Invariant manifolds

Marc R. Roussel

October 4, 2005

1 Introduction

In our earlier discussion of stability analysis, we emphasized the significance of the eigenvalues of the Jacobian matrix. What of the eigenvectors?

Let us briefly recall why the Jacobian is important. In the vicinity of an equilibrium point $\mathbf{x}^*$ of the autonomous ordinary differential equation (ODE)

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}), \tag{1}$$

provided that none of the eigenvalues are zero, the differential equation can be approximated by

$$\dot{\delta\mathbf{x}} = \mathbf{J}^*\,\delta\mathbf{x}, \tag{2}$$

where $\delta\mathbf{x} = \mathbf{x} - \mathbf{x}^*$, and $\mathbf{J}^*$ is the Jacobian evaluated at $\mathbf{x}^*$. $\mathbf{J}^*$ is a constant, so equation 2 is a linear ordinary differential equation. The solutions of this equation can be written in the form

$$\delta\mathbf{x}(t) = \sum c_i \mathbf{e}_i e^{\lambda_i t},$$

where $\mathbf{e}_i$ are the right eigenvectors and λ_i the corresponding eigenvalues of $\mathbf{J}^*$, and c_i are some coefficients chosen to satisfy the initial conditions. The eigenvalues therefore tell us something about how fast we approach or recede from the equilibrium point, and the eigenvectors tell us something about the directions along which this motion will occur.

To see this more clearly, consider a very simple example: Suppose that the equilibrium point is stable, i.e. that all the eigenvalues are negative, and that λ_1 is much smaller than any of the other eigenvalues in absolute value. Then the terms involving $e^{\lambda_i t}$ for $i > 1$ decay much faster than $e^{\lambda_1 t}$ so that, eventually, $\delta\mathbf{x}(t) \to c_1 e^{\lambda_1 t}\mathbf{e}_1$. Geometrically, this corresponds to approach to the equilibrium point along $\mathbf{e}_1$, the "slow" eigenvector.

That's nice, but what happens farther away from equilibrium, where equation 2 is no longer valid? Can we extend the linear picture away from the equilibrium point? In a way, the answer is yes, but the extension requires several new ideas, and of course some new and highly interesting complications arise.

2 Flow dynamics away from the equilibrium point

Suppose that $\mathbf{x}_1$ is an arbitrary point in the phase space of equation 1, and $\mathbf{x}_2$ is a second point close to $\mathbf{x}_1$. Let $\delta\mathbf{x}$ be the vector connecting $\mathbf{x}_1$ to $\mathbf{x}_2$, i.e.

$$\delta\mathbf{x} = \mathbf{x}_2 - \mathbf{x}_1.$$

The evolution in time of $\delta\mathbf{x}$ is therefore given by

$$\dot{\delta\mathbf{x}} = \dot{\mathbf{x}}_2 - \dot{\mathbf{x}}_1 = \mathbf{f}(\mathbf{x}_2) - \mathbf{f}(\mathbf{x}_1) = \mathbf{f}(\mathbf{x}_1 + \delta\mathbf{x}) - \mathbf{f}(\mathbf{x}_1).$$

We can Taylor expand $\mathbf{f}(\mathbf{x}_1 + \delta\mathbf{x})$ about $\delta\mathbf{x} = 0$:

$$\mathbf{f}(\mathbf{x}_1 + \delta\mathbf{x}) = \mathbf{f}(\mathbf{x}_1) + \mathbf{J}_1\delta\mathbf{x} + \dots$$

Here, $\mathbf{J}_1$ is the Jacobian evaluated at $\mathbf{x}_1$. To lowest order, the difference vector $\delta\mathbf{x}$ therefore obeys the differential equation

$$\dot{\delta\mathbf{x}} \approx \mathbf{J}_1\delta\mathbf{x}. \tag{3}$$

This looks an awful lot like the linearized equation for evolution near the equilibrium point. There's a major difference however. If we have two systems governed by the differential equation 1, they both move through phase space as time unfolds. Thus, the Jacobian matrix in equation 3 changes as time goes on, except in the special case of linear differential equations. The eigenvectors and eigenvalues therefore also change, so we have to be careful in interpreting the eigenvector decomposition of $\mathbf{J}_1$. Nevertheless, we can say that *locally*, we will observe convergence or divergence of trajectories either toward or away from each other at rates controlled by the eigenvalues of $\mathbf{J}_1$ and along the directions indicated by its eigenvectors.

Now suppose that we start very close to the equilibrium point on one of the eigenvectors. (For the sake of argument, consider the case of a real eigenvector corresponding to a unique eigenvalue.) If we follow this eigenvector backward in time (i.e. evolve the point under the action of the dynamical system $\dot{\mathbf{x}} = -\mathbf{f}(\mathbf{x})$), it will trace out a trajectory which, run forward in time, runs into the equilibrium point along the specified eigenvector. As long as the eigenvalue spectrum[1] doesn't change in any important way (eigenvalues don't change order, change signs, become complex, etc.), the trajectories near this "eigentrajectory" will behave much as they did near the corresponding eigenvector in the vicinity of the equilibrium point. The extension of the slow eigenvector of a stable equilibrium point should, for instance, attract nearby trajectories faster than systems move along it.

