale separation is merely a necessary condition.

The takeaway from our analysis is that timescale separation is really a local metric when it comes to assessing the validity of a QSS reduction. This is primarily because timescales are dependent on initial conditions, and the vector field can change dramatically once the trajectory leaves a small neighborhood containing the initial conditions. Roussel and Fraser [41] seem to have been the first to really understand this in the context of the Michaelis-Menten reaction mechanism (1), as their work focused on approximating the actual invariant slow manifold. In fact, to the best of the authors' knowledge, Roussel and Fraser (see [61], pp. 42-47) were the first to provide a solid geometrical argument that η is in fact the "global" qualifier for the validity of the sQSSA instead of ε_{SS} . By approximating the slow manifold to a high degree of accuracy, Roussel and Fraser [62, 42] were able to demonstrate that $\eta \ll 1$ is necessary for the long-time validity of the sQSSA.

6.2. Tightening the limit supremum

It is worth pointing out that a long-time bound smaller than ε_L can be obtained when $e_0 < s_0$. This is to be expected, since $\underline{\varepsilon}$ does not vanish if $s_0 > e_0$ as $K_M \rightarrow 0$. To obtain a tighter bound, note that the upper bound of $c - h^+(p)$ given in (119) is rather liberal. What we want to do is compute $\max(c - h^+(p))$ with more care. First, since $c < h^+(p)$ and $0 < h^+(p)$, it follows that $c - h^+(p) < 0$. Second, $h^+(p)$ is a decreasing function of p , which means $\min h^+(p)$ occurs when $p = s_0$. However, if $c = 0$, then $p = s_0 - c$. Thus, we want to maximize:

$$c - h^+(s_0 - c) \equiv \theta(c) < 0. \quad (127)$$

Since $\theta'(c) > 0$ and $\theta(c) < 0$, the maximum of $\theta(c)$ occurs when $c = \lambda$, from which it is straightforward to show

$$\frac{d\varepsilon^2}{dt} \leq -\varepsilon^2 k_1 |\theta(\lambda)| + \frac{\left(\max \left| \frac{dh^-}{dp} \right| \right)^2 \left(\sup \left| \frac{dp}{dt} \right| \right)^2}{k_1 |\theta(\lambda)|}, \quad (128)$$

where $\theta(\lambda)$ is given by:

$$\theta(\lambda) = \frac{\lambda}{2} - \frac{1}{2} \left(e_0 + K_M + \sqrt{(e_0 + K_M + \lambda)^2 - 4e_0\lambda} \right). \quad (129)$$

An even tighter bound on $\theta(c)$ can be obtained by noting that $\lambda = e_0$, and thus $\theta(\lambda) = \theta(e_0)$:

$$\theta(c) \leq \theta(e_0) = -\frac{1}{2} \left(K_M + \sqrt{4e_0 K_M + K_M^2} \right). \quad (130)$$

The application of Gronwall's Lemma reveals

$$\limsup_{t \rightarrow \infty} (c - h^-(p))^2 \leq \lambda^2 \left(\frac{\eta}{1 + \eta} \right)^2 \left(\frac{1}{1 + \kappa} \right)^2 \left(\frac{2K_M}{K_M + \sqrt{K_M^2 + 4e_0 K_M}} \right)^2 \equiv \lambda^2 \varepsilon_{LT}^2, \quad (131)$$

which vanishes as $K_M \rightarrow 0$, even if $e_0 < s_0$. Thus, the long-time bound “ ε_{LT} ” vanishes if $e_0 \rightarrow 0$, $K_M \rightarrow 0$, $e_0 \rightarrow \infty$, $K_M \rightarrow \infty$, or $k_2 \rightarrow 0$. Finally, note that

$$\frac{K_M}{e_0 - \lambda + K_M} = \frac{K_M}{-\theta(e_0)} = \frac{K_M}{|\theta(e_0)|} \quad (132)$$

when $e_0 = s_0$, so there is a nice “overlap” of small parameters. Note also that the rate of convergence as $K_M \rightarrow 0$ is $\mathcal{O}(\sqrt{K_M/e_0})$ when $e_0 \approx s_0$. This is as expected. In fast/slow systems that contain a transcritical point, trajectories remain at a distance that is $\mathcal{O}(\sqrt{\varepsilon})$ from the transcritical point once $0 < \varepsilon \ll 1$ (again, see [53] for details).

7. Discussion

The primary goal of this work was to illustrate how a combination of scaling and non-scaling methods can be employed to determine the validity of the sQSSA, extended QSSA, rQSSA, and tQSSA. Collectively, our work provides an analysis that admits a considerably clearer understanding of each QSSA, its corresponding origin, and its validity. Using energy methods, we have recovered a clear set of associated “small parameters” that correspond to the three types of fast/slow dynamics considered under the QSSA: the sQSSA, the rQSSA, and the tQSSA (see Table 1).

The closure of the rQSSA problem for the MM reaction mechanism (1) is a testament to the power of using a combination of scaling and non-scaling methods, as we were able to derive a “small parameter” that ensures the validity of the rQSSA, $\underline{\varepsilon} \equiv K_M(e_0 - \lambda)$. In previous studies [30, 31], it was assumed that the rQSSA is valid only when $s_0 \ll e_0$. The small parameter, $\underline{\varepsilon}$, is a major improvement on the small parameters previously reported in the literature, because it quantifies and illustrates the validity of the rQSSA in regimes where $e_0 \lesssim s_0$. Moreover, for the first time, we have shown that the critical manifold associated with the rQSSA at finite initial enzyme concentration contains a transcritical singularity that significantly influences the validity of the rQSSA whenever $e_0 < s_0$ and $K_M \ll e_0$, since the critical (sub) manifolds undergo an exchange of stability at the singular point. On the other hand, this singularity is removed from the physical domain of interest if one takes $e_0 \rightarrow \infty$. This is the distinguishing difference between our analysis and the previous analysis of Schnell and Maini [31].

Our analysis has practical implications in the laboratory for the use of mathematical approximations under the rQSSA. Experimental assays are typically implemented to quantify the enzyme activity of a specific reaction. This is done by fitting experimental data to a particular model equation derived from a QSS reduction (i.e., either the sQSSA, the rQSSA, or the tQSSA). When the rQSSA is valid, progress curve experiments for p can be fitted to the product formation rate expression (23), which will allow the estimation of the catalytic rate constant, k_2 . In mathematical terms, the procedure of fitting experimental data to a mathematical model so as to estimate kinetic parameters is an *inverse problem* [15, 16]. Contrary to what has been reported, experimental assays can be carried out when initial substrate and enzyme ratios are of similar magnitude, and the rQSSA can be used to estimate k_2 , as long as $\underline{\varepsilon} \ll 1$.

The mathematical analysis of the Michaelis–Menten reaction mechanism (1) has a long and rich history. We believe that our analysis has deepened the overall understanding of the QSSA, and that our approach provides an important improvement to the implementation of this approximation in enzyme kinetics. The methods we employed can be adapted to analyze complex enzyme catalyzed reactions, or multiscale dynamical systems. This will be the focus of our future work.

Acknowledgement

We are grateful to Sebastian Walcher (RWTH Aachen University) and Wylie Stroberg (University of Michigan Medical School) for reading and providing comments to previous versions of this manuscript. Justin Eilertsen is supported by the University of Michigan Postdoctoral Pediatric Endocrinology and Diabetes Training Program “Developmental Origins of Metabolic Disorder” (NIH/NIDDK Grant: K12 DK071212).

Appendix A. Appendix

Here we present a detailed outline of the the method of slowly-varying Lyapunov functions applied to the governing equations of the Michaelis–Menten reaction mechanism as discussed in Section 4.3

Appendix A.1. Gronwall’s lemma

This differential version can be found in [63]. Suppose that $u(t) \in C^1([0, T]; \mathbb{R}_0^+)$ satisfies¹⁶

$$\frac{du}{dt} \leq \lambda u + f(t), \quad u(0) = u_0. \quad (\text{A.1})$$

Then, for $\lambda \in \mathbb{R}$ and $f(t) \in L^1([0, T]; \mathbb{R}_0^+)$, it holds that:

$$u(t) \leq e^{\lambda t} u_0 + e^{\lambda t} \int_0^t e^{-\lambda \tau} f(\tau) d\tau, \quad \forall t \in [0, T]. \quad (\text{A.2})$$

Appendix A.2. The “contractivity” condition: Existence of a Lyapunov function

It will help to briefly review the concept of a Lyapunov function before we introduce the idea of a “contractivity” condition. Consider a simple dynamical system of the form

$$x' = f(x), \quad (\text{A.3})$$

where “ $'$ ” denotes differentiation with respect to time, t . The vector field $f: \mathbb{R}^n \mapsto \mathbb{R}^n$ generates a flow, and any scalar-valued function of x (we will call this function $L(x): \mathbb{R}^n \mapsto \mathbb{R}$) will change with respect to time according to:

$$L' = f(x) \cdot \nabla L \equiv \mathcal{L}L, \quad (\text{A.4})$$

where the operator “ \mathcal{G} ” is known as a “generator.” Now suppose that the dynamical system (A.3) has an equilibrium point at $x = x^*$, so that $f(x^*) = 0$. Linearization is typically the preferred method of choice to determine the stability of x^* . However, it is well-known that if

$$0 < L, \ \& \ \mathcal{G}L < 0, \quad \forall x \in \mathcal{B} \setminus \{x^*\}, \quad (\text{A.5})$$

where x^* is an interior point of the neighborhood (ball) \mathcal{B} , then the equilibrium point x^* is asymptotically stable, and “ L ” is called a *Lyapunov* function.

