1

What's it all about?

Mathematical models of nature almost always involve solving equations (often differential equations) and these models, and their equations, frequently depend on external parameters. Examples of parameters might be the temperature of a chemical reaction, or the load on a bridge, or the tension in a rope (will it snap?), or the temperature of the ocean for plankton populations. In such models, it is important to understand how phenomena associated to that model can change as the parameters are varied. Usually one finds that a small change in the value of the parameters produces a corresponding small change in the (set of) solutions of the equation. But occasionally, for particular values of the parameter there is a more radical change, and such changes are called *bifurcations*. Often these bifurcations involve simply a change in the number of solutions. This chapter illustrates these ideas with a few examples.

Bifurcation theory is the (mathematical) study of such qualitative changes arising as parameters are varied. In this book, we consider a subset of this very general theory, namely *local bifurcations*, which excludes for example, routes to chaos in dynamical systems and other global bifurcations: everything we study can be described by local questions and local changes.

The majority of applications of mathematics involve differential equations (ordinary or partial), and the theory of bifurcations can be applied to these in a straightforward manner, as we will see in the first example below. However, the ideas are more general, and can be applied to other systems that depend on parameters, not just differential equations.

The general approach is to consider an equation $g(x) = 0$, where g may have several components,

$$g(x) = (g_1(x),\ g_2(x), \dots,\ g_p(x)),$$

and indeed so may x, that is $x = (x_1, x_2, \dots, x_n)$. Then introduce a parameter λ (also possibly multi–dimensional), writing

$$g_\lambda(x) = 0,$$

or $G(x;\lambda) = 0$. This is called a ***family*** of equations, depending on the parameter λ. We shall always assume our families are smooth as functions of (x, λ) (i.e. of class C^∞). The basic question of bifurcation theory is, how do solutions in x of these equations change as λ varies? And a bifurcation occurs when the change is in some sense qualitative.

There are many applications where the equation is a so-called *variational problem*, which means that the equation $g(x) = 0$ is in fact of the form $\nabla V(x) = 0$ for some scalar function V, usually called the *potential*. Then solutions of the equations $g(x) = 0$ are critical points of the function V. Zeeman's catastrophe machine described in Section 1.3 is one such physical example. A more geometric example is described in Section 1.4.

1.1 The fold or saddle-node bifurcation

The simplest mathematical example exhibiting a bifurcation is provided by the ordinary differential equation (ODE),

$$\dot{x} = x^2 + \lambda. \tag{1.1}$$

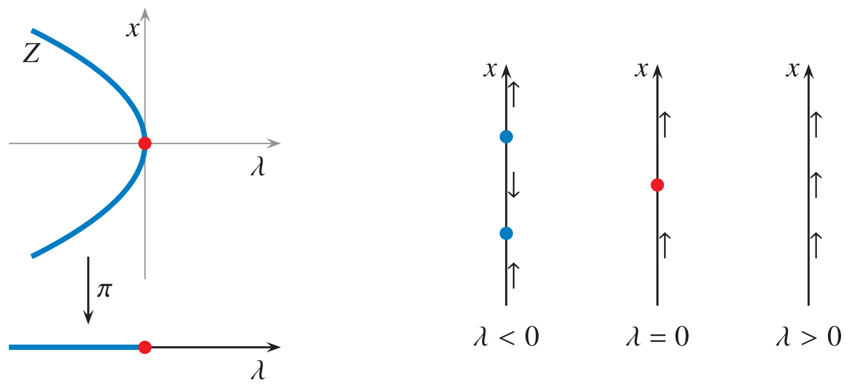
Here $\lambda \in \mathbb{R}$ is the parameter, and one often refers to $x \in \mathbb{R}$ as the ***state variable***. The dot over the x denotes the time-derivative, and a solution to the equation would be a function $x(t)$. Since this is a first order ODE, an equilibrium point occurs wherever the right-hand side vanishes. The equilibria therefore occur where

$$x^2 + \lambda = 0.$$

Define $g_\lambda(x) = x^2 + \lambda$. Then we are interested in solutions of $g_\lambda(x) = 0$, that is in the zeros of g. We call this set Z. Thus,

$$Z = \left\{(x, \lambda) \in \mathbb{R}^2 \mid x^2 + \lambda = 0\right\}.$$

The question we address is how the number of points in Z depends on λ. In this example, the curve Z is a parabola in the left half of the plane, as illustrated in Figure 1.1A. For $\lambda < 0$ there are two solutions (two equilibrium points), at $x = \pm\sqrt{-\lambda}$, and as λ increases to 0 these coalesce and then for $\lambda > 0$ there are no solutions (or they become complex, but we are just interested in real solutions). The transition, or *bifurcation*, occurs when $\lambda = 0$ (marked in red). The set of parameter values where such a bifurcation occurs is called the ***bifurcation set*** or ***discriminant***. The map π shown in the diagram is simply the projection taking $(x, \lambda) \in Z$ to the parameter value λ.