We can generalize the picture given above for trajectories which extend single eigenvectors to cover the case of sets of trajectories which extend several eigenvectors with a given property. The extension of a pair of eigenvectors would be a surface, the extension of three eigenvectors would be a three-dimensional hypersurface, and so on. These extensions, which are called **invariant manifolds**, turn out to be theoretically quite important. Here are some formal definitions:

[1] This is exactly what it sounds like: the collection of all the eigenvalues. It's called a spectrum because we sometime image it by drawing a line at the value of (e.g.) the real part of each eigenvalue. If there's more than one eigenvalue with the same real part, we draw twice as tall a line.

Definition 1 *A* ***differentiable manifold*** *is a continuously and smoothly parameterizable geometric object. In other words, it is a geometric object for which it is possible to establish a system under which every point within the object can be labeled with a unique identifier (e.g. coordinates), and for which the labels vary continuously and smoothly as we move across the object.*

Comment 1 *The word "smooth" implies that the parameterization of a differentiable manifold has at least a few continuous derivatives. I'm trying to avoid getting too ridiculously technical, so I'm being deliberately vague about exactly how many continuous derivatives are enough to make a function smooth. In general in what follows, every function we see will have at least two continuous derivatives.*

Comment 2 *Since differentiable manifolds are continuously parameterizable, then in any given smooth coordinate system (e.g. in our phase-space coordinates), it is possible to express the manifold* locally *(i.e. in some neighborhood of any specified point) as the graph of a function. In other words, if we have a d-dimensional manifold in an n-dimensional space, then at every point in the manifold, we can write*

$$\mathbf{z} = \mathbf{h}(\mathbf{y}),$$

where $\mathbf{y}$ *is a set of* d *of the coordinates in our space,* $\mathbf{h}$ *is a differentiable function of* d *variables, and* $\mathbf{z}$ *is the remaining set of* $n-d$ *coordinates. Because this is only guaranteed locally, we might have to use a different set of variables* $\mathbf{y}$ *and a different* $\mathbf{h}$ *in different parts of the manifold.*

Definition 2 *An* ***invariant set*** *is any set of points in a dynamical system which are mapped into other points in the same set by the evolution operator.*

Example 2.1 Any equilibrium point or set of equilibrium points is an invariant set since each of these points is mapped into itself by the evolution operator.

Example 2.2 A trajectory is an invariant set because each point in the trajectory evolves into another point in the same trajectory under the action of the evolution operator.

Definition 3 *An* ***invariant manifold*** *is exactly what it sounds like: an invariant set that happens to be a differentiable manifold.*

Example 2.3 A single equilibrium point is an invariant manifold: It's clearly invariant, and it's a trivial (zero-dimensional) manifold. A set of equilibrium points on the other hand is not an invariant manifold because it lacks continuity.

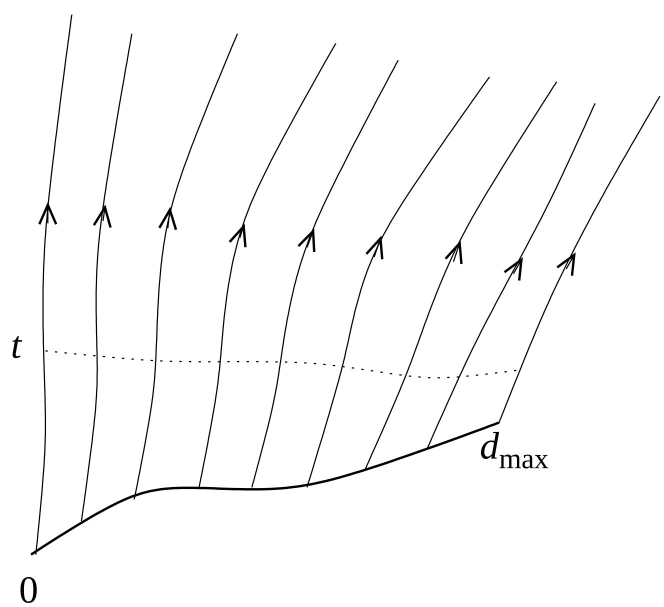


Figure 1: Sketch of a two-dimensional invariant manifold produced by the prescription of example 2.4. The initial curve is shown in bold. This is an arbitrary smoothly parameterizable curve. The dotted curve connects points corresponding to the same time t along each of the trajectories. Each point on the manifold can be labeled by a pair of values: The distance along the initial curve from which the trajectory was started, and the time along the trajectory.

