A Geometric Approach to Systems with Multiple Time Scales*

Christopher K. R. T. Jones

Division of Applied Mathematics
Brown University
Providence, RI 02912

December 18, 1997

1 Introduction

Physical systems often involve several processes that are evolving on different time scales. The resulting equations have a specific structure which can be exploited to great effect and the kind of results that can be obtained, at least with a certain goal in mind, are the subject of this paper. On account of the physical motivation for the different scales, there are many areas in which applications can be found.

Apart from their physical relevance, there is another important reason for developing a theory tailored to multiple time scale systems. Systems in low dimensions, here one should think of dimensions 1 and 2, can often be analyzed by well-established techniques. Indeed, the asymptotic behavior is determined by either periodic orbits and/or critical points, as follows from the celebrated Poincaré-Bendixson Theorem in 2 dimensions and purely elementary considerations in 1 dimension. If the underlying dimension of some system under consideration is higher then we are faced with two difficulties: first, the techniques are largely unavailable for performing the analysis and, secondly, the phenomena occurring can be considerably more complicated, involving, for instance, chaotic behavior. To resolve this apparent intractability of higher-dimensional problems, techniques involving reductions to lower dimensional phase spaces are very attractive. Such techniques are often found through the presence of conserved quantities or symmetries. While these techniques do make certain higher dimensional problems susceptible to analysis, the resulting behavior must necessarily be characteristic of the lower dimensional system and thus the richness of motion available in higher dimensional systems is not reflected in systems for which this approach works. Systems with multiple time scales, however, offer a method of reduction which preserves the higher dimensional nature of the behavior. This idea seems counter-intuitive but is not unreasonable when one considers the fact that at least two separate reductions are being used and, although each leads to a lower dimensional system, the combined effect is to allow behavior characteristic of the full dimensional space.

It is a consequence of the Rectification Theorem, see Arnold [1], which states that a flow is locally trivial except near a critical point, that interesting behavior in a dynamical system will either occur near a critical point or be a result of recurrent motion. It is then not surprising

*To appear in The Bulletin of JSIAM (Ouyou-Suri), Vol. 7, No. 4, 1997 (Japanese).

that homoclinic orbits, which combine both the possibility of interesting behavior near a critical point and recurrent motion, are so important in the understanding of dynamical systems in general. Much of the work described in this paper is aimed at the development of techniques for constructing homoclinic orbits. The reduction afforded by the fast/slow decomposition of multiple time scale systems gives a way of determining homoclinic orbits in (possibly) high-dimensional spaces.

2 Formulation of Problem

The general form of the equations to be considered is

$$\begin{aligned} x' &= f(x, y, \epsilon), \\ y' &= \epsilon g(x, y, \epsilon), \end{aligned} \tag{1}$$

where $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^l$ and ϵ is a real parameter that is assumed to be small. The independent variable will be denoted t and so $' = \frac{d}{dt}$. The functions f and g are taken to be as smooth as needed. The key point to note about the system (1) is that the variable x is varying very rapidly compared with y. This can be viewed, in a way more clearly, by changing the time scale and setting

$$\tau = \epsilon t,$$

which has the effect of transforming (1) to

$$\begin{aligned} \epsilon \dot{x} &= f(x, y, \epsilon), \\ \dot{y} &= g(x, y, \epsilon). \end{aligned} \tag{2}$$

It is instructive to consider the limits of both (1) and (2) as $\epsilon \to 0$. These are

$$\begin{aligned} x' &= f(x, y, 0), \\ y' &= 0, \end{aligned} \tag{3}$$

and

$$\begin{aligned} f(x, y, 0) &= 0, \\ y' &= g(x, y, 0), \end{aligned} \tag{4}$$

respectively. While the limiting fast form (3) remains a differential equation, albeit with a trivialized part, the limiting slow form (4) has become a lower order differential equation coupled with a constraint. The motion on the fast time scale will satisfy the first of the equations in the system (3) with the role of y being reduced to that of a parameter. The motion on the slow time scale will be restricted to trajectories satisfying, at least approximately, the constraining condition that is the first equation in the system (4). These slow motions will then be governed by the second equation of that limiting system. Of course, both scalings agree on the level of phase space structure when $\epsilon \neq 0$ but offer very different perspectives since they differ radically in the limit when $\epsilon = 0$. The main goal of singular perturbation theory is to use these limits to understand structure in the full system when $\epsilon \neq 0$.

