

APPROPRIATE INITIAL CONDITIONS FOR ASYMPTOTIC DESCRIPTIONS OF THE LONG TERM EVOLUTION OF DYNAMICAL SYSTEMS

A. J. ROBERTS¹

(Received 21 April 1988; revised 28 November 1988)

Abstract

A centre manifold or invariant manifold description of the evolution of a dynamical system provides a simplified view of the long term evolution of the system. In this work, I describe a procedure to estimate the appropriate starting position on the manifold which best matches an initial condition off the manifold. I apply the procedure to three examples: a simple dynamical system, a five-equation model of quasi-geostrophic flow, and shear dispersion in a channel. The analysis is also relevant to determining how best to account, within the invariant manifold description, for a small forcing in the full system.

1. Introduction

The behaviour of many systems can be quantified in terms of the evolution of the descriptors of the state of the system. The evolution is then typically governed by a set of differential equations, which are ordinary differential equations if the system has a finite number of degrees of freedom, or often partial differential equations if the system has an infinite number of degrees of freedom. These statements refer to an enormously general class of problems and their mathematical formulation. It is this vast range of problems which is addressed in this paper; I shall refer to them as dynamical systems and use ideas developed for dynamical systems.

¹Applied Mathematics Department, University of Adelaide, S.A. 5000, Australia.

© Copyright Australian Mathematical Society 1989, Serial-fee code 0334-2700/89

My particular concern is that the full evolution equations, which I will take to be the exact equations, for many systems of interest are far too complicated for any solution to be calculated. The aim of much applied mathematics is to develop tractable approximations to the equations of the full system. However, many approximations are developed heuristically with little mathematical understanding of precisely what it is that is being done. Recently, centre manifold theory (see Carr [2]) has been extended (Roberts [12, 13]) to provide a formal procedure to calculate invariant manifolds of dynamical systems, and the evolution on them. The importance of this is that many practical approximations, such as incompressible and irrotational fluid flow, can immediately be seen to have the characteristics of a leading approximation to an invariant manifold; see the discussion in [12] for more details. Thus many useful approximations should be able to be put into a centre manifold or invariant manifold framework. This would immediately clarify their basic nature and lead to mechanistic improvements in the approximations, should they be needed; see the papers by Roberts [10, 11, 13] and Mercer & Roberts [7]. For example, the multi-mode analysis of Smith [16] for shear dispersion is being reworked and extended using the invariant manifold view. Moreover, there are proved results about the nature and approximation of centre manifolds, so at least for these cases the approximation of the full problem can be fully understood.

Previously, centre manifold theory has primarily been used to answer questions about the stability of fixed points of dynamical systems. However, starting with the work of Coulet & Spiegel [3] on amplitude equations in thermo-haline and triple convection, and with the work of Muncaster [8] on “coarse theories” of elastic bodies and of the kinetic theory of gases, invariant manifold theory has more recently been seen as a vital tool of mathematical analysis. It provides a simplified description of the evolution of a dynamical system by concentrating upon only a small subset of the solutions to the full problem. In the usual case of a dissipative system this can be justified rigorously as a long term asymptotic description of the system’s evolution. See the proofs by Foias *et al* [5] for the Kuramoto-Sivashinsky equation, and by Doering *et al* [4] for the complex Ginzburg-Landau equation.

The formal procedure inspired by the invariant manifold viewpoint provides a description of a set of states of the system, and also a description of the evolution among those states by a set of differential equations; see Carr [2] or Roberts [12] for some examples. However, it is not enough just to know the differential equations governing the evolution on the centre or invariant manifold. To form a complete problem we also need initial conditions; if the evolution on the invariant manifold is governed by partial differential equations then boundary conditions are also needed, see Smith [17].

The simple geometric picture of an invariant manifold gives a way of deriving the appropriate initial conditions for an approximation. Previously the initial conditions have been either guessed or derived by model-specific heuristic arguments. The aim of this paper is to illustrate the derivation of initial conditions for an approximation obtained through the invariant manifold viewpoint. Because most invariant manifolds are only known asymptotically (see Roberts [12]), the appropriate initial condition will also be known only asymptotically. The results in some example applications are impressive: in a very simple model of the atmosphere (due to Lorenz [6]) the calculations of an initial condition on the analogue of the quasi-geostrophic invariant manifold is analogous to the process of balancing which is essential in numerical meteorological models; in shear dispersion in a channel the calculation of the appropriate initial condition for a given discharge (which differs from the usual assumption, see the end of Section 3 in Smith [16]) leads naturally to the prediction of centroid displacements and variance deficits in the dispersal of a contaminant.

In Section 2, I discuss a simple two-variable dynamical system which is (nearly) solvable in closed form. The resulting analytic results about the long-term behaviour show that there is a favoured starting location on the invariant manifold to correspond to any given initial condition of the full system. The two-dimensional geometric picture, which is drawn exactly, then inspires the derivation of the projection of any initial condition onto the invariant manifold. In the case of a centre manifold, the existence of such a projection for initial conditions sufficiently close to the centre manifold is assured by Theorem 2(b) in Carr [2]. However, no one has previously derived an explicit formula for the projection. The analytic solution for this problem then confirms the veracity of the derived formal procedure.

By a simple (in theory) change of basis, the full equations for a dynamical system may be transformed into a convenient standard form. The procedure to project initial conditions onto an invariant manifold is described for the standard form in Section 3. The resultant procedure is then applied directly to a five-mode primitive equation model of the atmosphere introduced by Lorenz [6]. The results described in Section 4 give the planes of fast gravity-wave like oscillations around the quasi-geostrophic invariant manifold. These planes govern the projection of initial conditions, or the balancing, onto quasi-geostrophy. Another observation is that the projection onto the invariant manifold conserves invariants of the full system.

However, in practical applications, the change of basis needed to put the governing equations into standard form is typically undesirable as it reduces the physical understanding of the details at hand. The most general procedure for projecting initial conditions onto an invariant manifold is described in

Section 5. It deals with the case of a possibly infinite dimensional system, not in the standard form, which possesses a finite-dimensional invariant manifold of interest which represents the approximation. In Section 6 the derived procedure is applied to the problem of contaminant being dispersed by a shear flow in a channel. The consequent derivation of appropriate initial conditions for the Taylor model of shear dispersion then shows that effects such as the variance deficit in the dispersion of contaminant may be predicted within the Taylor model simply by using the correct initial condition.

Another useful consequence of the invariant manifold viewpoint is that a time dependent forcing of the full system can be incorporated systematically into the approximation. The procedure to do this is described in Section 7, and is based directly upon the projection of initial conditions which is the topic of this paper.

2. A simple dynamical system

The purpose of this section is to illustrate simply many of the concepts and details which will be derived in general in later sections. I do this by examining the simple but nontrivial dynamical system

$$\dot{x} = -xy, \quad (2.1a)$$

$$\dot{y} = -y + x^2 - 2y^2, \quad (2.1b)$$

(in which $\dot{(\)}$ denotes a derivative with respect to time) for which much analysis can be done exactly. The trajectories of this two dimensional system are given by the solid lines on Figure 1.

The centre manifold of this system is easily found by standard analysis, see Carr [2], to be exactly $y = x^2$, which is plotted as the bold solid line on Figure 1. It is convenient to use a new variable s to parameterise the location of the system on the centre manifold (later n will be used as a complementary coordinate to s). Thus the center manifold description is

$$x = s, \quad y = s^2, \quad \text{on which } \dot{s} = -s^3, \quad (2.2)$$

which is soon valid as the system tends exponentially quickly to this state.

We view (2.2) as an approximate description of the behaviour of solutions to (2.1). The question we address is: what starting value of s , say s_0 , should be used for (2.2) to best match the long term behaviour of the solution of (2.1) which is initially at the point (x_0, y_0) not on the centre manifold?

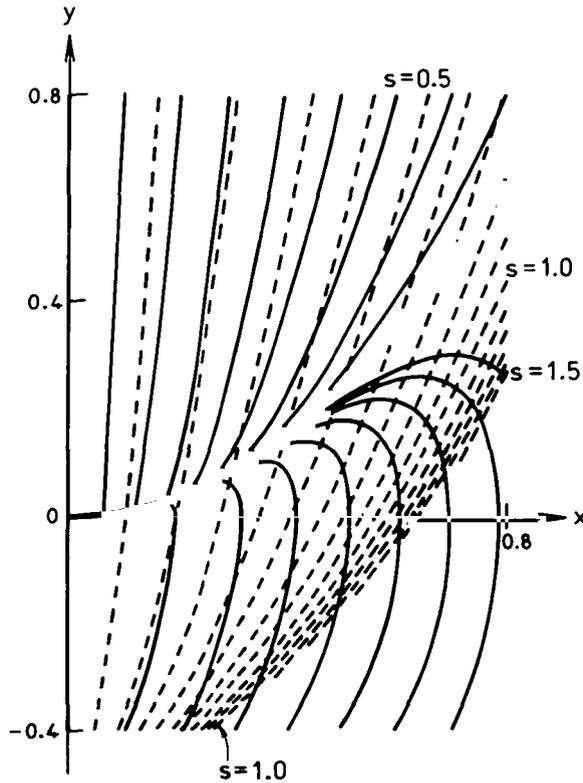


FIGURE 1. Trajectories of the dynamical system (2.1), with the bold solid line being the centre manifold. The dashed lines are those initial points whose long-term behaviour is identical (to an exponentially decaying difference).