Example 2.4 Suppose that we draw a smooth curve in phase space which we can parameterize in some way, e.g. by distance from the end of the curve where we started drawing. We then take each point on this curve as the initial condition for the differential equation 1 and compute trajectories for all of these initial conditions. (Obviously, we can't do this in practice since there are an infinite number of points on the curve...) Figure 1 shows a sketch of what we might get. Provided the vector field is smooth, connecting points corresponding to the same point in time along each trajectory will generate a smooth curve. We therefore have a geometric object which is invariant, since trajectories are individually invariant, and which can be smoothly parameterized by the distance from the end of our curve at which we started the trajectory and by the time along the trajectory. The result is therefore a two-dimensional invariant manifold, which is just a piece of a two-dimensional surface in phase space.

Since invariant manifolds are differentiable manifolds, then at each point in a d-dimensional manifold we can write

$$\mathbf{z} = \mathbf{z}(\mathbf{y}),$$

where, again, $\mathbf{y}$ is a set of d of the phase-space coordinates, and $\mathbf{z}$ represents the remaining $n-d$ coordinates. Take a time derivative of this equation with respect to time. According to the chain rule,

$$\frac{d\mathbf{z}}{dt} = \sum_{i=1}^{d} \frac{\partial \mathbf{z}}{\partial y_i}\frac{dy_i}{dt}.$$

The time derivatives are just the rate equations of our dynamical system. Since we are, for the moment, only considering autonomous ODEs, these rates only depend on $\mathbf{x} \equiv \cup\{\mathbf{y},\mathbf{z}\}$, so this equation is a partial differential equation for $\mathbf{z}(\mathbf{y})$. This very important equation is known in dynamical systems theory as the **manifold equation**. We shall return to it shortly.

3 Special eigenspaces of equilibrium points

The idea of extending eigenvectors into invariant manifolds will only be useful if we can locate sets of eigenvectors and, by extension, invariant manifolds, which have interesting properties. Based on our knowledge of linearized stability analysis, the following three eigenspaces (subspaces of the full phase space spanned by eigenvectors) are obvious candidates:

Definition 4 *The **stable eigenspace** E^s is the space spanned by the eigenvectors whose corresponding eigenvalues have negative real parts.*

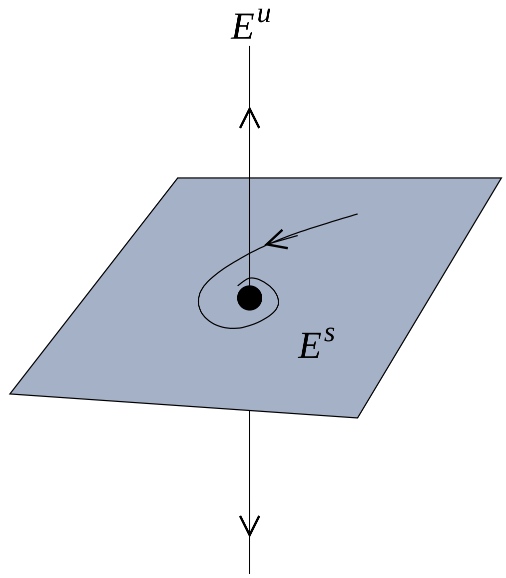
Definition 5 *The **unstable eigenspace** E^u is the space spanned by the eigenvectors whose corresponding eigenvalues have positive real parts.*

Definition 6 *The **centre eigenspace** E^c is the space spanned by the eigenvectors whose corresponding eigenvalues have a real part of zero.*

Comment 3 *Complex eigenvectors and their eigenvalues come in complex-conjugate pairs (identical real parts and imaginary parts of opposite sign). In this case, the appropriate eigenspace is spanned by the real and imaginary parts of the vector.*

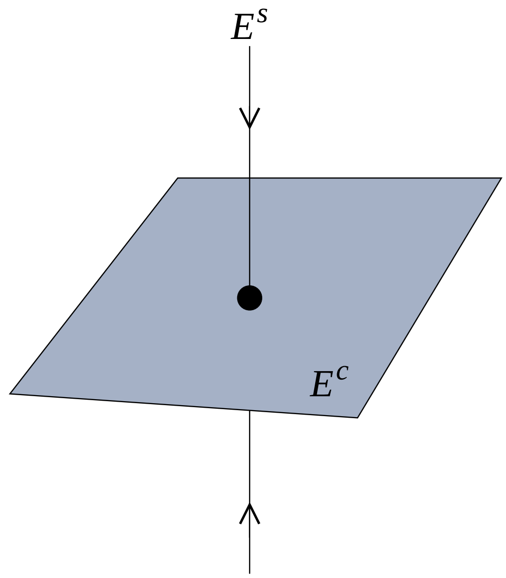
The following examples illustrate the importance of these eigenspaces in understanding the flow near an equilibrium point.

Example 3.1 Suppose that an equilibrium point in a three-dimensional system has a pair of complex-conjugate eigenvalues with negative real parts, and a positive eigenvalue. Then the flow near the equilibrium point has to look qualitatively as follows:



E^u is, in this case, just the unstable eigenvector. Any system on E^u will escape the equilibrium point along this vector. Trajectories which start in the stable eigenspace E^s spiral down to the equilibrium point. Trajectories which start in neither eigenspace will execute a motion which is a combination of the two: They spiral toward E^u while moving away from the equilibrium point along this vector. (Think of a corkscrew around E^u, and you'll have a fair picture of what happens to these trajectories. Note that, eventually, and neglecting the fact that we are moving away from equilibrium into a region where things might get more complicated, the motion corresponding to the stable spiral will bring the system onto the unstable eigenspace so that the latter represents the long-term evolution of the system.

