GENERIC BIFURCATION IN THE OBSTACLE PROBLEM*

By

JOHN MALLET-PARET

Brown University

Abstract We consider a class of singularities, locally of the form $y^2 = p(x)$ near the origin in R^2, describing the shape of a free boundary curve arising from an elliptic free boundary value problem. The point of view taken is that of generic bifurcation, in particular with more than one parameter present. Of prime interest is a description of the unfoldings of such singularities, their normal forms, and generic conditions for one- and two-parameter unfoldings. The two simplest cases corresponding to perturbations of singularities $y^2 = x^n + O(x^{n+1})$, $n = 4, 5$ are treated in greater detail and the bifurcation diagram for a generic two-parameter unfolding is given.

Our results do not rigorously concern the free boundary problem itself, but rather set down a formal framework, or model, for studying this problem in terms of bifurcation theory. We prove theorems describing this model. Nevertheless, our results have a bearing on any rigorous analysis of this problem since they form the necessary first step to such an analysis. The theory for computing the normal forms of solutions up to first order, for example, is given here.

1. Introduction. Consider the following obstacle problem in two dimensions: in a bounded domain $\Omega \subseteq R^2$ are given functions

$$\psi: \bar{\Omega} \to R, \qquad \Delta\psi < 0$$

representing an obstacle and

$$g: \partial\Omega \to R, \qquad g(w) > \psi(w)$$

representing the boundary position of a membrane to be stretched over the obstacle. The membrane is given as the graph of a function $u: \bar{\Omega} \to R$ which minimizes the energy

$$J(u) = \frac{1}{2} \int_\Omega |\nabla u|^2$$

over an admissible class of functions

$$u: \bar{\Omega} \to R, \qquad u = g \quad \text{on} \quad \partial\Omega, \qquad u \geq \psi \quad \text{in} \quad \Omega.$$

The set of contact

$$I = \{w \in \Omega \mid u(w) = \psi(w)\}$$

between the membrane and obstacle is of particular interest.

* Received September 12, 1978. This research was supported in part by the National Science Foundation under NSF-MCS 78-18858 and in part by the U.S. Army Research Office under ARO-D-31-124-73-G130. The author would like to thank David Schaeffer for several helpful conversations on this problem.

Such problems, and more general ones, have been studied by many authors including Brézis, Caffarelli, Kinderlehrer, Lewy, Nirenberg, Rivière, Schaeffer, and Stampacchia. It is known [1] that the above problem has a unique solution $u \in C^1(\bar{\Omega})$ with Du Lipschitz in $\bar{\Omega}$, provided ψ is sufficiently smooth. Under quite general conditions [3, 9, 10, 11] the free boundary ∂I consists of smoothly parameterized arcs, possibly with cusps; and the parameterization is analytic if ψ is. Schaeffer [13, 14] has studied how the set I changes as the data ψ and g vary, and has pointed out the need for a generic theory of such variations. Such a theory presumabably could take the form of a bifurcation theory or unfolding theory for the singularities of ∂I. A significant point here is that the unfoldings one would encounter would not be generic in the sense of singularity theory as developed by Thom, Mather, Arnold and others; only very special types of singularities can occur. This is seen, for example, from a result of Kinderlehrer and Nirenberg [10]: for $n \geq 1$ odd the free bounary can never have a cusp of the form

$$(y - y_0)^2 \sim K(x - x_0)^{2n+1}, \qquad K \neq 0$$

near some (x_0, y_0), whereas for n even such a cusp can occur. (In fact, the case $n = 1$ was first noted by Schaeffer [15].) Their proof involves first straightening out the cusp to a line segment by means of a conformal mapping; then an analysis of several terms of the Taylor series of u near (x_0, y_0), based on the equation $\Delta u = 0$ governing u in $\Omega - I$, gives a contradiction to $u \geq \psi$.

In this paper, we make a detailed study of a class of singularities of a free boundary, and their bifurcations, as the data of the system varies parametrically. Of particular interest is the multi-parameter case, with more than one scalar parameter present. Generic conditions, describing how the parameters enter the system in a non-degenerate manner, are derived; in such a case the bifurcation set in the parameter space can be described, at least up to first order.

Our results do not rigorously concern the above obstacle problem, but rather set down a formal basis, or model, for studying this problem from the point of view of generic bifurcations. We prove theorems describing this model. We believe such a study of the formal calculations is worthwhile for a number of reasons.

(1) In many bifurcation problems the analysis divides naturally into two parts: formal analysis (which typically can involve equating coefficients in some small-parameter expansion and solving for unknown coefficients) and rigorous justification (employing, for example, the Lyapunov-Schmidt method, and other uses of the implicit function theorem coupled with various scaling techniques). Often the formal calculations involve rather straightforward ideas such as solving a linear equation with the Fredholm alternative, or solving a system of polynomials in one or two real variables. Here, however, the formal calculations themselves involve significant mathematical questions. For example, they deal with multivalued functions which must be considered on a Riemann surface, a careful study of which must be made. The formal theory, in itself, gives rise to a number of mathematically intriguing questions.

(2) The formal calculations are a necessary first step before corresponding rigorous results can be obtained. In particular, they are needed to compute normal forms for singularities and their deformations, and to compute bifurcation sets in the parameter space. For problems arising in specific applications these, indeed, may be the questions of central interest.

(3) For multi-parameter bifurcation problems the appropriate generic hypotheses describing how the parameters enter the system follow naturally from the formal analysis.

This analysis, in fact, can provide a general framework in which multi-parameter problems, generic or not, can be placed.

(4) A general method for making a rigorous bifurcation analysis of such problems involves obtaining rather sharp upper and lower *a priori* estimates u^+ and u^- for the true solution u, and corresponding estimates for the location of the free boundary ∂I. This was carried out in [6] for a particular two-parameter problem using "restricted unfolding" techniques, as described below. To use these techniques in other such problems it is necessary to have at hand candidates for u^+ and u^-, preferably in some closed analytical form. The formal theory supplies such candidates as first-order approximations to the true solution u.