2.1 Trajectories near the centre manifold

If the system is near but not on the centre manifold, then it follows a trajectory which asymptotes to the centre manifold. In this problem, trajectories can be found exactly (see Roberts [9]) to be the curves of constant ψ where

$$\psi = \left(\frac{y}{x^2} - 1\right) \exp\left(\frac{1 + 2y}{2x^2}\right). \tag{2.3}$$

The trajectory $\psi = 0$ is the centre manifold (2.2). Near the centre manifold, ψ is small, so that nearby trajectories are given as

$$y = x^2 + \psi x^2 \exp\left(-\frac{1}{2x^2} - 1\right) + O(\psi^2). \tag{2.4}$$

On such a trajectory the system evolves such that

$$x = [C + 2t - 2\psi \exp(-1 - C - t)]^{-1/2} + O(\psi^2), \tag{2.5}$$

where C is a constant.

Now if the system is initially at (x_0, y_0) at time $t = 0$, then the constant C in (2.5) is determined to be

$$C = 1/x_0^2 + 2\tau \quad \text{where } \tau = (y_0/x_0^2 - 1) \exp(y_0/x_0^2 - 1), \quad (2.6)$$

whence the evolution (2.5) on the particular trajectory may be rewritten as

$$x = \left[1/x_0^2 + 2(t + \tau) - \tau e^{-t}/x_0^2 \right]^{-1/2} + O(\psi^2). \quad (2.7)$$

From this form of the evolution along the trajectories it is apparent that there are two effects on the system if it is initially off the centre manifold; that is, if $\tau \neq 0$. Firstly, there is an exponential transient e^{-t} which is negligible as it has no long term effect. Secondly, there exists a time shift τ in the long term algebraic decay to the stable fixed point at the origin. This time shift is a significant long term effect of initially being off the centre manifold. I deduce that if the system is started on the centre manifold at the same x -coordinate as the initial point (x_0, y_0) , then the evolution on the centre manifold does not approach the true evolution exponentially quickly.

Instead of starting the centre manifold evolution at the same value of the x -coordinate as the initial condition, we should find a starting location on the centre manifold which best corresponds to the full system initially being at a point (x_0, y_0) off the centre manifold. Neglecting the exponential transients, (2.7) can be rewritten in the form

$$x = \left[1/s_0^2 + 2t \right]^{-1/2} \quad \text{where } s_0 = x_0 - x_0(y_0 - x_0^2) + O(\psi^2). \quad (2.8)$$

Thus the long term behaviour of a solution which is off the centre manifold is the same as a solution on the centre manifold, provided that the starting point on the centre manifold is $s = s_0$ as given in (2.8).

2.2 Numerical computation of initial conditions

The task of this subsection is to find numerically a start point on the centre manifold which best corresponds to a given initial condition. Given an initial point, I integrated the original ordinary differential equation (2.1) until, at some position, the solution came within some tolerance of the centre manifold. The appropriate start point on the centre manifold is that which reaches this same position after the same time.

Thus there exists a function $s_0(x_0, y_0)$ which gives the best start point on the centre manifold as a function of the initial point. Contours of this function are plotted in Figure 1 as the dashed lines. The intercept of these lines with the centre manifold shows the appropriate start point on the centre manifold to be used for any given initial condition. Thus this defines an appropriate projection of initial conditions onto the centre manifold. In meteorology, this is known as *balancing* (see Section 4). In the later analysis

of more general systems, I shall define the contours displayed in Figure 1 to be the projection manifolds.

In Figure 1 it is apparent that for initial conditions below and too far away from the centre manifold, there is no appropriate start point on the centre manifold. In this problem, this is just a reflection of the fact that it only takes a finite amount of time to come from infinity on the centre manifold to a finite position on the centre manifold. Thus the concept of projecting the initial conditions is only necessarily valid near the centre manifold.

2.3 The local direction of the projection manifolds

In this subsection, I derive an equation which could (but will not) be solved to give the projection manifolds displayed in Figure 1. Suppose s is a pseudo-time along the trajectories (related to the time by $\dot{s} = -s^3$ as in (2.2)), and n parameterises the family of trajectories (hence n is a function of ψ). I treat s and n as a new set of coordinates for the phase space. The coordinate lines of the (s, n) system are the trajectories and contours plotted in Figure 1. Let $x = X(s, n)$ and $y = Y(s, n)$ describe the coordinate transform between the new system and Cartesian coordinates, and let $\mathbf{x}^T = (X, Y)$. Represent the governing differential equation as $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ where $\mathbf{x}^T = (x, y)$, and let the Jacobian of the right-hand side be $\mathcal{F} = \mathbf{f}_x = [\partial f_i / \partial x_j]$.

Consider Figure 2 in which two neighbouring trajectories, n and $n + \Delta n$, are displayed for a small time interval, $\Delta s = \dot{s} \Delta t$, in which the system evolves

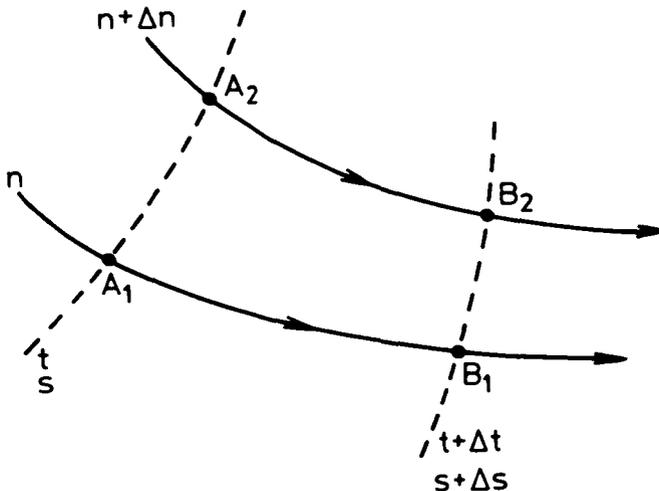


FIGURE 2. A small part of the phase plane with two neighbouring trajectories and two neighbouring projection manifolds.

from the points A_i to the corresponding point B_i . Now, from the coordinate transformation

$$\overrightarrow{A_1 B_1} = \Delta s \mathbf{X}_s + (\Delta s^2 / 2) \mathbf{X}_{ss} + O(\Delta^3)$$

and

$$\overrightarrow{A_2 B_2} = \Delta s \mathbf{X}_s + (\Delta s^2 / 2) \mathbf{X}_{ss} + \Delta s \Delta n \mathbf{X}_{ns} + O(\Delta^3),$$

where the subscripts n and s denote partial differentiation with respect to the subscript, and where all the above quantities are evaluated at the point A_1 .

But also $\overrightarrow{A_i B_i}$ is part of a trajectory, and it follows that

$$\overrightarrow{A_1 B_1} = \Delta t \mathbf{f} + (\Delta t^2 \mathcal{F} / 2) \mathbf{f} + O(\Delta^3)$$

and

$$\overrightarrow{A_2 B_2} = \Delta t (\mathbf{f} + \Delta n \mathcal{F} \mathbf{X}_n) + (\Delta t^2 \mathcal{F} / 2) \mathbf{f} + O(\Delta^3),$$

where all of the above quantities are also evaluated at the point A_1 . Upon considering $\overrightarrow{A_2 B_2} - \overrightarrow{A_1 B_1}$ we find

$$\Delta s \Delta n \mathbf{X}_{ns} = \Delta t \Delta n \mathcal{F} \mathbf{X}_n + O(\Delta^3),$$

from which we readily deduce that

$$\dot{s} \mathbf{X}_{ns} = \mathcal{F} \mathbf{X}_n. \tag{2.9}$$

Geometrically \mathbf{X}_n is the local direction of the n -coordinate curves, and thus the \mathbf{X}_n appearing in this equation gives the local direction of those initial points which have the same long term behaviour; that is they will have the same best starting point on the centre manifold. Hence \mathbf{X}_n is the local tangent to the local projection manifold. Remembering that \dot{s} is simply a function of s , given by (2.2), this equation is solved by integration with respect to s , that is by integrating along trajectories. A difficulty is that (2.9) is very hard to solve, even approximately, the reason being that each value of n is tied to a particular trajectory, and so solving (2.9) also gives the shape of nearby trajectories as well as the local direction of an s_0 -contour. Fortunately we do not need \mathbf{X}_n precisely, just its direction, and this removes the difficulty.