So what does this have to do with a contraction condition? Keeping the idea of a Lyapunov function in mind, consider a fast/slow system:

$$x' = \varepsilon f(x, y), \quad x(0) = x_0 \quad (\text{A.6a})$$

$$y' = g(x, y), \quad y(0) = y_0, \quad (\text{A.6b})$$

where the slow variable, $x \in \mathbb{R}^n$, the fast variable $y \in \mathbb{R}^m$, and thus $f(x, y) : \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}^n$ and $g(x, y) : \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}^m$. The “fast subsystem” associated with (A.6a)-(A.6b) is

$$y' = g(x_0, y). \quad (\text{A.7})$$

Assume the fast subsystem has an equilibrium point $y^* = h_0(x_0)$, where $h : \mathbb{R}^n \mapsto \mathbb{R}^m$, and $g(x_0, h_0(x_0)) = 0$. Since x_0 is a constant with respect to the fast subsystem, g can be treated as a map from \mathbb{R}^m to \mathbb{R}^m . Once again, the stability of y^* can be found via linearization. However, suppose we want to determine an appropriate Lyapunov function. Generally, finding a Lyapunov function can be difficult, and so it is common practice to write down a simple polynomial function and “check” to see if it satisfies the requirements given in (A.5). The simplest choice is a quadratic function, $L(y) = \|y - h_0(x_0)\|^2$, which is locally positive-definite since $L(y) > 0, \forall y \neq y^* = h_0(x_0)$. The second property we need to check is whether or not L' is locally negative-definite; thus, we need to show that

$$L' = 2\langle (y - y^*), g(x_0, y) \rangle < 0, \quad \forall y \in \mathcal{B} \setminus \{y^*\}, \quad (\text{A.8})$$

where $\langle \cdot, \cdot \rangle$ denotes the usual scalar product between two vectors, and $\|v\| = \sqrt{\langle v, v \rangle}$. It is not entirely clear how to go about showing that L' is negative definite, since $g(x_0, y)$ and $(y - y^*)$ can change signs. However, a little “trick” can go a long way. First, rewrite (A.8) as

$$2\langle (y - y^*), g(x_0, y) \rangle = 2\langle (y - y^*), g(x_0, y) - g(x_0, y^*) \rangle. \quad (\text{A.9})$$

Now suppose we know that there is some positive number “ ζ ” such that

$$\langle (y - y^*), g(x_0, y) - g(x_0, y^*) \rangle \leq -\zeta \|y - y^*\|^2. \quad (\text{A.10})$$

If such a number “ ζ ” exists, then we can bound L'

$$L' = 2\langle (y - y^*), g(x_0, y) \rangle \leq -2\zeta \|y - y^*\|^2 \leq 0, \quad (\text{A.11})$$

and thus we have shown that L' is locally negative-definite. Consequently, $L = \|y - y^*\|^2 \equiv \|y - h_0(x_0)\|^2$ is a Lyapunov function. Why then, do we refer to (A.10) as a “contractivity” condition? Well, notice that (A.11) defines a differential inequality

$$L' \leq -2\zeta L, \quad (\text{A.12})$$

and integrating this inequality (courtesy of Gronwall’s lemma) yields

$$L \leq L(0)e^{-2\zeta t}. \quad (\text{A.13})$$

Consequently, the distance between y_0 and $h_0(x_0)$ decays to zero exponentially; hence, the distance “contracts.”

In the subsection that follows, we will investigate the growth of $L = \|y - h_0(x)\|^2$, but we will allow x to change in time, since x will vary slowly in comparison to y when $0 < \varepsilon \ll 1$.

Thus, we will be interested in finding an upper bound on

$$\frac{dL}{dT} \equiv \dot{L}, \quad (\text{A.14})$$

where $T = \varepsilon t$ is the slow timescale. The motivation for doing this should be obvious: we want to compute an estimate on how well the reduced system, $\dot{x} = f(x, h_0(x))$, approximates the full system. Because we will take into account the temporal variation of x , L is no longer a Lyapunov function. However, because x varies slowly, we have chosen to refer to L in this case as a *slowly-varying Lyapunov function*. One could also refer to this procedure as simply an “energy method,” since we are ultimately bounding the energy (norm) of the error $y - h_0(x)$.

Appendix A.3. General procedure: Estimation of bounds via energy methods

Let us start by estimating bounds on the enslavement of the fast variable. The calculation outlined below follows from the original work of Tikhonov [13] and Gradshtein [14], and the particular calculation we outline here can be found in [63]. This procedure will then be used to develop bounds for enslavement in the (p, c) coordinate system. We will express the procedure in terms with respect to a the fast/slow system (A.6a)-(A.6b) and, since our specific problem of interest is two-dimensional, we will assume $x \in \mathbb{R}$ and $y \in \mathbb{R}$. However, the reader should keep in mind that this procedure extends to higher-dimensional problems.

With respect to the slow time, the system (A.6a)-(A.6b) is expressed in form

$$\dot{x} = f(x, y), \quad (\text{A.15a})$$

$$\varepsilon \dot{y} = g(x, y), \quad (\text{A.15b})$$

where “ $\dot{\cdot}$ ” denotes differentiation with respect to the slow time, T . Let $g(\cdot, y)$ be contractive in the sense that

$$(g(x, y_1) - g(x, y_2)) \cdot (y_1 - y_2) \leq -\zeta \cdot (y_1 - y_2)^2, \quad (\text{A.16})$$

where the “ \cdot ” in (A.16) and in what follows denotes scalar multiplication. Next, suppose that $h_0(x)$ satisfies $g(x, h_0(x)) = 0$. Defining $y = h_0(x) + z$, where $z = y - h_0(x)$, it follows that

$$\dot{z} = \dot{y} - D_x h_0(x) \cdot \dot{x}.$$

Then, the differential equation for the energy, $L \equiv z^2 = (y - h_0(x))^2$, is

$$\frac{1}{2} \dot{z}^2 = \frac{1}{\varepsilon} z \cdot g(x, z + h_0(x)) - z \cdot D_x h_0(x) \cdot \dot{x}. \quad (\text{A.17})$$

Next, we want to get this expression into a useful form so that we can apply Gronwall’s lemma. There are essentially two “tricks” that can be utilized to manipulate this expression: (i.) we can “add/subtract” zero, or (ii.) we can “multiply” by one. We will choose (i). The term $\varepsilon^{-1} z \cdot g(x, z + h_0(x))$ can be expanded by “subtracting” zero,

$$\frac{1}{\varepsilon} z \cdot g(x, z + h_0(x)) = \frac{1}{\varepsilon} [z \cdot g(x, z + h_0(x)) - z \cdot g(x, h_0(x))], \quad (\text{A.18})$$

which holds since $g(x, h_0(x)) = 0$. Thus, we obtain

$$\frac{1}{2} \dot{z}^2 = \frac{1}{\varepsilon} [z \cdot g(x, z + h_0(x)) - z \cdot g(x, h_0(x))] - z \cdot D_x h_0(x) \cdot \dot{x}. \quad (\text{A.19})$$

Our primary interest is not in finding an actual solution to (A.19). Instead, we want to compute an upper bound on the growth of z^2 , which means we need to convert the differential equation (A.19) into a differential *inequality*. To do this, we will take advantage of the “contractivity” condition. First rewrite the first term on the right hand side of (A.19) as:

$$[z \cdot g(x, z + h_0(x)) - z \cdot g(x, h_0(x))] = [g(x, z + h_0(x)) - g(x, h_0(x))] \cdot [z + h_0(x) - h_0(x)]. \quad (\text{A.20})$$

Let $z + h_0(x)$ denote “ y_1 ” in (A.16), and let $h_0(x)$ denote “ y_2 ” in (A.16). Then

$$(g(x, z + h_0(x)) - g(x, h_0(x))) \cdot (z + h_0(x) - h_0(x)) \leq -\zeta \cdot z^2 \quad (\text{A.21})$$

It follows from (A.21) that

$$\frac{1}{2}\dot{z}^2 \leq -\frac{\zeta}{\varepsilon}z^2 - z \cdot D_x h_0(x) \cdot \dot{x}. \quad (\text{A.22})$$

Maximizing the inequality (A.22) yields

$$\frac{1}{2}\dot{z}^2 \leq -\frac{\zeta}{\varepsilon}z^2 + |z| \max |D_x h_0(x)| \max |\dot{x}|. \quad (\text{A.23})$$

In order to use Gronwall's lemma, we want the right hand side of (A.23) to be in terms of z^2 instead of $|z|$. To do this, let $a \equiv |z|$ and $b \equiv \max |D_x h_0(x)| \max |\dot{x}|$. Then, from Cauchy's inequality, $(a \cdot b \leq \delta a^2 + b^2/4\delta)$, it holds that

$$|z| \max |D_x h_0(x)| \max |\dot{x}| \leq \delta z^2 + \frac{(\max |D_x h_0(x)| \max |\dot{x}|)^2}{4\delta}, \quad (\text{A.24})$$

$\forall \delta > 0$.

