# Nonlinear Dynamics and Chaos: part 1. One-dimensional Flows Ch 3 Bifrucation

Charlie Seager

4/16/2024

Chapter 3.0 Introduction Given the triviality of the dynamics, what's interesting about one-dimensional systems? Answer: Dependence on parameters. The qualitative structure of the flow can change as parameters are varied. In particular, fixed points can be created or destroyed, or their stability can change. These qualitative changes in the dynamics are called bifurcations, and the parameter values at which they occur are called bifurcation points.
Bifurcations are important scientifically – they provide models of transitions and instabilities as some control parameter is varied. For example, consider the buckling of a beam. If a small weight is placed on top in Fig 3.0.1 the beam can support the load and remain vertical. But if the load is too heavy, the vertical position becomes unstable and the beam may buckle
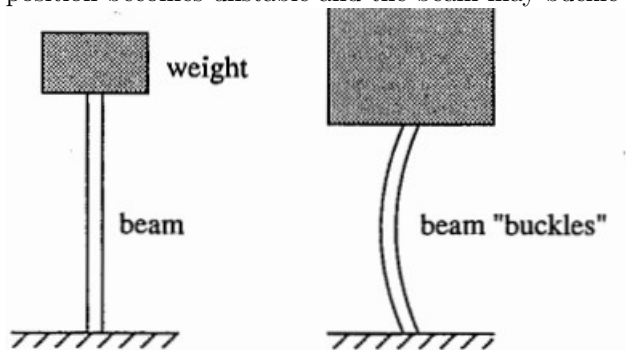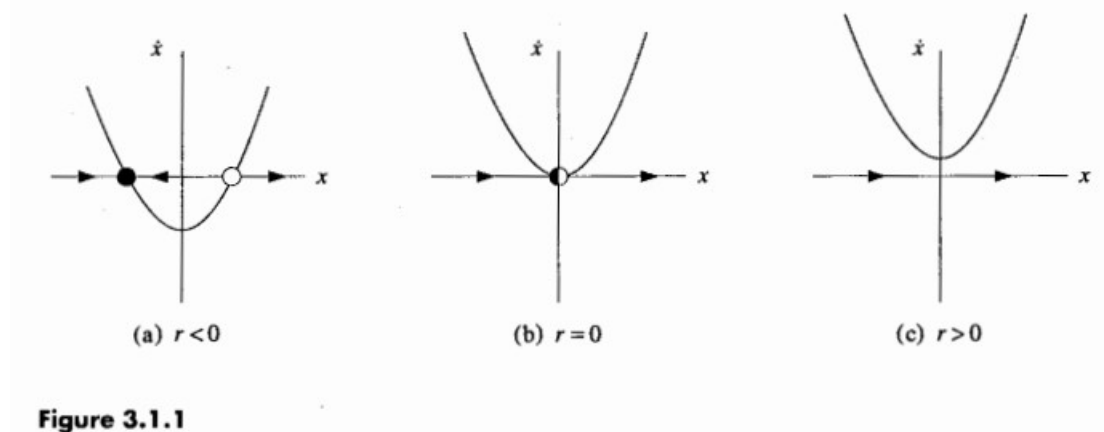

Figure 3.0.1

Here the weight plays the role of the control parameter, and the deflection of the beam from vertical plays the role of the dynamical variable x.

Chapter 3.1 Saddle-Node Bifurcation
The saddle-node bifurcation is the basic mechanism by which fixed points are created and destroyed. As a parameter is varied, two fixed points move toward each other, collide, and mutually annihilate. The prototypical example of a saddle node bifurcation is given by the first order system

$$\dot{x} = r + x^2$$

where r is a parameter, which may be positive, negative, or zero. When r is negative, there are two fixed points, one stable and one unstable (Fig 3.1.1.a)


(a) $r < 0$ (b) $r = 0$ (c) $r > 0$

Figure 3.1.1

As r approaches 0 from below, the parabola moves up and the two fixed points move toward each other. When r=0, the fixed points coalesce into a half-stable fixed point at $x^* = 0$ (Fig 3.1.1b). This type of fixed point is extremely delicate – it vanishes as soon as $r > 0$ and now there are no fixed points at all (Fig 3.1.1c).

In this example, we say that a bifurcation occured at r = 0, since the vector fields for $r < 0$ and $r > 0$ are qualitatively different.

Graphical Conventions

There are several other ways to depicit a saddle-node bifurcation. We can show a stack of vector fields for discrete values of r (Fig 3.1.2). This representation emphasizes the dependence of the fixed poitns on r. In the limit of a continuous stack of vector fields, we have a picture like Fig 3.1.3. The curve shown is $r = -x^2$, i.e., $\dot{x} = 0$, which gives the fixed points for different r. To distinguish between stable and unstable fixed points, we use a solid line for stable poitns and a
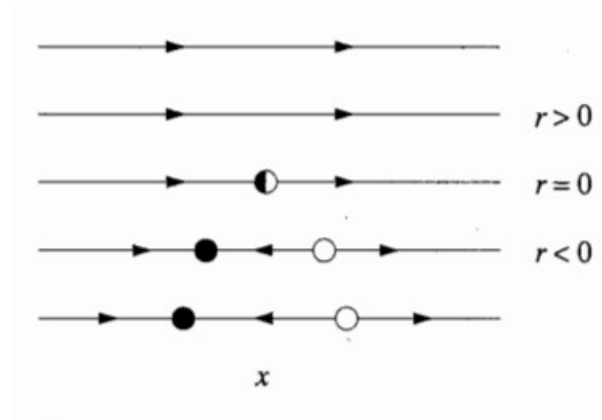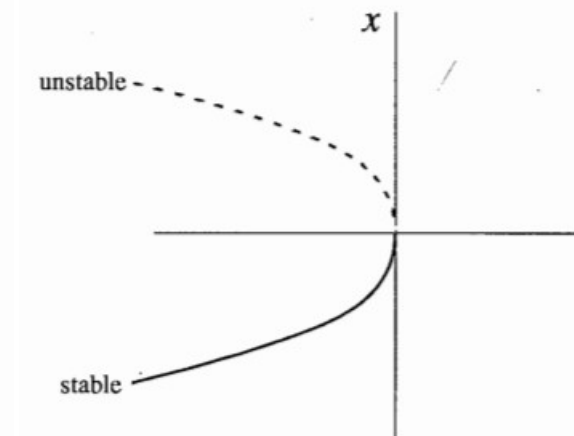
Figure 3.1.2

broken line for unstable ones.

However the most common way to depict the bifurcation is to invert the axes of Figure 3.1.3. The rationale is that r plays the role of an independent variable, and so should be plotted horizontally (Figure 3.1.4). The drawback is that now the x-axis has to be plotted vertically, which looks strange at first. Arrows are sometimes included in the picture but not always. This picture is called the bifurcation diagram for the saddle-node bifurcation.
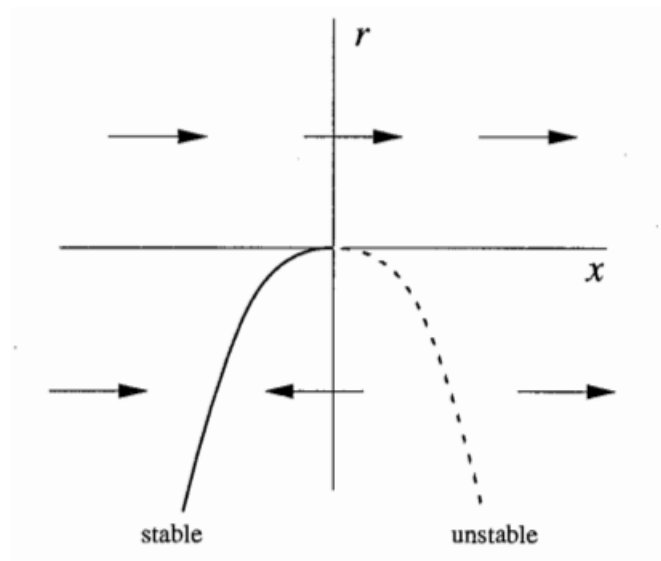

Figure 3.1.3

Figure 3.1.4

Terminology Bifurcation theory is rife with conflicting terminology. The subject really hasnt settled down yet and different people use different words for the same thing. For example, the saddle-node bifurcation is sometimes called a fold bifurcation (because the curve in Fig 3.1.4 has a fold in it) or a turning point bifurcation (because the point $(x,r) = (0,0)$ is a "turning point")

3

Admittedly, the term saddle-node doesnt make much sense for vector fields on the line. The name derives from a completely analogous bifurcation seen in a higher-dimensional context, such as vector fields on the plane, where fixed poitns known as saddles and nodes can collide and annihilate (see section 8.1). The prize for most inventive teminology must go to Abraham and Shaw (1988) who write of a blue sky bifurcation. This term comes from viewing a saddle-node bifurcation in the other direction: a pair of fixed points appears "out of the clear blue sky" as a parameter is varied. For example, the vector field

$$\dot{x} = r - x^2$$

has no fixed poitns for $r < 0$ but then one materializes when r $=$ 0 and splits into two when $r > 0$ (Figure 3.1.5). Incidentally, this example also explains why we use the word "bifurcation": it means "splitting into two branches."