3 Travelling Pulses on Nerve Axons

For putting the analysis into context, it will be helpful to consider an application. Hodgkin-Huxley derived their celebrated equations in [10] in order to determine the relevant physical properties of a nerve axon and its environment that are responsible for the propagation of a voltage excitation. Their successful prediction of a travelling pulse has given us the explanation of nerve impulse propagation. The equations have the general form of a diffusion equation, for the voltage, coupled to subsidiary ordinary differential equations which govern the behavior of phenomenological variables related to chemical concentration differences across the membrane. Assuming that the voltage varies on a faster time scale than the other variables, the equations can then be written in the following form

$$\begin{aligned} u_t &= u_{xx} + f(u,w), \\ w_t &= \epsilon g(u,w), \end{aligned} \tag{5}$$

where f and g are nonlinearities to be specified. The dimension of w depends on how many chemicals are being taken into account. Here, it will be either 1 or 2 dimensional. The equations for a travelling wave of (5) are

$$\begin{aligned} u' &= v, \\ v' &= \theta v - f(u,w), \\ w' &= \frac{\epsilon}{\theta} g(u,w), \\ \theta' &= 0, \end{aligned} \tag{6}$$

where the wave has the form $(u(x+\theta t), w(x+\theta t))$ and the speed parameter θ has been added as a slow variable. The independent variable in (6) is $\zeta = x + \theta t$. The idea is that the nerve activation and recovery are governed by the fast variables (u and v) and that these are well separated by slow changes in chemical composition differences (governed by w.)

A picture can then be built as to how a travelling wave of (5) can be constructed as a homoclinic orbit of (6) to a fixed critical point which corresponds to the rest state. The nerve should initially be activated, which is represented in (6) as an orbit of the limiting fast subsystem jumping from the rest state to another slow manifold. This is then, formally, followed by a transition on the slow manifold. This motion on the slow manifold sets the system up for a further jump back to the slow manifold containing the original rest state. Finally, there is motion back to the rest state on this slow manifold. This singular solution is depicted in Figure 1. It is not a true trajectory of the system as it involves orbits of systems that are incompatible, namely the fast and slow limits of (6) in the form of (3) and (4) respectively. The question is whether there is a real trajectory that lies near this formal orbit when $\epsilon > 0$ but sufficiently small.

The first step is to describe the results on which the elucidation of the flow on and near the slow manifolds is founded. These results are due to Fenichel [4, 5, 6] and were first proved by him in a more general context. He then later specialized the results to singularly perturbed problems with great effect [7].

4 Slow Manifolds

Fenichel proved three important theorems that, under certain conditions, give a complete description of the $\epsilon \neq 0$ flow in the neighborhood of the sets where the constraint

$$f(x,y,0) = 0$$

holds. First note that if $\epsilon = 0$ and $(\hat{x},\hat{y}) \in \{f(x,y,0)=0\}$ then $(\hat{x},\hat{y})$ is a critical point of (3). Let $M_0 \subset \{f(x,y,0)=0\}$ and suppose that it is an l-dimensional manifold. We say that M_0 is *normally hyperbolic* if the linearization of (3) at each point $(\hat{x},\hat{y}) \in \hat{M}_0$ has exactly l eigenvalues