(A) The fold, or saddle-node, bifurcation diagram

(B) Phase diagram for the saddle-node bifurcation (1.1)

FIGURE 1.1 (A) shows the equilibrium points, forming a smooth curve in the (x, λ)-plane. (B) shows whether x is increasing or decreasing (the sign of $\dot{x}$) for different values of λ; the dots represent the equilibrium points and correspond to points on the curve in (A).

The behaviour of the differential equation is illustrated in Figure 1.1B. There are two equilibrium points when $\lambda < 0$ and none for $\lambda > 0$. In differential equations, this transition is often called a ***saddle-node bifurcation*** because in two dimensions, when $\lambda < 0$, one of the equilibria would be a saddle and the other a node. In singularity theory, where the specific application is not of concern, it is more generally called a ***fold bifurcation***, because of the shape of the curve Z folding over with respect to the parameter space (the λ-axis).

Remark In this simple example, the differential equation is a standard one and can be solved explicitly (by separation of variables, the type of solution depends on the sign of λ). However, more generally, bifurcation theory can be used to study equilibria (and neighbouring dynamics) of systems of ODEs where this is not the case, such as for example the ODE $\dot{x} = x^2 \mathrm{e}^x + \lambda$, which does not have a closed form solution but still exhibits a saddle-node bifurcation. ❞

The beauty of these ideas is that while the example above is so simple (g is quadratic), it contains essentially all that is expected to occur if there is just one parameter and no other restrictions. Imagine a small perturbation of the curve Z shown in Figure 1.1A; it seems reasonable to think that there will still be a single point where the curve 'folds over', with two solutions on the left and none on the

Cambridge University Press
978-1-107-15164-2 — Singularities, Bifurcations and Catastrophes
James Montaldi
Excerpt
More Information

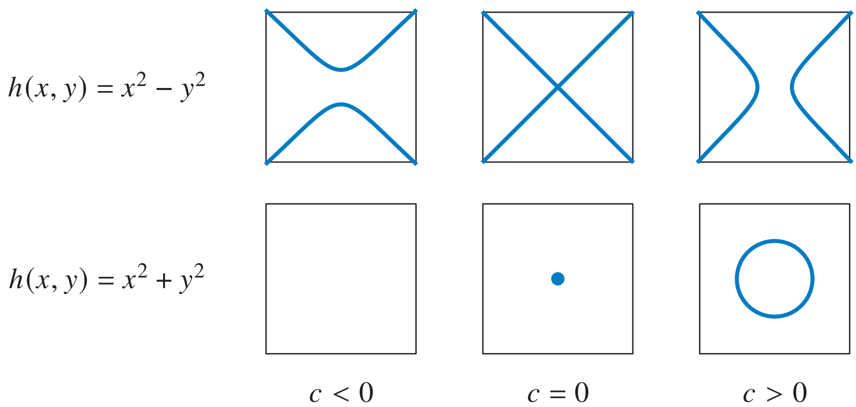


FIGURE 1.2 Two examples of bifurcation of contours $h(x, y) = c$ as c varies across the critical value of the function h.

right of this fold point (and one can prove this using the implicit function theorem; see Problem 1.6). This illustrates the *robustness* of the saddle-node bifurcation. In Section 1.5 below, we look briefly at an important bifurcation with one parameter, the pitchfork bifurcation, but where small perturbations do change its form. But first we look at two places bifurcations occur, the contours of a function as a parametrized set of equations, and a mechanical example with two parameters.