2.4 The projection of initial conditions

Given the earlier warning about the local nature of projecting the initial conditions onto the centre manifold, it is appropriate to find the direction of the projection only on the centre manifold, especially as in many situations this is all that will be known about the solutions to the full problem. That is, all I am seeking is the direction of the projection manifolds, displayed in Figure 1, on the centre manifold. From the results expressed in (2.8), we may readily deduce that the desired slope of these contours is just $2s + 1/s$.

My aim here is to derive this sort of result without knowing all the details of the trajectories adjacent to the centre manifold.

Near the origin we expect the projection to be vertical, as this is the direction of the eigenvector corresponding to the exponential decay; more generally this is the direction of the stable manifold. Thus I seek the direction of the projection on the centre manifold in the form $(p(s), 1)^T$. Suppose that

$$\mathbf{X}_n = q(s) \begin{bmatrix} p(s) \\ 1 \end{bmatrix}, \quad (2.10)$$

where we do not need to find $q(s)$. Substituting this into (2.9) and eliminating $q(s)$, I find that $p(s)$ must satisfy the Riccati equation

$$\dot{s}p' = [1 \ -p]\mathcal{F} \begin{bmatrix} p \\ 1 \end{bmatrix}. \quad (2.11)$$

For the particular problem (2.1), and its centre manifold (2.2), (2.11) reduces to

$$-s^3p' = -s + (1 + 3s^2)p - 2sp^2, \quad (2.12)$$

to be solved such that $p(0) = 0$, as the projection manifold should be vertical at the origin (from a linear analysis). It is easy to check that $p(s) = s/(1+2s^2)$ is an exact solution to (2.12) and hence the local direction of initial conditions which should be projected to the same starting point on the centre manifold is $(s/(1+2s^2), 1)^T$. This is the same direction as found earlier.

Thus, for any given initial condition off the centre manifold the appropriate starting point on the centre manifold is approximately the position such that the initial condition is in the above direction from the starting position.

3. Standard projection of initial conditions

I now derive a way to project initial conditions, which are near an invariant manifold, onto the invariant manifold when the governing differential equations have been transformed into a convenient form. Suppose the differential equations are in the standard form

$$\dot{\mathbf{x}} = A\mathbf{x} + \mathbf{f}(\mathbf{x}, \mathbf{y}), \quad \mathbf{x} \in \mathbb{R}^m \quad (3.1a)$$

$$\dot{\mathbf{y}} = B\mathbf{y} + \mathbf{g}(\mathbf{x}, \mathbf{y}), \quad \mathbf{y} \in \mathbb{R}^n \quad (3.1b)$$

in which $A\mathbf{x}$ and $B\mathbf{y}$ contain all the linear terms of the equations and the functions \mathbf{f} and \mathbf{g} are strictly nonlinear, of $O(|\mathbf{x}|^2 + |\mathbf{y}|^2)$, and are sufficiently well behaved near the origin. In principle, this standard form can always be obtained by a change of basis.

If the real part of the eigenvalues of A are all zero and the real part of the eigenvalues of B are all negative, then rigorous results exist about the

existence and the calculation of the m -dimensional centre manifold $y = h(x)$ for the system, see Carr [2]. However, my interest lies in the more general situation where I suppose there exists an m -dimensional invariant manifold described parametrically by

$$x = s, \quad y = h(s) \quad \text{on which } \dot{s} = As + f(s, h(s)), \quad (3.2)$$

where $s \in \mathbb{R}^m$ parameterises the locations on the invariant manifold (see Roberts [12] for some examples). This forms a valid long term description of the evolution of the full system if the real parts of the eigenvalues of B are all large and negative. In this general situation I now seek to find the appropriate projection onto the invariant manifold of an initial condition which is off the manifold.

For any starting point on the invariant manifold, there exists a set of points which, if the system starts at one of the points, has the same long term behaviour. The set of points form an n -dimensional projection manifold (for example, the dashed lines in Figure 1). I aim to find a set of vectors spanning the tangent space of such a projection manifold at the invariant manifold. This set of vectors will serve to define an approximately correct projection of initial conditions onto the invariant manifold.

3.1 An equation governing the direction

I start by investigating how a given infinitesimal vector at an angle to the invariant manifold evolves in time with the flow of the system (3.1). By similar arguments to that employed in Section 2 it is easy to see that such a vector n (analogous to X_n in Section 2) evolves according to

$$\dot{n} = \mathcal{F} n, \quad \text{where } \mathcal{F} = \begin{bmatrix} A + f_x & f_y \\ g_x & B + g_y \end{bmatrix} \quad (3.3)$$

is the Jacobian of the right-hand side of (3.1).

Let N be a matrix composed of n such vectors which are linearly independent. Then they span the desired tangent space and satisfy $\dot{N} = \mathcal{F}N$. However, just as in Section 2, this equation is generally too difficult to solve. Since I am only interested in the space spanned by the columns of N , let

$$N = \begin{bmatrix} P \\ I_n \end{bmatrix} Q, \quad (3.4)$$

where P is an $m \times n$ matrix and Q is an invertible $n \times n$ matrix. Thus the columns of $\begin{bmatrix} P \\ I_n \end{bmatrix}$ form a new basis for the tangent space and hence for the projection. Substituting (3.4) into (3.3), premultiplying by $[I_m - P]$, and considering P as a function of s (the parameter of the invariant manifold) I obtain

$$\dot{s} \bullet \nabla_s P = [I_m - P] \mathcal{F} \begin{bmatrix} P \\ I_n \end{bmatrix}, \quad (3.5)$$

where \dot{s} is given in (3.2) and the matrix

$$\dot{s} \bullet \nabla_s P = \left[\sum_{k=1}^m \frac{\partial P_{ij}}{\partial s_k} \dot{s}_k \right]. \tag{3.6}$$

Equation (3.5) is to be solved such that P at $s = 0$ is zero as, for equations in the form of (3.1), the projection manifold at the origin is normal to the invariant manifold.

3.2 Asymptotic solution

Exact solutions of (3.5) will be very rare, especially as the invariant manifold is typically only known approximately. It is usually most convenient to rewrite (3.5) in a form suitable for iteration (or for a formal power series solution), namely

$$PB - AP + (As) \bullet \nabla_s P = -f \bullet \nabla_s P + [I_m \ -P] \begin{bmatrix} \mathbf{f}_x & \mathbf{f}_y \\ \mathbf{g}_x & \mathbf{g}_y \end{bmatrix} \begin{bmatrix} P \\ I_n \end{bmatrix}. \tag{3.7}$$

The right-hand side of this equation contains all the nonlinear terms in s , and the left-hand side contains all the linear terms. The usual iteration scheme is to start with $P = 0$, and then solve for the left-hand side P given the right-hand side evaluated for the previous estimate of P . The result is usually a multinomial in s which, near the origin, is asymptotic to an exact solution.

If the matrix A is zero, and thus the invariant manifold is a simple centre manifold, then the iteration process is almost trivial, it just involves inverting B . If A is in Jordan form, then each iteration step may be done easily, so long as the rows of P are found in an appropriate order.

If both A and B are diagonal, then each iteration step is straightforward, and explicitly displays general features of which to be wary. When A and B are diagonal, the equation for a particular element of the matrix P is decoupled from the other elements. Consider the equation for the (i, j) th element, and suppose that there exists a term in the right-hand side of (3.7) of the form $c_{ij} \prod_k s_k^{l_k}$ where l_k are a set of integral exponents. Letting λ_i and μ_j be the diagonal elements of A and B respectively (i.e. their eigenvalues), the equation is then of the form

$$P_{ij}\mu_j - \lambda_i P_{ij} + \sum_k \frac{\partial P_{ij}}{\partial s_k} \lambda_k s_k = c_{ij} \prod_k s_k^{l_k},$$

which has the particular solution

$$P_{ij} = \frac{c_{ij} \prod_k s_k^{l_k}}{\mu_j - \lambda_i + \sum_k l_k \lambda_k}. \tag{3.8}$$

The embarrassing possibility in (3.8) is that of dividing by zero; it is impossible if the invariant manifold is the centre manifold. Such a division by

zero generally indicates that the object to be described is not unique, due to nonlinear resonance, for example. This also often occurs in the actual asymptotic description of an invariant manifold (see Roberts [12]) and there it serves to warn of the non-uniqueness of the invariant manifold. The practical effect is that there is no purpose in computing any further refinement to the approximate description, as such a refinement is immaterial. I expect a similar conclusion here, but this needs further research.