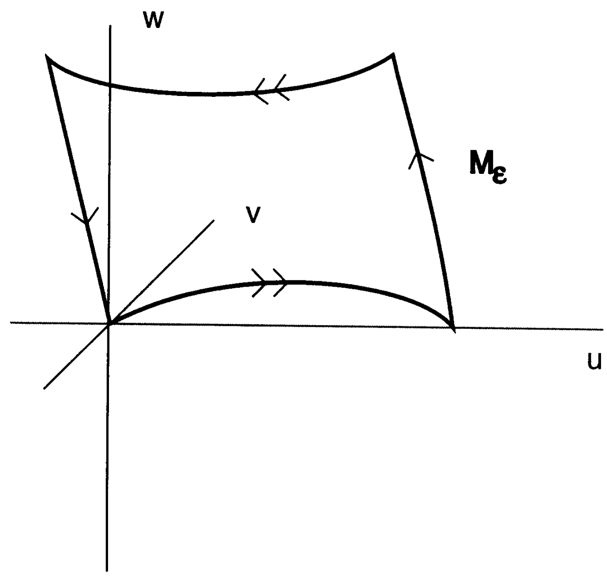

Figure 1: The singular orbit of the FitzHugh-Nagumo system starts with a fast jump (marked by a double arrow) followed by a slow transition on M_0, which is the limit of the depicted M_ϵ. There is then a fast transition back to the left slow manifold and a slow return to rest.

with zero real part, where $\hat{M}_0$ is an l-dimensional manifold in $\{f(x,y,0)=0\}$ containing M_0 in its interior. Note that, since M_0 is l-dimensional, there must be, at least, l eigenvalues with zero real part; the requirement of normal hyperbolicity is then that all directions normal to M_0 correspond to stable or unstable directions, in other words eigenvalues of non-zero real part. Fenichel's First Theorem [7] can then be stated .

Theorem 4.1 *(Fenichel's First Theorem) If M_0 is a normally hyperbolic manifold then there is, for ϵ sufficiently small, a manifold that is locally invariant relative to (1), or equivalently (2), which we name M_ϵ that is within $O(\epsilon)$ of, and diffeomorphic to, M_0.*

Since M_ϵ is locally invariant and l-dimensional it carries a lower dimensional flow than that of the full equations, and hence can be regarded as a reduction of the full system. An equation for the flow on M_ϵ can easily be calculated if it is represented in terms of x being given as a function of y. Let that function be expressed as $x = h_\epsilon(y)$, then the system governing the flow on M_ϵ can be written as

$$\dot{y} = g\left(h_\epsilon(y), y, \epsilon\right). \tag{7}$$

Note that $h_\epsilon \to h_0$ as $\epsilon \to 0$ and thus the limit of this system is given by the equation for the limiting flow

$$\dot{y} = g\left(h_0(y), y, 0\right). \tag{8}$$

The great advantage of this formulation is that (7) is a regular perturbation of (8) and thus the singular nature of the perturbation has been suppressed. If there is some structure existing in the flow of (8) that is robust then we can reasonably expect it to perturb to that of (7). There are many interesting examples in which the flow on the slow manifold is sufficient for the determination of the desired features. Such cases include the work of Ercolani et al. [3], Gardner and Jones [8], Ogawa [16] and Rubin, Jones and Maxey [17]. However this kind of reduction is direct and the phenomena found are typical of the lower dimensional slow system. I am concerned mostly in this paper, as stated earlier, with phenomena that exploit the reductions afforded by both fast and slow time scales.

5 Behavior Near Slow Manifolds

The analysis of the previous section i.e., Fenichel's First Theorem, renders a description of the flow on the slow manifold. It is important, however, to understand the flow in a neighborhood of the slow manifold.