Example 3.2 Suppose that an equilibrium point has two eigenvalues with zero real parts and that all the rest are negative. A three-dimensional picture of the flow near the equilibrium point might look as follows:



Any trajectory which starts off the centre manifold E^c is taken there along a path roughly parallel to E^s. Accordingly, the flow eventually lies in the centre eigenspace. Note that we can't say exactly what happens when systems reach this eigenspace, which is why I didn't draw any trajectories on E^c. This means however that the problem of analyzing the behavior in a high-dimensional space has now been reduced to the simpler problem of analyzing the motion on a two-dimensional surface.

4 Some important classes of invariant manifolds

Since invariant manifolds are just extensions of eigenspaces, they have similar names, with similar meanings. The following technical definitions can be given for the stable and unstable manifolds:

Definition 7 *The **stable manifold** W^s of an equilibrium point $\mathcal{P}$ is a set of points in phase space with the following two properties:*

1. *For $\mathbf{x} \in W^s$, $\varphi^t(\mathbf{x}) \to \mathcal{P}$ as $t \to \infty$.*
2. *W^s is tangent to E^s at $\mathcal{P}$.*

Definition 8 *The **unstable manifold** W^u of an equilibrium point is a set of points in phase space with the following two properties:*

1. *For $\mathbf{x} \in W^u$, $\varphi^t(\mathbf{x}) \to \mathcal{P}$ as $t \to -\infty$.*
2. *W^u is tangent to E^u at $\mathcal{P}$.*

Comment 4 *It is possible to generalize the definitions of stable and unstable manifolds to arbitrary attractors, and not just equilibria.*

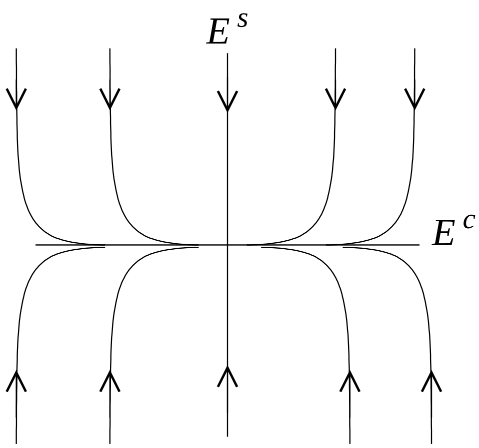
Comment 5 *Note that in both of the above definitions, we talk of manifolds* of an equilibrium point. *Each equilibrium point or, in general, attractor of a dynamical system, will have its own set of manifolds. It is, technically, meaningless to speak of the stable or unstable manifold without specifying which attractor it belongs to, although people sometimes do that, letting the reader infer from context exactly what is meant.*

Comment 6 *Note that the above definitions don't invoke the properties of invariance or of differentiability directly. However, stable and unstable manifolds defined this way must be invariant, differentiable manifolds. I'll let you think about that one.*

The definition of a centre manifold is trickier. We want this manifold to become tangent to the centre eigenspace near the equilibrium point, but we don't know how the system behaves in the centre manifold any more than we knew what was supposed to happen in the centre eigenspace, so we can't use stability to define this manifold as we did with the stable and unstable manifolds.

Definition 9 *The* ***centre manifold*** *of an equilibrium point $\mathcal{P}$ is an invariant manifold of the differential equations with the added property that the manifold is tangent to E^c at $\mathcal{P}$.*

This is a weaker definition than those of the stable and unstable manifolds. Accordingly, *the centre manifold is not unique*. To see why this is, consider the following sketch of the flow for a planar system with a stable and a centre manifold:



The stable manifold is the unique trajectory which runs into the stable eigenspace. However, our definition allows *any* of the other trajectories to be called a centre manifold: They are invariant manifolds since they are trajectories, and they all become tangent to E^c as they approach the equilibrium point. That being said, they all collapse onto a particular trajectory, which is shown here as being identical to E^c but which could curve away from this line. In some sense, the other trajectories are "ordinary" centre manifolds while the central one is "The Centre Manifold". The problem with the classical definition given above is that it applies generally to systems with any manifold structure. However, the centre manifold is mostly useful in systems in which there are no eigenvalues with positive real parts, i.e. in systems with no unstable eigenspace. If there is an unstable eigenspace, the equilibrium point is unstable, and it probably doesn't matter that there are zero eigenvalues. However, as argued earlier, if the system only has stable and centre eigenspaces, then all the interest is on the behavior in the centre eigenspace.