3.3 The projection

Suppose $\mathbf{u}_0 = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$ is a given initial point in \mathbb{R}^{m+n} . I want to find the starting position on the invariant manifold, $\mathbf{u}_0^{\#} = \begin{bmatrix} x_0^{\#} \\ \mathbf{h}(x_0^{\#}) \end{bmatrix}$, such that $\mathbf{u}_0 - \mathbf{u}_0^{\#}$ is a linear combination of the columns of $\begin{bmatrix} P(x_0^{\#}) \\ I_n \end{bmatrix}$. It is elementary to see that this will be the case if and only if

$$\mathbf{x}_0^{\#} = \mathbf{x}_0 - P(\mathbf{x}_0^{\#})(y_0 - \mathbf{h}(\mathbf{x}_0^{\#})). \quad (3.9)$$

This is a set of m nonlinear equations to be solved for the m unknowns in $\mathbf{x}_0^{\#}$. Near the origin it may be solved efficiently by iteration (starting with $\mathbf{x}_0^{\#} = \mathbf{x}_0$ for example).

3.4 Discussion

For many dynamical systems, for example that discussed in Section 2, the projection of the initial conditions onto the invariant manifold will be accurate in the long term. However, if the flow on the invariant manifold has a positive Lyapunov exponent, see Schuster [14], then the projection would have to be exponentially accurate to guarantee the same long term behaviour on and off the manifold. In general, this cannot be achieved. In such a situation, there would be two time scales: on the time scale of attraction to the manifold, the approximate solution on the manifold would be approached exponentially-quickly by the true solution; on the time scale of the loss of information on the manifold, the true solution and the approximation would diverge, although the true solution would still be approaching the manifold. However, this latter divergence is not a serious limitation as, due to the divergence characterised by the positive Lyapunov exponent, it would occur for almost every initial condition on the manifold.

Also, I have taken the full system (3.1) to be an exact description of the full problem; rarely will this be precisely true. Thus, to be of relevance, the invariant manifold must be structurally stable with respect to typical perturbations of the full system. There is very little known about the structural stability of invariant manifolds, although Beyn [1] has investigated the time discretisation of a set of ordinary differential equations and its effect on the centre-unstable manifold. I observe that if the eigenvalues of B are not near

zero then the formal construction of invariant manifolds, as discussed in [12], is largely insensitive to the details of the ordinary differential equations (3.1). The exception to this is that the construction is sensitive at orders of approximation corresponding to the non-uniqueness of the invariant manifold. However, these orders should not be calculated, see Sections 3 and 5 in [12], as they depend upon the initial conditions. Thus the significant orders of approximation of an invariant manifold are typically structurally stable.

Although derived in the context of an exponentially attracting manifold, the above derivation of the projection is not restricted solely to this situation. The derivation is based upon the flow in the neighbourhood of the invariant manifold, namely the evolution of \mathbf{n} as given by (3.3). But because I am only interested in the space spanned by the possible vectors \mathbf{n} , it is immaterial whether the invariant manifold is attracting, is repelling, or is a centre for the neighbouring flow. What the derivation guarantees is that, if a solution on the manifold and a solution just off the manifold start on the same projection manifold, then they will continue to be on the same projection manifold (although different to the one they started on) as time progresses. Thus, as seen in Section 2, if the invariant manifold is attracting then the projection along a projection manifold gives the appropriate initial condition. But furthermore, if the invariant manifold acts as a “centre” for the neighbouring flow then this will still be true. In particular, if the invariant manifold contains periodic cycles then the two solutions (one on and one off the invariant manifold) will maintain the same phase relationship as they evolve around the cycle. This property is used in the next section in an example of the projection of initial conditions.

4. Example: a model of balancing onto quasi-geostrophy

Lorenz [6] proposed the following set of five coupled ordinary differential equations,

$$\dot{u} = -vw + buz, \quad (4.1a)$$

$$\dot{v} = uw - buz, \quad (4.1b)$$

$$\dot{w} = -uv, \quad (4.1c)$$

$$\dot{x} = -z, \quad (4.1d)$$

$$\dot{z} = x + buv, \quad (4.1e)$$

to act as a model describing coupled Rossby waves and gravity waves in the atmosphere. If the parameter $b = 0$ then the last two equations decouple from the first three. In this case x and z just undergo rapid sinusoidal oscillation

analogous to fast gravity waves in the atmosphere, while u, v and w undergo slow nonlinear elliptic function oscillation analogous to Rossby waves. For finite values of b the picture is qualitatively the same near the origin.

The geostrophic approximation in meteorology is simply to neglect x and z and thus neglect the unwanted fast gravity waves; this is a tangent-space approximation to an invariant manifold of (4.1). The particular invariant manifold which is approximated represents a state of quasi-geostrophy; in meteorology this manifold is often known as the slow manifold. One of the problems in weather forecasting is that field data is often noisy, and numerical models using the field data become swamped with large amplitude gravity waves which make long range predictions impossible. Thus the data is “balanced”, before being used in a numerical model, by filtering out the fast gravity waves. This balancing is equivalent to projecting the given initial conditions onto the invariant manifold corresponding to quasi-geostrophic flow. Thus it is precisely the process addressed in this paper.

The model equations (4.1) are in the standard form (3.1), and so I seek a 3-dimensional invariant manifold of the model which is parameterised by $\mathbf{u} = (u, v, w)^T$. Thus the 3×3 matrix A is identically zero, the matrix B is

$$B = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad \text{and} \quad \begin{bmatrix} \mathbf{f} \\ \mathbf{g} \end{bmatrix} = \begin{bmatrix} -vw + b v z \\ uw - b u z \\ -uv \\ 0 \\ b u v \end{bmatrix} \quad (4.2)$$

are the nonlinear terms. Firstly, I find the quasi-geostrophic invariant manifold. Secondly, I use the results of Section 3 to find the projection onto this invariant manifold. One thing to note is that this invariant manifold is not approached exponentially by nearby trajectories. Because the eigenvalues of B are pure imaginary, representing neglected oscillations, the invariant manifold acts more like a centre for the nearby trajectories; indeed this invariant manifold is an example of what is termed a sub-centre manifold (see Sijbrand [15]). Geometrically the plane which is projected onto a particular point on the invariant manifold is the plane in which rapid gravity wave oscillations take place if the system starts off the invariant manifold.

4.1 The quasi-geostrophic invariant manifold

Since u, v and w are to parameterise locations on the invariant manifold, I pose that

$$x = X(u, v, w) \quad \text{and} \quad z = Z(u, v, w) \quad (4.3)$$

describes the invariant manifold. Substituting into (4.1) and using the chain rule, X and Z must satisfy

$$X = -buv - Z_u(vw - bvZ) + Z_v(uw - buZ) - Z_w uv, \quad (4.4a)$$

$$Z = X_u(vw - bvZ) - X_v(-uw + buZ) + X_w uv. \quad (4.4b)$$

This may be solved by iteration or substitution of a formal power series (see Roberts [9, 12, 13] for some examples) to find that the first nontrivial curving shape of the invariant manifold is given by

$$x = X(u, v, w) = -buv + O(|\mathbf{u}|^4), \quad (4.5a)$$

$$z = Z(u, v, w) = b(u^2 - v^2)w + O(|\mathbf{u}|^5). \quad (4.5b)$$

On this invariant manifold the system evolves according to

$$\dot{u} = -vw \left[1 - b^2(u^2 - v^2) \right] + O(|\mathbf{u}|^6), \quad (4.6a)$$

$$\dot{v} = uw \left[1 - b^2(u^2 - v^2) \right] + O(|\mathbf{u}|^6), \quad (4.6b)$$

$$\dot{w} = -uv. \quad (4.6c)$$

The expressions in (4.5) and (4.6) are the first few terms of an asymptotic series which is likely to be divergent. Indeed, Lorenz [6] numerically constructed an invariant manifold for the system (4.1), and found that it contained weak singularities due to nonlinear resonances. An asymptotic expansion for such a manifold would necessarily be divergent. Furthermore, practical algorithms for balancing data also appear to be divergent if too many iterations are carried out. A divergent series is often of practical use.

4.2 Balancing

I now use this as an example of the application of the equations derived in Section 3. I seek a 3×2 matrix $P(\mathbf{u})$ such that the two columns of $\begin{bmatrix} P \\ J_2 \end{bmatrix}$ span the plane of gravity wave oscillations about the quasi-geostrophic invariant manifold (4.5). The task of balancing some given initial data is then the same as finding the position on the invariant manifold whose projection plane passes through the given data. That this is not exactly true is due to the curving nature of the projection manifolds, but for practical purposes it will probably suffice.