$r<0$ $r=0$ $r>0$

Figure 3.1.5
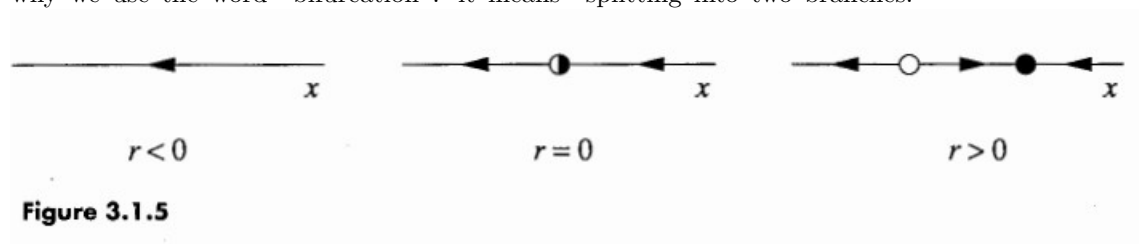
Normal Forms In a certain sense, the examples $\dot{x} = r - x^2$ or $\dot{x} = r + x^2$ are representative of all saddle node bifurcations: that's why we called them "prototypical". The idea is that, close to a saddle-node bifurcation, the dynamics typically look like $\dot{x} = r - x^2$ or $\dot{x} = r + x^2$
For instance, consider Example 3.1.2 near the bifurcation at x $=$ 0 and r $=$ 1 using the Taylor expansion for e^{-x} about x=0, we find

$$\dot{x} = r - x - e^{-x} = r - x = [1 - x + \tfrac{x^2}{2!} + \ldots] = (r-1) - \tfrac{x^2}{2} + \ldots$$

to leading order in x. This has the same algebraic form as $\dot{x} = r - x^2$, and can be made to agree exactly by appropiate rescalings of x and r.
It's easy to understand why saddle-node bifurcations typically have this algebraic form. We just ask ourselves: how can two fixed points occur where the graph of f(x) intersects the x-axis. For a saddle-node bifurcation ot be possible, we need two nearby roots of f(x): this means f(x) must look locally "bowl-shaped" or parabolic (Figure 3.1.7)

4
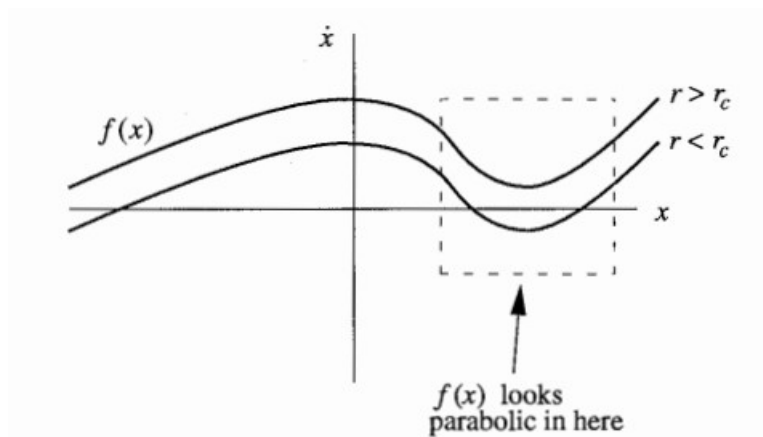
Figure 3.1.7

Now we use a microscope to zoom in on the behavior near the bifurcation. As r varies, we see a parabola intersecting the x-axis, then becoming tangent to it, and then failing to intersect. This is exactly the scenario in the prototypical Figure 3.1.1.

Here's a more algebraic version of the same argument. We regard f as a function of both x and r, and examine the behavior of $\dot{x} = f(x, r)$ near the bifurcation at $x = x^*$ and r = r. Taylor's expansion yields.

$$\dot{x} = f(x, r) =$$
$$f(x^*, r_c) + (x - x^*)\frac{\partial f}{\partial x}\big|_{(x^*, r_c)} + (r - r_c)\frac{\partial f}{\partial r}\big|_{(x^*, r_c)} + \frac{1}{2}(x - x^*)^2\frac{\partial^2 f}{\partial x^2}\big|_{(x^*, r_c)} + \ldots$$

where we have neglected quadratic terms in $(r - r_c)$ and cubic terms in $(x - x^*)$. Two of the terms in this equation vanish: $f(x^*, r_c) = 0$ since x^* is a fixed point, and $\partial f/\partial x\big|_{(x^*, r_c)} = 0$ by the tangency condition of a saddle-node bifurcation. Thus

$$\dot{x} = a(r - r_c) + b(x - x^*)^2 + \ldots \qquad (3)$$

where $a = \partial f/\partial r\big|_{(x^*, r_c)}$. Equation (3) agrees with the form of our prototypical examples. (We are assuming that $a, b \neq 0$ which is the typical case: for instance it would be a very special situation if the second derivative $\partial^2 f/\partial x^2$ also happened to vanish at the fixed point).

What we have been calling prototypical examples are more conventionally known as normal fomrs for the saddle-node bifurcation. There is much, much more to normal forms than we have indicated here. We will be seeing their improtance throughout this book. For a more detailed and precise discussion see Guckenheimer and Holmes (1983) or Wiggins (1990).

Chapter 3.2 Transcritical Bifurcation

There are certain scientific situations where a fixed point must exist for all values of a parameter and can never be destroyed. For example, in the logistic equation and other simple models for the growth of a single species, there is a

fixed point at zero population, regardless of the value of the value of the growth rate. However, such a fixed point may change its stability as the parameter is varied. The transcritical bifurcation is the standard mechanism for such changes in stability.

The normal form for a transcritical bifurcation is

$$\dot{x} = rx - x^2$$

This looks like the logistic equation of Section 2.3, but now we allow x and r to be either positive or negative.

Figure 3.2.1 shows the vector field as r varies. Note that there is a fixed point at $x^* = 0$ for all values of r.


(a) $r < 0$ (b) $r = 0$ (c) $r > 0$
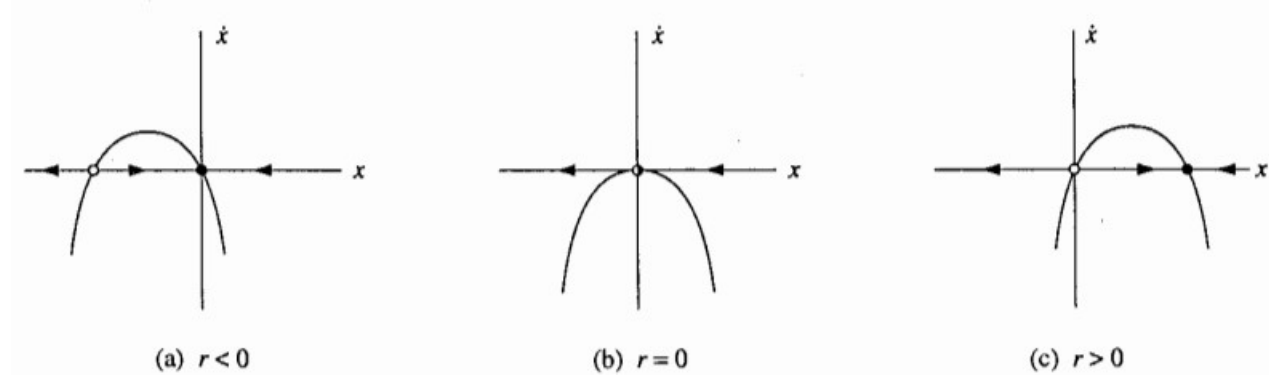
Figure 3.2.1

For $r < 0$, there is an unstable fixed point at $x^* = r$ and a stable fixed point at $x^* = 0$. As r increases, the unstable fixed point approaches the origin has become unstable and $x^* = r$ is now stable. Some people say that an exchange of stabilities has taken place between the two fixed points.

Please note the important difference between the saddle node and transcritical bifurcations: in the transcritical case, the two fixed points don't disappear after the bifurcations- instead they just switch their stability.

Figure 3.2.2 shows the bifurcation diagram for the transcritical bifurcation. As in Figure 3.1.4, the parameter r is regarded as the independent variable, and the fixed points $x^* = 0$ and $x^* = r$ are shown as dependent variables.
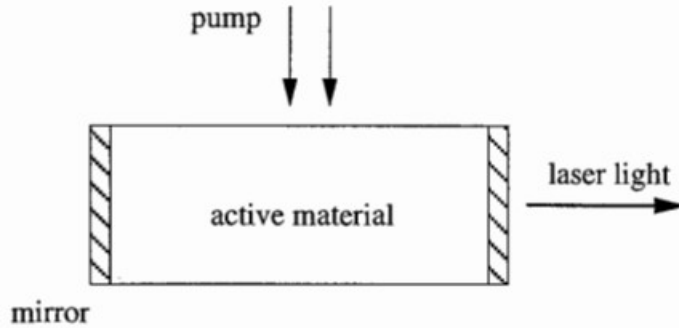
Figure 3.2.2
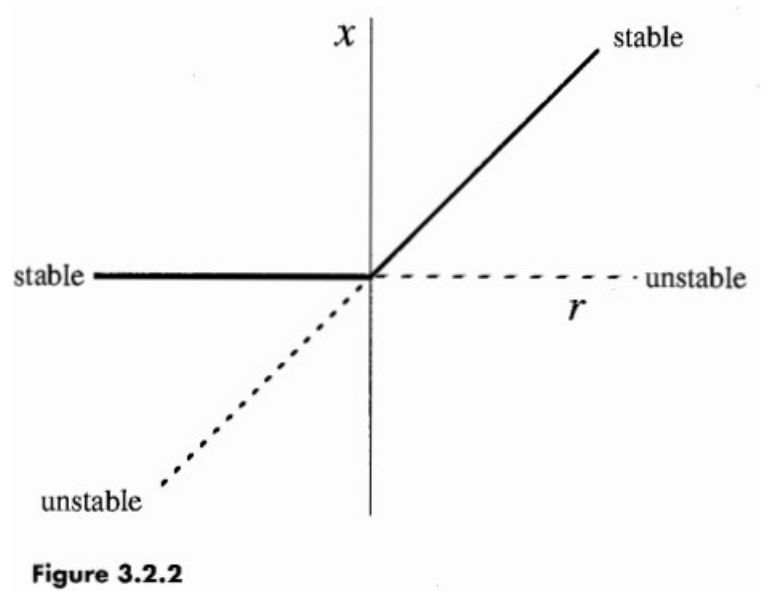
Chapter 3.3 Laser Threshold

Now its time to apply our mathematics to a scientific example. We analyze an extremely simplified model for a laser, following the treatment given by Haken (1983).

Physical Background

We are going to consider a particular type of laser known as a solid-state laser which consists of a collection of special "laser-active" atoms embedded in a solid state matrix, bounded by partially reflecting mirrors at either end. An external energy source is used to excite or "pump" the atoms out of their ground states (Figure 3.3.1)

pump

active material

laser light

mirror

Figure 3.3.1

Each atom can be thought of as a little antenna radiating energy. When the pumping is relatively weak, the laser acts just like an ordinary lamp: the excited atoms oscillate independently of one another and emit randomply phased light waves.

Now suppose we increase the strength of the pumping. At first nothing different happens, but then the suddenly, when the pump strength exceeds a certain threshold, the atoms begin to oscillate in phase-the lap has turned into a laser. Now the trillions of little atennas act like one giant antenna and produce a beam of radiation that is much more coherent and intense than that produced below the laser threshold.

This sudden onset of coherence is amazing, considering that the atoms are being excited completely at random by the pump! Hence the process is self-organizing: the coherence develops because of a cooperative interaction among the atoms themselves. **Model**
A proper explanation of the laser phenomenon would require us to delve into quantum mechanics. See Milonni and Eberly (1988) for an intuitive discussion.

Instead we consider a simplified model of the essential physics (Haken 1983 p. 127). The dynamical variable is the number of photons n(t) in the laser field. Its rate of change is given by

$$\dot{n} = gain - loss = GnN - kn$$

The gain term comes from the process of stimulated emission in which photons stimulate excited atoms to emit additional photons. Because this process occurs via random encounters between photons and excited atoms, it accurs at a rate proportional to n and to the number of excited atoms, denoted by N(t). The parameter $G > 0$ is known as the gain coefficient. The loss term models the escape of photons through the endfaces of the laser. The parameter $k > 0$ is a

rate constant: its reciprocal $\tau = 1/k$ represents the typical lifetime of a photon in the laser.