Choosing $\delta = \zeta/2\varepsilon$ yields

$$|z| \max |D_x h_0(x)| \max |\dot{x}| \leq \frac{\zeta}{2\varepsilon}z^2 + \varepsilon \frac{(\max |D_x h_0(x)| \max |\dot{x}|)^2}{2\zeta}. \quad (\text{A.25})$$

Substitution of (A.25) into (A.23) reveals the differential inequality

$$\dot{z}^2 \leq -\frac{\zeta}{\varepsilon}z^2 + \varepsilon \frac{(\max |D_x h_0(x)| \max |\dot{x}|)^2}{2\zeta}. \quad (\text{A.26})$$

The expression (A.26) is now in a form that be integrated directly, and it follows from Gronwall's lemma that,

$$z^2 \leq z^2(0)e^{-\zeta T/\varepsilon} + \varepsilon^2 \frac{(\max |D_x h_0(x)| \max |\dot{x}|)^2}{\zeta^2} (1 - e^{-\zeta T/\varepsilon}), \quad (\text{A.27a})$$

$$\leq z^2(0)e^{-\zeta T/\varepsilon} + \varepsilon^2 \frac{(\max |D_x h_0(x)| \max |\dot{x}|)^2}{\zeta^2}. \quad (\text{A.27b})$$

Finally, from the triangle inequality, we obtain:

$$|z| \equiv |y - h_0(x)| \leq |y_0 - h_0(x_0)| e^{-\zeta T/2\varepsilon} + \varepsilon \frac{\max |D_x h_0(x)| \max |\dot{x}|}{\zeta}. \quad (\text{A.28})$$

Appendix A.4. Computation of the upper bound for enslavement of c in (p, c) coordinates

Here we will directly calculate the constants ζ , $\max |D_x h(x)|$ and $\max |\dot{x}|$ need for the upper bound. Note that the computation of ζ is calculated over the domain $(p, c) \in \Omega \equiv [0, s_0] \times [0, \lambda]$, since trajectories that start on $(p, c)(0) = (0, 0)$ are confined to this domain due to conservation.

In the (p, c) coordinate system, the c -nullcline is given by:

$$h(p) = \frac{1}{2} \left(e_0 + K_M + s_0 - p - \sqrt{(e_0 + K_M + s_0 - p)^2 - 4e_0(s_0 - p)} \right). \quad (\text{A.29})$$

Finding $\max |h'(p)|$ is relatively straightforward. First,

$$\frac{d^2 h}{dp^2} = - \frac{2e_0 K_M}{((e_0 + K_M + s_0 - p)^2 - 4e_0(s_0 - p))^{3/2}} < 0, \quad (\text{A.30})$$

and therefore the derivative is monotonically decreasing on the interval $p \in [0, s_0]$. Thus, we obtain:

$$\max \left| \frac{dh}{dp} \right| = - \frac{dh}{dp} \Big|_{p=s_0} = \frac{e_0}{K_M + e_0}. \quad (\text{A.31})$$

For contraction, we define

$$G(c, p) \equiv \dot{c} = k_1 c^2 - k_1 (K_M + e_0 + s_0 - p)c + k_1 e_0 (s_0 - p). \quad (\text{A.32})$$

We will use the Mean Value Theorem to find ζ . The derivative of $G(c, p)$ is given by

$$\frac{\partial G}{\partial c} = -k_1(e_0 + s_0 + K_M - 2c - p) \leq -k_1(e_0 + K_M + s_0) + \sup k_1(2c + p). \quad (\text{A.33})$$

We need to find $\sup(2c + p)$. A straightforward bound is $2\lambda + s_0$. However, we want to find the supremum subject to the constraints: $s_0 = s + c + p$ and $c \leq \lambda$. Employing the conservation law " $s_0 = s + c + p$ yields

$$2c + p = s_0 - s + c \leq s_0 + \lambda. \quad (\text{A.34})$$

Consequently, since $2c + p = s_0 + c - s \leq s_0 + \lambda$, it follows that

$$|c - c_0(p)| \leq \lambda e^{-t\zeta_T/2} + \frac{\eta\lambda}{(1+\eta)(1+\kappa)} \left(\frac{K_M}{e_0 - \lambda + K_M} \right), \quad (\text{A.35})$$

where ζ_T is given by

$$\zeta_T = k_1(e_0 - \lambda + K_M). \quad (\text{A.36})$$

Diving both sides of (A.35) by λ and expressing $p = s_0 \bar{p}$ yields the dimensionless form of the inequality

$$|\underline{c} - \bar{c}_0(\bar{p})| \leq e^{-\tilde{T}/2\varepsilon\nu} + \frac{\underline{\varepsilon}}{(1+\underline{\varepsilon})} \frac{\eta}{(1+\eta)} \frac{1}{(1+\kappa)}. \quad (\text{A.37})$$

Since $\nu \gg 1$, we obtain

$$|\underline{c} - \bar{c}_0(\bar{p})| \leq e^{-\tilde{T}/2\varepsilon} + \underline{\varepsilon}. \quad (\text{A.38})$$

Furthermore, since $\zeta_T \approx k_1 K_M$, it holds that

$$|\underline{c} - \bar{c}_0(\bar{p})| \leq e^{-\tilde{T}/2\varepsilon S S} + \eta. \quad (\text{A.39})$$

Let us define a “long-time epsilon, ε_L ,”

$$\varepsilon_L \equiv \frac{\underline{\varepsilon}}{(1+\underline{\varepsilon})} \frac{\eta}{(1+\eta)} \frac{1}{(1+\kappa)}. \quad (\text{A.40})$$

Expressing the exponential in terms of Tzafiriri’s timescale, $T_z = k_2 \lambda t / s_0$, we obtain:

$$|\underline{c} - \bar{c}_0(\bar{p})| \leq e^{-s_0 T_z / 2\lambda\varepsilon\nu} + \varepsilon_L. \quad (\text{A.41})$$

Moreover, we can define a “decay epsilon,” $\varepsilon_D \equiv s_0^{-1} \lambda \varepsilon \nu$, and write

$$|\underline{c} - \bar{c}_0(\bar{p})| \leq e^{-T_z / 2\varepsilon_D} + \varepsilon_L, \quad (\text{A.42})$$

where $\varepsilon_D \gg \varepsilon_L$.

Appendix A.5. The tQSSA as used in practice

The c -nullcline is rarely used for parameter estimation in progress curve experiments. Typically, at high enzyme concentrations, one uses

$$c \approx \frac{e_0(s_0 - p)}{e_0 + K_M + s_0 - p} \quad (\text{A.43})$$

in order to employ

$$\frac{dp}{dt} \approx \frac{k_2 e_0 (s_0 - p)}{e_0 + K_M + s_0 - p} \quad (\text{A.44})$$

for parameter estimation (see [16] for a good discussion on the tQSSA and parameter estimation). Calculating an upper bound on

$$|\mathcal{E}_T| \equiv \left| c - \frac{e_0 (s_0 - p)}{e_0 + K_M + s_0 - p} \right| \quad (\text{A.45})$$

is done using the methods previously. First,

$$\frac{1}{2} \frac{d\mathcal{E}_T^2}{dt} = -k_1 (e_0 + K_M + s_0 - p) \mathcal{E}_T^2 + \left[k_1 c^2 + \frac{k_2 c e_0 (e_0 + K_M)}{(e_0 + K_M + s_0 - p)^2} \right] \mathcal{E}_T, \quad (\text{A.46a})$$

$$\leq -k_1 (e_0 + K_M) \mathcal{E}_T^2 + \left[k_1 \lambda^2 + \frac{k_2 \lambda e_0}{e_0 + K_M} \right] |\mathcal{E}_T|. \quad (\text{A.46b})$$

Applying the Cauchy- δ inequality with $\delta = k_1 (e_0 + K_M)/2$, followed by integration, yields

$$\mathcal{E}_T^2 \leq \mathcal{E}_T^2(0) e^{-k_1 (K_M + e_0)t} + \frac{1}{(k_1 (e_0 + K_M))^2} \left[k_1 \lambda^2 + \frac{k_2 \lambda e_0}{e_0 + K_M} \right]^2. \quad (\text{A.47})$$

Finally, after scaling and algebraic simplification, it holds that:

$$|\mathcal{E}_T| \leq |\mathcal{E}_T(0)| e^{-k_1 (K_M + e_0)t/2} + \lambda \left[\frac{\lambda}{e_0 + K_M} + v \left(\frac{e_0}{e_0 + K_M} \right) \left(\frac{K_M}{e_0 + K_M} \right) \right]. \quad (\text{A.48})$$

In terms of T_z , the inequality (A.48) is

$$|\bar{\mathcal{E}}_T| \leq |\bar{\mathcal{E}}_T(0)| e^{-s_0 T_z / 2\lambda\tilde{e}} + \frac{\eta(1 + v\tilde{e})}{1 + \eta}, \quad \text{if } e_0 \leq s_0, \quad (\text{A.49a})$$

$$|\bar{\mathcal{E}}_T| \leq |\bar{\mathcal{E}}_T(0)| e^{-s_0 T_z / 2\lambda\tilde{e}} + \frac{\lambda}{e_0 + K_M} + \eta v\tilde{e}, \quad \text{if } s_0 < e_0 \quad (\text{A.49b})$$

where $\bar{\mathcal{E}}_T = \lambda^{-1} \mathcal{E}_T$ and $\tilde{e} \equiv K_M / (e_0 + K_M)$. Thus, the tQSSA as used in practice is valid at very low e_0 as well as very high e_0 .