For this problem (3.7) becomes

$$PB = - \begin{bmatrix} -vw + bvZ \\ uw - buZ \\ -uv \end{bmatrix} \bullet \nabla_{\mathbf{u}} P + [I_3 - P] \begin{bmatrix} 0 & -w + bZ & -v & 0 & bv \\ w - bZ & 0 & u & 0 & -bu \\ -v & -u & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 \\ bv & bu & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} P \\ I_2 \end{bmatrix}, \tag{4.7}$$

where Z is given approximately by (4.5b). Solving this by iteration, starting with $P^{(0)} = 0$ on the right-hand side of (4.7), I find that

$$P^{(1)} = \begin{bmatrix} -bv & 0 \\ bu & 0 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad P^{(2)} = \begin{bmatrix} -bv & 0 \\ bu & 0 \\ 0 & -b(u^2 - v^2) \end{bmatrix} \tag{4.8}$$

where some third-order terms in $P^{(2)}$ have been neglected, as they may be changed in the next iteration. Thus the plane of gravity-wave oscillations at some location \mathbf{u} on the invariant manifold is approximately spanned by the two columns of $\begin{bmatrix} P^{(2)} \\ I_2 \end{bmatrix}$, namely by

$$\mathbf{e}_1 = \begin{bmatrix} -bv \\ bu \\ 0 \\ 1 \\ 0 \end{bmatrix} + O(|\mathbf{u}|^3) \quad \text{and} \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 0 \\ -b(u^2 - v^2) \\ 0 \\ 1 \end{bmatrix} + O(|\mathbf{u}|^3). \tag{4.9}$$

It is interesting to note that, at this order of accuracy, the curving nature of the invariant manifold has no effect on these results; it is only at fourth order that it has an effect. To obtain the above results the quasi-geostrophic invariant manifold could have been approximated simply by $x = z = 0$.

4.3 Conserved quantities

A further interesting feature of the differential equations (4.1) is that they possess the two invariants

$$H = u^2 + v^2, \quad \text{and} \quad I = v^2 + w^2 + x^2 + z^2. \tag{4.10}$$

These invariants arise from conservation of analogies to energy and potential enstrophy. It is clearly desirable that these quantities be conserved in the projection, although in practice it is enough for them to be approximately conserved in the approximate projection.

Here, such conservation of the invariants during the projection may be shown by establishing that the normals, ∇H and ∇I , to the invariant hyper-surfaces are orthogonal to the vectors defining the tangent plane of the projection manifold, \mathbf{e}_1 and \mathbf{e}_2 given by (4.9), where they are all evaluated on

the quasi-geostrophic invariant manifold. It is elementary algebra to find that this is indeed the case.

5. General form of the projection of initial conditions

In principle, the standard form of the projection derived in Section 3 would always suffice, as a given dynamical system could always be transformed into the standard form (3.1). In practice, such a transformation is often inconvenient to do and furthermore it often gives a set of equations for which a meaning, in terms of the original problem, cannot be assigned to any of the terms. This is clearly undesirable. Instead it is preferable to work with the original equations as much as possible.

In this section suppose that the system of equations is in the general form

$$\dot{\mathbf{u}} = \mathcal{L}\mathbf{u} + \mathbf{N}(\mathbf{u}), \quad (5.1)$$

such that $\mathbf{u} = 0$ is a fixed point, where \mathcal{L} is a linear operator, and where $\mathbf{N}(\mathbf{u})$ is strictly nonlinear, of $O(|\mathbf{u}|^2)$, and sufficiently well-behaved near the origin. I shall allow the system to be infinite dimension (although I shall still refer to \mathbf{u} as a vector), but for ease of exposition I shall assume that the invariant manifold is finite dimensional. Thus, let $\langle \cdot, \cdot \rangle$ be the inner product defined on the (possibly complex) phase space \mathcal{U} , and suppose that (5.1) possess the m -dimensional invariant manifold \mathcal{M} , parameterised by $\mathbf{s} \in \mathbb{R}^m$, which is described by

$$\mathbf{u} = \mathbf{v}(\mathbf{s}) \quad \text{such that} \quad \dot{\mathbf{s}} = \mathcal{S}\mathbf{s} + \mathbf{g}(\mathbf{s}), \quad (5.2)$$

where $\mathcal{S}\mathbf{s}$ is linear in \mathbf{s} and $\mathbf{g}(\mathbf{s})$ is strictly nonlinear. Asymptotic approximations to such an invariant manifold description may be found by a formal procedure, see Roberts [12], which is a generalisation of one proposed for centre manifolds by Coulet & Spiegel [3] (and explained more simply in Roberts [9]).

Basic entities which are needed in the following analysis are the following. At the origin the tangent space, denoted by $E^{\mathcal{M}}$, to the invariant manifold \mathcal{M} is spanned by a set of m linearly independent vectors in \mathcal{U} ; let these vectors be denoted by $\mathbf{e}_i^{\mathcal{M}}$. The tangent space $E^{\mathcal{M}}$ is spanned by a set of m eigenmodes of the linear operator \mathcal{L} . Let the space spanned by the corresponding eigenmodes of the adjoint linear operator, \mathcal{L}^\dagger , be denoted by $Z^{\mathcal{M}}$; and let m linearly independent vectors in this space be denoted by $\mathbf{z}_i^{\mathcal{M}}$. For example, in a finite dimensional system the vectors $\mathbf{e}_i^{\mathcal{M}}$ and $\mathbf{z}_i^{\mathcal{M}}$ could be corresponding pairs of right- and left-eigenvectors.

5.1 An equation for the projection

Consider an infinitesimal vector \mathbf{n} lying at an angle to the invariant manifold \mathcal{M} and evolving with the flow of the system. By similar arguments to that used in Section 2 \mathbf{n} evolves in time according to (3.3), but here the Jacobian is

$$\mathcal{F} = \mathcal{L} + \mathcal{N} = \mathcal{L} + \mathbf{N}_{\mathbf{u}}, \tag{5.3}$$

in which both \mathcal{L} and \mathcal{N} are linear operators but \mathcal{N} varies over the invariant manifold, being zero at the origin, while \mathcal{L} is constant. Thus, considering \mathbf{n} to be a function of position on the invariant manifold, the governing equation is

$$\dot{\mathbf{s}} \bullet \nabla_{\mathbf{s}} \mathbf{n} = \mathcal{F} \mathbf{n}, \tag{5.4}$$

where $\dot{\mathbf{s}}$ is the function of \mathbf{s} given in (5.2). Letting \mathbf{n}_{α} be a family of vectors (α need not be discrete) which span the tangent space to the projection manifolds, they all evolve according to (5.4). Analogous with the argument given in Section 3.3, what is really needed to calculate the projection onto the invariant manifold is a set of m linearly independent vectors, \mathbf{r}_i say, which are orthogonal to the \mathbf{n}_{α} . The condition that \mathbf{s}_0 be the appropriate starting point on the invariant manifold corresponding to a given initial condition \mathbf{u}_0 is that $\mathbf{u}_0 - \mathbf{v}(\mathbf{s}_0)$ be in the space spanned by the \mathbf{n}_{α} . This is then simply the set of m equations

$$\langle \mathbf{r}_i(\mathbf{s}_0), \mathbf{u}_0 - \mathbf{v}(\mathbf{s}_0) \rangle = 0 \quad \text{for all } i. \tag{5.5}$$

This is important to realise, because it is much easier to find the m vectors \mathbf{r}_i than the possibly infinite number of vectors \mathbf{n}_{α} .

At the origin the projection manifold must be tangent to the space spanned by all the eigenmodes of \mathcal{L} which are not included in the invariant manifold tangent space $E^{\mathcal{M}}$. Thus by the orthogonality of “left-” and “right-eigenvectors” the tangent space to the projection manifold must be orthogonal to $Z^{\mathcal{M}}$. Hence at the origin the vectors \mathbf{r}_i should be m linearly independent vectors in $Z^{\mathcal{M}}$, and hence must tend to $\mathbf{z}_i^{\mathcal{M}}$ at the origin. Away from the origin it is straightforward to show that the vectors \mathbf{r}_i should satisfy

$$\dot{\mathbf{s}} \bullet \nabla_{\mathbf{s}} \mathbf{r}_i = -\mathcal{F}^{\dagger} \mathbf{r}_i, \tag{5.6}$$

in order to stay orthogonal to the projection manifolds, where \mathcal{F}^{\dagger} is the adjoint of the Jacobian \mathcal{F} .

As before, equation (5.6) is too hard to solve as it contains information about the shape of neighbouring trajectories as well as the desired information about the projection. Since I just want the space spanned by the vectors \mathbf{r}_i , I seek to find the vectors $\mathbf{z}_i^{\mathcal{M}} - \mathbf{p}_i$ where

$$\mathbf{r}_i = q_{ij}(\mathbf{z}_j^{\mathcal{M}} - \mathbf{p}_j) \quad \text{such that } \langle \mathbf{p}_i, \mathbf{e}_j^{\mathcal{M}} \rangle = 0 \text{ and } \mathbf{p}_i(0) = 0, \tag{5.7}$$

where q_{ij} are the elements of an $m \times m$ invertible matrix $Q(\mathbf{s})$, and the summation convention has, and will, be used. The vectors \mathbf{p}_i are analogous to the rows of the $m \times n$ matrix P used in Section 3 (and Q is analogous to that used in Section 3). Substituting (5.7) into (5.6), applying the operator $\langle \cdot, \mathbf{e}_k^\mathcal{H} \rangle$ to be able to eliminate the q_{ij} , I find that the vectors \mathbf{p}_i must satisfy

$$\langle \mathcal{F}^\dagger(\mathbf{z}_i^\mathcal{H} - \mathbf{p}_i), \mathbf{e}_k^\mathcal{H} \rangle^* A_{kj}^\mathcal{H} (\mathbf{z}_j^\mathcal{H} - \mathbf{p}_j) + (\mathcal{G}\mathbf{s} + \mathbf{g}(\mathbf{s})) \bullet \nabla_s \mathbf{p}_i = \mathcal{F}^\dagger(\mathbf{z}_i^\mathcal{H} - \mathbf{p}_i), \quad (5.8)$$

where the superscript $*$ denotes complex conjugation, and where the $m \times m$ constant matrix $A^\mathcal{H}$ is the inverse of $\langle \mathbf{z}_i^\mathcal{H}, \mathbf{e}_j^\mathcal{H} \rangle^*$ (often the vectors are chosen so that this matrix is the identity matrix). This is analogous to equations (3.5).