Now comes the key physical idea: after an excited atom emits a photon, it drops down to a lower energy level and is no longer excited. Thus N decreases by the emission of photons. To capture this effect, we need to write an equation relating N to n. Suppose that in the absence of laser action, the pump keeps the number of excited atoms fixed at N_0. Then the actual number of excited atoms will be reduced by the laser process. Specifially, we assume $N(t) = N_0 - \alpha n$ where $\alpha > 0$ is the rate at which atoms drop back to their ground states. Then

$$\dot{n} = Gn(N_0 - \alpha n) - kn = (GN_0 - k)n - (\alpha G)n^2$$

We're finally on familiar ground- this is a first order system for n(t). Figure 3.3.2 shows the corresponding vector field for different values of the pump strength N_0. Note that only positive values of n are physically meaningful.

$N_0 < k/G$ $N_0 = k/G$ $N_0 > k/G$
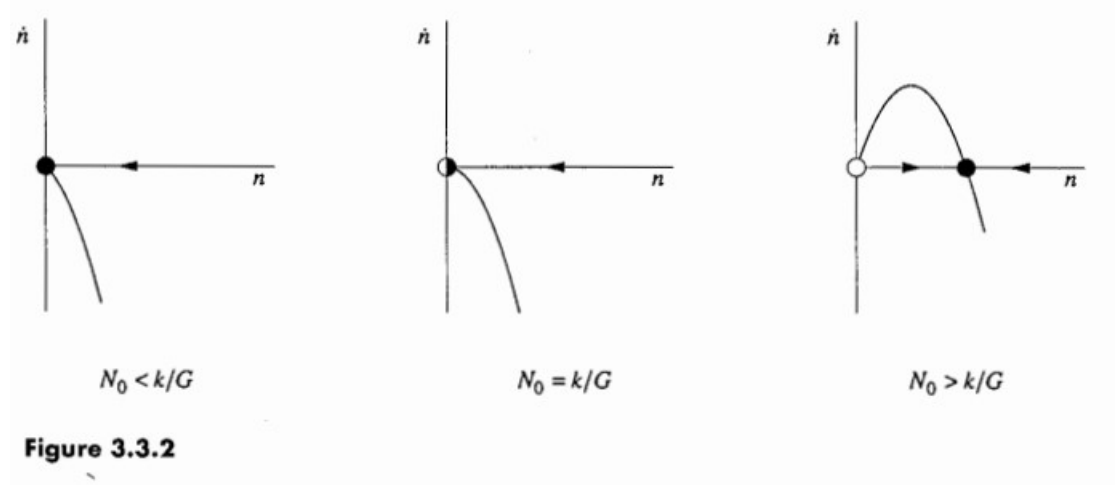
Figure 3.3.2

When $N_0 < k/G$, the fixed point at $n^* = 0$. This means that there is no stimulated emission and the laser acts like a lamp. As the pump strength N_0 is increased, the system undergoes a transcritical bifurcation when $N_0 = k/G$. For $N_0 > k/G$, the origin loses stability and a stable fixed point appears at $n^* = (GN_0 - k)/\alpha G > 0$, corresponding to spontaneous laser action. Thus $N_0 = k/G$ can be interpreted as the laser threshold in this model. Figure 3.3.3 summarizes
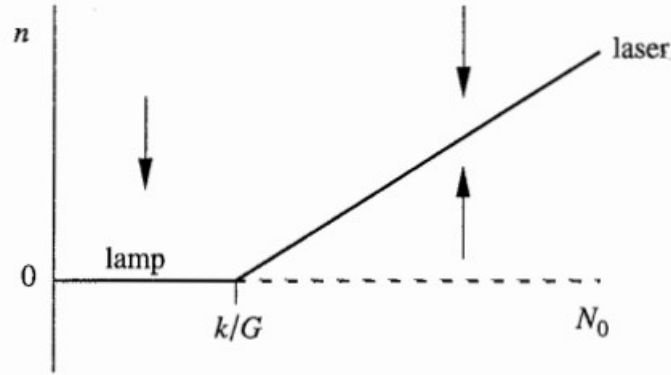
Figure 3.3.3

our results   Although this model correctly predicts the existene of a threshold, it ignores the dynamics of the excited atoms, the existence of spontaneous emission, and several other complications.

Chapter 3.4 Pitchfork Bifurcation

We turn now to a third kind of bifurcation, the so-called pitchfork bifurcation. This bifurcation is common in physical problems that have a symmetry. For example, many problems have a spatial symmetry between left and right. In such cases, fixed points tend to appear and disappear in symmetrical pairs. In the buckling example of Figure 3.0.1, the beam is stable in the vertical position if the load is small. In this case there is a stable fixed point corresponding to zer deflection. But if the load exceeds the buckling threshold, the beam may buckle to either the left or right. The vertical positions has gone unstable, and two new symmetrical fixed points, corresponding to left- and right-buckled configurations, have been born.

There are two very different types of pitchfork bifurcation. The simpler type is called supercritical, and will be discussed first.

Supercritical Pitchfork Bifurcation

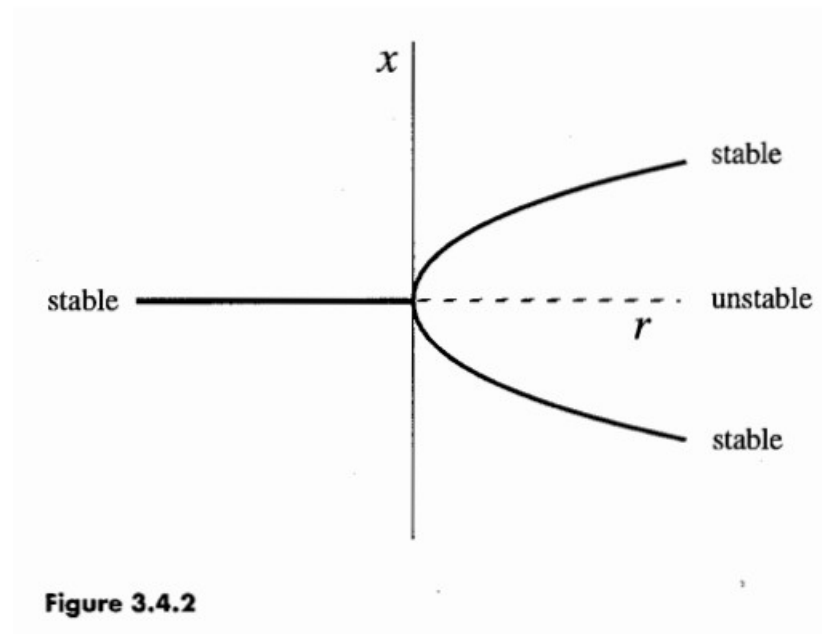The normal form of the supercritical pitchfork bifurcation is

$$\dot{x} = rx - x^3$$

Note that this equation is invariant under the change of variables $x \to -x$. That is, if we replace x by -x and then cancel the resulting minus signs on both sides of the equation, we get (1) back again. This invariance is the mathematical expression of the left-right symmetry mentioned earlier. (More technically, one says that the vector field is equivalen, but we'll use the more familiar language.)

10

(a) $r<0$ (b)
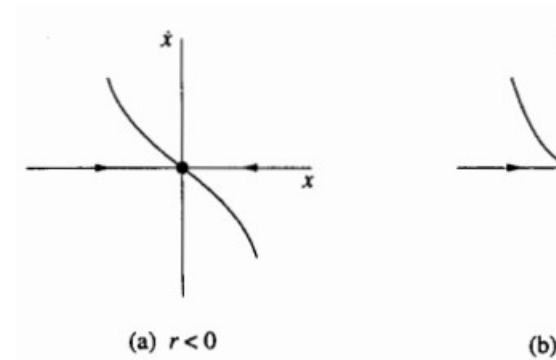
Figure 3.4.1

Figure 3.4.1 shows the vector field for different values of r. When $r < 0$, the origin is the only fixed point, and it is stable. When r=0, the origin is still stable, but much more weakly so, since the linearization vanishes. Now solutions no longer decay exponentially fast – instead the decay is much slower algebraic function of time. This lethargic decay is called critical slowing down in the physics literature Finally, when $r > 0$, the origin has become unstable. Two new stable fixed points appear on either side of the origin, symmetrically located at $x^* = \pm\sqrt{r}$.

The reason for the term "pitchfork becomes clear when we plot the bifurcation diagram(Figure 3.4.2). Actually, pitchfork trifurcation might be a better word!
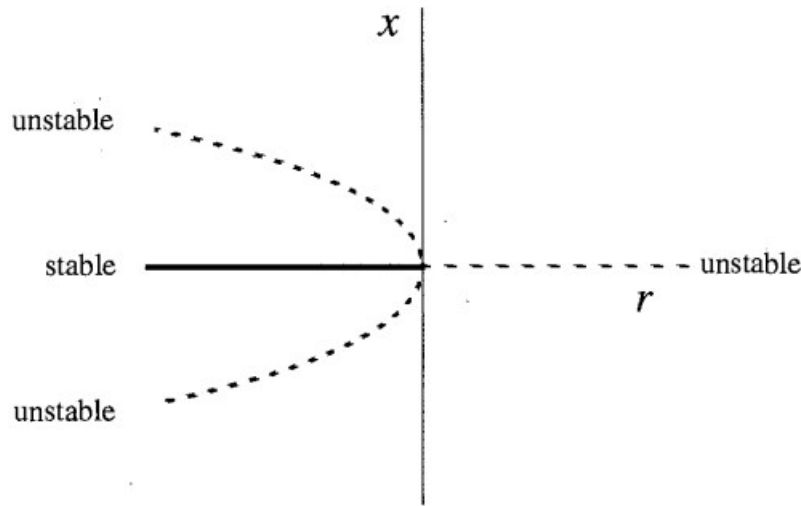

Figure 3.4.2

11

Subcritical Pitchfork Bifurcation

In the supercritical case $\dot{x} = rx - x^3$ discussed above, the cubic term is stabilizing: it acts as a restoring force that pulls x(t) back toward x=0. If instead the cubic term were destabilizing as in

$$\dot{x} = rx + x^3$$

then we'd have a subcritical pitchfork bifurcation. Figure 3.4.6 shows the bifurcation diagram

Figure 3.4.6

Compared to Figure 3.4.2, the pitchfork is invented. The nonzero fixed points $x^* = \pm\sqrt{-r}$ are unstable and exist only below the bifurcation ($r < 0$), which motivates the term "subcritical". More importantly, the origin is stable for $r < 0$ and unstable for $r > 0$, as in the supercritical case, but now the instability for $r > 0$ is not opposed by the cubic term–in fact the cubic term lends a helping hand in driving the trajectories out to infinity! This effect leads to blow-up: one can show that $x(t) \rightarrow \pm\infty$ in finite time, starting from any initial condition $x_0 \neq 0$.