The first step is to linearize in the fast limiting system (3) at an arbitrary critical point in the limiting slow manifold M_0. By the condition of normal hyperbolicity, this linearization has exactly l eigenvalues of zero real part and these correspond to eigenvectors that span the slow manifold. The normal directions are either stable or unstable. Thus, attached to each point $p \in M_0$ there is a stable manifold, $W^s(p)$, consisting of points which render trajectories tending exponentially to p as $t \to +\infty$, and an unstable manifold, $W^u(p)$, with points whose trajectories tend to similarly to p but as $t \to -\infty$. The dimensions of these manifolds, say m and k respectively, add up to $n = m + k$. These manifolds can be collected together to make manifolds for the full (limiting) slow manifold M_0:

$$\begin{aligned} W^s(M_0) &= \cup_{p \in M_o} W^s(p), \\ W^u(M_0) &= \cup_{p \in M_o} W^u(p). \end{aligned}$$

If now ϵ is turned on and the full flow is considered, the perturbation of these manifolds is addressed in Fenichel's Second Theorem.

Theorem 5.1 *(Fenichel's Second Theorem) If M_0 is a normally hyperbolic manifold then there are, for ϵ sufficiently small, manifolds that are locally invariant relative to (1) or, equivalently, (2) which we name $W^s(M_\epsilon)$ and $W^u(M_\epsilon)$ that are within $O(\epsilon)$ of, and diffeomorphic to, $W^s(M_0)$ and $W^s(M_0)$ respectively.*

The use of the terms stable and unstable manifold is justified by the connection with the corresponding objects when $\epsilon = 0$, but also can be more directly related to uniform exponential decay estimates on the manifolds, see [11]. This begins to give a clearer picture of the flow in the neighborhood of the slow manifold. There is, however, still more information to be gleaned from the $\epsilon = 0$ limit.

As referred to above, each individual point $p \in M_0$ has an attendant stable manifold $W^s(p)$ as well as unstable manifold $W^u(p)$. It seems absurd to suggest that these individual stable and unstable manifolds perturb as their base points do not remain as critical points when $\epsilon > 0$. However, this turns out to be not only a meaningful question, but one that leads to an extremely useful elucidation of the local structure of the flow. These individual stable and unstable manifolds attached to points in M_ϵ contain points that flow along in unison. The picture then is that, when $\epsilon \neq 0$, attached to each of the trajectories flowing in M_ϵ are manifolds of points that are carried with that trajectory. Moreover, one can determine that on $W^s(p \cdot t)$ ($p \cdot t$ here denotes the trajectory through p evolved after time t), the orbits will get closer to $p \cdot t$ as t increases, with an analogous statement holding for the unstable manifold.

Theorem 5.2 *(Fenichel's Third Theorem)If M_0 is a normally hyperbolic manifold then there are, for ϵ sufficiently small and for each point $p \in M_\epsilon$, manifolds that form an invariant family* *relative to (1), or equivalently (2), which we name $W^u(p)$ and $W^s(p)$, that are within $O(\epsilon)$ of, and diffeomorphic to, the corresponding manifolds when $\epsilon = 0$.*

The statement that they form an invariant family means that

$$\begin{aligned} W^u(p)\cdot t &\subset W^u(p\cdot t), \\ W^s(p)\cdot t &\subset W^s(p\cdot t), \end{aligned}$$

which says exactly that these stalks are carried along by the flow. These "invariant" manifolds are often called Fenichel fibers.

All of the above theorems can be put to use in clarifying the local structure near a slow manifold by successive changes of coordinates, each of which straighten out certain of the manifolds; where straightening out here is interpreted as mapping the manifold in question onto a coordinate plane. The result is a form of the equations which reflects the structure exposed by these theorems. We name the resulting normal form after Fenichel, but it was written down in this way first by Jones, Kopell and Kaper [12].