5.2 Asymptotic solution

Exact solutions of (5.8) will be very rare, especially as the invariant manifold is typically known only approximately. To find an asymptotic solution of (5.8) near the origin it is convenient to rewrite it as

$$\begin{aligned} \mathcal{L}^\dagger \mathbf{p}_i - \langle \mathcal{L}^\dagger \mathbf{z}_i^\mathcal{H}, \mathbf{e}_k^\mathcal{H} \rangle A_{kj}^\mathcal{H} \mathbf{p}_j + (\mathcal{G}\mathbf{s}) \bullet \nabla_s \mathbf{p}_i \\ = -\mathbf{g} \bullet \nabla_s \mathbf{p}_i + \mathcal{N}^\dagger(\mathbf{z}_i^\mathcal{H} - \mathbf{p}_i) - \langle (\mathbf{z}_i^\mathcal{H} - \mathbf{p}_i), \mathcal{N} \mathbf{e}_k^\mathcal{H} \rangle^* A_{kj}^\mathcal{H} (\mathbf{z}_j^\mathcal{H} - \mathbf{p}_j) \end{aligned} \quad (5.9)$$

where the expression

$$\mathcal{L}^\dagger \mathbf{z}_i^\mathcal{H} - \langle \mathcal{L}^\dagger \mathbf{z}_i^\mathcal{H}, \mathbf{e}_k^\mathcal{H} \rangle^* A_{kj}^\mathcal{H} \mathbf{z}_j^\mathcal{H}$$

has been omitted as it is zero. This can be seen by first realising that it must be in the space $Z^\mathcal{H}$, and second, by showing that it is orthogonal to the space $E^\mathcal{H}$ and hence must lie in the space spanned by those eigenmodes of \mathcal{L}^\dagger which are not in $Z^\mathcal{H}$. These are mutually contradictory unless the expression is zero.

Equation (5.9) is directly analogous to (3.7) and can be solved approximately in the same manner.

6. Example: shear dispersion in a channel

Taylor [18] considered dispersion of a solute in a channel or pipe and argued that after sufficient time had elapsed, the cross-sectionally averaged solute concentration satisfied a one-dimensional advection-diffusion equation. His analysis has been the basis of many mathematical predictions about dispersion in the environment. The approximate solution he derived can be rigorously supported by the application of centre manifold theory. This is discussed in Mercer & Roberts [7] who also give a formal extension of the theory which allows effects of variations in the geometry and the flow parameters to be quantitatively included within the asymptotic approximation.

With the theory developed in the previous section, I now derive what the appropriate starting condition for the asymptotic approximation is, in terms of the initial condition for the full problem.

To make things definite, but non-dimensional, consider a two-dimensional channel extending to infinity in the x -direction and with straight sides enclosing $-1 \leq y \leq 1$. Let the concentration of some solute in the channel be $c(x, y, t)$. The solute is advected by the plane-Poiseuille flow

$$u(y) = (3/2)U(1 - y^2), \quad (6.1)$$

and is diffused across the channel (but, for simplicity, not down the channel). Thus the governing differential equation for the concentration c is

$$\frac{\partial c}{\partial t} = -u(y)\frac{\partial c}{\partial x} + \frac{\partial^2 c}{\partial y^2} \quad \text{such that} \quad \frac{\partial c}{\partial y} = 0 \quad \text{on} \quad y = \pm 1. \quad (6.2)$$

The centre manifold analysis will give the long term behaviour of the concentration c in terms of the evolution of the cross-sectional average concentration, denoted by $C(x, t)$. The theory developed earlier will give the appropriate starting condition, $C = C_0(x)$ at time $t = 0$, in terms of the initial distribution of the solute, $c = c_0(x, y)$ at time $t = 0$. In particular, I show that the standard practice of assuming that $C_0(x)$ is the cross-stream average of $c_0(x, y)$ is not quite correct.

6.1 The centre manifold

To apply centre manifold theory most clearly, it is best to take the Fourier transform of (6.2), although a more direct procedure is discussed in [10]. Denoting transformed quantities by $\hat{\ }^$'s the governing differential equation (6.2) becomes

$$\frac{\partial \hat{c}}{\partial t} = \frac{\partial^2 \hat{c}}{\partial y^2} - ik u(y) \hat{c} \quad \text{such that} \quad \frac{\partial \hat{c}}{\partial y} = 0 \quad \text{on} \quad y = \pm 1, \quad (6.3a)$$

where k is the wavenumber of the Fourier transform. This transformation decouples all the components at the different wavenumbers, and so they can be considered separately. Thus adjoining the equation

$$\partial k / \partial t = 0, \quad (6.3b)$$

I can write the two equations in (6.3) in the form of (5.1) where

$$\mathbf{u} = \begin{bmatrix} k \\ \hat{c} \end{bmatrix}, \quad \mathcal{L} = \begin{bmatrix} 0 & 0 \\ 0 & \frac{\partial^2}{\partial y^2} \end{bmatrix}, \quad \text{and} \quad \mathbf{N}(\mathbf{u}) = \begin{bmatrix} 0 \\ -iu(y)k\hat{c} \end{bmatrix}. \quad (6.4)$$

The linear operator \mathcal{L} has two zero eigenvalues and all the rest are negative (the largest of these being $-\pi^2/4$). Thus (6.3) has a 2-dimensional centre

manifold, at least near the origin of \mathbf{u} -space, which is approached exponentially quickly and can be parameterised by the two variables k and \hat{C} . I find the centre manifold to be given approximately by

$$\hat{c} \sim \hat{C} - ikUv_1(y)\hat{C} \text{ in which } v_1(y) = (7 - 30y^2 + 15y^4)/120, \quad (6.5)$$

where k and \hat{C} evolve according to

$$\partial \hat{C} / \partial t \sim -ikU\hat{C} - k^2(2/105)U^2\hat{C} \quad \text{and} \quad \partial k / \partial t = 0. \quad (6.6)$$

These results are essentially power series in k and are thus valid for small wavenumber; that is, for concentrations which are slowly varying in space. See Roberts [10] for a detailed discussion of the slowly-varying approximation.

In terms of the original physical variables, the centre manifold description is

$$c \sim C - Uv_1(y)\frac{\partial C}{\partial x} \quad \text{such that} \quad \frac{\partial C}{\partial t} \sim -U\frac{\partial C}{\partial x} + \frac{2U^2}{105}\frac{\partial^2 C}{\partial x^2}, \quad (6.7)$$

which is precisely Taylor's approximation for this problem. Higher orders in the above approximation may be easily calculated through this approach; this has been done by Mercer & Roberts [7], who showed that the resultant series converges in some sense.

6.2 The projection of initial conditions

Having described the centre manifold for this problem, I now can solve (5.9) to find the vectors defining the projection of initial conditions onto the centre manifold.

Firstly, I need to define an inner product on the phase space, namely

$$\langle \mathbf{u}_1, \mathbf{u}_2 \rangle = k_1^* k_2 + \frac{1}{2} \int_{-1}^1 \hat{c}_1^*(y) \hat{c}_2(y) dy. \quad (6.8)$$

Secondly, the adjoint of the Jacobian evaluated on the centre manifold is $\mathcal{J}^\dagger = \mathcal{L}^\dagger + \mathcal{N}^\dagger$ where

$$\mathcal{L}^\dagger = \mathcal{L} \quad \text{and} \quad \mathcal{N}^\dagger \mathbf{u}_1 = \left[\begin{array}{c} \frac{i}{2} \int_{-1}^1 u(y) \hat{c}(y)^* \hat{c}_1(y) dy \\ ik^* u(y) \hat{c}_1(y) \end{array} \right], \quad (6.9)$$

where \hat{c} and k are on the centre manifold as given in (6.5). Thirdly, I need a basis for the tangent space, $E^\mathcal{K}$, to the centre manifold and a basis for the corresponding space, $Z^\mathcal{K}$, of the adjoint. Since the linear operator \mathcal{L} is self-adjoint these may as well be the same, and so I take

$$\mathbf{e}_1^\mathcal{K} = \mathbf{z}_1^\mathcal{K} = \begin{bmatrix} 1 \\ 0(y) \end{bmatrix} \quad \text{and} \quad \mathbf{e}_2^\mathcal{K} = \mathbf{z}_2^\mathcal{K} = \begin{bmatrix} 0 \\ 1(y) \end{bmatrix}. \quad (6.10)$$