References

- [1]. Crampin EJ, Schnell S, McSharry PE, Mathematical and computational techniques to deduce complex biochemical reaction mechanisms, *Prog. Biophys. Mol. Biol* 86 (2004) 77–112. [PubMed: 15261526]
- [2]. Sauro HM, Kholodenko BN, Quantitative analysis of signaling networks, *Prog. Biophys. Mol. Biol* 86 (2004) 5–43. [PubMed: 15261524]

- [3]. Turner T, Schnell S, Burrage K, Stochastic approaches for modelling in vivo reactions, *Comput. Biol. Chem* 28 (3) (2004) 165–178. [PubMed: 15261147]
- [4]. Grima R, Leier A, Exact product formation rates for stochastic enzyme kinetics, *J. Phys. Chem. B* 121 (1) (2017) 13–23. [PubMed: 27959536]
- [5]. Bertram R, Rubin JE, Multi-timescale systems and fast-slow analysis, *Math. Biosci* 287 (2017) 105–121, ISSN 0025-5564. [PubMed: 27424950]
- [6]. Rubin J, Wechselberger M, The selection of mixed-mode oscillations in a Hodgkin-Huxley model with multiple timescales, *Chaos* 18 (1) (2008) 015105, 12, ISSN 1054-1500. [PubMed: 18377086]
- [7]. Booth V, Diniz Behn CG, Physiologically-based modeling of sleep-wake regulatory networks, *Math. Biosci* 250 (2014) 54–68, ISSN 0025-5564. [PubMed: 24530893]
- [8]. Couillet PH, Spiegel EA, Amplitude equations for systems with competing instabilities, *SIAM J. Appl. Math* 43 (4) (1983) 776–821, ISSN 0036-1399.
- [9]. Roberts AJ, Low-dimensional modelling of dynamics via computer algebra, *Comput. Phys. Commun* 100 (3) (1997) 215–230.
- [10]. Roberts AJ, The utility of an invariant manifold description of the evolution of a dynamical system, *SIAM J. Math. Anal* 20 (6) (1989) 1447–1458, ISSN 0036-1410.
- [11]. Crawford JD, Knobloch E, Symmetry and Symmetry-Breaking Bifurcations in Fluid Dynamics, *Annu. Rev. Fluid Mech* 23 (1) (1991) 341–387.
- [12]. O’Malley RE Jr., Singular perturbation methods for ordinary differential equations, vol. 89 of *Applied Mathematical Sciences*, Springer-Verlag, New York, ISBN 0-387-97556-X, 1991.
- [13]. Tikhonov A, Systems of differential equations containing small parameters in their derivatives, *Mat. Sb. (N.S.)* 31 (1952) 575–586.
- [14]. Gradšteĭn IS, Application of AM Lyapunov’s theory of stability to the theory of differential equations with small coefficients in the derivatives, *Mat. Sbornik N. S* 32 (1953) 263–286.
- [15]. Stroberg W, Schnell S, On the estimation errors of K_M and V from time-course experiments using the Michaelis–Menten equation, *Biophys. Chem* 219 (2016) 17–27. [PubMed: 27677118]
- [16]. Choi B, Rempala GA, Kim JK, Beyond the Michaelis–Menten equation: Accurate and efficient estimation of enzyme kinetic parameters, *Sci. Rep* 7 (2017) 17018. [PubMed: 29208922]
- [17]. Roberts AJ, Backwards theory supports modelling via invariant manifolds for non-autonomous dynamical systems, arXiv:1804.06998v4 2018.
- [18]. Heineken FG, Tsuchiya HM, Aris R, On the mathematical status of the pseudosteady state hypothesis of biochemical kinetics, *Math. Biosci* 1 (1967) 95–113.
- [19]. Laidler KJ, Theory of the transient phase in kinetics with special reference to enzyme systems, *Can. J. Chem* 33 (1955) 1614–1624.
- [20]. Reich JG, Sel’kov EE, Mathematical analysis of metabolic networks, *FEBS Lett.* 40 (1974) S112–S118.
- [21]. Schauer M, Heinrich R, Analysis of the quasi-steady-state approximation for an enzymatic one-substrate reaction, *J. Theor. Biol* 79 (1979) 425–442. [PubMed: 513815]
- [22]. Segel LA, On the validity of the steady state assumption of enzyme kinetics, *Bull. Math. Biol* 50 (1988) 579–593. [PubMed: 3219446]
- [23]. Burke MA, Maini PK, Murray JD, On the kinetics of suicide substrates, *Biophys. Chem* 37 (1990) 81–90. [PubMed: 2285805]
- [24]. Burke MA, Maini PK, Murray JD, Suicide substrate reaction-diffusion equations: Varying the source, *IMA J. Math. Appl. Med. Biol* 10 (1993) 97–114. [PubMed: 8370995]
- [25]. Schnell S, Mendoza C, Enzyme kinetics of multiple alternative substrates, *J. Math. Chem* 27 (1-2) (2000) 155–170, ISSN 0259-9791.
- [26]. Schnell S, Mendoza C, Time-dependent closed form solutions for fully competitive enzyme reactions, *Bull. Math. Biol* 62 (2000) 321–336. [PubMed: 10824432]
- [27]. Eilertsen J, Stroberg W, Schnell S, A theory of reactant-stationary kinetics for a mechanism of zymogen activation, *Biophys. Chem* 242 (2018) 34–44. [PubMed: 30218978]
- [28]. Eilertsen J, Stroberg W, Schnell S, Phase-plane geometries in coupled enzyme assays, *Math. Biosci* 306 (2018) 126–135, ISSN 0025-5564. [PubMed: 30261179]

- [29]. Eilertsen J, Stroberg W, Schnell S, Characteristic, completion or matching timescales? An analysis of temporary boundaries in enzyme kinetics, *J. Theoret. Biol* 481 (2019) 28–43, ISSN 0022-5193. [PubMed: 30615881]
- [30]. Segel LA, Slemrod M, The quasi-steady-state assumption: A case study in perturbation, *SIAM Rev.* 31 (1989) 446–477.
- [31]. Schnell S, Maini PK, Enzyme kinetics at high enzyme concentration, *Bull. Math. Biol* 62 (2000) 483–499. [PubMed: 10812718]
- [32]. Borghans JA, Boer RJD, Segel LA, Extending the quasi-steady state approximation by changing variables, *Bull. Math. Biol* 58 (1996) 43–63. [PubMed: 8819753]
- [33]. Schnell S, Maini PK, Enzyme kinetics far from the standard quasi-steady-state and equilibrium approximations, *Math. Comput. Modelling* 35 (1-2) (2002) 137–144, ISSN 0895-7177.
- [34]. Tzafiriri AR, Michaelis-Menten kinetics at high enzyme concentrations, *Bull. Math. Biol* 65 (2003) 1111–1129. [PubMed: 14607291]
- [35]. Goeke A, Walcher S, Zerz E, Determining “small parameters” for quasi-steady state, *J. Differ. Equations* 259 (2015) 1149–1180.
- [36]. Goeke A, Walcher S, A constructive approach to quasi-steady state reductions, *J. Math. Chem* 52 (2014) 2596–2626.
- [37]. Noethen L, Walcher S, Tikhonov’s theorem and quasi-steady state, *Dis. and Cont. Dyn. Sys. B* 16 (2011) 945–961.
- [38]. Fenichel N, Geometric singular perturbation theory for ordinary differential equations, *J. Differ. Equations* 31 (1979) 53–98.
- [39]. Hanson SM, Schnell S, Reactant stationary approximation in enzyme kinetics, *J. Phys. Chem. A* 112 (2008) 8654–8658. [PubMed: 18714952]
- [40]. Nguyen AH, Fraser S, Geometrical picture of reaction in enzyme kinetics, *J. Chem. Phys.* 91 (1989) 186–193.
- [41]. Roussel MR, Fraser SJ, Geometry of the steady-state approximation: Perturbation and accelerated convergence methods, *J. Chem. Phys* 93 (1990) 1072–1081.
- [42]. Roussel MR, Fraser SJ, Accurate steady-state approximations: Implications for kinetics experiments and mechanism, *J. Phys. Chem* 95 (1991) 8762–8770.
- [43]. Goeke A, Schilli C, Walcher S, Zerz E, Computing quasi-steady state reductions, *J. Math. Chem* 50 (2012) 1495–1513.
- [44]. Pedersen MG, Bersani AM, Bersani E, The total Quasi-Steady-State Approximation for fully competitive enzyme reactions, *Bull. Math. Biol* 69 (2006) 433–457. [PubMed: 16850351]
- [45]. Tzafiriri AR, Edelman ER, Quasi-steady-state kinetics at enzyme and substrate concentrations in excess of the Michaelis-Menten constant, *J. Theoret. Biol* 245 (4) (2007) 737–748, ISSN 0022-5193. [PubMed: 17234216]
- [46]. Dingee JW, Anton AB, A new perturbation solution to the Michaelis-Menten problem, *AIChE J.* 54 (2008) 1344–1357.
- [47]. Dell’Acqua G, Bersani AM, A perturbation solution of Michaelis-Menten kinetics in a “total” framework, *J. Math. Chem* 50 (5) (2012) 1136–1148, ISSN 0259–9791.
- [48]. Dhatt S, Bhattacharyya K, Single-substrate enzyme kinetics: the quasi-steady-state approximation and beyond, *J. Math. Chem* 51 (5) (2013) 1467–1477, ISSN 0259–9791.
- [49]. Pedersen MG, Bersani AM, Bersani E, Cortese G, The total quasi-steady-state approximation for complex enzyme reactions, *Math. Comput. Simulation* 79 (4) (2008) 1010–1019, ISSN 0378-4754.
- [50]. Bersani AM, Bersani E, Dell’Acqua G, Pedersen MG, New trends and perspectives in nonlinear intracellular dynamics: one century from Michaelis-Menten paper, *Contin. Mech. Thermodyn* 27 (4-5) (2015) 659–684, ISSN 0935-1175, doi: \let\@tempa\@bibinfo@X@doi10.1007/s00161-014-0367-4, URL 10.1007/s00161-014-0367-4.
- [51]. Battelli F, Lazzari C, Singular perturbation theory for open enzyme reaction networks, *IMA J. Math. Appl. Med. Biol* 3 (1) (1986) 41–51, ISSN 0265–0746. [PubMed: 3453832]
- [52]. Verhulst F, Singular perturbation methods for slow-fast dynamics, *Nonlinear Dynam.* 50 (2007) 747–753.