We may broadly distinguish two viewpoints which one can take in analyzing singularitites. In the *universal unfolding* (u.u.) approach one begins with a specific singularity and tries to describe all nearby singularities up to some equivalence relation such as nonsingular coordinate change. For example, it is intuitively clear that any smooth map near $t \in R \rightarrow t^2 \in R$ must, near the origin, resemble $t \rightarrow t^2 + \alpha_1 t + \alpha_0$ for some (α_1, α_0) near (0, 0). This can be made precise in a number of ways; for example, if $F(t, \lambda)$ is real analytic for $(t, \lambda) \in R \times R^k$ near (0, 0) and $F(t, 0) = t^2$, then by the Weierstrass preparation theorem F can be factored as

$$F(t, \lambda) = E(t, \lambda)(t^2 + \alpha_1(\lambda)t + \alpha_0(\lambda))$$

where E, α_1 and α_0 are analytic and satisfy

$$E(t, 0) = 1, \qquad \alpha_1(0) = \alpha_0(0) = 0.$$

The normal form $t^2 + \alpha_1 t + \alpha_0$ is called the universal unfolding of the singularity $t \rightarrow t^2$; the number of parameters in the universal unfolding, two here, is called the codimension of the singularity.

In the *restricted unfolding* (r.u.) approach one begins with a specific parameterized family, say $F(t, \lambda)$, and attempts to analyze this directly. Even though it may be possible to invoke theorems, such as the preparation theorem, which reduce F to a normal form, this may not be the best course of action. For one thing, the normal form itself must be analyzed: suppose, for example, F: $R \times R \rightarrow R$ has the form

$$F(t, 0) = t^{2n} + O(t^{2n+1}), \qquad \partial F/\partial\lambda(0, 0) = K > 0$$

and we wish to describe the zeros of F near $t = 0$, for a given λ. The preparation theorem reduces this to an analysis of

$$t^{2n} + \alpha_{2n-1}(\lambda)t^{2n-1} + \cdots + \alpha_0(\lambda) = 0, \;\; \alpha_j(0) = 0, \qquad d\alpha_0/d\lambda(0) = K$$

which is no simpler a problem; it is just as easy to study F directly. This may be done by first establishing the *a priori* estimate $|t| \leq$ (constant) $|\lambda|^{1/2n}$ for zeros of F. This estimate justifies the scaling $t \rightarrow \mu t$ where $\mu = |\lambda|^{1/2n}$, and leads to analysis of the zeros of

$$\begin{aligned} G(t, \mu) &= \mu^{-2n} F(\mu t, (\operatorname{sgn} \lambda)\mu^{2n}) \\ &= t^{2n} + (\operatorname{sgn} \lambda)K + O(\mu). \end{aligned}$$

It is easily seen that there are no zeros when $\lambda > 0$; for $\lambda < 0$ an easy application of the implicit function theorem to G shows that F has exactly two zeros and they have the asymptotic forms $t \sim \pm(-\lambda K)^{1/2n}$.

The r.u. method thus has the advantage of using relatively elementary techniques (such as scaling and the implicit function theorem) to obtain quite precise information about specific singularities. A disadvantage is that there is no general approach; the techniques are often applied in an *ad hoc* manner and are not easily used at all if more then two or three parameters are present. The u.u. approach, on the other hand, is not limited by the number of parameters or the way they appear. But, perhaps its most significant feature is the way it illuminates and unifies the various families of singularities and hypotheses encountered in the r.u. case. In the above example the hypothesis $\partial F/\partial\lambda\ (0, 0) \neq 0$ translates to $d\alpha_0/d\lambda(0) \neq 0$ and in this form makes the choice of scaling factors more obvious. For a two-parameter family, with parameters $(\lambda_1, \lambda_0) \in R^2$, a natural hypothesis on F would be $|\partial(F, F_t)/\partial(\lambda_1, \lambda_0)| \neq 0$, as this is equivalent to $|\partial(\alpha_1, \alpha_0)/\partial(\lambda_1, \lambda_0)| \neq 0$. Under such conditions the zeros of F can be analyzed by scalings suggested by the normal form. In fact, if $F(t, 0) = t^2$, $\lambda \in R^k$ and $\partial(F, F_t)/\partial\lambda$ has rank two at the origin, it is always possible to select two coordinates (λ_i, λ_j) for which $|\partial(F, F_t)/\partial(\lambda_i, \lambda_j)| \neq 0$, and give a fairly complete picture of the zeros of F by r.u. techniques. Such questions and, in general, the relation between the u.u. and r.u. approaches were investigated by Chow, Hale and Mallet-Paret [4, 5]. See also Hale [7, 8], Mallet-Paret [12], and Golubitsky and Schaeffer [16].

For the obstacle problem above Schaeffer [13] proved an implicit function theorem which suggested analyzing singularities of the free boundary ∂I using r.u. techniques. Such an analysis was carried out by Chow and Mallet-Paret [6] for two specific cases: (1) a one-parameter system in which the obstacle ψ touches the membrane at one point, in a non-degenerate manner, and varying the parameter causes the obstacle at that point to move up or down with non-zero speed; and (2) a similar situation but with two parameters and two points of contact, where the two parameters describe independent vertical variations of the obstacle near the two points. This analysis involves two steps: (1) obtaining upper and lower *a priori* estimates u^+ and u^- of the solution u, and estimates of the location of the free boundary; and (2) a scaling of the system, justified by these estimates, and application of Schaeffer's implicit function theorem. (In fact, since the independent variable here is two-dimensional one can use the classical reflection techniques of Lewy and Stampacchia [11] in place of the more complicated implicit function theorem. An advantage of this is that the reflection arguments are local.) The scaling yields an obstacle problem which is a perturbation of a particularly simple one, in which ∂I is an ellipse and the obstacle is a paraboloid. Indeed, u^+ and u^- can also be chosen to be of such a simple form.