In real physical systems, such an explosive instability is usually opposed by the stabilizing influence of higher-order terms. Assuming that the system is still symmetric under $x \rightarrow -x$, the first stabilizing term must be x^5. Thus the canonical example of a system with a subcritical pitchfork bifurcation is

$$\dot{x} = rx + x^3 - x^5 \qquad (3)$$

There's no loss in generality in assuming that the coefficients of x^3 and x^5 are 1.

The detailed anlysis of (3) is left to you. But we will summarize the main results here. Figure 3.4.7 shows the bifurcation diagram for (3). For small x, the picture looks just like Figure 3.4.6: the origin is locally stable for $r < 0$

and two backward-bending branches of unstable fixed points bifurcate from the origin when r=0. The new feature, due to the x^5 term, is that the unstable branches turn around and become stable at $r = r_s$, where $r_s < 0$. These stable large-amplitude branches exist for all $r > r_s$
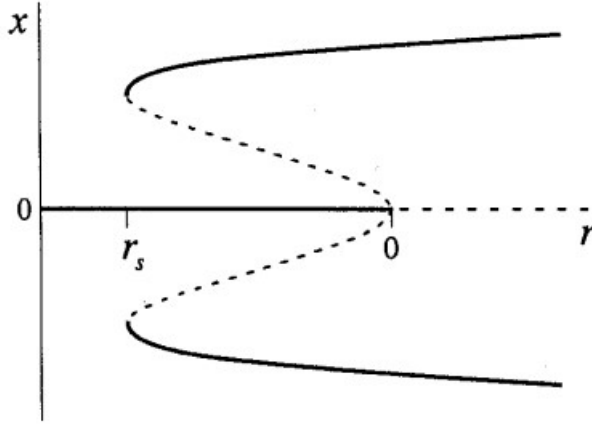
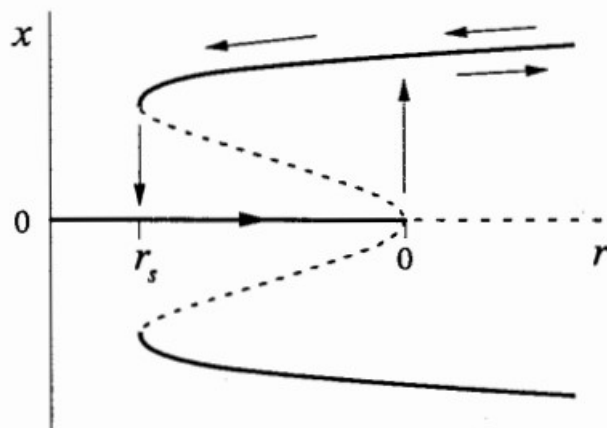**Figure 3.4.7**

There are several things to note about Figure 3.4.7   1. In the range $r_s < r < 0$, two qualitatively different stable states coexist namely the origin and the large-amplitude fixed points. The initial condition x_0 determines which fixed point is approached as $t \to \infty$. One consequence is that the origin is stable to small pertubations, but not to large ones– in this sense the origin is locally stable, but not globally stable.

2. The existence of different stable states allows for the possibility of jumps and hysteresis as r is varied. Suppose we start the system in the state $x^* = 0$, and then slowly increase the parameter r (indicated by an arrow along the r-axis of Figure 3.4.8)

13

Figure 3.4.8

Then the state remains at the origin until r=0, when the origin loses stability. Now the slightest nudge will cause the state to jump to one of the large-amplitude branches. With further increases of r, the state moves out along the large-amplitude branch. If r is now decreased, the state remains on the large-amplitude branch, even when r is decreased state to jump back to the origin. This lack of reversibility as a parameter is varied is called hysteresis.

3. The bifurcation at r_s is a saddle node bifurcation, in which stable and unstable fixed points are born "out the clear blue sky" as r is increased. **Terminology** As usual in bifurcation theory, there are several other names for the bifurcations discussed here. The supercritical pitchfork is sometimes called a forward bifurcation, and is closely related to a continuous or second-order phase transition in statistical mechanics. The subcritical is sometimes called an inverted or backward bifurcation, and is related to discontinuous or first-order phase transitions. In the engineering literature, the supercritical bifurcation is sometimes called soft or safe, because the nonzero fixed points are born at small amplitude; in contrast the subcritical bifurcation is hard or dangerous, because of the jump from zero to large amplitude.

Chapter 3.5 Overdamped Bead on a Rotating Hoop In this section we analyze a classic problem from first-year physics, the bead on a rotating hoop. This problem provides an example of a bifurcation in a mechanical system. It also illustrates the subtleties involved in replacing Newton's law, which is a second-order equation, by a simpler first-order equation.

The mechanical system is shown in Figure 3.5.1. A bead of mass m slides along a wire hoop of radius r. The hoop is constrained to rotate at a constant angular velocity ω about its vertical axis. The problem is to analyze the motion of the bead, given that it is acted on by both gravitational and centrifugal forces. This is the usual statement of the problem, but now we want to add a new twist: suppose that theres also a frictional force on the bead that opposes its motion. To be specific, imagine that the whole system is immersed in a vat of

14

molasses or some other very viscous fluid, and that the friction is due to viscous damping.
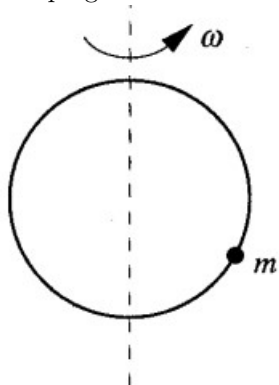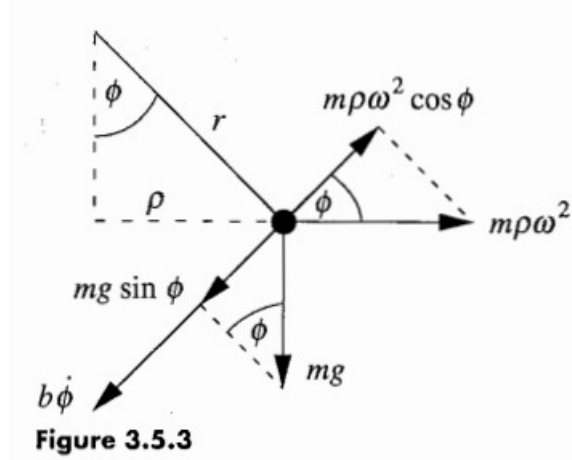

Figure 3.5.1

Let ϕ be the angle between the bead and the downward vertical direction. By convention, we restrict ϕ to the range $-\pi < \phi \leq \pi$, so there's only one angle for each point on the hoop. Also, let $p = r\sin\phi$ denote the distance of the bead from the vertical axis. Then the coordinates are as shown in Figure 3.5.2.
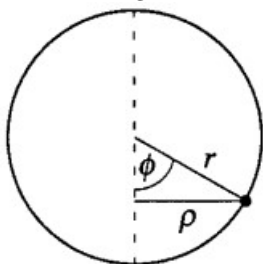
Figure 3.5.2

Now we write Newton's law for the bead. There's a downward gravitational force mg, a sideways centrifugal force $mp\omega^2$, and a tangential damping force $b\dot{\phi}$ (The constants g and b are taken to be positive; negative signs will be added later as needed). The hoop is assumed to be rigid so we only have to resolve the forces along the tangential direction, as have to resolve the forces along the tangential direction as shown in Figure 3.5.3. After substituting $p = r\sin\phi$ in the centrifugal term, and recalling that the tangential acceleration is $r\ddot{\phi}$ we obtain the governing equation

$$mr\ddot{\phi} = -b\dot{\phi} - mg\sin\phi + mr\omega^2\sin\phi\cos\phi$$

15

Figure 3.5.3

This is a second-order differential equation, since the second derivative $\ddot{\phi}$ is the highest one that appears. We are not yet equipped to analyze second-order equations, so we would like to find some conditions under which we can safely neglect the $mr\ddot{\phi}$ term. Then (1) reduces to a first-order equation, and we can apply our machinery to it.

Of course, this is a dicey business: we can't just neglect terms because we feel like it! But we will for now, and then at the end of this section we'll try to find a regime where our approximation is valid.

Analysis of the First-order system

Our concern now is with the first-order system

$$b\dot{\phi} = -mgsin\phi + mr\omega^2 sin\phi cos\phi = mgsin\phi(\tfrac{r\omega^2}{g}cos\phi - 1)$$

The fixed points of (2) correspond to equilibrium positions for the bead. What's your intuition about where such equilibria can occur? We would expect the bead to remain at rest if placed at the top or the bottom of the hoop. Can other fixed points occur? And what about stability? Is the bottom always stable?