Let us therefore consider the case which interests us most. Suppose that none of the eigenvalues of the Jacobian matrix of an equilibrium point have positive real parts, and that some of the eigenvalues have zero real parts. Then the following **centre manifold theorem** holds:

Theorem 1 *In some neighborhood U of the equilibrium point, there exists a unique centre manifold W^c such that, for any $\mathbf{x} \in U$, $\varphi^t(\mathbf{x}) \to W^c$ as $t \to \infty$.*

Comment 7 *In this case, $t \to \infty$ is just a sneaky way of avoiding having to make estimates of how long it takes for trajectories to collapse to the manifold. In practice, it's often relatively quick.*

The centre manifold theorem guarantees the kind of behavior seen in the diagram above for systems which have a centre and no unstable manifold. Following the reasoning of example 3.2, we can conclude that we just need to analyze the behavior in W^c in order to figure out how the system behaves at long times.

Finally, we sometimes see dynamical systems lacking a centre manifold, but which still behave a bit as the centre manifold theorem describes in that there is a low-dimensional manifold which attracts the trajectories faster than systems move toward the equilibrium point. This is usually due to the existence of a **slow manifold**. Slow manifolds occur in systems where the equilibrium point is stable (all eigenvalues of the Jacobian negative) but where some eigenvalues are much smaller than others in absolute value. The small eigenvalues lead to relatively slow movement toward the equilibrium point, while the fast eigenvalues are responsible for a rapid decay toward the slow manifold.

5 Applications of invariant manifolds

Both of the following examples involve centre manifolds, but the principles for calculating at least local approximations to other kinds of invariant manifolds are much the same.

5.1 The Lindemann mechanism revisited

The rate equations for the Lindemann mechanism are

$$\begin{aligned} \dot{a} &= -a^2 + \alpha ab, \\ \dot{b} &= a^2 - \alpha ab - b. \end{aligned}$$

We found previously that the Jacobian at the equilibrium point (0,0) was

$$\mathbf{J}^* = \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix}.$$

The eigenvalues are $\lambda_0 = 0$ and $\lambda_1 = -1$. This is thus a system with a centre manifold. We will now use centre manifold theory to determine the stability of the equilibrium point, a task which is beyond the reach of stability analysis.

First, we determine the eigenvectors of $\mathbf{J}^*$. The eigenvectors satisfy $\mathbf{J}^*\mathbf{e}_i = \lambda_i \mathbf{e}_i$, or $(\lambda_i \mathbf{I} - \mathbf{J}^*)\mathbf{e}_i = 0$. For λ_0, we get

$$\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} e_{01} \\ e_{02} \end{bmatrix} = 0,$$

from which we conclude that $e_{02} = 0$, i.e. that $\mathbf{e}_0 = (1,0)$ (or any multiple thereof). Similarly, for λ_1,

$$\begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} e_{11} \\ e_{12} \end{bmatrix} = 0,$$

which implies that $\mathbf{e}_1 = (0,1)$. The two eigenvectors in this case happen to be the two coordinate axes. This is, of course, not generally true, nor is it normally the case that the eigenvectors are orthogonal to each other.

The centre manifold will therefore approach the equilibrium point along the a axis. It should therefore be possible to write an equation for the centre manifold of the form $b = b(a)$. On the other hand, we would run into trouble if we wrote $a = a(b)$, since the manifold would be vertical at the equilibrium point. Given such an equation, we could write

$$\dot{a} = -a^2 + \alpha a b(a) \tag{4}$$

for the slow evolution along the centre manifold. Because the equilibrium point is at (0,0) and the centre manifold enters the equilibrium point at a slope of zero, we know that the Taylor expansion of $b(a)$ is of the form

$$b = b_2 a^2 + b_3 a^3 + O(a^4). \tag{5}$$

The symbol $O(a^4)$ indicates that the next term in the expansion, which we aren't writing down, would be proportional to a^4. This means that $\alpha a b(a) = O(a^3)$. Thus, on the centre manifold near the equilibrium point,

$$\dot{a} = -a^2 + O(a^3). \tag{6}$$

Let us review what we have learned. First, our equilibrium point has a stable and a centre manifold. Motion parallel to the stable manifold rapidly leads to the centre manifold, because components off the manifold are damped out at a rate proportional to $e^{\lambda_1 t} = e^{-t}$. Once a system gets to the manifold, its motion is governed by equation 4, which reduced to 6 near the equilibrium point. For $a > 0$, it follows that the equilibrium point is stable. Amusingly, the equilibrium point is only **semi-stable**, as we see quite clearly here: If we take small, negative (unphysical) values of a, equation 6 still applies, but now trajectories move *away* from the equilibrium point. In other words, the equilibrium point is only stable from the right. This is pretty peculiar behavior, but then just about everything about the irreversible Lindemann mechanism is peculiar.

When we use a centre manifold argument to determine the stability of an equilibrium, we generally need to do more work than we have done here. It is not typically the case that we can determine the behavior on the manifold by inspection. Usually in fact, we have to use the manifold equation to determine the coefficients of the Taylor expansion 5, then substitute the resulting equation into 4 and simplify to determine the leading-order behavior of the rate. As an exercise, let's work out the coefficients b_2 and b_3. The manifold equation for a planar system is

$$\dot{b} = \frac{db}{da}\dot{a}.$$

Note that the partial derivatives in the general manifold equation become an ordinary derivative here since we're computing a one-dimensional manifold.