- [53]. Kuehn C, Multiple time scale dynamics, Springer, Cham, 2015.
- [54]. Hek G, Geometric singular perturbation theory in biological practice, *J. Math. Biol* 60 (3) (2010) 347–386, ISSN 0303-6812. [PubMed: 19347340]
- [55]. Roussel MR, Heineken, Tsuchiya and Aris on the mathematical status of the pseudosteady state hypothesis: A classic from Volume 1 of *Mathematical Biosciences*, *Math. Biosci* 318 (2019) 108274, 8, ISSN 0025-5564. [PubMed: 31697965]
- [56]. Boulter F, Lemaire F, Sedoglavic A, Ürgüplü A, Towards an automated reduction method for polynomial ODE models of biochemical reaction systems, *Math. Comput. Sci* 2 (3) (2009) 443–464, ISSN 1661-8270.
- [57]. Goeke A, Walcher S, Quasi-steady state: searching for and utilizing small parameters, in: *Recent trends in dynamical systems*, vol. 35 of *Springer Proc. Math. Stat*, Springer, Basel, 153–178, 2013.
- [58]. Goeke A, Walcher S, Zerz E, Classical quasi-steady state reduction—A mathematical characterization, *Phys. D* 345 (2017) 11–26.
- [59]. Roberts AJ, Appropriate initial conditions for asymptotic descriptions of the long term evolution of dynamical systems, *J. Austral. Math. Soc. Ser. B* 31 (1) (1989) 48–75, ISSN 0334-2700.
- [60]. Lax C, Walcher S, Singular perturbations and scaling, *Discrete Contin. Dyn. Syst. Ser. B* 25 (1) (2020) 1–29, ISSN 1531-3492.
- [61]. Roussel MR, A rigorous approach to steady-state kinetics applied to simple enzyme mechanisms, ProQuest LLC, Ann Arbor, MI, ISBN 978-0315-97252-0, thesis (Ph.D.)—University of Toronto (Canada), 1994.
- [62]. Fraser SJ, The steady state and equilibrium approximations: A geometrical picture, *J. Chem. Phys* 88 (8) (1988) 4732–4738.
- [63]. Pavliotis GA, Stuart AM, *Multiscale Methods: Averaging and homogenization*, Springer, New York, 2008.

Highlights

- A non-scaling method is utilized to derive conditions for the quasi-steady-state approximations.
- A parameter determining the validity of the reverse quasi-steady-state approximation is derived
- A dynamical transcritical bifurcation is shown to occur in the phase-plane of the model
- Our analysis provides a deeper understanding of the quasi-steady-state approximation

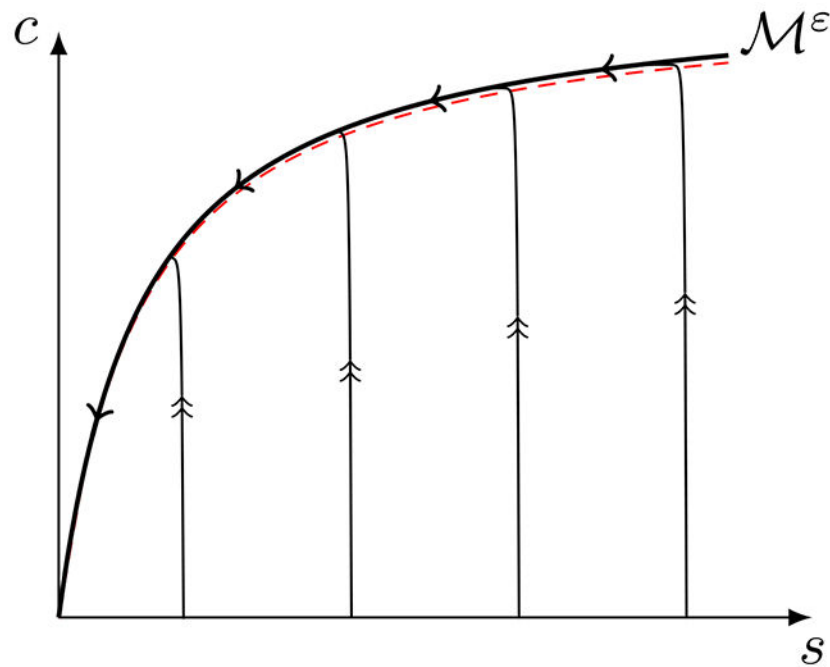


Figure 1: The RSA ensures the validity of the sQSSA: Fast/slow dynamics occur when $\epsilon_{SS} \ll 1$ for the MM reaction mechanism (1).

This figure is an illustration of the phase-plane dynamics when $\epsilon_{SS} \ll 1$ and the sQSSA is valid; it serves to convey conceptual aspects of the phase-plane when the RSA is valid. The red dashed curve is the c -nullcline. The black curve the lies just above the c -nullcline is the invariant slow manifold. The thin lines with double arrows are illustrative of typical phase-plane trajectories when $\epsilon_{SS} \ll 1$: trajectories starting on the s -axis approach the c -nullcline *almost* vertically. Once the trajectory reaches the c -nullcline it closely follows \mathcal{M}^ϵ during the QSS phase of the reaction. Double arrows indicate fast dynamics, and single arrows represent slow dynamics. The reader should bear in mind that, slightly contrary to what is illustrated, the tangent vector to the trajectory is perfectly horizontal when it intercepts the c -nullcline, since its vertical velocity component vanishes at the point of interception.

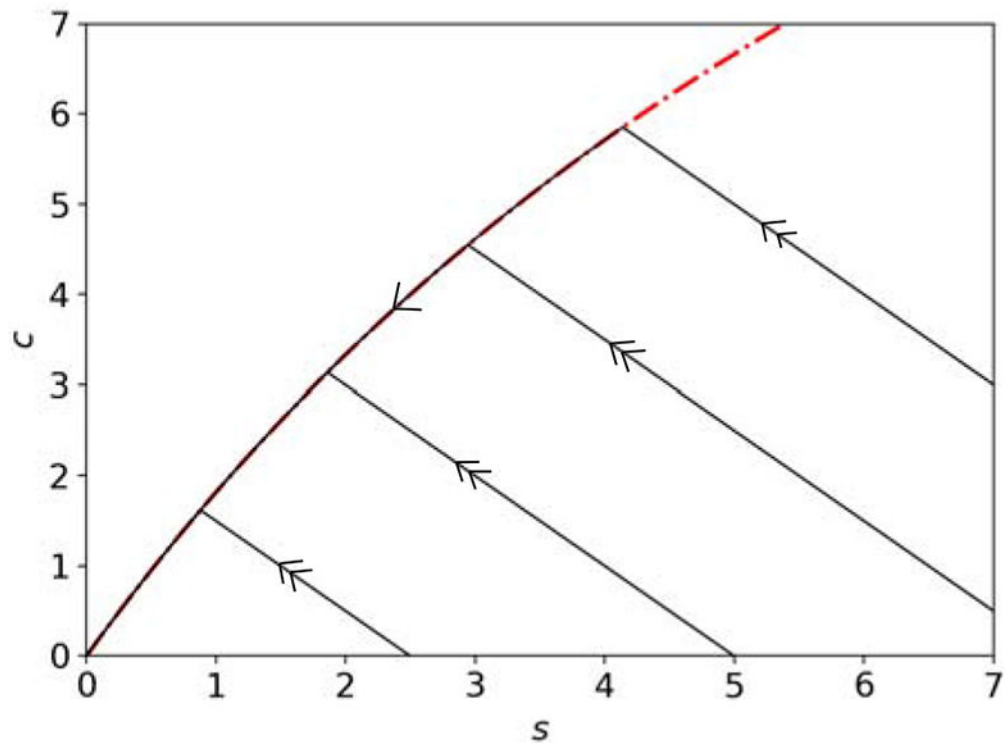


Figure 2: The extended QSSA: Fast/slow dynamics still occur when $\epsilon_{SS} \sim 1$, but $\nu \ll 1$ for the MM reaction mechanism (1).