1.2 Bifurcations of contours

Landscape is determined in part by the height above sea level of each point of some region of the Earth. A ***contour*** is a curve on the landscape along which the height is constant; that is, for a given height the associated contour is the set of all points with that particular height. Let x, y be coordinates in the region in question, and $h(x, y)$ the height function. Then a contour at height c is the set of solutions of the equation

$$h(x, y) - c = 0.$$

Here we have a fixed function h and we can consider c as a parameter. Of course, height is only one example; another is the atmospheric pressure as a function on the surface, in which case the 'contours' are the familiar isobars from weather maps (although atmospheric pressure is best expressed as a function of three variables $P(x, y, z)$ as it varies with altitude z).

Consider a function $h(x, y)$ and the resulting equation $h(x, y) = c$. Most of the contours are curves, and a natural question to ask is, as c is varied, how can these curves change? The contours of a function are also called its ***level sets***.

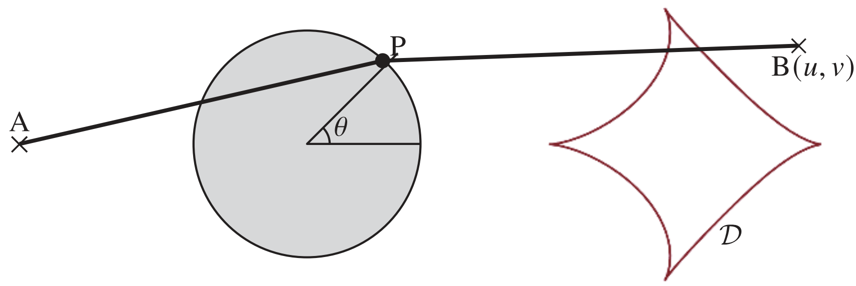


FIGURE 1.3 Schematic diagram of the Zeeman catastrophe machine. The red curve marked $\mathcal{D}$ is the discriminant or bifurcation set; notice the four cusps. Its precise size and position depend on the physical characteristics of the elastics and the position of the point A.

For example, suppose $h(x, y) = x^2 - y^2$. Then the contours are either hyperbolae or a pair of lines and the transition is depicted in the top row of Figure 1.2. For $h(x, y) = x^2 + y^2$ the contour is a circle for $c > 0$, a single point for $c = 0$ and is empty for $c < 0$. See the lower figures in Figure 1.2. In both cases a change occurs as one crosses the level $c = 0$, and one can show in general that qualitative changes only occur at *critical values* of the function; that is, the value the function takes at a critical point. We will study this in greater depth in later chapters.

A similar example in more variables is provided by $h(x, y, z) = x^2 + y^2 - z^2$. The zero level of this function is a circular cone in $\mathbb{R}^3$, while $h = 1$ is a one-sheeted hyperboloid and $h = -1$ is a two-sheeted hyperboloid.

1.3 Zeeman catastrophe machine

Conceived by Christopher Zeeman to illustrate the ideas of catastrophe theory, the Zeeman catastrophe machine consists of a wheel free to rotate about its centre, with a peg P attached at a point of its circumference. To the peg are attached two elastics: the other end of the first is pinned at a fixed point A in the plane of the wheel, while the other end of the second elastic is held by hand at a second point $B(u, v)$ in that plane. See Figure 1.3. The question is, how many equilibrium states are there of the wheel?

The answer will depend on where the end B is held; that is on the values of u and v, so these are the parameters. For each choice of point (u, v), the total elastic potential $V_{(u,v)}(\theta)$ is a function of θ, the position of the wheel (see Figure 1.4), and the equilibrium points are the points θ where V has a critical point: $\frac{\mathrm{d}}{\mathrm{d}\theta}V_{(u,v)}(\theta) = 0$.

The computation of the potential is straightforward but lengthy (and not relevant here), but the conclusion can be described simply. In the (u, v)-plane, there is a curve with four cusps, marked $\mathcal{D}$ in the figure. If the point B is within the curve, the wheel has four equilibrium points, two of which are stable (where $V'' > 0$) and two are unstable (where $V'' < 0$). On the other hand, if B lies outside this curve, then the wheel has only one stable and one unstable equilibrium point. The transition from four to two critical points happens when B approaches the curve $\mathcal{D}$ from the inside, and two of the critical points get closer and coalesce becoming degenerate in the process, and then disappearing; this curve $\mathcal{D}$ is therefore the discriminant of this family. This transition is the same as that in the fold bifurcation described above, although something more involved happens at the cusp points of the discriminant.