Equation (2) shows that there are always fixed points where $sin\phi = 0$, namely $\phi^* = 0$ (the bottom of the hoop) and $\phi^* = \pi$ (the top). The more interesting result is that there are two additional fixed points if

$$\tfrac{r\omega^2}{g} > 1$$

that is, if the hoop is spinning fast enough. These fixed points satisfy $\phi^* = \pm cos^{-1}(g/r\omega^2)$. To visualize them, we introduce a parameter
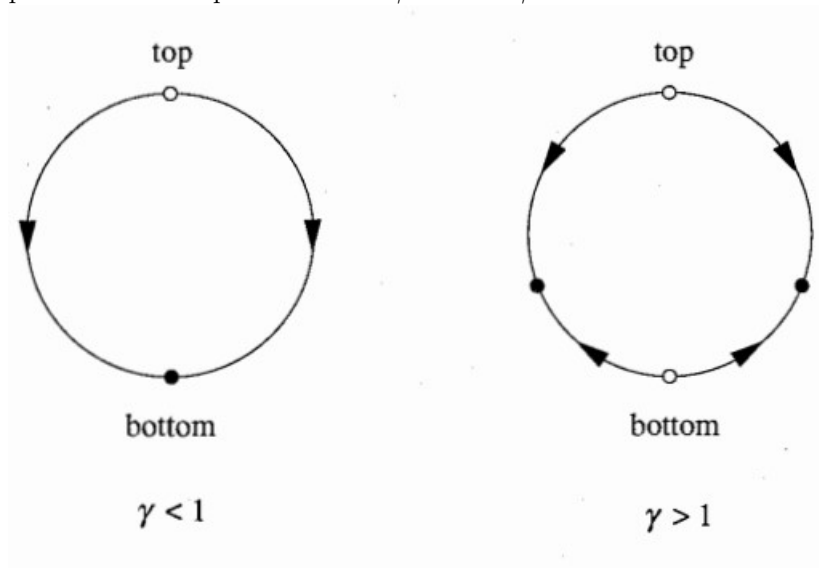
$$\gamma = \tfrac{r\omega^2}{g}$$

and solve $cos\phi^* = 1/\gamma$ graphically. We plot $cos\phi$ vs ϕ, and look for intersections with the constant function $1/\gamma$, shown as a horizontal line in fugre 3.5.4. For $\gamma < 1$ there are no intersections, whereas for $\gamma > 1$ there is a symmetrical pair of intersections to either side of $\phi^* = 0$.
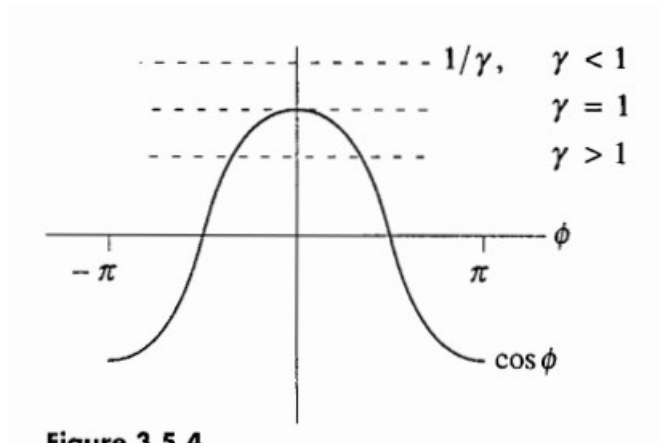
Figure 3.5.4
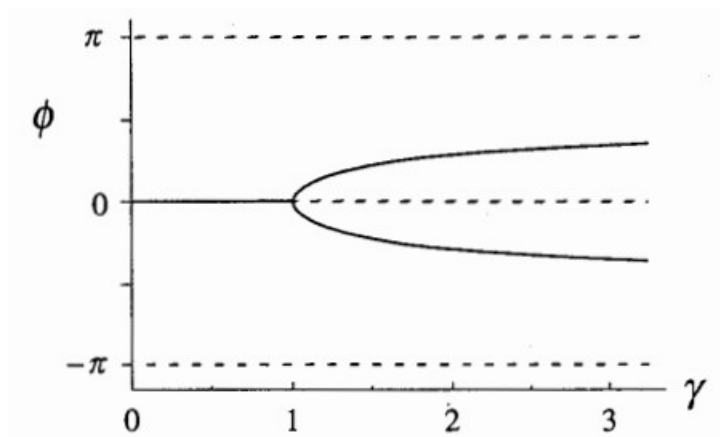
As $y \to \infty$ these intersections approach $\pm\pi/2$. Figure 3.5.5 plots the fixed points on the hoop for the cases $\gamma < 1$ and $\gamma > 1$

Figure 3.5.5

To sumarize our results so far, let's plot all the fixed points as a function of the parameter γ (Figure 3.5.6). As usual, solid lines denote stable fixed points and broken lines denote unstable fixed points.

17

Figure 3.5.6

We now see that a supercritical pitchfork bifurcation occurs at $\gamma = 1$. It's left to you to check the stability of the fixed points, using linear stability analysis or graphical methods.

Here's the physical interpretation of the results: When $\gamma < 1$, the hoop is rotating slowly and the centrifugal force is too weak to balance the force of gravity. Thus the bead slides down to the bottom and stays there. But if $\gamma > 1$ the hoop is spinning fast enough that the bottom becomes unstable. Since the centrifugal force grows as the bead moves farther from the bottom, any slight displacement of the grows as the bead moves farther from the bottom, any slight displacement of the bead will be amplified. The bead is therefore pushed up the hoop until gravity balances the centrifugal force: this balance occurs at $\phi^* = \pm cos^{-1}(g/r\omega^2)$. Which of these two fixed points is actually selected depends on the initial disturbance. Even though the two fixed points are entirely symmetrical, an assymmetry in the initial though the two fixed points are entirely symmetrical, an assymmetry in the inital conditions will lead to one of them being chose-physicist sometimes refer to these as symmetry-broken solutions. In other words, the solution has less symmetry than the governing equation.

What is the symmetry of the governing equation? Clearly the left and right halves of the hoop are physically equivalent-this is reflected by the invariance of (1) and (2) under the change of variables $\phi \rightarrow -\phi$. As we mentioned in Section 3.4 pitchfork bifurcations are to be expected in situations where such a symmetry exists. **Dimensional Analysis and Scaling** Now we need to address the question: When is it valid to neglect the inertia term $mr\ddot{\phi}$ in (1)? At first sight the limit $m \rightarrow 0$ looks promising, but then we notice that we're throwing out the baby with the bathwater: the centrifugal and gravitational terms vanish in this limit too! So we have to be more careful.

In problems like this, it is helpful to express the equation in dimensionless form (at present all the terms in (1) have the dimensions of force) The advantage of a dimensionless formulation is that we know how to define small-it

means "much less than 1". Furthermore, nondimensionalizing the equation reduces the number of parameters by lumping them together into dimensionless groups. The reductionalways simplifies the analysis. For an excellent introduction to dimensional analysis, see Lin and Segel (1988).

There are often several ways to nondimensionalize an equation, and the best choice might not be clear at first. Therefore we proceed in a flexible fashion. We define a dimensionless time τ by

$$\tau = \frac{t}{T}$$

where T is a characteristic time scale to be chosen later. When T is chosen correctly, the new derivatives $d\phi/d\tau$ and $d^2\phi/d\tau^2$ should be O(1), i.e. of order unity. To express these new derivatives in terms of the old ones, we use the chain rule:

$$\dot{\phi} = \frac{\phi}{dt} = \frac{d\phi}{d\tau}\frac{d\tau}{dt} = \frac{1}{T}\frac{d\phi}{d\tau}$$

and similarly

$$\ddot{\phi} = \frac{1}{T^2}\frac{d^2\phi}{d\tau^2}$$

(The easy way to remember these formulas is to formally substitute $T\tau$ for t). Hence (1) becomes

$$\frac{mr}{T^2}\frac{d^2\phi}{d\tau^2} = -\frac{b}{T}\frac{d\phi}{d\tau} - mgsin\phi + mr\omega^2 sin\phi cos\phi$$

Now since this equation is a balance of forces, we nondimensionalize it by dividing by a force mg. This yields the dimensionless equation

$$(\frac{r}{gT^2}\frac{d^2\phi}{d\tau^2} = -(\frac{b}{mgT})\frac{d\phi}{d\tau} - sin\phi + (\frac{r\omega^2}{g})sin\phi cos\phi$$

Each of the terms in parantheses is a dimensionless group. We recognize the group $r\omega^2/g$ in the last term-that's our old friend γ from earlier in the section.

We are interested in the regime where the left-hand side of (3) is negligible compared to all the other terms, and where all the terms on the right hand side are of comparable size. Since the derivatives are O(1) by assumption, and $sin\phi = O(1)$, we see that we need

$$\frac{b}{mgT} \approx O(1), \text{ and } \frac{r}{gT^2} << 1$$

The first of these requirements sets the time scale T: a natural choice is

$$T = \frac{b}{mg}.$$

Then the condition $r/gT^2 << 1$ becomes

$$\frac{r}{g}(\frac{mg}{b})^2 << 1$$

or equivalently

$$b^2 >> m^2 gr$$

19

This can be interpreted as saying that the damping is very strong, or that the mass is very small, now in a precise sense.

The condition (4) motivates us to introduce a dimensionless group

$$\epsilon = \frac{m^2 gr}{b^2}. \qquad (5)$$

Then (3) becomes

$$\epsilon \frac{d^2\phi}{d\tau^2} = -\frac{d\phi}{d\tau} - sin\phi + \gamma sin\phi cos\phi \qquad (6)$$

As advertised, the dimensionless Equation (6) is simpler than (1): the five parameters m, g, r ω and b have been replaced by two dimensionless groups γ and ϵ.

In summary, our dimensional analysis suggests that in the overdamped limit $\epsilon \to 0$, (6) should be well approximated by the first-order system

$$\frac{d\phi}{d\tau} = f(\phi) \qquad (7)$$

where

$$f(\phi) = -sin\phi + \gamma sin\phi cos\phi = sin\phi(\gamma cos\phi - 1) \qquad (7)$$

A paradox

Unfortunately, there is something fundamentally wrong with our idea of replacing a second-order equation by a first-order equation. The trouble is that a second order equation requires two initial conditions, whereas a first-order equation has only one. In our case, the beads motion is determined by its initial position and velocity. These two quantites can be chosen completely independent of each other. But that's not true for the first-order system: given the initial position, the initial velocity is dictated by the equation $d\phi/d\tau = f(\phi)$. Thus the solution to the first-order system will not, in general, be able to satisfy both initial conditions.

We seem to have run into a paradox. Is (7) valid in the overdamped limit or not? If it is valid, how can we satisfy the two arbitrary initial conditions demanded by (6)?

The resolution of the paradox requires us to analyze the second-order system (6). We haven't dealt with second-order systems before-thats the subject of Chapter 5. But continue reading if your interested.

Phase Plane Analysis   Throughout Chapter 2 and 3, we have exploited the idea that's a first-order system $\dot{x} = f(x)$ can be regarded as a vector field on a line. By analogy, the second-order system (6) can be regarded as a vector field on a plane, the so-called phase plane.