The solid black curves are the numerical solutions to the mass action equations (5a)-(5c), and the dashed/dotted red curve is the c -nullcline. In this simulation, $k_1 = k_{-1} = 10$, $k_2 = 0.01$ with $\kappa \equiv k_{-1}/k_2 = 100$, $\nu = 1/101$, and $\epsilon_{SS} = 1$ in arbitrary units for illustrative purposes. Note that the trajectory is not a vertical straight line during the initial phase of the reaction. As a consequence, the RSA does not hold [30, 39], and there is noticeable depletion of substrate during the transient phase. However, the trajectory still follows the c -nullcline after a brief transient and achieves the slow QSS dynamics. The double arrows indicate fast dynamics, and single arrows indicate slow dynamics.

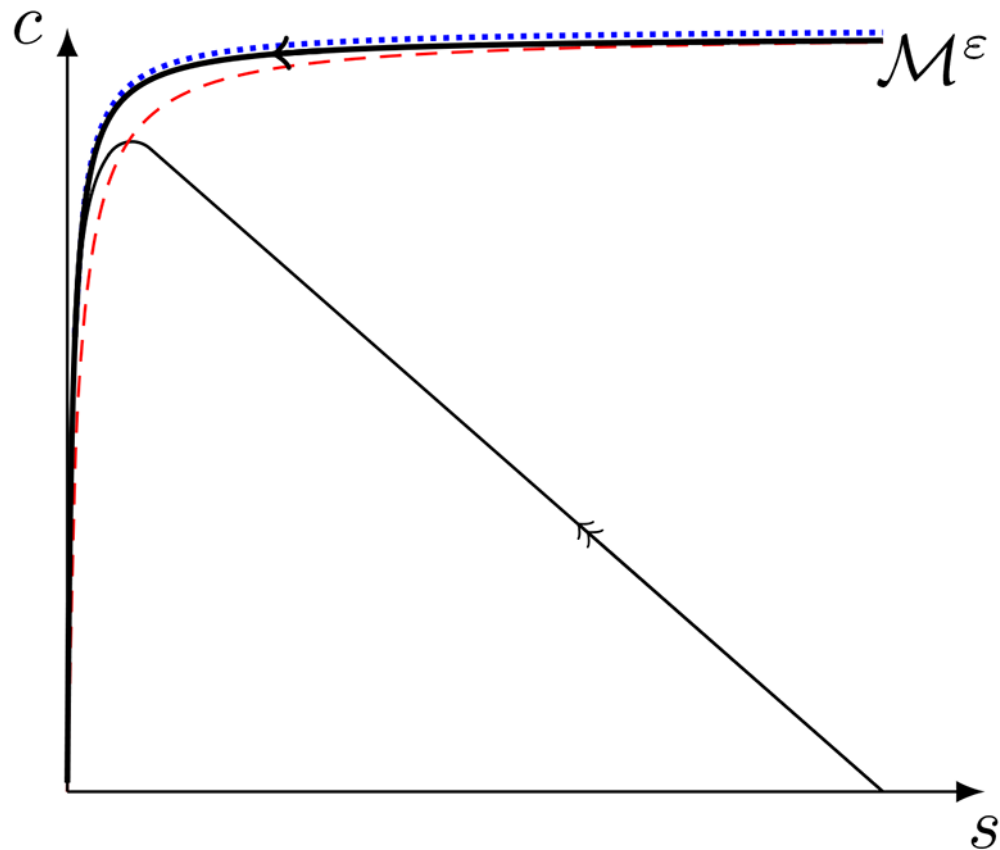


Figure 3: In the rQSSA, the substrate concentration is in a QSS phase in the MM reaction mechanism (1).

This figure is an illustration of the rQSSA. The thick black curve is an illustration of the slow manifold; the dashed red curve is the c -nullcline, and the dotted blue curve is the s -nullcline. The thin black curve illustrates a typical trajectory when the rQSSA is valid. During the transient phase, the complex concentration rapidly reaches its maximum value, and the initial substrate concentration nearly vanishes. After the transient phase, the trajectory follows the slow manifold, \mathcal{M}^ϵ , during which time s remains in a QSS phase. Double arrows represent fast dynamics, and single arrows represent slow dynamics.

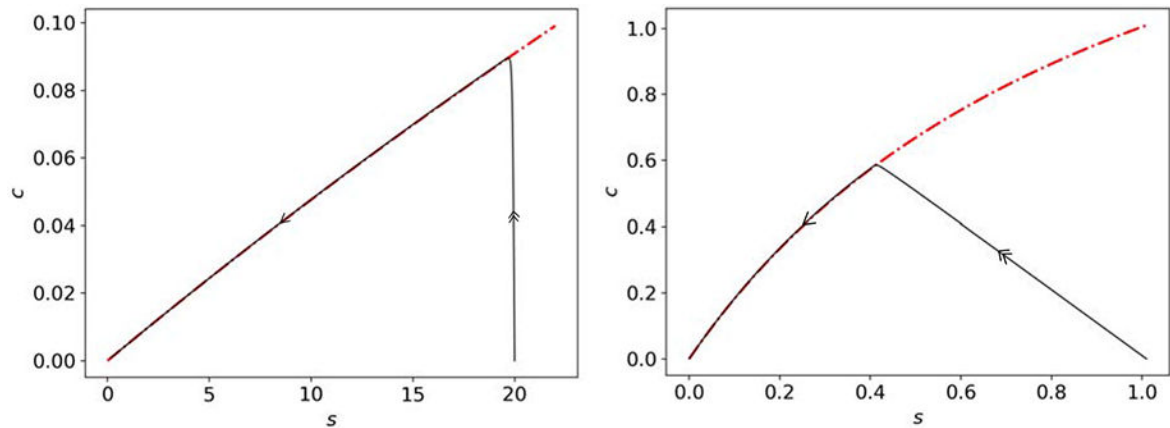


Figure 4: The extended QSSA equation (21) proposed by Segel and Slemrod [30] is invalid when $\nu \ll 1$ but $\epsilon_{SS} = \sigma = 1$.

In both panels, the solid black curve is the numerical solutions to the mass action equations (5a)-(5c), and the dashed/dotted red curve is the c -nullcline. **LEFT:** In this panel, $s_0 = 20$, $e_0 = 1$, $k_2 = k_{-1} = 10$, $k_1 = 0.1$, and the sQSSA is valid, and the trajectory moves almost vertically towards the c -nullcline, which is the hallmark of the sQSSA in the (s, c) phase-plane. **RIGHT:** This panel illustrates the path of a trajectory when $k_2 \ll k_{-1}$ but $\epsilon_{SS} = 1$. The parameters utilized in the numerical simulation are: $k_1 = k_{-1} = 1$, $k_2 = 0.01$ with $\kappa = 100$ and $\sigma = 1$. Notice the trajectory is no longer *vertical* during the initial phase of the reaction as there is substantial loss of substrate during the fast transient. Once the trajectory reaches the c -nullcline, the approximation given by (51) is valid, and can be equipped with the initial condition $(\sqrt{2} - 1)s_0$ if $\epsilon_{SS} = \sigma = 1$. Although the sQSSA *looks* valid in both panels based on the proximity of the trajectory to the c -nullcline, GSPT indicates it is only valid in the left panel. In this figure, double arrows represent fast dynamics, and single arrows represent slow dynamics. All units are arbitrary.

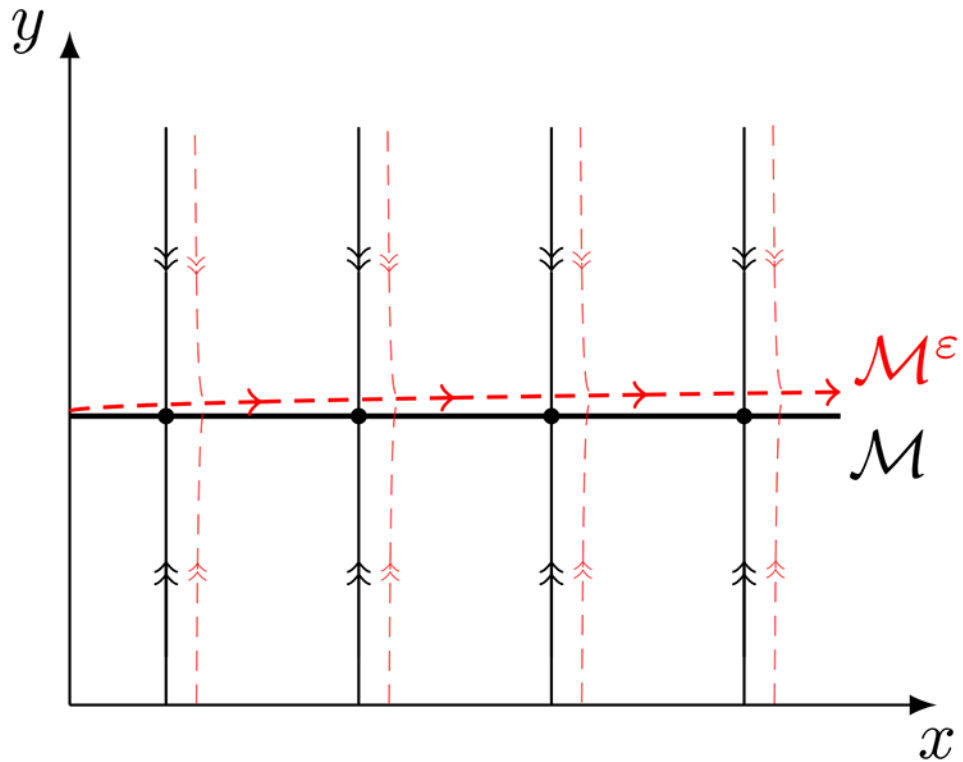


Figure 5: Illustration of a normally hyperbolic and invariant critical manifold, \mathcal{M} .