In this context then the philosophy of the present paper is to place the obstacle problem in a u.u. setting with a view to using r.u. techniques. We begin with a specific singularity $y^2 = x^n b_0(x)^2$ of ∂I and obtain necessary and sufficient conditions for a deformation of this singularity, at least locally, to give rise to valid solutions. Such solutions can presumably be used, with r.u. techniques, as estimates u^+ and u^- in proving bifurcation results. The precise conditions on the singularity are given in Theorems 4.4, 5.2 and 5.4. The resulting normal forms of the singularity, in the simpler cases $n = 4$ and 5, are given in Sec. 6. In Sec. 7, a first-order analysis of variations in the boundary condition g is made, and from this arise the generic conditions for multi-parameter bifurcation. Finally, in Secs. 8 and 9 generic one- and two-parameter unfoldings for the cases $n = 4$ and 5 are considered; normal forms and bifurcation diagrams are given.

Throughout, we assume the obstacle is a paraboloid

$$\psi(x, y) = -\tfrac{1}{2}(x^2 + y^2) \qquad (x, y) \in R^2.$$

This is no major restriction as we are concerned with a local analysis. If more generally only $\Delta\psi < 0$ holds, then by adding a harmonic function to ψ and multiplying by a positive constant we obtain an equivalent problem with

$$\psi(x, y) = -\tfrac{1}{2}(x^2 + y^2) + O(|x|^3 + |y|^3)$$

near the origin. The higher-order cubic terms in ψ should not affect the resulting theory significantly, at least for the simpler singularities.

2. Local solutions of the obstacle problem. The unique solution u of the obstacle problem in Sec. 1 is characterized by the following system of differential inequalities (see [6, 13]):

$$\begin{aligned} u&: \bar{\Omega} \to R, u \in C^1, \\ u &= g \quad \text{on} \quad \partial\Omega, \\ u &\geq \psi \quad \text{in} \quad \Omega, \\ \Delta u &= 0 \quad \text{in} \quad \Omega - I, \\ I &= \{(x, y) \in \Omega \mid u(x, y) = \psi(x, y)\}. \end{aligned}$$

As noted above, under very general conditions ∂I is a smoothly parameterized arc, possibly with cusps. In fact, if ψ is analytic then any part of ∂I with a C^1 parameterization has an analytic parameterization.

Suppose now ψ has the form

$$\psi(x, y) = -\tfrac{1}{2}(x^2 + y^2) = -\tfrac{1}{2}|w|^2$$
$$(x, y) \in R^2, \qquad w = x + iy \in C$$

near the origin $(0, 0) \in \Omega$. Let ∂I pass through the origin as a non-singular arc $\mathscr{C}$ with I lying to one side; then we may assume $\mathscr{C}$ and I locally have the form

$$\mathscr{C}: y = a(x), \qquad a(0) = 0, a \text{ is real analytic,}$$
$$I: y \leq a(x).$$

In I we have $u = \psi$; since $u \in C^1$ we have

$$\left\{ \begin{aligned} & u = -\frac{1}{2}(x^2 + y^2) \\ & \operatorname{grad} u = \left(\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}\right) = -(x, y) \end{aligned} \right\} \quad (x, y) \in C. \tag{2.1}$$

For $y > a(x)$, u is harmonic so it is uniquely determined there from the data (2.1) on $\mathscr{C}$ by the Cauchy-Kowalewski theorem. In fact, an explicit formula for u can be given (see Schaeffer [14]): let $h(z)$ be the holomorphic function

$$w = h(z) = z + ia(z) \tag{2.2}$$

taking values in the w-plane. Clearly, h conformally maps a neighborhood of $z = 0$ to a neighborhood of $w = 0$, taking the real axis to the curve $\mathscr{C}$. Let, for w near 0,

$$f(w) = 2h^{-1}(w) - w \tag{2.3}$$

and note that

$$f(w) = \bar{w}, \qquad w \in \mathscr{C}. \tag{2.4}$$

It is now a simple matter, using the Cauchy-Riemann equations, to see that u is the harmonic function

$$u(w) = -\mathrm{Re} \int_0^w f(\alpha)\, d\alpha, \qquad w \notin I.$$

In terms of the z variable this formula is

$$\begin{aligned} u(w) &= \frac{1}{2} \mathrm{Re}\,(h(z)^2) - 2\,\mathrm{Re} \int_0^z \zeta h'(\zeta)\, d\zeta, \\ \mathrm{Re}\, z &> 0, \qquad w = h(z). \end{aligned} \tag{2.5}$$

If n is the unit normal for $\mathscr{C}$, then along $\mathscr{C}$

$$\frac{\partial^2(u - \psi)}{\partial n^2} = \Delta(u - \psi) = -\Delta\psi = 2.$$

Hence, $u > \psi$ in some one-sided neighborhood $0 < y - a(x) < \varepsilon$ of $\mathscr{C}$.

Thus, given a non-singular analytic curve in the plane, there is defined a unique harmonic function u (2.5). If the curve varies analytically with a parameter λ, say as

$$\mathscr{C}_\lambda : y = a(x, \lambda), \qquad \lambda \in R^k \quad \text{near} \quad 0$$

then $u = u(w, \lambda)$ is well-defined in a uniform neighborhood of $\mathscr{C}_\lambda$, and satisfies $u > \psi$ in the neighborhood. Below, we study the case where $\mathscr{C}_0$ has a singularity and $\mathscr{C}_\lambda$ describes some unfolding of this singularity—then u may not be defined in a uniform neighborhood; and even if it is $u > \psi$ may not hold. Our object is to describe perturbations $\mathscr{C}_\lambda$ for which these properties do hold.