The plane is spanned by two axes, one for the angle ϕ and one for the angular velocity $d\phi/d\tau$. To simplify the notation, let

$$\Omega = \phi^{'} = d\phi/d\tau$$

where prime denotes differentiation with respect to τ. Then an initial condition for (6) corresponds to a point (ϕ_0, Ω_0) in the phase plane (Figure 3.5.7). As

time evolves the phase point $(\phi(t), \Omega(t))$ moves around in the phase plane along a trajectory determined by the solution to (6)
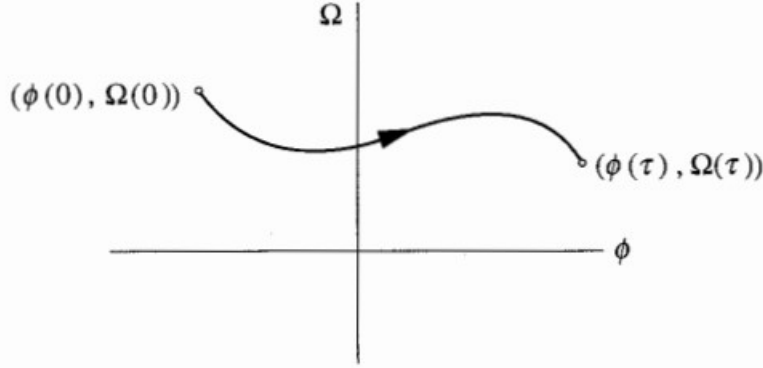
Figure 3.5.7

Our goal now is to see what those trajectories actually look like. As before, the key idea is that the differential equation can be interpreted as a vector field on the phase space. To convert (6) into a vector field, we first rewrite it as

$$\epsilon\Omega' = f(\phi) - \Omega$$

Along with the definition $\phi' = \Omega$, this yields the vector field

$$\phi' = \Omega$$

$$\Omega' = \tfrac{1}{\epsilon}(f(\phi) - \Omega)$$

We interpret the vector (ϕ', Ω') at the point (ϕ, Ω) as the local velocity of a phase fluid flowing steadily on the plane. Note that the velocity vector now has two components, one in the ϕ-direction and one in the Ω-direction. To visualize the trajectories, we just imagine how the phase point would move as it is carried along by the phase fluid.

In general, the pattern of trajectories would be difficult in picture, but the present case is simple because we are only interested in the limit $\epsilon \to 0$. In this limit, all trajectories slam straight up or down onto the curve C defined by $f(\phi) = \Omega$ and then slowly ooze along this curve until they reach a fixed point (Figure 3.5.8).

21
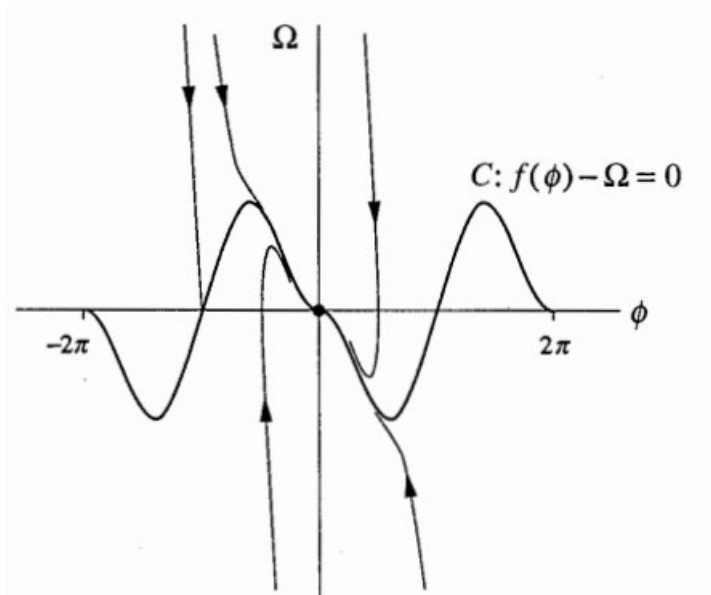
Figure 3.5.8

To arrive at this striking conclusion, let's do an order of magnitude calculation. Suppose that the phase point lies off the curve C. For instance, suppose (ϕ, Ω) lies an O(1) distance below the curve C, i.e. $\Omega < f(\phi)$ and $f(\phi) - \Omega \approx O(1)$. Then (8b) shows that Ω' is enormously positive: $\Omega' \approx O(1/\epsilon) >> 1$. Thus the phase point zaps like lightning up to the region where $f(\phi) - \Omega \approx O(\epsilon)$. In the limit $\epsilon \to 0$, this region is indistinguishable from C. Once the phase point is on C, it evolves according to $\Omega = f(\phi)$: that is, it approximately satisfies the first-order equation $\phi' = f(\phi)$

Our conclusion is that a typical trajectory is made of two parts: a rapid initial transient during which the phase point zaps onto the curve where $\phi' = f(\phi)$, followed by a much slower drift along this curve.

Now we see how the paradox is resolved: The second-order system (6) does behave like the first-order system (7) but only after a rapid initial transient. During this transient, it is not correct to neglect the term $\epsilon d^2\phi/d\tau^2$. The problem with our earlier approach is that we used only a single time scale T=b/mg: this time scale is characteristic of the slow drift process, but not of the rapid transient.

A singular limit

The difficulty we have encountered here occurs throughout science and engineering. In some limit of interest (here, the limit of strong damping), the term containing the highest order derivative drops out of the governing equation. Then the initial conditions or boundary conditions can't be satisfied. Such a limit is often called singular. For example, in fluid mechanics, the limit of high Reynolds number is a singular limit: it accounts for the presence of extremely

22

thin "boundary layers" in the flow over airplane wings. In our problem, the rapid transient played the role of a boundary layer - it is a thin layer of time that occurs near the boundary t=0.

The branch of mathematics that deals with singular limits is called singular pertubation theory. See Jordan and Smith (1987) or Lin and Segel (1988) for an introduction. **Chapter 3.6 Imperfect Bifurcations and Catastrophes** As we mentioned earlier, pitchfork bifurcations are common in problems that have a symmetry. For example, in the problem of the bead on a rotating hoop (Section 3.5), there was a perfect symmetry between the left and right sides of the hoop. But in many real world circumstances, the symmetry is only approximate-an imperfection leads to a slight difference between left and right. We now want to see what happens when such imperfections are present.

For example, consider the system

$$\dot{x} = h + rx - x^3$$

If h=0, we have the normal form for a supercritical pitchfork bifurcation, and there's a perfect symmetry between x and -x. But this symmetry is broken when $h \neq 0$; for this reason we refer to h as an imperfection parameter.

Equation (1) is a bit harder we have two independent parameters to worry about (h and r). To keep things straight, we'll think of r as fixed and then examine the effects of varying h. The first step is to analyze the fixed points of (1). These can be found explicitly, but we'd have to invoke the messy formula for the roots of a cubic equation. It's clearer to use a graphical appraoch, as in Example 3.1.2. We plot the graphs of $y = rx - x^3$ and y=-h on the same axes and look for intersection (Figure 3.6.1). These intersections occur at the fixed points of (1). When $r \leq 0$, the cubic is monotonically decreasing, and so it intersects the horizontal line y=-h in exactly one point (Figure 3.6.1a). The more intersecting case is $r > 0$; then one, two or three intersections are possible, depending on the value of h (Figure 3.6.1b).

(a) $r \le 0$ (b) $r > 0$
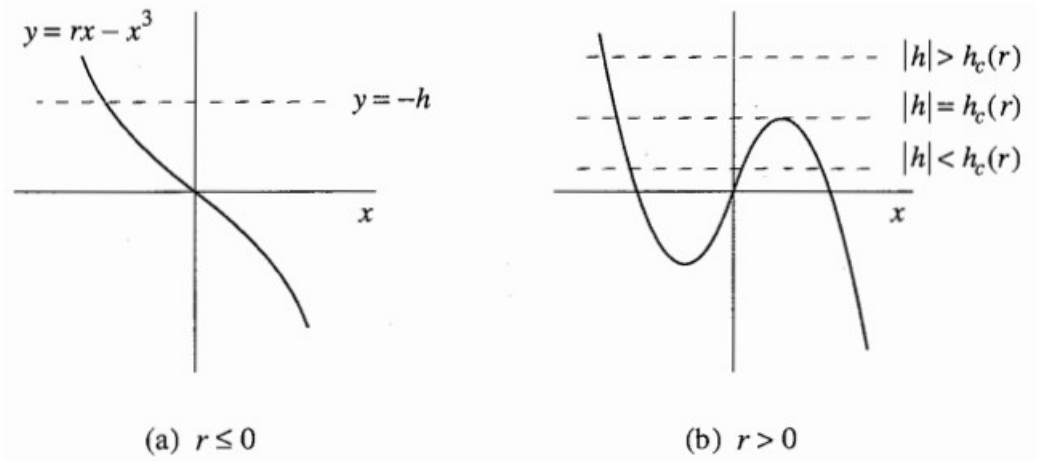
Figure 3.6.1

The critical case occurs when the horizontal line is just tangent to either the local minimum or maximum of the cubic; then we have a saddle node bifurcation. To find the values of h at which this bifurcation occurs, note that the cubic has a local maximum when $\frac{d}{dx}(rx - x^3) = r - 3x^2 = 0$. Hence

$$x_{max} = \sqrt{\tfrac{r}{3}},$$

and the value of the cubic at the local maximum is $rx_{max} - (x_{max})^3 = \frac{2r}{3}\sqrt{\frac{r}{3}}$ Similarly, the value at the minimum is the negative of this quantity. Hence saddle-node bifurcations occur when $h = \pm h_c(r)$, where

$$h_c(r) = \tfrac{2r}{3}\sqrt{\tfrac{r}{3}}$$

Equation (1) has three fixed points for $|h| < h_c(r)$ and one fixed point for $|h| > h_c(r)$. To summarize the results so far, we plot the bifurcation curves $h = \pm h_c(r)$ in the (r, h) plane (Figure 3.6.2). Note that the two bifurcation curves meet tangentially at (r,h)=(0,0); such a point is called a cusp point. We also label the regions that correspond to different numbers of fixed points. Saddle-node bifurcation occur all along the boundary of the regions, except at the cusp point, where we have a codimension-2 bifurcation. (This fancy terminology essentially means that we have had to tune two parameters, h and r, to achieve this type of bifurcation. Until now, all our bifurcation could be achieved by tuning a single parameter, and were therefore codimension-1 bifurcations).
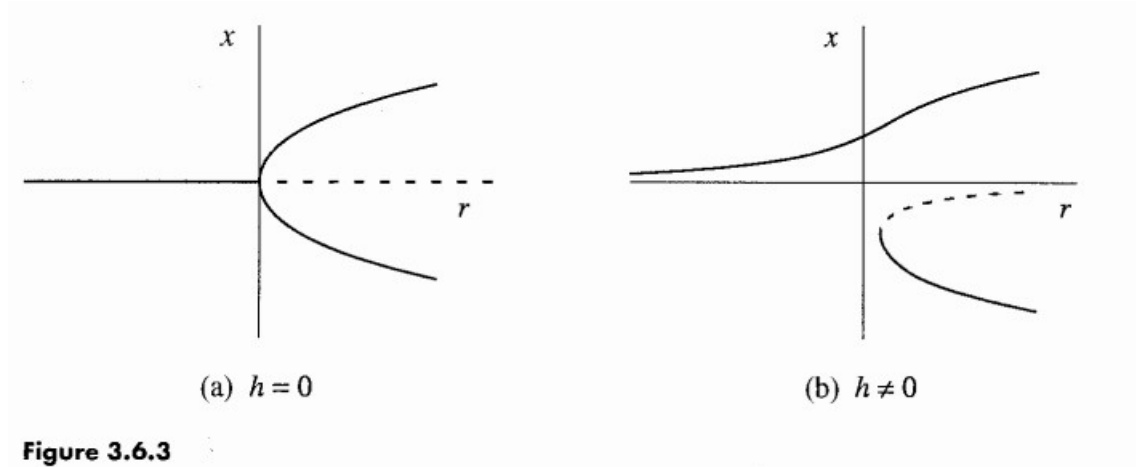
24
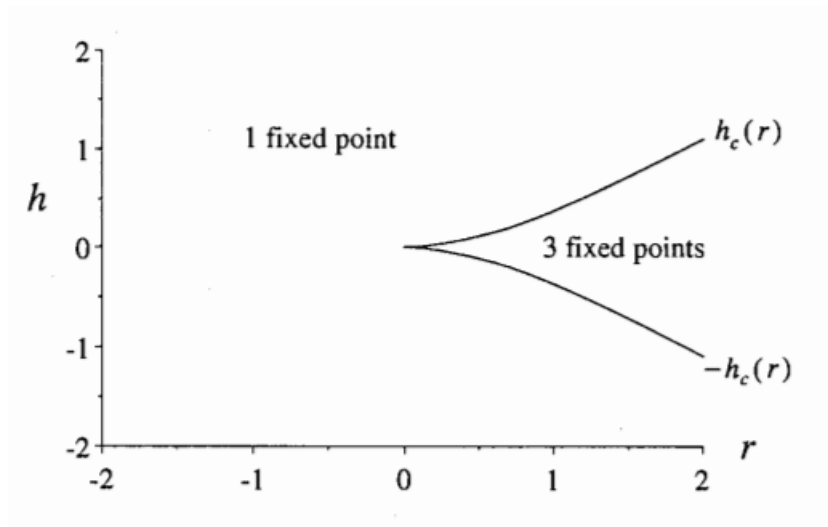
Figure 3.6.2