The black lines with arrows and fixed points represent trajectories within the unperturbed vector field; the red dashed lines represent trajectories in the perturbed vector field. The straight black lines with double arrowheads represent trajectories moving towards the fixed points (filled black circles) that lie on the normally hyperbolic, attracting and invariant manifold \mathcal{M} , depicted by the thick black line. The flow on \mathcal{M} is trivial. However, once ε is non-zero and the vector field is smoothly perturbed, an invariant manifold \mathcal{M}^ε emerges (red), on which the flow is slow, but no longer trivial. Trajectories that start off of \mathcal{M}^ε quickly approach it, then follow it tangentially. Obviously, the fixed points drawn on this manifold are isolated, so the reader should bear in mind that \mathcal{M} is filled with non-isolated fixed points.

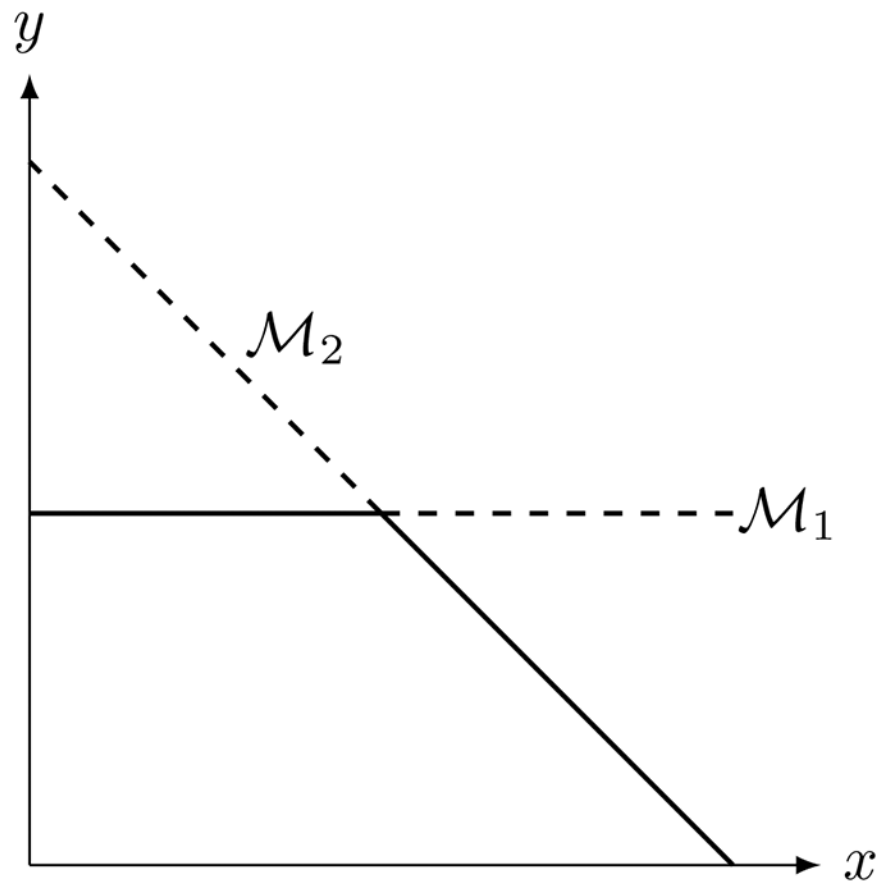


Figure 6: The transcritical bifurcation point corresponds to a loss of normal hyperbolicity and an exchange of stability.

Here we illustrate a transcritical bifurcation. Thick dashed lines corresponds to unstable fixed points in which (58) is greater than zero. Thick solid lines represent stable branches of fixed points for which (58) is negative. Two manifolds, \mathcal{M}_1 and \mathcal{M}_2 , cross at a singular point. At the crossing point (i.e., the point of intersection), normal hyperbolicity is lost, and (58) fails to hold. Each manifold undergoes a change in stability at the point of intersection (i.e., the singular point delimits the point at which the stability of each of the manifolds changes from attractive to repulsive or vice-versa).

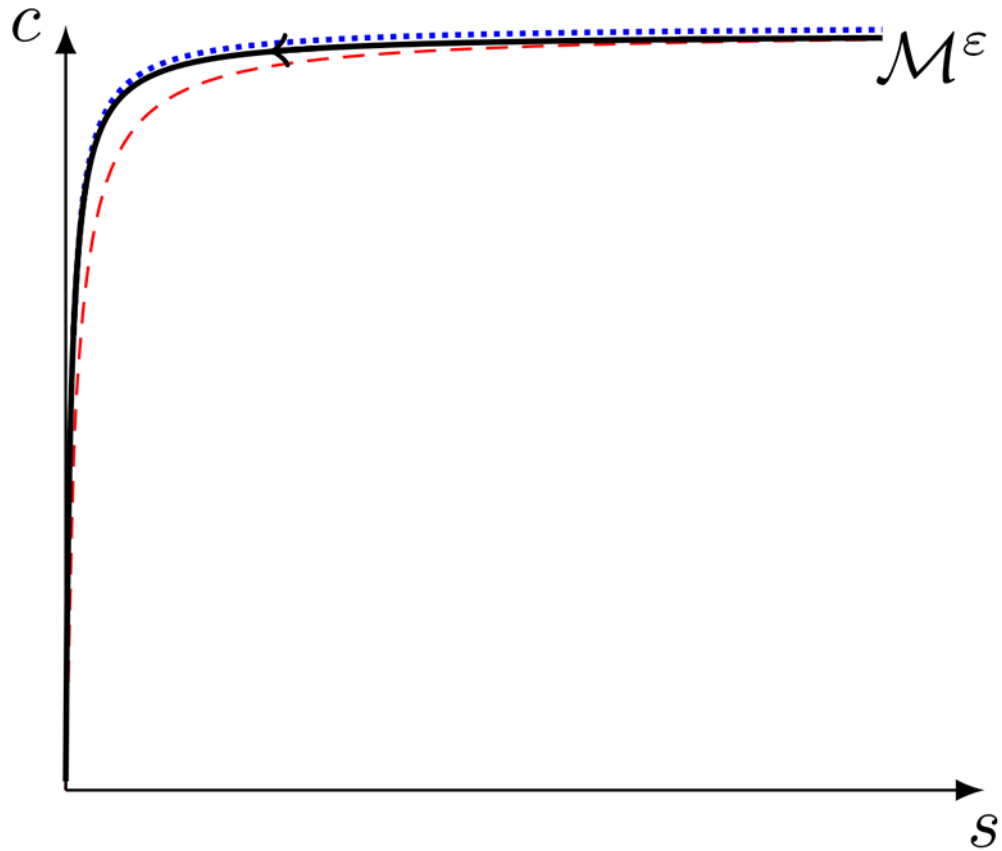


Figure 7: Near the origin, the slow manifold lies close to the s -nullcline in the (s, c) phase-plane when $1 \ll \eta$.

This figure is an illustration and is meant to convey conceptual features that occur in the phase-plane when $\eta \gg 1$. The solid black curve is the slow manifold; the dotted blue curve is the s -nullcline, and the dashed red curve is the c -nullcline. When η is sufficiently large, the slow manifold lies very close to the s -nullcline at sufficiently small s . Comparing with figure 3, this is exactly the condition we need in order to apply the rQSSA

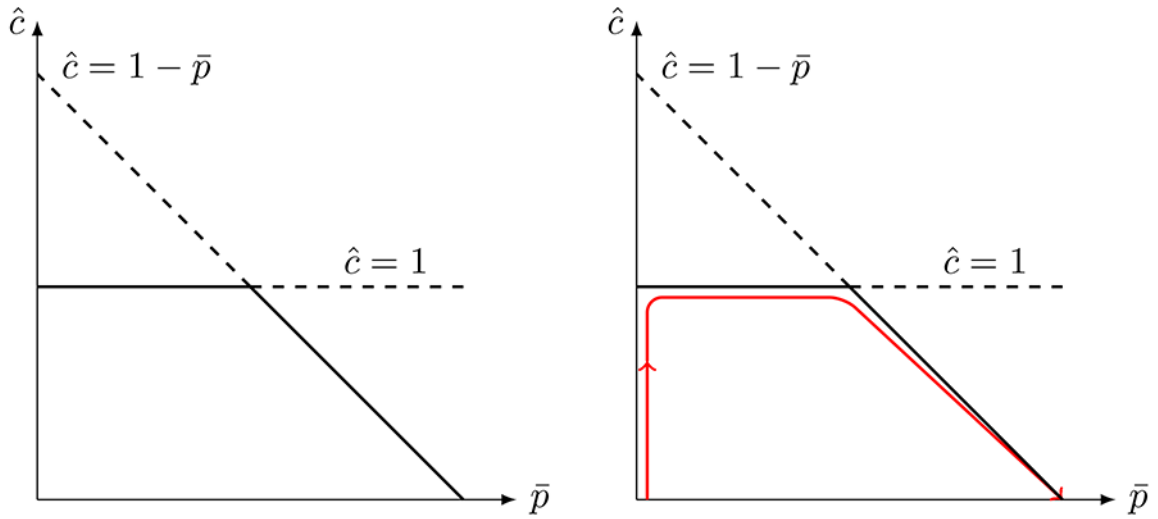


Figure 8: The critical set of the rQSSA when $e_0 < s_0$ and $K_M = 0$.