We can evaluate db/da, the derivative on the manifold, directly from our ansatz 5:

$$\frac{db}{da} = 2b_2a + 3b_3a^2 + O(a^3).$$

Note that I'm keeping track of the truncation order. This is always a good idea. If we're careful about this, we can be reasonably sure at the end that we haven't left out any terms which would contribute to the coefficients we're trying to calculate.

We now go back to the manifold equation, substituting in the rate equations, our ansatz for the manifold, and the derivative of b on the manifold:

$$\begin{aligned} a^2 - \alpha a\left[b_2a^2 + b_3a^3 + O(a^4)\right] - \left[b_2a^2 + b_3a^3 + O(a^4)\right] \\ = \left[2b_2a + 3b_3a^2 + O(a^3)\right]\left\{-a^2 + \alpha a\left[b_2a^2 + b_3a^3 + O(a^4)\right]\right\}. \end{aligned}$$

We now collect on one side in powers of a:

$$a^2(1-b_2) + a^3(2b_2 - \alpha b_2 - b_3) + O(a^4) = 0.$$

Since this equation must be valid for any value of a, the coefficients of each term must individually be equal to zero. Thus we get

$$\begin{aligned} b_2 &= 1. \\ b_3 &= (2-\alpha)b_2 = 2-\alpha. \end{aligned}$$

5.2 A simple AIDS model

Consider the following set of dimensionless ODEs, which describe a very simple model for the spread of AIDS through two populations which have infrequent high-risk contacts:[2]

$$\begin{aligned} \dot{c}_1 &= -\alpha c_1 + p_1(c_1 + \beta_1 c_2), && \text{(7a)} \\ \dot{p}_1 &= p_1(1 - c_1 - \beta_1 c_2), && \text{(7b)} \\ \dot{c}_2 &= -\alpha c_2 + p_2(c_2 + \beta_2 c_1), && \text{(7c)} \\ \dot{p}_2 &= p_2(1 - c_2 - \beta_2 c_1). && \text{(7d)} \end{aligned}$$

The details of this model can be studied from the original reference. Briefly, p_j represents the number of healthy individuals in subpopulation j, while c_j represents the number of contagious individuals. Terms in αc_j represent increased mortality in the contagious group due to the disease. Terms of the form p_jc_k (whether $j = k$ or not) represent the transmission of the disease from contagious individuals. Due to the assumptions of the model, $\alpha > 0$ and $0 < \beta_j < 1$.

It is convenient to use Maple from here on. We start by defining the rate equations:

[2] M. R. Roussel, SIAM Rev. **39**, 106 (1997).

```
> c1dot := -alpha*c1 + p1*(c1+beta1*c2):
> p1dot := p1*(1-c1-beta1*c2):
> c2dot := -alpha*c2 + p2*(c2+beta2*c1):
> p2dot := p2*(1-c2-beta2*c1):
```

We now solve for the steady states of the model:

```
> steady_states :=
  solve({c1dot=0,p1dot=0,c2dot=0,p2dot=0},{c1,p1,c2,p2});
```

$$steady\_states := \{c1 = 0, p1 = 0, p2 = 0, c2 = 0\}, \{c1 = 0, c2 = 1, p1 = 0, p2 = \alpha\}, \{p2 = 0, c2 = 0, p1 = \alpha, c1 = 1\}, \{p2 = \frac{\alpha(-1+\beta 2)}{-1+\beta 2\,\beta 1}, c2 = \frac{-1+\beta 2}{-1+\beta 2\,\beta 1}, p1 = \frac{\alpha(-1+\beta 1)}{-1+\beta 2\,\beta 1}, c1 = \frac{-1+\beta 1}{-1+\beta 2\,\beta 1}\}$$

There are four steady states. The first one corresponds to extinction of the entire population, the second and third to extinction of one subpopulation or the other, and the last to coexistence of both subpopulations.

We will now determine the stability of these four steady states. Note that in the `linalg` routines, you have to always keep the same order for the variables. I choose to list them in the order (c_1, p_1, c_2, p_2). Any other order would do, provided you're consistent.

```
> with(linalg):
Warning, the protected names norm and trace have been redefined and
unprotected
> J := jacobian([c1dot,p1dot,c2dot,p2dot],[c1,p1,c2,p2]);
```

$$J := \begin{bmatrix} -\alpha + p1 & c1 + \beta 1\, c2 & p1\,\beta 1 & 0 \\ -p1 & 1 - c1 - \beta 1\, c2 & -p1\,\beta 1 & 0 \\ p2\,\beta 2 & 0 & -\alpha + p2 & c2 + \beta 2\, c1 \\ -p2\,\beta 2 & 0 & -p2 & 1 - c2 - \beta 2\, c1 \end{bmatrix}$$

```
> J1 := map(x->subs(steady_states[1],x),J);
```