To solve (5.9) to find the projection, I first observe that, because the invariant manifold is actually a centre manifold and because of the above choices, the form of (5.9) simplifies to

$$\begin{aligned} \mathcal{L}^\dagger \mathbf{p}_i = & - \left[-ikU\hat{C} - k^2 \frac{2U^2}{105} \hat{C} + \dots \right] \frac{\partial p_i}{\partial \hat{C}} \\ & + \mathcal{N}^\dagger(\mathbf{z}_i^\# - \mathbf{p}_i) - \langle (\mathbf{z}_i^\# - \mathbf{p}_i), \mathcal{N} \mathbf{e}_j^\# \rangle^* (\mathbf{z}_j^\# - \mathbf{p}_j). \end{aligned} \tag{6.11}$$

Substituting $\mathbf{p}_i = 0$ into the right-hand side of (6.11), we can then solve for the first nontrivial approximations to \mathbf{p}_i , which are, neglecting some not-yet-significant quadratic terms in k ,

$$\mathbf{p}_1^{(1)} = 0 \quad \text{and} \quad \mathbf{p}_2^{(1)} = \begin{bmatrix} 0 \\ -ikUv_1(y) \end{bmatrix}. \tag{6.12}$$

Further iterations may be carried out to provide more terms in the expansion for the projection vectors. However, for my illustrative purposes, the approximation (6.12) will suffice.

The aim of the projection is to find the starting point on the centre manifold, parametrically given by k_0 and \hat{C}_0 say, which best corresponds to the initial condition, given by k and $\hat{c}_0(y)$ say. The earlier analysis shows that the displacement from one point to the other should be orthogonal, with respect to the inner product $\langle \cdot, \cdot \rangle$, to the two derived vectors $\mathbf{z}_i^\# - \mathbf{p}_i$. Firstly, $\mathbf{z}_1^\# - \mathbf{p}_1^{(1)} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ with the immediate consequence that $k_0 = k$, that is, the wavenumber of the initial condition is unchanged in the projection onto the centre manifold. This is eminently reasonable from a physical point of view. Secondly, $\mathbf{z}_2^\# - \mathbf{p}_2^{(1)} = \begin{bmatrix} 0 \\ 1+ikUv_1(y) \end{bmatrix}$, which gives the appropriate initial condition for the Fourier component of the cross-sectionally averaged concentration to be

$$\hat{C}_0 \sim \overline{\hat{c}_0} - ikU\overline{v_1\hat{c}_0}, \tag{6.13}$$

where the overbar indicates the average over y , namely $\frac{1}{2} \int_{-1}^1 \dots dy$.

Returning to the original physical variables, by taking the inverse Fourier transform, (6.13) becomes

$$C_0 \sim \overline{c_0} - U \frac{\partial}{\partial x} (\overline{v_1 c_0}). \tag{6.14}$$

Because the leading term on the right-hand side of (6.14) is just $\overline{c_0}$, this justifies the usual approximation of taking the given initial condition and averaging over the channel to give the starting condition for the asymptotic approximation. However, the analysis here shows how this simple approach should be modified if the given initial condition involves significant spatial variations; for example, a point release. The modification will improve the long term agreement between the results of the centre manifold model and

the solution of the full problem. Furthermore, by doing a bit more work to solve (6.11) more accurately, further improvements may be made to the expression for the starting condition.

6.3 A point release at the bank

To illustrate the sort of predictions that can be made from the above analysis I now consider the idealisation of a point release of solute adjacent to the side of the channel $y = 1$. The centre manifold is only valid for slowly-varying concentrations, which is certainly not true for this initial condition. Thus the predictions for small time by the centre manifold model will be unphysical, containing negative concentrations for example. Nonetheless the long term predictions are improved remarkably.

Suppose the initial condition for the full problem is

$$c_0(x, y) = 2\delta(x)\delta(y - 1^-), \quad (6.15)$$

then the starting condition for the centre manifold model of the full problem is, from (6.14) and (6.5), simply

$$C_0(x) = \delta(x) + U\delta'(x)/15. \quad (6.16)$$

The new term from this analysis is the $\delta'(x)$ term; this has long-term effects on the solution of the one-dimensional advection-diffusion equation in (6.7). For brevity denote the effective diffusivity and the coefficient of the new term in the initial condition by

$$D = 2U^2/105 \quad \text{and} \quad \alpha = U/15 \quad (6.17)$$

respectively. The solution to (6.7) for the initial condition (6.16) is

$$C(x, t) = \frac{1}{2\sqrt{\pi Dt}} \left[1 - \alpha \left(\frac{x - Ut}{2Dt} \right) \right] \exp \left[-\frac{(x - Ut)^2}{4Dt} \right]. \quad (6.18)$$

Thus the effect of the $\alpha\delta'(x)$ term is to bias the cross-sectional average concentration C towards smaller x . This is the expected result, as solute released at the side of the channel needs time to diffuse into the centre of the channel where the bulk of the downstream advection takes place. Thus the advection of the solute should be delayed somewhat, as (6.18) now shows.

The effects can be simply observed in the moments of the resulting distribution. The mean location of the solute as predicted by (6.18) is

$$\bar{x} = \int_{-\infty}^{\infty} xC(x, t) dx = Ut - \alpha = U(t - 1/15). \quad (6.19)$$

This increases linearly by the advection downstream, but there is a delay of $\Delta t = 1/15$ due to the release at the side. The spread of the solute, as measured by the variance, predicted by (6.18) is

$$\sigma^2 = \int_{-\infty}^{\infty} x^2 C dx - (\bar{x})^2 = 2Dt - \alpha^2 = 2D(t - 7/60). \quad (6.20)$$

As is conventional, the variance increases linearly due to the shear dispersion, but again there is a delay in the effect of the shear dispersion due to side release. These effects are well known in the study of dispersion in rivers and channels. Previously there has been no known systematic way to include these delays within the framework of the Taylor model.

One thing to observe in the variance (and in the mean, but to a physically less disturbing extent) is that for small time the predicted variance is actually negative! This is attributable to the negative concentrations which exist and are significant for small time in the solution (6.18). It does not refute the validity of the model; instead it acts as a useful warning that the concentration gradients are too steep for the model to be valid. Because the original problem is linear in c , this does not destroy the validity of the model in the long term. Previously there has been no possibility of such a warning occurring in the model.

7. Time dependent forcing on an invariant manifold

So far, all applications of centre manifold and invariant manifold descriptions of the long term evolution of dynamical systems have been to autonomous sets of differential equations. Having to ignore the possibility of some time dependent forcing in the differential equations is a serious limitation. However, with the projective tools developed in the earlier sections, I can now show how to incorporate a weak forcing of the full system into a forcing which is within the invariant manifold description.

7.1 The appropriate forcing

Suppose that a dynamical system of the form of (5.1) has some forcing applied, say $\mathbf{F}(t)$, so that it is modified to

$$\partial \mathbf{u} / \partial t = \mathcal{L} \mathbf{u} + \mathbf{N}(\mathbf{u}) + \mathbf{F}(t). \quad (7.1)$$

Further, suppose that the system without any applied forcing possesses an invariant manifold \mathcal{M} described by (5.2). To see the effects of the applied forcing, I suppose that the forcing can be approximated sufficiently well by a series of isolated impulses, namely

$$\mathbf{F}(t) \approx \sum_l \mathbf{F}_l \delta(t - t_l), \quad (7.2)$$

where t_l is the time of the l th impulse \mathbf{F}_l . In each force-free time interval between impulses, the system will approach the invariant manifold exponentially quickly. Thus suppose that at a time $t = t_l^-$, which is just before the l th impulse, the system is on the invariant manifold at a location given by

$\mathbf{u}_l^- = \mathbf{v}(\mathbf{s}_l^-)$. Assuming the differential operator \mathcal{L} and \mathbf{N} are smooth enough, (7.1) may be integrated across the time of the l th impulse to give that

$$\mathbf{u}_l^+ = \mathbf{u}_l^- + \mathbf{F}_l. \quad (7.3)$$

The problem is that \mathbf{u}_l^+ is generally off the invariant manifold \mathcal{M} and so has to be projected back onto \mathcal{M} in order for the invariant manifold description to be applied until the time of the next impulse. This consideration will give the forcing $\mathbf{F}^{\mathcal{M}}(t)$, which it is appropriate to use within the invariant manifold description to best match the forcing of the full system.