After these changes of coordinates, see [11] and [18] for details, the equations have the form

$$a' = \Lambda(a,b,y,\epsilon)a, \tag{9}$$
$$b' = \Gamma(a,b,y,\epsilon)b, \tag{10}$$
$$y' = \epsilon\left(h(y,\epsilon) + H(a,b,y,\epsilon)ab\right). \tag{11}$$

where $a \in \mathbb{R}^m, b \in \mathbb{R}^k$ and $y \in \mathbb{R}^l$. The variable a is in the unstable direction, which can be seen from the fact that $\Lambda(0,0,y,0)$ has all eigenvalues being of positive real part. Similarly b corresponds to the stable directions with $\Gamma(0,0,y,0)$ having eigenvalues of negative real part. The set $\{a = 0\}$ is then the stable manifold $W^s(M_\epsilon)$ and the set $\{b = 0\}$ is the unstable manifold $W^u(M_\epsilon)$. Thus the set $\{a = b = 0\}$ is the slow manifold M_ϵ itself. The form of the first two equations of (10) follows merely from transforming so that the manifolds lie in their respective coordinate planes. This part comes from straightening the stable and unstable manifolds.

The fibers, however, give extra information which is reflected in the third equation of (10). In either the stable or the unstable manifold, the fibers can be straightened out with the result that the flow on the slow variables in these sets is decoupled from the fast variables. This is exactly what happens in the third equation of (10) where the second term vanishes (the product written should be interpreted as a tensor product) when either $a = 0$ or $b = 0$.

This seems to be the most complete normal form governing the vector field near a slow manifold. It reflects the key points about the structure there and is central in the analysis of the passage of invariant manifolds near a slow manifold.

6 Passage Near a Slow Manifold

The example of nerve impulse propagation will be considered again to see how the need arises for techniques that facilitate the tracking of invariant manifolds through the phase space of a singularly perturbed system. In particular, the question as to whether a true homoclinic orbit lives near the formal singular orbit described in Section 3 will be the focus of the following.

There have been many approaches to answering this question, ranging form topological [2] to (hard) analytical [9]. In the work under discussion here, a geometric view is taken. The idea is first to note that the desired homoclinic orbit is constructed by finding an intersection of the unstable manifold of the rest state, W^u, and its stable manifold, W^s, at some value of the speed θ. One should then think of the unstable manifold as varying in θ and follow this manifold

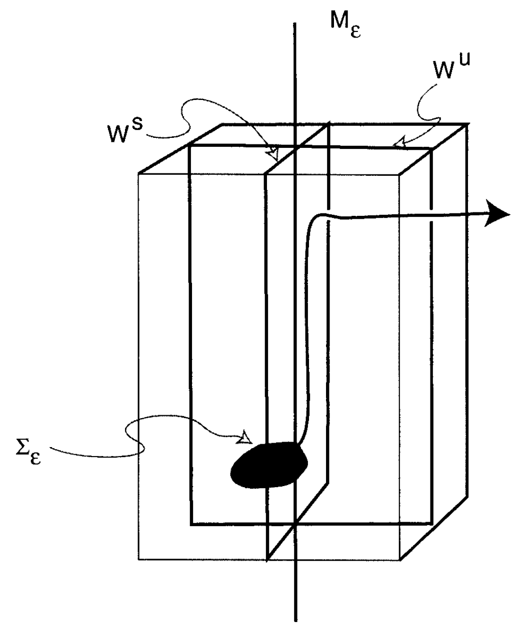

Figure 2: The slow manifold is depicted as a vertical line at the center of the box. Its stable and unstable manifolds are coordinate planes (in Fenichel coordinates). The Exchange Lemma addresses the passage of an invariant manifold, Σ_ϵ, near M_ϵ. A trajectory is shown coming from Σ_ϵ and it is the configuration of the forward iterate of Σ_ϵ as the trajectory exits the box near W^u that is determined by the Exchange Lemma.

around the full phase space (augmented by the variable θ.) This manifold that is being tracked also depends on ϵ and so is denoted Σ_ϵ. This should be tracked around the phase space, using the formal singular orbit as a guide, until it is forced to meet W^s. The key part of this strategy is to understand the passage near the slow manifold, on the right of Figure 1, that it comes close to but must depart from eventually. I shall denote this slow manifold M_ϵ. Indeed, the tracked manifold Σ_ϵ must be seen to come out of the neighborhood of this right slow manifold ready to be guided along the fast recovery jump back to the left slow manifold. A key point to note here is that the final meeting will need to be achieved via a transversal intersection as the Implicit Function Theorem is used to make the construction precise. Thus, at each stage along the guiding of the orbit, we seek transversal intersections.