Pictures like Figure 3.6.2 will prove very useful in our future work. We will refer to such pictures as stability diagrams. They show the different types of behavior that occur as we move around in parameter space (here, the (r,h) plane).

Now let's present our results in a more familiar way by showing the bifurcation diagram of x^* vs r, for fixed h (Figure 3.6.3)

(a) $h = 0$ (b) $h \neq 0$

Figure 3.6.3

When h=0 we have the usual pitchfork diagram (Figure 3.6.3a). The upper piece consists entirely of stable fixed points, whereas the lower piece has both stable and unstable branches. As we increase r from negative values, there's no longer a sharp transition at r=0; the fixed point simply glides smoothly along the upper branch. Furthermore, the lower branch of stable points is not accessible unless we make a fairly large disturbance.

Alternatively, we could plot x^* vs. h, for fixed r (Figure 3.6.4)

(a) $r \leq 0$ **(b)** $r > 0$
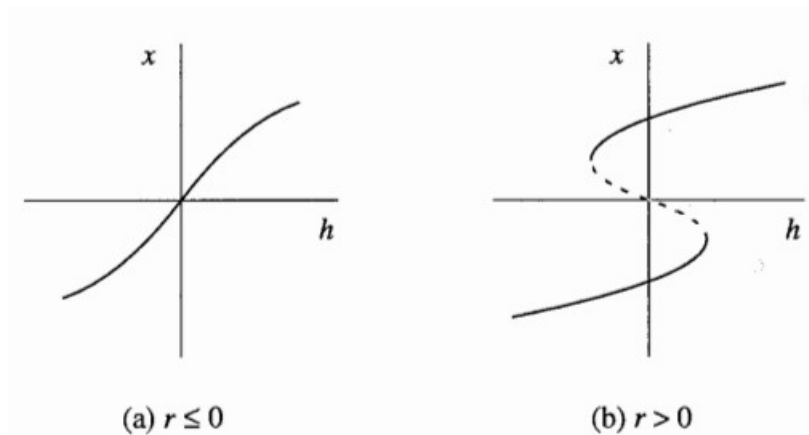
Figure 3.6.4

When $r \leq 0$ there's one stable fixed point for each h (Figure 3.6.4a). However when $r > 0$ there are three fixed points when $|h| < h_c(r)$ and one otherwise (Figure 3.6.4b). In the triple-valued region, the middle branch is unstable and the upper and lower branches are stable. Note that these graphs look like Figure 3.6.1 rotated by 90°.

There is one last way to plot the results, which may appeal to you if you like to picture things in three dimensions. This method of presentation contains all of the others as cross section or projections. If we plot the fixed points x^* above the (r,h) plane, we get the cusp catastrophe surface shown in Figure 3.6.5. The surface folds over on itself in certain places. The projection of these folds onto the (r,h) plane yields the bifurcation curves shown in Figure 3.6.2. A cross section at fixed h yields Figure 3.6.3 and a cross section at fixed r yields Figure 3.6.4.

The term catastrophe is motivated by the fact that as parameters change the state of the system can be carried over the edge of the upper surface, after which it drops discontinuously to the lower surface (Figure 3.6.6) This jump could be truly catastrophic for the equilibrium of a bridge or a building. We will see scientific examples of catastrophes in the context of insect outbreaks (next section 3.7) and in the following example from mechanics. For more about catastrophe theory, see Zeeman (1977). Incidentally, there was a violent controversy about this subject in the late 1970s. If you like watching fights, have a look at Zahler and Sussman (1977) and Kolata (1977).

26

Figure 3.6.5
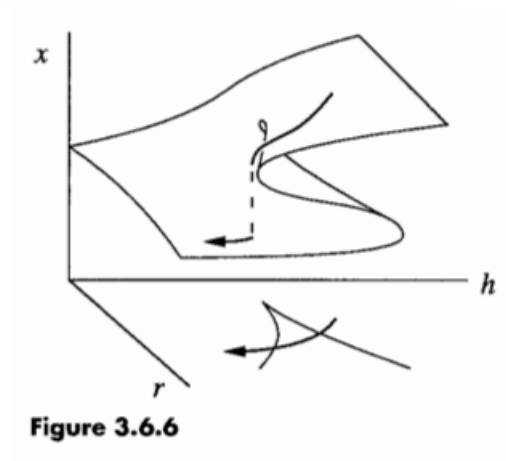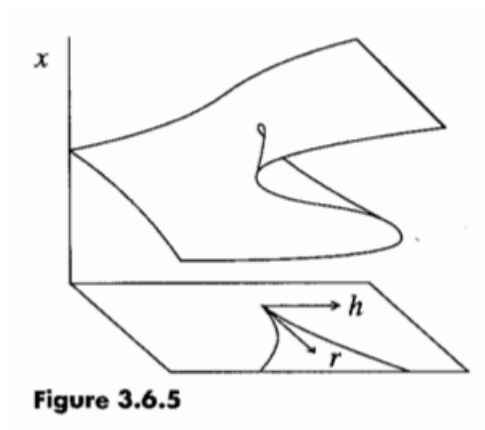
Figure 3.6.6

Bead on a Tilted Wire

As a simple example of imperfect bifurcation and catastrophe, consider the following mechanical system (Figure 3.6.7)
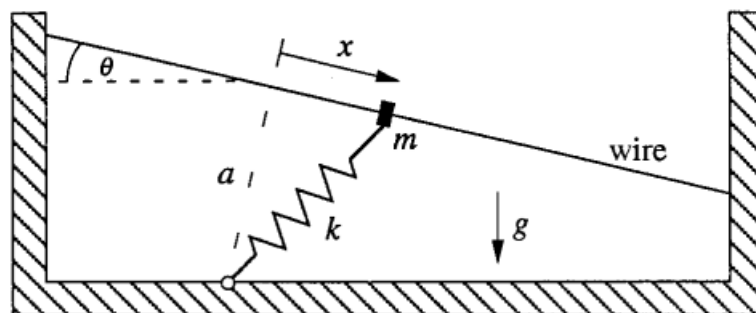
Figure 3.6.7

27

A bead of mass m is constrained to slide along a straight wire inclined at an angle θ with respect to the horizontal. The mass is attached to a spring of stiffness k and relaxed length L_0 and is also acted on by gravity. We choose coordinates along the wire so that x=0 occurs at the point closest to the support point of the spring: let a be the distance between this support point and the wire.

But first let's get some physical intuition. When the wire is horizontal ($\theta = 0$), there is perfect symmetry between the left and right sides of the wire, and x=0 is always an equilibrium position. The stability of this equilibrium depends on the relative sizes of L_0 and a: if $L_0 < a$, the spring is in tension and so the equilibrium should be stable. But if $L_0 > a$, the spring is compressed and so we expect an unstable equilibrium x=0 and a pair of stable equilibria to either side of it.

The problem becomes more interesting when we tilt the wire ($\theta \neq 0$) For small tilting, we expect that there are still three equilibria if $L_0 > a$. However if the tilt becomes too steep perhaps you can see intuitively that the uphill equilibrium might suddenly disappear, causing the bead to jup catastrophically to the downhill equilibrium. You might even want to build this mechanical system and try it.

Chapter 3.7 Insect outbreak For a biological example of bifurcation and catastrophe, we turn now to a model for the sudden outbreak of an insect called the spruce budworm. This insect is a serious pest in eastern Canada, where it attacks the leaves of the balsam fir tree. When an outbreak occurs, the budworms can defoliate and kill most of the fir trees in the forest in about four years.

Ludwig et al. (1978) proposed and analyzed an elegant model of the interaction between budworms and the forest. They simplified the problem by exploiting a separation of time scales: the budworm population evolves on a fast time scale (they can increase their density fivefold in a year, so they have a characteristic time scale of moths), whereas the trees grow and die on a slow time scale (they can completely replace their foliage in about 7-10 years, and their life span in the absence of budworms is 100-150 years). Thus as far as the budworm dynamics are concerned, the forest variables may be treated as constants. At the end of the anlysis, we will allow the forest variables may be treated as constants. At the end of the analysis, we will allow the forest variables to drift very slowly-this drift ultimately triggers an outbreak. **Model** The proposed model for the budworm population dynamic is

$$\dot{N} = RN(1 - \tfrac{N}{K}) - p(N)$$

In the absence of predators, the budworm population N(t) is assumed to grow logistically with growth rate R and carrying capacity K. The carrying capacity depends on the amount of foliage left on the trees and so it is a slowly drifting parameter; at this stage we treat it as fixed. The term p(N) represents the death rate due to predation, chiefly by birds, and is assumed to have the

shape shown in Figure 3.7.1. There is almost no predation when budworms are scarce; the birds seek food elsewhere. However, once the population exceeds a certain critical level N=A, the predation turns on sharply and then saturates (the birds are eating as fast as they can). Ludwig et al. (1978) assumed the specific form

$$p(N) = \frac{BN^2}{A^2+N^2}$$

where $A, B > 0$. Thus the full model is

$$\dot{N} = RN(1 - \tfrac{N}{K}) - \frac{BN^2}{A^2+N^2}$$

We now have to clarify some things. What do we mean by an "outbreak" in the context of this model? The idea must be that, as parameters drift, the budworm population suddenly jumps from a low to a high level. But what do we mean by "low" and "high" and are there solutions with this character? To answer these questions, it is convenient to recast the model into a dimensionless form, as in Section 3.5

Dimensionless Formulation

The model (1) has four parameters: R, K, A and B. As usual there are various ways to nondimensionalize the system. For example, both A and K have the same dimension as N, and so either N/A or N/K could serve as a dimensionless population level. It often takes some trial and error to find the best choice. In this choice our heuristic will be to scale the equation so that all the dimensionless groups are pushed into the logistic part of the dynamics, with none of the predation part. This turns out to ease the graphical analysis of the fixed points.