1.4 An example from geometry: the evolute

Consider a smooth simple closed curve C in the plane (e.g. an ellipse: a curve is said to be *smooth* if it has a parametrization whose derivative is nowhere zero). Let $P(u, v)$ be a point in the plane (possibly on C) with coordinates (u, v). The geometric question is, can you draw a perpendicular to the curve from the point P, and if so how many? (If P lies on the curve then we allow that the 'segment' (of zero length) from P to P is perpendicular to the curve.)

For example, if C is an ellipse, and P is at its centre, then it is not hard to see that there are 4 such perpendiculars – one to each of the points on the axes of the ellipse. What happens to those 4 points if P is perturbed? The feet of the perpendiculars will move, but can there be a different number of them? Imagine instead a point P' on the major axis of the ellipse, but outside the ellipse. It is easy to see that there are now only 2 perpendiculars from P' to C. See Figure 1.5. The bifurcation question is, how does 4 change to 2 as P is moved? And more completely, what is this number for all possible points P?

One observation is that for any P there are at least two such perpendiculars, and these arise at the nearest and furthest points of the curve to P as some thought should convince you (and which we prove below). This suggests defining the function on C which is the distance of each point of C to P. In fact we use the square of the distance which leads to simpler expressions after differentiating.

Let $\mathbf{r}(t)$ be a regular parametrization of the plane curve C, where 'regular' means that its derivative $\dot{\mathbf{r}}(t)$ is never zero, and for each point $\mathbf{c} = (u, v)$ in the

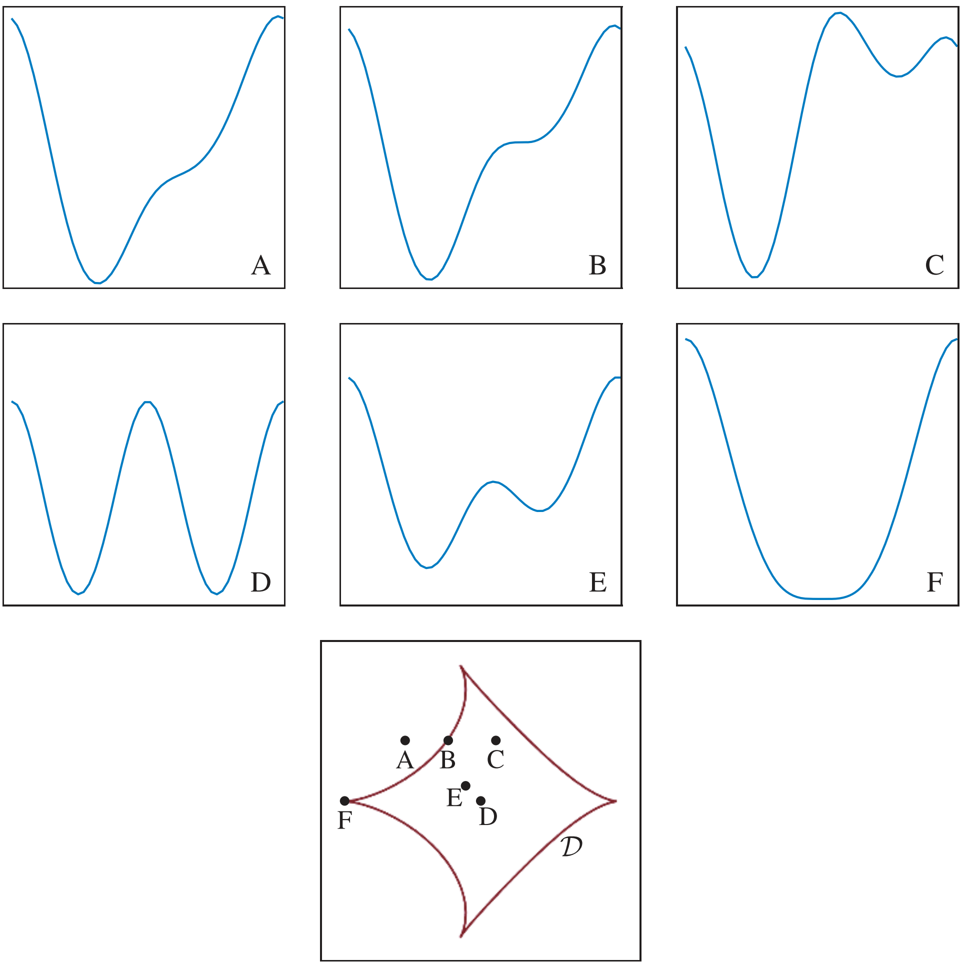


FIGURE 1.4 Graphs of the potential V in Zeeman's catastrophe machine for the six different parameter values shown in the bottom figure. Note that Figures B and F have degenerate critical points, and the corresponding points in the bottom diagram lie on the discriminant $\mathcal{D}$. The horizontal axis in diagrams A–F is $\theta \in [0, 2\pi]$.