The result of tracking Σ_ϵ along the fast activation jump is that it enters a neighborhood of M_ϵ transversally intersecting $W^s(M_\epsilon)$. The key result in determining its fate upon exit from a neighborhood of M_ϵ is the Exchange Lemma, see [13, 11]. In the following, the set J is the trajectory of the limiting slow flow on the manifold M_0 through the end point of the fast jumping orbit when $\epsilon = 0$. In order to be precise about entrance and exit of a neighborhood of M_ϵ, the set B is taken to be a fixed neighborhood of M_ϵ, which can be based on a box in Fenichel coordinates $\{a, b \leq \Delta\}$. The set-up can be seen in Figure 2; however, this figure is slightly deceptive in that the slow manifold M_ϵ can only be depicted as 1-dimensional and it is only in cases that it is higher-dimensional and Σ_ϵ is different in dimension to $W^u(M_\epsilon)$ that the subtlety arises.

Lemma 6.1 *(Exchange Lemma) If Σ_ϵ transversally intersects $W^s(M_\epsilon)$ in ∂B (boundary of B) then at an exit point of a trajectory having stayed exponentially long ($O(e^{-\frac{c}{\epsilon}})$) in B, Σ_ϵ is $C^1 - O(\epsilon)$ close to $W^u(J)$ upon exit.*

The set $W^u(J)$ is the union of the unstable manifolds (Fenichel fibers) at each point in J. In a practical sense this lemma allows one to track the "shooting" manifold through B and set it up to be tracked along the recovery fast jump. It turns out that it contains exactly the information needed to afford the correct transversality arguments that make this work. It also has an appealing conceptual interpretation, which gives the origin of its name. The manifold Σ_ϵ enters B with some information, where "information" is measured in terms of its tangent vectors as, ultimately, transversality of manifolds is the key to the construction. This information in the example considered here comes from the variation of the speed c. The passage near the slow manifold causes this information to be lost and exchanged. The information with which it emerges from this neighborhood comes directly from the slow flow as it is encoded in the set J. If one imagines successive passages near slow manifolds in this fashion, each passage would cause a divesting of information acquired from the previous slow manifold and its replacement by the information from the new slow manifold.

Clearly too few details can be given here to even sketch the proof of the Exchange Lemma, but some remarks are in order. The Fenichel Normal Form given in Section 5 is the key to the proof. Following tangent spaces is most readily achieved by the use of differential forms. These forms are coordinates for the tangent spaces and a central part of the proof involves calculations of equations for the evolution of these forms based on the equation of variations. The equation of variations is calculated in the coordinates of the Fenichel Normal Form and the proof lies in determining which of the forms are dominant.

The Exchange Lemma can be applied to great effect in constructing homoclinic orbits and this has been done in many examples, either using the form presented here under the guise of various generalizations, see [14, 12, 13, 11]. I want to point out here, however, how it can be used to illustrate the point made at the beginning of this paper that fast/slow systems can retain higher dimensional effects while only lower-dimensional analyses are invoked.

A well-known scenario for chaos is the so-called Silnikov mechanism. If a homoclinic orbit is found that connects an unstable manifold to a 2-dimensional stable manifold on which a spiral action occurs, then a horseshoe map can be found nearby, see, for instance, [19]. Such a situation can be easily engendered in the current context of equation (6). Indeed if the recovery (slow) variable w is 2-dimensional, then the rest state can be, under certain circumstances, a stable spiral. In this case the decay to rest would be oscillatory (and it has been suggested that this is physiologically realistic.) The homoclinic orbit constructed above would then be a Silnikov orbit. There are eigenvalue conditions needing to be satisfied but these are straightforward to verify due to the singular nature of the problem.