$$J1 := \begin{bmatrix} -\alpha & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -\alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The eigenvalues of a diagonal matrix are just the diagonal elements. In this case, the eigenvalues are therefore $-\alpha$ (twice) and 1 (twice). The positive eigenvalues make the extinction fixed point unstable. (Phew!)

```
> J2 := map(x->subs(steady_states[2],x),J);
```

$$J2 := \begin{bmatrix} -\alpha & \beta 1 & 0 & 0 \\ 0 & 1 - \beta 1 & 0 & 0 \\ \alpha\,\beta 2 & 0 & 0 & 1 \\ -\alpha\,\beta 2 & 0 & -\alpha & 0 \end{bmatrix}$$

```
> eigenvals(J2);
```

$$-\alpha, 1-\beta 1, \sqrt{-\alpha}, -\sqrt{-\alpha}$$

Since $\beta_1 < 1$, the second eigenvalue is positive, so this steady state, in which population 1 becomes extinct, is also unstable. I won't do the calculation for the third steady state, but it shouldn't surprise you, given the symmetry of the model, that it's unstable too.

The final steady state leads to a slightly more complex, but still tractable problem. I won't show the matrix J_4, the Jacobian evaluated at the fourth steady state. It's computed like the others, and its form doesn't suggest anything in particular. I'll just go straight to calculating the characteristic polynomial:

```
> J4 := map(x->subs(steady_states[4],x),J):
> factor(charpoly(J4,lambda));
```

$$\frac{(\alpha+\lambda^2)\,(2\,\lambda\,\alpha\,\beta 2\,\beta 1+\alpha\,\beta 1-\lambda\,\alpha\,\beta 1+\alpha\,\beta 2-\lambda\,\alpha\,\beta 2-\alpha\,\beta 2\,\beta 1-\alpha-\lambda^2+\lambda^2\,\beta 2\,\beta 1)}{-1+\beta 2\,\beta 1}$$

The eigenvalues are roots of this equation. Thus, either the first term in the numerator is zero, or the second. The first term gives us a pair of purely imaginary eigenvalues:

$$\lambda^c_\pm = \pm\sqrt{-\alpha}.$$

We will therefore have a two-dimensional centre manifold associated with these eigenvalues. To see what the second term in the characteristic equation implies, collect it in powers of λ:

```
> collect(2*lambda*alpha*beta2*beta1+alpha*beta1-lambda*alpha*beta1
  +alpha*beta2-lambda*alpha*beta2-alpha*beta2*beta1-alpha-lambda^2
  +lambda^2*beta2*beta1,lambda);
```

$$(-1+\beta 2\,\beta 1)\,\lambda^2+(2\,\alpha\,\beta 2\,\beta 1-\alpha\,\beta 2-\alpha\,\beta 1)\,\lambda+\alpha\,\beta 1-\alpha+\alpha\,\beta 2-\alpha\,\beta 2\,\beta 1$$

Again, we're trying to find out the values of λ at which this equation equals zero. It won't change anything to multiply the whole equation by -1. I'll do that, and rearrange by hand a bit:

$$\lambda^2(1-\beta_1\beta_2)+\lambda\alpha\,[\beta_1(1-\beta_2)+\beta_2(1-\beta_1)]+\alpha(1-\beta_1)(1-\beta_2)=0.$$

It should now be clear that all the coefficients in this equation are positive. You should convince yourself that, if a quadratic equation has only positive coefficients, the real parts of the roots are necessarily negative. Thus, these two eigenvalues (and their eigenvectors) are associated with the stable manifold.

The fourth equilibrium point therefore has a two-dimensional stable manifold and a two-dimensional centre manifold. We can apply the centre manifold theorem again, which says that, eventually, the system will collapse onto the centre manifold so that we need only concern ourselves with the dynamics in this plane.

Let's compute the eigenvectors corresponding to the centre-manifold eigenvalues:

```
> eigenvectors(J4);
```

This command calculates all the eigenvectors and eigenvalues. I'll only show the output corresponding to $\lambda_{\pm}^{c}$:

$$[\sqrt{-\alpha}, 1, \{\left[-\frac{\sqrt{-\alpha}\,(-1+\beta 1)}{\alpha\,(-1+\beta 2)}, \frac{-1+\beta 1}{-1+\beta 2}, -\frac{\sqrt{-\alpha}}{\alpha}, 1\right]\}],$$
$$[-\sqrt{-\alpha}, 1, \{\left[\frac{\sqrt{-\alpha}\,(-1+\beta 1)}{\alpha\,(-1+\beta 2)}, \frac{-1+\beta 1}{-1+\beta 2}, \frac{\sqrt{-\alpha}}{\alpha}, 1\right]\}]$$

In each line above, the first value is the eigenvalue, and the second is its multiplicity (how many different eigenvectors there are for this eigenvalue). The eigenvector itself appears after these two quantities. Note that the eigenvectors are of the form

$$\mathbf{e}_{\pm}^{c} = \left(\mp\frac{i\sqrt{\alpha}(1-\beta_1)}{\alpha(1-\beta_2)}, \frac{1-\beta_1}{1-\beta_2}, \mp\frac{i\sqrt{\alpha}}{\alpha}, 1\right).$$