3. Curves with cusps. Let $\mathscr{C}_0$ denote the solution set of

$$y^2 = p_0(x) \qquad (x, y) \in R^2 \quad \text{near} \quad (0, 0)$$

where

$$p_0(x) = x^n b_0(x)^2, \qquad n \geq 3, \qquad b_0(0) > 0 \tag{3.1}$$

is real analytic near zero, and let

$$I_0 = \{(x, y) \mid y^2 \leq p_0(x)\}.$$

We regard $\mathscr{C}_0$ as a free boundary and I_0 as the contact set of the obstacle problem; we subject these sets to perturbations to obtain I_λ and $\mathscr{C}_\lambda = \partial I_\lambda$, and wish to study those perturbations giving local solutions $u(w, \lambda)$ as in (2.5), defined in a uniform region of the complement of I_λ, satisfying $u > \psi$ in that region. For simplicity we assume symmetry of I_λ with respect to the x-axis, although non-symmetric perturbations could just as well be studied. Also, the choice of $y^2 \leq p_0(x)$ to describe I_0, rather than the opposite inequality, was made in view of a result of Caffarelli [2] that points of ∂I which are singularities must be points of zero density of I. (For this reason, we consider only $n \geq 3$; $n = 2$ cannot occur.)

The family of perturbations of I_0 is motivated from the following considerations. Let

$$F_0(x, y) = y^2 - p_0(x)$$

so I_0 is the region where $F_0 \leq 0$. Any real analytic function, even in y, near F_0 must have the form

$$F(x, y) = E(x, y^2)(y^2 - p(x))$$
$$|E(0, 0) - 1| \ll 0$$

by the Weierstrass preparation theorem. By "near" we mean, for example, obtained through an analytically parameterized family $F(x, y, \lambda)$ for $\lambda \in R^k$ near the origin. In such a case p is also near p_0, so again by the Preparation Theorem

$$p(x) = (x^n + \alpha_{n-1}x^{n-1} + \cdots + \alpha_0)b(x)^2 \tag{3.2}$$

$$\alpha_j \in R, \qquad |\alpha_j| \ll 1. \tag{3.3}$$

$$b(0) > 0, \qquad b \quad \text{is near} \quad b_0 .$$

As we are interested in the region I where $F \leq 0$, we may neglect the factor $E(x, y^2)$. Therefore, we shall study solutions $u(w)$ arising from

$$I = \text{cl int}\{(x, y) \mid y^2 \leq p(x)\}, \qquad \mathscr{C} = \partial I \tag{3.4}$$

with $p(x)$ as in (3.2), (3.3), and $b(x)$ real analytic and uniformly near $b_0(x)$ for x in some disc in the complex plane:

$$\sup|b(x) - b_0(x)| \leq \varepsilon, \qquad x \in C, \qquad |x| \leq \delta.$$

Note that I does not have the isolated points $(x_0, 0)$ at which $p(x_0) = 0$ and p has a local maximum for x real; such points may not actually lie in the region of contact between the membrane and obstacle, as we shall see in Sec. 5.

4. Restrictions on the singularity. With $p(x)$ as above three conditions must be met in order to construct the function u in the complement of I, in a uniform neighborhood of zero:

(1) the function f in (2.3), (2.4) must be holomorphic (i.e. have no branch points) and single-valued in the complement of I;
(2) the integral (2.5) defining u must be independent of the path of integration (since I may have several connected components) and must equal ψ on each component of $\mathscr{C}$ (i.e., there must be a unique choice of a constant of integration, or lower limit of the integral, valid for all components of $\mathscr{C}$); and
(3) with (1) and (2) satisfied we must have $u \geq \psi$ in the complement of I.

We shall give necessary and sufficient conditions for (1), (2) and (3) to hold.

Let us first analyze (1). The mapping $w = h(z)$ (2.2) takes the form

$$w = h(z) = z \pm ip(z)^{1/2} \tag{4.1}$$

or, equivalently,

$$G(z, w) = 0$$

where

$$G(z, w) = (z - w)^2 + p(z). \tag{4.2}$$

Therefore, h and h^{-1} must be interpreted as mappings between Riemann surfaces R_z and R_w for z and w respectively. Since $|p''(z)| \ll 1$ (because $n \geq 3$) we see that by Rouché's theorem and the Weierstrass preparation theorem the following result holds.

LEMMA 4.1. The Riemann surfaces R_z and R_w for the zeros of G (4.2) are, near $(z, w) = (0, 0)$, two-sheeted coverings of the z- and w-planes respectively, with finitely many branch points. For each fixed z near zero $G(z, \cdot)$ has two roots near zero (counting multiplicity), and similarly for each fixed w. These roots near some point (z_0, w_0) where $G(z_0, w_0) = 0$ can be expressed by giving w as a power series in one of $z - z_0$ or $(z - z_0)^{1/2}$; similarly, z can be given as a power series in one of $w - w_0$ or $(w - w_0)^{1/2}$.

It is an easy matter to determine the location of the branch points of R_z and R_w. Those of R_z are clearly the points z_0 which are zeros of p of odd order. A necessary condition that w_0 be a branch point of R_w is that there exist z_0 with

$$G(z_0, w_0) = \frac{\partial G}{\partial z}(z_0, w_0) = 0; \tag{4.3}$$

this condition is sufficient if, in addition,

$$\frac{\partial G}{\partial w}(z_0, w_0) \neq 0. \tag{4.4}$$

The formula for G shows (4.3) is equivalent to

$$w_0 = z_0 + \tfrac{1}{2}p'(z_0), \qquad 4p(z_0) + (p'(z_0))^2 = 0$$

and that (4.4), in addition, holds if and only if $p'(z_0) \neq 0$. (This last condition is equivalent to saying z_0 is a simple zero of $4p + (p')^2$, since $|p''| \ll 1$.)