This figure provides a visualization of the invariant set recovered by setting $e_0 < s_0$ and $k_{-1} = k_2 = 0$. LEFT: The critical set contains two transversely intersecting branches of fixed points. Thick solid curves correspond to non-isolated *stable* fixed points, and thick dashed lines correspond to *unstable* fixed points. The horizontal curve corresponds to the critical set $\hat{c} = 1$, and the diagonal curve corresponds to the critical set $\bar{p} + \hat{c} = 1$. The trajectory rapidly approaches the curve $\hat{c} = 1$, then reaches the curve $1 = \hat{c} + \bar{p}$ once $t \sim t_g$ and begins descending towards the origin. Clearly, each set constitutes a normally hyperbolic manifold everywhere except where the branches intersect, which corresponds to a transcritical singularity. RIGHT: A typical trajectory (red solid curve) closely follows the attracting critical submanifolds once the perturbation is turned on.

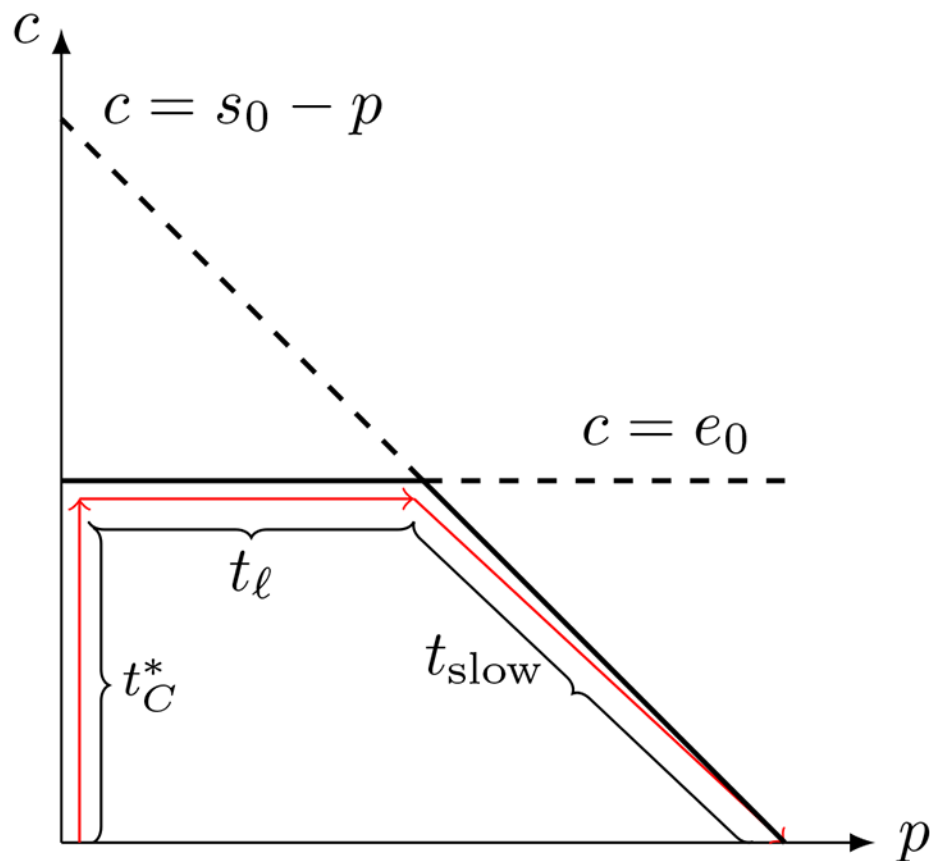


Figure 9: There are three stages in the dynamics of the MM reaction mechanism (1) when $e_0 < s_0$ and $K_M \approx 0$.

The red arrows demarcate the approximate path of a typical trajectory during each stage. Stage one is the fast transient: the trajectory rapidly approaches the curve $c = e_0$, and the duration of this stage is roughly t_C^* . During the second stage, the trajectory closely follows the curve $c = e_0$; the duration of this stage is roughly t_l . The final stage corresponds to the rQSSA, as the trajectory follows the curve $c = s_0 - p$. The timescale corresponding to stage three is $t_{\text{slow}} = 1/k_2$. If $t_C^* \ll t_l \ll 1/k_2$, then the rQSSA is at least heuristically valid.

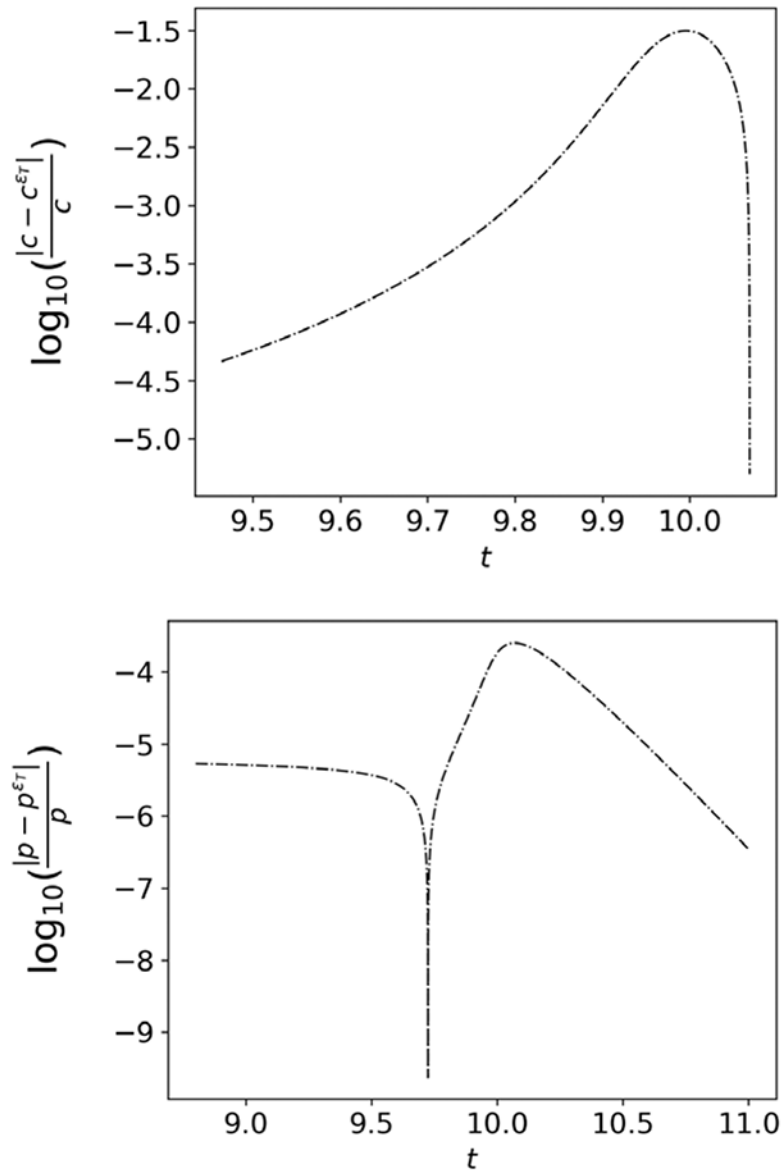


Figure 10: The tQSSA can worsen as $t \rightarrow \infty$ if ε_L is not sufficiently small.

In both panels, the numerical solution is obtained with parameter values: $\varepsilon_0 = 10$, $s_0 = 1000$, $k_1 = 20$, $k_{-1} = 10$, $k_2 = 10$. Thus, $\varepsilon_T \sim 10^{-6}$, but $\varepsilon_L \approx 0.45$. TOP PANEL: The \log_{10} of the relative error between the numerical solution to the mass action equations (27a)-(27b) for c , and the numerical solution to the asymptotic approximation (29), denoted c^{ε_T} , is demarcated by the dashed-dotted line. BOTTOM PANEL: The dashed-dotted line demarcates the \log_{10} of the relative error between the numerical solution to the mass action equations (27a)-(27b) for p , and the numerical solution to the asymptotic approximation (30), denoted p^{ε_T} . Even though timescale separation is quite large, there are regions in the numerically-obtained time course data where c^{ε_T} approximates c only to within one digit of accuracy. While the loss in accuracy may not be outright disastrous in terms of parameter estimation, it is an observation that is worth pointing out in order to gain a deeper quantitative understanding of

the various QSS approximations utilized for Michaelis–Menten reaction mechanism (1). The parameters have been assigned arbitrary units for illustrative purposes.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1:
Small parameters and critical sets for the standard, reverse, and total QSSA.

Each of the critical manifolds that appear when either η , $\underline{\varepsilon}$ or ν are identically zero correspond to zero product formation. Hence, the tQSSA is valid in all three cases. The notation $\bar{c}_0(\bar{p}; K_S)$ denotes the c -nullcline with K_M replaced with K_S , which occurs when $k_2 = 0$.

Parameter	Approximation	Critical Manifold (Set)
η	sQSSA	$1 - \bar{s} - \bar{p} = 0$
$\underline{\varepsilon}$	rQSSA	$\{(\hat{c}, \bar{p}) \in \mathbb{R}^+ \mid (1 - \hat{c})(1 - \hat{c} - \bar{p}) = 0\}$
ν	tQSSA	$\bar{c}_0(\bar{p}; K_S)$