As was briefly mentioned in comment 3, the basis of the centre eigenspace can be obtained simply by taking the real and imaginary parts of one of these vectors:

$$\begin{aligned}\mathbf{e}_1 &= \left(0, \frac{1-\beta_1}{1-\beta_2}, 0, 1\right), \\ \text{and } \mathbf{e}_2 &= \left(\frac{1-\beta_1}{\sqrt{\alpha}(1-\beta_2)}, 0, \frac{1}{\sqrt{\alpha}}, 0\right).\end{aligned}$$

If we let $\mathbf{x} = (c_1, p_1, c_2, p_2)$, and call the steady state we're currently analyzing $\mathbf{x}_4$, then the centre eigenspace of this steady state can be written in the form

$$\mathbf{x} = \mathbf{x}_4 + a_1\mathbf{e}_1 + a_2\mathbf{e}_2.$$

Working through it in detail, we have

$$\begin{aligned}c_1 &= \frac{1-\beta_1}{1-\beta_1\beta_2} + a_2\frac{1-\beta_1}{\sqrt{\alpha}(1-\beta_2)}. \\ p_1 &= \frac{\alpha(1-\beta_1)}{1-\beta_1\beta_2} + a_1\frac{1-\beta_1}{1-\beta_2}. \\ c_2 &= \frac{1-\beta_2}{1-\beta_1\beta_2} + \frac{a_2}{\sqrt{\alpha}}. \\ p_2 &= \frac{\alpha(1-\beta_2)}{1-\beta_1\beta_2} + a_1.\end{aligned}$$

We normally like to express manifolds and eigenspaces as explicit rather than parametric functions. In other words, we would prefer to write $c_2 = g(c_1, p_1)$ and $p_2 = h(c_1, p_1)$. To do this, all we have to do is to eliminate a_1 and a_2 from the above equations. If we do, we get a very simple result:

$$c_2 = \frac{1-\beta_2}{1-\beta_1}c_1, \tag{8a}$$

$$\text{and } p_2 = \frac{1-\beta_2}{1-\beta_1}p_1. \tag{8b}$$

We could now think about expanding the centre manifold in a series by adding quadratic terms (in c_1^2, c_2^2 and c_1c^2) to the above expressions, and then figuring out the coefficients using the manifold equation. However, for this particular model, the centre eigenspace turns out to be *exactly* the centre manifold, i.e. there are no quadratic or higher-order correction terms.

To prove this, we start by writing down the manifold equations:

$$\begin{aligned}
\dot{c}_2 &= \frac{\partial c_2}{\partial c_1}\dot{c}_1 + \frac{\partial c_2}{\partial p_1}\dot{p}_1, \\
\text{and } \dot{p}_2 &= \frac{\partial p_2}{\partial c_1}\dot{c}_1 + \frac{\partial p_2}{\partial p_1}\dot{p}_1.
\end{aligned}$$

On the centre eigenspace,

$$\frac{\partial c_2}{\partial p_1} = \frac{\partial p_2}{\partial c_1} = 0,$$

which simplifies our manifold equations considerably. Now substitute the relevant equations into the remaining terms. For the c_2 equation, we get

$$-\alpha c_1 \frac{1-\beta_2}{1-\beta_1} + p_1 \frac{1-\beta_2}{1-\beta_1}\left(c_1 \frac{1-\beta_2}{1-\beta_1} + \beta_2 c_1\right) = \frac{1-\beta_2}{1-\beta_1}\left[-\alpha c_1 + p_1\left(c_1 + \beta_1 c_1 \frac{1-\beta_2}{1-\beta_1}\right)\right].$$

$$\therefore -\alpha + p_1\left(\frac{1-\beta_2}{1-\beta_1} + \beta_2\right) = -\alpha + p_1\left(1 + \beta_1 \frac{1-\beta_2}{1-\beta_1}\right).$$

$$\therefore -\alpha + p_1 \frac{1-\beta_1\beta_2}{1-\beta_1} = -\alpha + p_1 \frac{1-\beta_1\beta_2}{1-\beta_1}.$$

The equation is identically satisfied by the centre eigenspace. It is easy to verify that the p_2 equation is also identically satisfied. This means that the centre eigenspace is actually an invariant manifold, and thus that it is the centre manifold for this problem.

If we now substitute our centre manifold equations 8 into equations 7a and 7b, we get the ODEs governing motion on the manifold:

$$\dot{c}_1 = -\alpha c_1 + c_1 p_1 \frac{1-\beta_1\beta_2}{1-\beta_1}, \tag{9a}$$

$$\dot{p}_1 = p_1\left(1 - c_1 \frac{1-\beta_1\beta_2}{1-\beta_1}\right). \tag{9b}$$

These equations turn out to be those of a very famous model from population ecology known as the Lotka-Volterra model. We'll save a discussion of the behavior of this model for another day. For now, let us just note that we have reduced the problem of understanding the long-term dynamics of the four-dimensional system 7 to the much simpler problem of studying the planar system 9.