If z_0 satisfies (4.3) with $p(z_0) = p'(z_0) = 0$, when is $w_0 = z_0$ a branch point of R_w? Let

$$p(z) = (z - z_0)^m q(z)^2, \qquad m \geq 2, \qquad q(z_0) \neq 0,$$
$$|q(z_0)| \ll 1 \quad \text{if} \quad m = 2$$

near z_0. If m is even then $G = 0$ is equivalent to the two equations

$$z - w = \pm i(z - z_0)^{m/2} q(z) \tag{4.5}$$

which each have holomorphic solutions of z as a function of w near (z_0, w_0); therefore w_0 is not a branch point. If m is odd, however, the solution of (4.5) is given as

$$z = w \pm i(w - w_0)^{m/2} q(z_0) + O(|w - w_0|^{(m+1)/2})$$

which is not homomorphic; therefore w_0 is a branch point. Note here that the order $m \geq 2$ of the zero z_0 of p is also the order of z_0 as a zero of $4p + (p')^2$. We summarize these results in a lemma.

LEMMA 4.2. The branch points of R_z are precisely the zeros of p of odd order. The branch points w_0 of R_w are given as

$$w_0 = z_0 + \tfrac{1}{2} p'(z_0)$$

where z_0 is an odd-order zero of the function

$$r(z) = 4p(z) + (p'(z))^2.$$

A point z_0 is a zero of p of order $m \geq 2$ if and only if it is a zero of r of order $m \geq 2$.

Locally $z = h^{-1}(w)$ is a holomorphic function of w away from the branch points of R_w, but never at the branch points. It follows that a necessary condition for $f(w)$ to be holomorphic in the complement of I, and for (1) to hold, is that all branch points of R_w lie in I. If this is so then in each component of the complement of I there are locally two holomorhpic choices for $h^{-1}(w)$, and so for $f(w)$, corresponding to the two unramified sheets of R_w there. Along the non-singular part of $\mathscr{C}$(i.e. except at points $w_0 \in \mathscr{C}$which are real non-simple zeros of p) there is locally a unique choice of f satisfying $f(w) = \bar{w}$ on $\mathscr{C}$; but this may not hold on all of $\mathscr{C}$(possibly if I is not connected, for example). This means that the branch points of R_w lying in I is not a sufficient condition for (1) to hold.

LEMMA 4.3. All branch points of R_w lie in I if and only if all odd-order zeros of $r(z) = 4p(z) + (p'(z))^2$ are real. This is true if and only if all odd-order zeros of $p(z)$ are real.

Proof. Let $w_0 = x_0 + iy_0$, x_0, $y_0 \in R$, be a zero of order $m \geq 3$ of p where m is odd. Then w_0 is a branch point of R_w. If w_0 is real then clearly $w_0 \in I$, so suppose $w_0 \notin R$. Recall here that $n \geq 3$ so we may assume in sime disc in C there is a sufficiently small bound for p'', say $|p''(z)| \leq \alpha \ll 1$. Now estimate:

$$|p(x_0)| = |p(x_0) - p(w_0)| \leq \tfrac{1}{2}\, \alpha^2 |x_0 - w_0|^2 = \tfrac{1}{2}\, \alpha^2 y_0^2 < y_0^2.$$

The last inequality shows $w_0 \notin I$, as required.

Now suppose $z_0 = z_1 + iz_2$, z_1, $z_2 \in R$, is a simple zero of $r(z)$ and $w_0 = z_0 + \tfrac{1}{2}\, p'(z_0)$ is the corresponding branch point of R_w. Note that $z \to z + \tfrac{1}{2}p'(z)$ is a local diffeomorphism preserving the real axis, so z_0 is real if and only if w_0 is real. First suppose z_0 is real; then

$$\begin{aligned} p(w_0) &= p(z_0) + (p(w_0) - p(z_0)) \\ &= -\tfrac{1}{4}(p'(z_0))^2 + (p(w_0) - p(z_0)) \\ &= -\tfrac{1}{4}(p'(z_0))^2 + p'(z_0)(w_0 - z_0) + O(\alpha^2|w_0 - z_0|^2) \\ &= \tfrac{1}{4}(p'(z_0))^2 + O(\alpha^2(p'(z_0))^2). \end{aligned} \tag{4.6}$$

Since $p'(z_0) \neq 0$ we conclude $p(w_0) > 0$; hence $w_0 \in I$.

If now z_0 is a non-real simple zero of r we must show $y_0^2 > |p(x_0)|$ to conclude $w_0 \notin I$. From the estimate (4.6) immediately above we obtain

$$\begin{aligned} p(x_0) &= p(w_0) + (p(x_0) - p(w_0)) \\ &= p(x_0) - p(w_0) + O(|p'(z_0)|^2) \\ &= -ip'(w_0)y_0 + O(\alpha y_0^2 + |p'(z_0)|^2) \\ &= -ip'(z_0)y_0 + O(\alpha y_0^2 + |p'(z_0)|^2). \end{aligned}$$

In order to show y_0^2 is greater than the norm of this last quantity it is sufficient to show

$$p'(z_0) = O(\alpha y_0). \tag{4.7}$$

To this end consider

$$\begin{aligned} 0 = r(z_0) &= r(z_1 + iz_2) \\ &= 4p(z_1 + iz_2) + (p'(z_1 + iz_2))^2 \\ &= 4p(z_1) + 4ip'(z_1)z_2 + (p'(z_1))^2 \\ &\quad + O(\alpha z_2^2 + |\alpha z_2 p'(z_1)|). \end{aligned}$$

Taking the imaginary part and cancelling $z_2 \neq 0$ gives $p'(z_1) = O(\alpha z_2)$ and hence

$$\begin{aligned} p'(z_0) &= p'(z_1) + O(\alpha|z_0 - z_1|) \\ &= O(\alpha z_2). \end{aligned} \tag{4.8}$$

Finally, we estimate

$$\begin{aligned} |z_2 - y_0| = |\mathrm{Im}(z_0 - w_0)| &\leq |z_0 - w_0| \\ &= O(p'(z_0)) \\ &= O(\alpha z_2) \end{aligned}$$

which together with (4.8) implies (4.7), as required.