The appropriate location on the invariant manifold corresponding to \mathbf{u}^+ (for simplicity I will, for a while, omit the subscript l referring to the l th impulse), parameterised by \mathbf{s}^+ , is found by requiring the displacement $\mathbf{u}^+ - \mathbf{v}(\mathbf{s}^+)$ to be orthogonal to the vectors $\mathbf{z}_i^{\mathcal{M}} - \mathbf{p}_i$ derived for the projection of initial conditions. If the impulsive forcing is not too big then the appropriate location on \mathcal{M} after the impulse will not be too far from its location before the impulse. Thus, $\mathbf{s}^+ = \mathbf{s}^- + \Delta\mathbf{s}$ where $\Delta\mathbf{s}$ is small and can be found by solving $\langle \mathbf{z}_i^{\mathcal{M}} - \mathbf{p}_i, \mathbf{u}^- + \mathbf{F} - \mathbf{v}(\mathbf{s}^- + \Delta\mathbf{s}) \rangle = 0$. Linearising this equation for $\Delta\mathbf{s}$, I find that

$$\Delta s_i = B_{ij} \langle \mathbf{z}_j^{\mathcal{M}} - \mathbf{p}_j, \mathbf{F} \rangle \text{ where } B_{ij} = \text{inverse of } \left\langle \mathbf{z}_i^{\mathcal{M}} - \mathbf{p}_i, \frac{\partial \mathbf{v}}{\partial \mathbf{s}_j} \right\rangle. \quad (7.4)$$

Thus the evolution on the invariant manifold \mathcal{M} may be given by the forced equation

$$\dot{\mathbf{s}} = \mathcal{G}\mathbf{s} + \mathbf{g}(\mathbf{s}) + \mathbf{F}^{\mathcal{M}}, \quad (7.5)$$

where the forcing (upon putting the subscript l back in) is in the form of the isolated impulses

$$\mathbf{F}^{\mathcal{M}} \approx \sum_l B_{ijl} \langle \mathbf{z}_j^{\mathcal{M}} - \mathbf{p}_{jl}, \mathbf{F}_l \rangle \delta(t - t_l). \quad (7.6)$$

This particular forcing just models the continuous forcing,

$$\mathbf{F}^{\mathcal{M}}(\mathbf{s}, t) = B_{ij}(\mathbf{s}) \langle \mathbf{z}_j^{\mathcal{M}} - \mathbf{p}_j(\mathbf{s}), \mathbf{F} \rangle, \quad (7.7)$$

which is the appropriate one to be used in the invariant manifold description of the evolution. The dominant contribution to the forcing (7.7) is just the component of the original forcing which is aligned in the invariant manifold, the components of \mathbf{F} being determined by decomposing it into a component in the local projection manifold plus a component in \mathcal{M} .

The above argument will be valid provided that the forcing of the full system is never so big that it pushes the system far from the invariant manifold \mathcal{M} . In application this would be hard to know. Perhaps the best measure of its validity would be to compare the neglected components of the forcing

with the rate of approach to the invariant manifold, as these are the competing influences. The approximation is valid if the neglected components of the forcing are small in the comparison.

7.2 Example: a simple dynamical system

Consider the dynamical system (2.1) upon which some forcing $\mathbf{F} = \begin{bmatrix} F_x \\ F_y \end{bmatrix}$ is acting. The invariant manifold is one-dimensional and so I find that $\mathbf{z}^s - \mathbf{p} = [1 \ -p]$ where $p = s/(1 + 2s^2)$ is the solution to (2.12). It is straightforward to calculate that $B = 1/([1 \ -p], [1 \ 2s]) = (1 + 2s^2)$. Upon substituting into (7.7) I find that the appropriate evolution equation for s , modified from (2.2), is

$$\dot{s} = -s^3 + (1 + 2s^2)F_x - sF_y. \quad (7.8)$$

The obvious intuitive result follows: if $F_x \neq 0$ and $F_y = 0$, then the forcing just pushes the stable fixed point to one side of the origin. However, what is not obvious is the effect of non-zero F_y : if $F_x = 0$ and $F_y = -\varepsilon^2$, then the evolution equation for s becomes

$$\dot{s} = -s^3 + \varepsilon^2 s, \quad (7.9)$$

which, instead of the one fixed point at the origin, predicts three fixed points at $s = 0, \pm\varepsilon$. Furthermore, the fixed point at the origin is now predicted to be unstable. This is an extremely surprising result; that a small force which is essentially at right-angles to the invariant manifold (and that component along the invariant manifold is always towards the stable origin) actually destabilises the fixed point at the origin!

Nonetheless these predictions are correct. For small ε the relevant fixed points of the full system (2.1) with the above forcing are: two stable fixed points at $(\pm\varepsilon, 0)$; and an unstable fixed point at approximately $(0, -\varepsilon^2 - 2\varepsilon^4)$. I surmise that this example demonstrates that the evolution on a centre manifold, as it is very slow, is extremely sensitive to external forces. It is remarkable that this projection of external forces onto the invariant manifold is indeed appropriate and accurate.

8. Summary

Invariant manifold theory (and the simpler centre manifold theory) provides a long-term, lower-dimensional, asymptotic description of the evolution of a dynamical system. The main principle is that the dynamical system will tend exponentially quickly to a state on the invariant manifold and then evolve according to the simpler set of differential equations which apply on the invariant manifold.

However, if an inappropriate starting point on the invariant manifold is chosen to correspond to a given initial condition for the full system, then there are long-term, nonexponentially decaying discrepancies between the asymptotic predictions and the correct solution. This can be seen in equation (2.7), and also in the shear dispersion results (6.18), (6.19) and (6.20). In this last case, it can be seen that the annoying time-lag discrepancies between the actual dispersion and the Taylor model of shear dispersion is not a weakness of the Taylor model, but is instead a failure to provide the proper initial conditions for the model. It is only with the simple geometric view provided by invariant manifold theory that this can be easily perceived and correct initial conditions found.

On another hand, these ideas can be applied to invariant manifolds which are not approached exponentially, but instead which act as centres for oscillatory behaviour. This was done in Section 4 to a primitive model of atmosphere flow, and was found to provide a systematic basis for the process of balancing which is essential in numerical meteorology. These sub-centre manifolds are also found in the derivation of continuum approximations to gas dynamics, see Muncaster [8].

Lastly, this approach has shown how to rationally transform an inhomogeneous time-dependent forcing of the full system into a corresponding forcing of an invariant manifold approximation.

References

- [1] W.-J. Beyn, "Centre manifolds of dynamical systems under discretisation", *Numer. Funct. Anal. & Opt.* **9** (1987) 381–414.
- [2] J. Carr, "Applications of centre manifold theory", *Appl. Math. Sci.* **35** (1981).
- [3] P. H. Coullet & E. A. Spiegel, "Amplitude equations for systems with competing instabilities", *SIAM J. Appl. Math.* **43** (1983) 776–821.
- [4] C. R. Doering, J. D. Gibbon, D. D. Holm & B. Nicolaenko, "Low-dimensional behaviour in the complex Ginzburg-Landau equation", preprint, (1987).
- [5] C. Foias, B. Nicolaenko, G. R. Sell & R. Temam, "Inertial manifolds for the Kuramoto-Sivashinsky equation and an estimate of their lowest dimension", *J. Math. Pure & Appl.* (1987) (to appear).
- [6] E. N. Lorenz, "On the existence of the slow manifold", *J. Atmos. Sci.* **43** (1986) 1547–1557.
- [7] G. N. Mercer & A. J. Roberts, "A centre manifold description of contaminant dispersion in channels and pipes with varying flow properties", submitted (1988).
- [8] R. G. Muncaster, "Invariant manifolds on mechanics I: The general construction of coarse theories from fine theories", *Arch. Rat. Mech.* **84** (1983) 353–373.
- [9] A. J. Roberts, "Simple examples of the derivation of amplitude equations for systems of equations possessing bifurcations", *J. Austral. Math. Soc.* **B 27** (1985) 48–65.
- [10] A. J. Roberts, "The application of centre manifold theory to the evolution of systems which vary slowly in space", *J. Austral. Math. Soc.* **B 29** (1988) 480–500.

- [11] A. J. Roberts, "A formal centre manifold description of the evolution of slowly-varying waves", preprint (1987).
- [12] A. J. Roberts, "The utility of an invariant manifold description of the evolution of a dynamical system", *SIAM J. of Math. Anal.* to appear (1989).
- [13] A. J. Roberts, "The construction of invariant manifolds for the Kuramoto-Sivashinsky equation", preprint (1988).
- [14] H. G. Schuster, *Deterministic chaos: an introduction* (Physik-Verlag, D-6940 Weinheim, F.R.G., 1984).
- [15] J. Sijbrand, "Properties of centre manifolds", *Trans. Amer. Math. Soc.* **289** (1985) 431–469.
- [16] R. Smith, "Diffusion in shear flows made easy: the Taylor limit", *J. Fluid Mech.* **175** (1987) 201–214.
- [17] R. Smith, "Boundary conditions for flow reactors", to appear *IMA J. Appl. Math.* (1987).
- [18] G. I. Taylor, "Dispersion of soluble matter in solvent flowing slowly through a tube", *Proc. Roy. Soc. Lond. A* **219** (1953) 186–203.