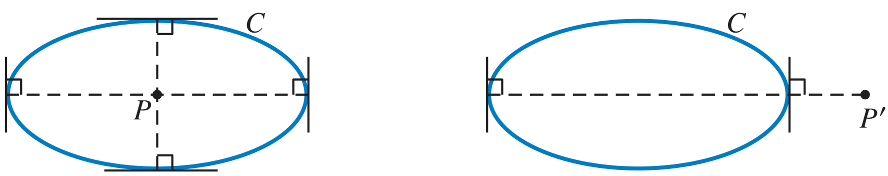


FIGURE 1.5 The dashed lines are perpendiculars from P and P' to the ellipse C.

plane, define the function

$$f_{\mathbf{c}}(t) = -\tfrac{1}{2}\,\|\mathbf{c} - \mathbf{r}(t)\|^2.$$

($\mathbf{c}$ is the position vector of the point P from the discussion above, and the factor $-\frac{1}{2}$ is for convenience.) This is a family of functions of t, with two parameters u and v. It measures the square of the distance from the point $\mathbf{c}$ to the point $\mathbf{r}(t)$; it's called the ***distance squared function***, or distance squared family.

Question: where does $f_{\mathbf{c}}$ have a critical point, and when is it degenerate?

First differentiate $f_{\mathbf{c}}$ (with respect to t),

$$f'_{\mathbf{c}}(t) = (\mathbf{c} - \mathbf{r}(t)) \cdot \dot{\mathbf{r}}(t). \tag{1.2}$$

Since $\dot{\mathbf{r}}(t)$ is the derivative of $\mathbf{r}(t)$, it represents a non-zero tangent vector to the curve. It follows that $f'_{\mathbf{c}}(t) = 0$, if $\mathbf{c}$ lies on the normal line to the curve at $\mathbf{r}(t)$. The set of critical points is therefore a very geometric object:

$$C = \{(t, u, v) \in \mathbb{R}^3 \mid (u, v) \text{ lies on the normal to the curve at } \mathbf{r}(t)\}. \tag{1.3}$$

Thus, for given P the original question is now, how many critical points does $f_{\mathbf{c}}$ have? In particular, the question of how many normals there are for a given point P is now cast as a variational problem.

Local changes in the number of critical points can only occur when a critical point is degenerate (as follows from the implicit function theorem). To see if the critical point is degenerate, we find the second derivative:

$$f''_{\mathbf{c}}(t) = (\mathbf{c} - \mathbf{r}(t)) \cdot \ddot{\mathbf{r}}(t) - \|\dot{\mathbf{r}}(t)\|^2. \tag{1.4}$$

Thus $f_{\mathbf{c}}$ has a degenerate critical point at t if both (1.2) and (1.4) are equal to zero. We can rewrite the two equations as,

$$\begin{cases} \dot{\mathbf{r}}(t) \cdot \mathbf{c} &= \mathbf{r}(t) \cdot \dot{\mathbf{r}}(t) \\ \ddot{\mathbf{r}}(t) \cdot \mathbf{c} &= \mathbf{r}(t) \cdot \ddot{\mathbf{r}}(t) + \|\dot{\mathbf{r}}(t)\|^2 \end{cases}$$

This is simply a pair of linear equations for $\mathbf{c}$, and if the coefficients $\ddot{\mathbf{r}}(t)$ and $\dot{\mathbf{r}}(t)$ are not parallel (they are both vectors), there is a unique solution $\mathbf{c}$, so giving a unique point[1] on that normal line. Call this point $\mathbf{e}(t)$: the resulting curve is called the ***evolute*** of the original curve. We have shown that the point $\mathbf{c} = \mathbf{e}(t)$ if and only if the function $f_{\mathbf{c}}$ has a degenerate critical point at t; the set of $\mathbf{e}(t)$ as t varies is therefore the discriminant of this family $f_{\mathbf{c}}$.