To justify the last statement of the lemma, recall first that multiple zeros of p and r coincide. Also note from the argument principle that p and r have the same number of zeros (counting multiplicity) near $z = 0$, and so have the same number of simple zeros. By Rouché's theorem we establish a one-to-one correspondence between these simple zeros so that real zeros of p correspond to real ones of r, and non-real ones to non-real ones. Let z_0 be a simple zero of r and set $\sigma = |r'(z_0)| > 0$. An easy estimate shows z_0 is the only zero of r in the σ-disc about z_0. One also sees easily that $p'(z_0)$ and $r'(z_0)$ are of the same order, in fact $4p'(z_0) = r'(z_0)(1 + O(\alpha))$. Now on the boundary of the σ-disc

$$\begin{aligned} r(z_0 + \sigma e^{i\theta}) - 4p(z_0 + \sigma e^{i\theta}) &= (p'(z_0 + \sigma e^{i\theta}))^2 = (p'(z_0) + O(\alpha\sigma))^2 \\ &= (\tfrac{1}{4}r'(z_0))^2(1 + O(\alpha)) \neq 0 \end{aligned}$$

so p also has exactly one zero in this disc. If $z_0 \in R$ then this zero of p must also be real. If $z_0 \notin R$ then the estimate (4.8) applies and the σ-disc is disjoint from the real line; hence the zero of p is not real. This completes the proof of the lemma.

As mentioned above, even if the conditions of Lemma 4.3 hold there may not exist a function f holomorphic in the complement of I and satisfying $f(w) = \bar{w}$ on $\mathscr{C} = \partial I$; if there is, it must have the form $2h^{-1}(w) - w$ for some branch of h^{-1}. We give necessary and sufficient conditions for f to exist, and therefore for (1) to be satisfied.

THEOREM 4.4. Assume all odd-order zeros of p are real, as in Lemma 4.3. Then there exists a branch of $h^{-1}(w)$ (4.1) holomorphic and single-valued in the complement of I such that $f(w) = 2h^{-1}(w) - w$ satisfies $f(w) = \bar{w}$ on $\mathscr{C}$, if and only if p has the following property:

$$\begin{aligned} &\text{if } z_1 < z_2 \text{ are real and } p(z_j) > 0, \\ &j = 1, 2, \text{ then the number of zeros of } p \text{ in the} \\ &\text{interval } (z_1, z_2) \subseteq R \text{ is a multiple of four.} \end{aligned} \tag{4.9}$$

Hence the hypothesis of Lemma 4.3 together with (4.9) are necessary and sufficient conditions for (1) to hold.

Proof. Suppose first all real zeros of p are simple and all non-real zeros have even order; note the hypothesis of Lemma 4.3 is satisfied. Then I consists of finitely many connected components $I_1, I_2, \cdots, I_d$ corresponding to the disjoint intervals in R where p is positive. The boundaries $\mathscr{C}_j = \partial I_j$ are smooth curves. Assume $\{I_j\}$ is ordered so that for increasing j the real parts of points in I_j are increasing.

For any $w \in \partial I$ with, say, Im $w > 0$ there are two distinct values of $h^{-1}(w)$, namely $h^{-1}(w) = \text{Re } w$ and one for which $h^{-1}(w) \notin R$. If $w \in R$ but $w \notin I$ (so $p(w) < 0$) again there are two distinct values: both are real, one satisfies $h^{-1}(w) < w$ and the other $w < h^{-1}(w)$. This can be seen, for example, by considering a line segment

$$(z_1, z_2) \subseteq R$$

where

$$p(z_1) = p(z_2) = 0, \qquad p(z) < 0 \text{ for } z \in (z_1, z_2). \tag{4.10}$$

These two branches of h^{-1} correspond to the choices $ip(z)^{1/2} > 0$ and $ip(z)^{1/2} < 0$ respectively.

Select points $w_j \in \mathscr{C}_j$ Im $w_j > 0$ for $j = 1, 2$ and let $h_j^{-1}(w)$ denote the branch of h^{-1}, in the upper half plane intersect the complement of I, for which $h_j^{-1}(w_j) = \text{Re } w_j$. We assume, in fact, h_j^{-1} is single-valued and holomorphic in a simply connected region $\mathscr{O}$ containing points in the upper half plane not in I. By analytic continuation $h_j^{-1}(w) = \text{Re } w$ and $f_j(w) = \bar{w}$ for $w \in \mathscr{C}_j \cap \mathscr{O}$ where $f_j(w) = 2h_j^{-1}(w) - w$. We claim that h_1^{-1} and h_2^{-1} are different branches of h^{-1}. To prove this let $z_1 \in R \cap \mathscr{C}_1$ be the rightmost point of I_1, i.e. the point with maximum real part, and $z_2 \in R \cap \mathscr{C}_2$ the leftmost point of I_2. Then $z_1 < z_2$ and (4.10) holds. Now $h_j^{-1}(z_j) = z_j$ so we may study h_j^{-1} near z_j by solving

$$G(h_j^{-1}(w), w) = 0$$

near $w = z_j$. A simple calculation gives

$$h_j^{-1}(w) = z_j + O((w - z_j)^2).$$

From this it follows that $(-1)^j(h_j^{-1}(w) - w) > 0$ for $w \in (z_1, z_2)$ near z_j, and hence for all $w \in (z_1, z_2)$. Thus, h_1^{-1} and h_2^{-1} take different values on (z_1, z_2), and so are different branches of h^{-1}. We conclude from this that if a branch of h^{-1} satisfies $h^{-1}(w) = \text{Re } w$ (equivalently $f(w) = \bar{w}$) on some $\mathscr{C}_j$, then this is satisfied only on alternate $\mathscr{C}_k$'s, that is, if and only if $k - j$ is even. Moreover, we see h^{-1} is single-valued in the complement of I, for by analytic continuation on a given $\mathscr{C}_j$ either $h^{-1}(w) = \text{Re } w$ identically holds, or does not.