In this example, an orbit can thus be constructed that entails the presence of complex motion and this array of orbits give a large class of interesting waves. However, the analysis of only 2-dimensional systems, given the Exchange Lemma, is used to find this homoclinic orbit which reveals a genuinely higher dimensional phenomenon.

References

[1] V.I.Arnol'd, *Ordinary Differential Equations*, MIT Press, Boston, MA, 1973.

[2] G.Carpenter, A geometric approach to singular perturbation problems with applications to nerve impulse equations, *J. Diff. Eq.* 23 (1977) 335-367.

[3] N.Ercolani, D.McLaughlin & H.Roitner, Attractors and transients for a perturbed periodic KdV equation: a nonlinear spectral analysis, *J. Nonlinear Sci.* 3 (1993) 477-539.

[4] N.Fenichel, Persistence and smoothness of invariant manifolds for flows, *Indiana Univ. Math. Journal*, 21 (1971) 193-226.

[5] N.Fenichel, Asymptotic stability with rate conditions, *Indiana Univ. Math. Journal*, 23 (1974) 1109-1137.

[6] N.Fenichel, Asymptotic stability with rate conditions II, *Indiana Univ. Math. Journal*, 26 (1977) 81-93.

[7] N.Fenichel, Geometric singular perturbation theory for ordinary differential equations, *J. Diff. Eq.* 31 (1979) 53-98.

[8] R.Gardner & C.Jones, Traveling waves of a perturbed diffusion equation arising in a phase field model, *Indiana U. Math. J.* 38 (1989) 1197-1222.

[9] S.Hastings, On travelling wave solutions of the Hodgkin-Huxley equations, *Arch. Rat. Mech. Anal.* 60 (1976) 229-257.

[10] A.L.Hodgkin & A.F.Huxley, A quantitative description of membrane current and its application to conduction and excitation in nerve, *J. Physiol.* 117 (1952) 500-544.

[11] *Geometric Singular Perturbation Theory* in Dynamical Systems: Montecatini, Terme 1994, Lecture Notes in Mathematics 1609, Springer-Verlag, Heidelberg, 1995.

[12] C.Jones, T.Kaper & N.Kopell, Tracking invariant manifolds up to exponentially small errors, *SIAM J. Math. Anal.* 27 (1996) 558-577.

[13] C.Jones & N.Kopell, Tracking invariant manifolds with differential forms in singularly perturbed systems, *J. Diff. Eq.*, 108 (1994) 64-88.

[14] C.Jones, N.Kopell & R.Langer, Construction of the FitzHugh- Nagumo pulse using differential forms, in *Patterns and Dynamics in Reactive Media*, H.Swinney, G.Aris, and D.Aronson eds., IMA Volumes in Mathematics and its Applications, 37 (1991), Springer-Verlag, New York.

[15] T.J.Kaper & G.Kovačič, Multi-bump orbits homoclinic to resonance bands, *Trans. Amer. Math. Soc.*, 348 (1996) 3835-3887.

[16] T.Ogawa, Travelling wave solutions to perturbed Korteweg-de Vries equations, *Hiroshima Math. J.* 24 (1994) 401-422.

[17] J.Rubin, C. Jones & M.Maxey, Settling and asymptotic motion of aerosol particles in a cellular flow field, *J. Nonlinear Sci.* 5 (1995) 337-358.

[18] S.-K.Tin, On the dynamics of tangent spaces near a normally hyperbolic manifold, Ph.D. Thesis, Brown University, 1994.

[19] S.Wiggins, *Global Bifurcations and Chaos*, Springer-Verlag, New York, 1988.