Example 1.1. As a specific example, consider the ellipse

$$\mathbf{r}(t) = (3\cos t,\ 2\sin t).$$

Then, with $\mathbf{c} = (u, v)$,

$$f_{\mathbf{c}}(t) = -\tfrac{1}{2}(u - 3\cos t)^2 - \tfrac{1}{2}(v - 2\sin t)^2. \tag{1.5}$$

The first two derivatives are $f_{\mathbf{c}}'(t) = -3u\sin t + 2v\cos t + 5\sin t\cos t$, and

$$f_{\mathbf{c}}''(t) = -3u\cos t - 2v\sin t + 10\cos^2 t - 5.$$

Solving $f'(t) = f''(t) = 0$ gives

$$u = \frac{5}{3}\cos^3 t, \quad v = -\frac{5}{2}\sin^3 t. \tag{1.6}$$

That is, $\mathbf{e}(t) = \left(\frac{5}{3}\cos^3 t,\ -\frac{5}{2}\sin^3 t\right)$; this curve is shown in Figure 1.6, together with the ellipse (notice that the ellipse is traversed anticlockwise, while the resulting parametrization of the evolute is clockwise). Note that this evolute or discriminant also has 4 cusps, like the ZCM above. We will see in later chapters that cusps occur very often on discriminants for 2 parameter families of functions, and using the theory of unfoldings we will explain why.

If $\mathbf{c}$ lies inside the evolute, the function $f_{\mathbf{c}}$ has 4 critical points, all nondegenerate, and if outside it has just 2. Indeed, using the symmetry of the ellipse if you take $\mathbf{c} = (0, 0)$ it is easy to see the 4 points of the curve for which the normal line passes through $\mathbf{c}$. If, on the other hand, $\mathbf{c}$ lies on the evolute but not at one of the cusps, then $f_{\mathbf{c}}$ has precisely 3 critical points, of which one is degenerate. Finally, if $\mathbf{c}$ lies at a cusp, $f_{\mathbf{c}}$ has a 'doubly' degenerate critical point and a nondegenerate one. As $\mathbf{c}$ varies from the interior of the evolute to the exterior, crossing at a regular point (ie, not at a cusp) then two of the critical points will coalesce and then disappear,

[1]it is in fact the *centre of curvature* of the curve at $\mathbf{r}(t)$; the evolute was originally defined by Huygens in the seventeenth century in his study of the pendulum, and it was later realised to be the locus of centres of curvature.

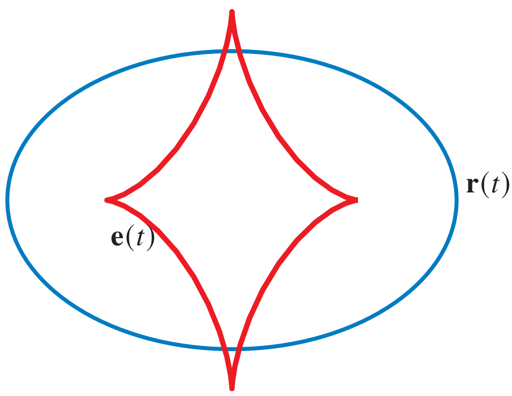


FIGURE 1.6 An ellipse with its evolute

just as they do for the fold family in Section 1.1 and the Zeeman Catastrophe Machine in Section 1.3. The 4 cusps are interesting geometrically: they are points on the evolute (centres of curvature) corresponding to points on the curve where the curvature has a local maximum or minimum.

If the major and minor axes of the ellipse were closer in value (here they are equal to 3 and 2 respectively), the evolute would be smaller, and in the limit as the ellipse tends to a circle, so the evolute tends to a single point: the centre of the circle. ✐

Applications of these ideas to the study of the geometry of curves and surfaces can be found in two books [18] and [61]; there is also a brief discussion in Chapter 15 in this book.

One question arising from the two very different examples, the evolute and Zeeman's catastrophe machine, is why do the bifurcation curves or discriminants have cusps? We will show in later chapters that this is very natural, given that we are studying a 2–parameter family of functions. The fact that in both cases there is only one variable θ or t turns out not to be important: it's the number of parameters that is central.

These two examples are both variational problems (arising from looking for critical points of functions), and such problems will be the study of the first part of this book. Later we will study more general (non-variational) bifurcation problems, but it will turn out that for two parameters, folds and cusps are still all that are to be expected.