Now to complete the proof of the theorem consider any $p(z)$ satisfying the hypothesis of Lemma 4.3 and suppose there exists a branch of f for which $f(w) = \bar{w}$ on all of I. Select real numbers $z_1 < z_2$ for which $p(z_j) > 0$ and let $w_j = z_j + ip(z_j)^{1/2}$ be the corresponding points of ∂I in, say, the upper half plane. Perturb the coefficients $\alpha_0, \cdots, \alpha_{n-1}$ (3.2) by a small amount to obtain

$$\tilde{p}(z) = (z^n + \tilde{\alpha}_{n-1}z^{n-1} + \cdots + \tilde{\alpha}_0)b(z)^2$$

$$\tilde{p}(z) > 0, \qquad \tilde{w}_j = z_j + i\tilde{p}(z_j)^{1/2}$$

so that all real roots of $\tilde{p}$ are simple and the hypothesis of Lemma 4.3 still holds. (One must take care, for example, to perturb a double real root of p into two simple real roots of $\tilde{p}$, not into simple complex conjugate roots.) Then there is a branch of $\tilde{f}$, uniformly near the above branch of f, which satisfies $\tilde{f}(w_j) = w_j$, $j = 1, 2$. This implies that w_1 and w_2 lie on curves $\tilde{\mathscr{C}}_{k_1}$ and $\tilde{\mathscr{C}}_{k_2}$ (in the above notation but with $\tilde{p}$ replacing p) with $k_1 - k_2$ even; or equivalently that the number of roots of $\tilde{p}$ on $(z_1, z_2) \subseteq R$ is a multiple of four. But then the form of the perturbation implies also that the number of roots of p on (z_1, z_2) is a multiple of four.

A similar argument gives the converse: if (4.9) holds then there exists f with $f(w) = \bar{w}$ on $\mathscr{C}$. This then completes the proof of the theorem.

Remark. A consequence of Theorem 4.4 is that for (1) to hold p cannot have a real zero of order 2 (mod 4) which is a local minimum. In particular, $n \equiv 2 \pmod 4$ in (3.1) cannot occur.

5. Further restrictions. Assume henceforth that p satisfies the hypotheses of Theorem 4.4 and condition (4.9) so there is defined a unique f, holomorphic in the complement of I, for which $f(w) = \bar{w}$ on $\mathscr{C} = \partial I$. We wish to study the integral

$$u(w) = -\mathrm{Re} \int^{w} f(\alpha) d\alpha, \qquad w \notin I \tag{5.1}$$

which represents, at least locally, a solution to the obstacle problem.

LEMMA 5.1. The integral (5.1) defines a single-valued harmonic function on the complement of I. On each connected component of $\mathscr{C}$, $u - \psi$ is constant.

Proof. Any closed curve Γ in the complement of I can be deformed to one Γ_0 consisting of arcs in $\mathscr{C}$ and arcs joining different connected components of $\mathscr{C}$. The arcs joining different components can be paired off so as to cancel; that is, Γ_0 is homologous to a union of closed curves $\Gamma_1, \cdots, \Gamma_d$ lying entirely in $\mathscr{C}$. Thus,

$$\int_{\Gamma_0} f(\alpha)\, d\alpha = \sum_{j=1}^{d} \int_{\Gamma_j} f(\alpha)\, d\alpha$$

so we must show

$$\mathrm{Re} \int_{\Gamma_j} f(\alpha) d\alpha = 0.$$

But $f(\alpha) = \bar{\alpha}$ on $\mathscr{C}$, so

$$\mathrm{Re} \int_{\Gamma_j} f(\alpha)\, d\alpha = \mathrm{Re} \int_{\Gamma_j} \bar{\alpha}\, d\alpha = \frac{1}{2} \int_{\Gamma_j} d|\alpha|^2 = 0.$$

To see that $u - \psi$ is constant on each component of $\mathscr{C}$ note that on $\mathscr{C}$

$$d(u - \psi) = \mathrm{Re}\,(\bar{w}\, dw) + \tfrac{1}{2}\, d|w|^2 = 0.$$

In order now for (2) to hold the constant values of $u - \psi$ on various components of $\mathscr{C}$ must be the same, namely zero. We calculate the difference of these constant values for two adjacent components of $\mathscr{C}$ as follows: let $z_1 < z_2$ be real numbers satisfying

$$p(z_1) = p(z_2) = 0 \text{ are zeros of odd order,}$$

$$p(z) \leq 0 \text{ for } z \in (z_1, z_2) \subseteq R.$$

Thus, $z_1, z_2 \in \mathscr{C}$ and $(z_1, z_2) \cap I = \phi$. For (2) to hold it is necessary and sufficient that $(u -$

$\psi)|_{z_1}^{z_2} = 0$ for all such z_1, z_2. We calculate this quantity using (2.5) and noting $h^{-1}(z_j) = z_j$ and $h(z)$ is real on (z_1, z_2).

$$\begin{aligned}(u - \psi)|_{z_1}^{z_2} &= \frac{1}{2}\operatorname{Re}(h(z)^2)|_{z_1}^{z_2} + \frac{1}{2}|h(z)|^2|_{z_1}^{z_2} - 2\operatorname{Re}\int_{z_1}^{z_2} zh'(z)\,dz \\ &= z_2^2 - z_1^2 - 2\int_{z_1}^{z_2} zh'(z)\,dz \\ &= -z_2^2 + z_1^2 + 2\int_{z_1}^{z_2} h(z)\,dz \\ &= -z_2^2 + z_1^2 + \int_{z_1}^{z_2} z + ip(z)^{1/2}\,dz \\ &= 2i\int_{z_1}^{z_2} p(z)^{1/2}\,dz. \end{aligned} \tag{5.2}$$

Therefore, we can state necessary and sufficient conditions for (2) to hold.