To get rid of the parameters in the predation term, we divide (1) by B and then let

$$x = N/A$$

which yields

$$\tfrac{A}{B}\tfrac{dx}{dt} = \tfrac{R}{B}Ax(1 - \tfrac{Ax}{K}) - \tfrac{x^2}{1+x^2} \qquad (2)$$

Equation (2) suggests that we should introduce a dimensionless time τ and dimensionless groups r and k, as follows:

$$\tau = \tfrac{Bt}{A}, \qquad r = \tfrac{RA}{B}, \qquad k = \tfrac{K}{A}$$

Then (2) becomes

$$\tfrac{dx}{dt} = rx(1 - \tfrac{x}{k}) - \tfrac{x^2}{1+x^2} \qquad (3)$$

which is our final dimensionless form. Here r and k are the dimensionless growth rate and carrying capacity, respectively.

Analysis of Fixed points

Equation (3) has a fixed point at $x^* = 0$ it is always unstable. The intuitive explanation is that the predation is extremely weak for small x, and

so the budworm population grows exponentially for x near zero.
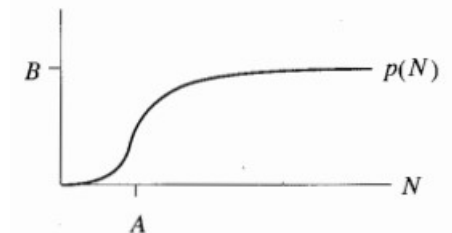
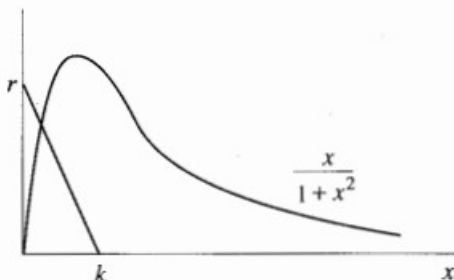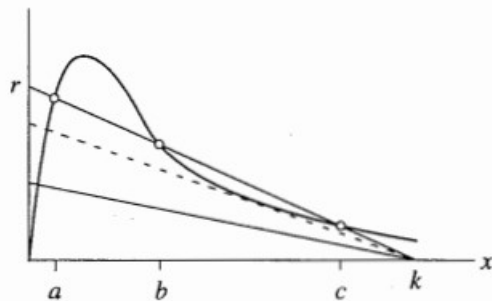**Figure 3.7.1**

The other fixed points of (3) are given by the solutions of

$$r(1 - \tfrac{x}{k}) = \tfrac{x}{1+x^2} \qquad (4)$$

This equation is easy to analyze graphically-we simply graph the right and left hand sides of (4), and look for intersections (Figure 3.7.2). The left hand side of (4) represents a straight line with x-intercept equal to r, and the right-hand side represents a curve that is independent of the parameters! Hence as we vary the parameters r and k, the line moves but the curve doesn't-this convenient property is what motivated our choice of nondimensionalization.

Figure 3.7.2

Figure 3.7.3

Figure 3.7.2 shows that if k is suficiently small, there is exactly one intersection for any $r > 0$. However for large k, we can have one, two or three intersections, depending on the value of r (Figure 3.7.3). Let's suppose that there are three intersections a, b and c. As we decrease r with k fixed, the line rotates counterclockwise about k. Then the fixed points b and c approach each other and eventually coalesce in a saddle-node bifurcation when the line intersects the curve tangentially (dashed line in Figure 3.7.3). After the bifurcation, the only remaining fixed point is a (in addition to $x^* = 0$ is unstable, and also observe that the stability type must alternate as we move along the x-axis. Hence a is stable, b is unstable, and c is stable. Thus, for r and k in the range corresponding to three positive fixed points, the vector field is qualitatively like that shown in Figure 3.7.4. The smaller stable fixed point a is called the refuge level of the budworm population, while the larger stable point c is the outbreak level. From the point of view of pest control, one would like to

30

keep the population at a and away from c. The fate of the system is determined by the initial condition x_0: an outbreak occurs if and only if $x_0 > b$. In this sense the unstable equilibrium b plays the role of a threshold.

An outbreak can also be triggered by a saddle node bifurcation. If the parameters r and k drift in such a way that the fixed point a disappears, then the population will jump suddenly to the outbreak level c. The situation is made worse by the hysteresis effect-even if the parameters are restored to their values before the outbreak, the populatoin will not drop back to the refuge level.

Calculating the Bifurcation Curves

Now we compute the curves in (k,r) space where the system undergoes saddle-node bifurcations. The calculation is somewhat harder than that in section 3.6: we will not be able to write r explicitly as a function of k, for example. Instead the bifurcation curves will be written in the parameric form (k(x), r(x)), where x runs through all positive values. (Please dont be confused by this traditional terminology-one would call x the "parameter" in these parametric equations, even though r and k are themselves parameters in a different sense).

As discussed earlier the condition for a saddle-node bifurcation is that the line $r(1 - x/k)$ intersects the curve $x/(1 + x^2)$ tangentially. Thus we require both

$$r(1 - \tfrac{x}{k}) = \tfrac{x}{1+x^2} \qquad (5)$$

and

$$\tfrac{d}{dx}[r(1 - \tfrac{x}{k})] = \tfrac{d}{dx}[\tfrac{x}{1+x^2}] \qquad (6)$$

After differentation (6) reduces to

$$-\tfrac{r}{k} = \tfrac{1-x^2}{(1+x^2)^2} \qquad (7)$$

We substitute this expression for r/k into (5), which allows us to express r solely in terms of x. The result is

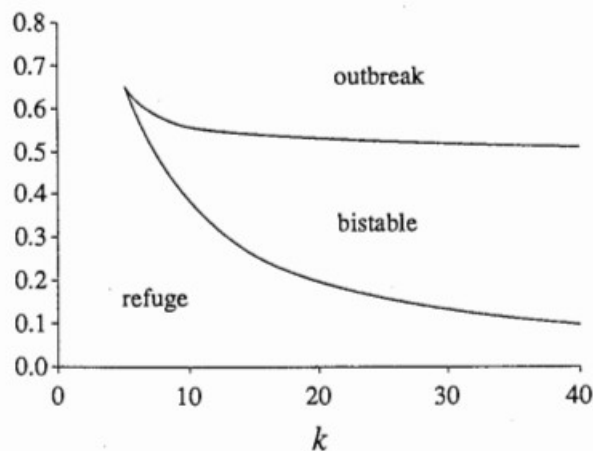$$r = \tfrac{2x^3}{(1+x^2)^2} \qquad (8)$$

Then inserting (8) into (7) yields

$$k = \tfrac{2x^3}{x^2-1} \qquad (9)$$

The condition $k > 0$ implies that x must be restricted to the range $x > 1$.

Together (8) and (9) define the bifurcation curves. For each $x > 1$, we plot the corresponding point $(k(x), r(x))$ in the (k, r) plane. The resulting curves are shown in Figure 3.7.5 (Excercise 3.7.2 deals with some of the analytical properties of these curves)

Figure 3.7.5
The different regions in Figure 3.7.5 are labeled according to the stable fixed point that exist. The refuge level a is the only stable state for low r, and the outbreak level c is the only stable state for large r. In the bistable region, both stable states exist.

The stability diagram is very similar to Figure 3.6.2. It too can be regarded as the projection of a cusp catastrophe surface, as schematically illustrated in Figure 3.7.6. You are hereby challanged to graph the surface accurately!
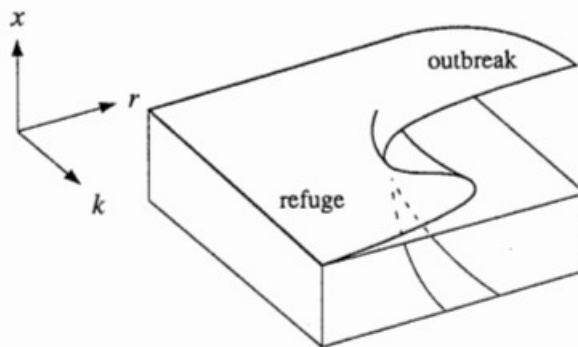
Figure 3.7.6

Comparison with Observation
Now we need to decide on biologically plausible values of the dimensionless groups $r = RA/B$ and $k = K/A$. A complication is that these parameters may drift slowly as the condition of the forrest changes. According to Ludwig et al. (1978), r increases as the forrest grows, while k remains fixed.

They reason as follows: let S denote the average size of the trees, interpreted as the total surface area of the branches in a stand. Then the carrying capacity K should be proportional to the available foilage, so $K = K'S$. Simi-

32

larly, the half saturation parameter A in the predation term should be proportional to S; predators such as birds search units of foilage, not acres of forest and so the relevant quantity A' must have the dimensions of budworms per unit of branch area. Hence $A = A'S$ and therefore

$$r = \frac{RA'}{B}S, \qquad k = \frac{K'}{A'} \qquad (10)$$

The experimental observations suggest that for a young forest, typically $k \approx 300$ and $r < 1/2$ so the parameters lie in the bistable region. The budworm population is kept down by the birds, which find it easy to search the small number of branches per acre. However, as the forest grows, S increases and therefore the point (k,r) drifts upward in parameter space toward the outbreak region of Figure 3.7.5. Ludwig et al. (1978) estimate that $r \approx 1$ for a fully mature forest, which lies dangeruously in the outbreak region. After an outbreak occurs, the fir trees die and the forest is taken over by birch trees. But they are less efficient at using nutrient and eventually the fir trees come back-this recovery takes about 50-100 years (Murray 1989).

We conclude by mention some of the approximations in the model presented here. The tree dynamics have been neglected; see Ludwig et al. (1979) and Murray (1989) for treatments of this aspect of the problem.