THEOREM 5.2. Let p satisfy the hypothesis of Theorem 4.4, and condition (4.9), so that (1) holds. Then with the right choice of a constant of integration (5.1) defines a single-valued harmonic function in the complement of I with $u = \psi$ and grad u = grad ψ on $\mathscr{C}$ (that is, (2) holds) if and only if p has the following property:

if $z_1 < z_2$ are real zeros of p of odd order, and $p(z) \leq 0$ for $z \in (z_1, z_2)$, then

$$\int_{z_1}^{z_2} (-p(z))^{1/2}\,dz = 0$$

where $(-p(z))^{1/2}$ is holomorphic in (z_1, z_2).

We have, thus far, obtained necessary and sufficient conditions for (1) and (2) to hold. With these conditions there is defined uniquely the function u as described in Theorem 5.2. Near points of $\mathscr{C}$ which are not cusps (say where Im $w \neq 0$) we have noted that $u > \psi$ holds in some one-sided neighborhood in the complement of I. But this may not hold near a cusp, and may not hold far away from $\mathscr{C}$—that is, the region where $u \geq \psi$ may not have a uniform size even for perturbations p of p_0 satisfying (1) and (2). The following result of Kinderlehner and Nirenberg describes the situation near cusps arising from odd order zeros of p.

THEOREM 5.3 (Kinderlehrer and Nirenberg [10]). Let $z_0 \in R$ be a zero of p of order $2m + 1$, $m \geq 0$. If m is even then $u \geq \psi$ in some neighborhood of z_0, in the complement of I. But if m is odd there are points in the complement of I arbitrarily near z_0 with $u < \psi$. In particular a necessary condition for (3) is that p have no real zeros of order 3 (mod 4).

Proof. We present only the case where m is odd, say $m = 2a + 1$. Assume $z_0 = 0$ and $p(z) = c^2 z^{4a+3} + O(z^{4a+4})$ where $c > 0$. Then for the branch of h^{-1} with $h^{-1}(w) = \operatorname{Re} w$, $w \in \mathscr{C}$ we have

$$\begin{aligned} h(z) &= z + (-p(z))^{1/2}, \qquad z < 0 \\ &= z + c(-z)^{2a+3/2} + O(z^{2a+2}). \end{aligned}$$

Thus, from (2.5) for $z < 0$ (or equivalently $w = h(z) < 0$)

$$\begin{aligned} u(w) - \psi(w) &= \frac{1}{2} \operatorname{Re} (h(z)^2) + \frac{1}{2}|h(z)|^2 - 2 \operatorname{Re} \int_0^z \zeta h'(\zeta)\, d\zeta \\ &= h(z)^2 - 2 \int_0^z \zeta h'(\zeta)\, d\zeta \\ &= - \frac{4c(-z)^{2a+5/2}}{4a+5} + O(z^{2a+3}) \\ &< 0. \end{aligned}$$

This proves the result. The essential difference when $m = 2a$ is even is the square root of $-p$ takes the opposite sign:

$$h(z) = z - (-p(z))^{1/2}, \qquad z < 0.$$

It thus follows from the above results that the values of n for which perturbations of the cusp

$$\begin{aligned} y^2 &= p_0(x) \\ &= x^n b_0(x)^2, \qquad b_0(0) > 0 \end{aligned} \tag{5.3}$$

can be studied are limited to the values $n \equiv 0$ or 1 (mod 4). If n is a multiple of four then (5.3) describes a double cusp, actually two analytic curves tangent at one point but not crossing. The complement of I consists of two regions, one above the upper curve, the other below the lower one. The function $u(w)$ defined in these regions satisfies $u > \psi$ there, so one may begin to study perturbations of the cusp as described in Sec. 3. Similarly, Theorem 5.3 guarantees that $u > \psi$ in a neighborhood of zero, in the exterior of I, if $n \equiv 1$ (mod 4).

Assume then $n \equiv 0$ or 1 (mod 4) and consider a disc Δ about zero with small radius; its circumference $\partial\Delta$ intersects the curve $\mathscr{C}_0$, corresponding to the unperturbed problem $y^2 = p_0(x)$, transversally. At the intersection points $u - \psi$ vanishes exactly to second order:

$$\left.\begin{aligned} &u - \psi = \frac{d}{ds}(u - \psi) = 0 \\ &\frac{d^2}{ds^2}(u - \psi) > 0 \end{aligned}\right\} \text{ on } \partial\Delta \cap \mathscr{C}_0 \tag{5.4}$$

where d/ds means differentiation along $\partial\Delta$. And because $n \equiv 0$ or 1 (mod 4) we have, in fact

$$u - \psi > 0 \quad \text{on} \quad \partial\Delta - I_0. \tag{5.5}$$

For the perturbed problem, with $\mathscr{C}$ described by $y^2 = p(x)$, it follows from transversality arguments that (5.4), (5.5) continue to hold, with $\mathscr{C}_0$ and I_0 replaced with $\mathscr{C}$ and I.

Therefore, if p is such that (1) and (2) hold, as described in Theorem 5.2, the only way (3) could fail is if

$$\min\,(u(w) - \psi(w)) \le 0, \qquad w \in \overline{\Delta - I}.$$

From the above remarks this cannot happen on $\partial\Delta - I$; and also $u - \psi = 0$ on $\mathscr{C} = \partial I$.

Therefore, any negative minimum w_0 of $u - \psi$ must occur in the interior of $\Delta - I$, and in particular

$$\operatorname{grad}(u - \psi) = 0 \tag{5.6}$$

at that point. From (5.1), (5.6) is equivalent to $f(w_0) = \bar{w}_0$, that is

$$h^{-1}(w_0) = \operatorname{Re} w_0 .$$

This implies $z_0 = h^{-1}(w_0)$ is real, and $p(z_0) \geq 0$; and this says that either $w_0 \in \mathscr{C}$ in which case $u(w_0) - \psi(w_0) = 0$ is not negative, or that

$$\begin{aligned} &p(z_0) = 0 \quad \text{and} \\ &z_0 \in R \text{ is a local maximum of } p. \end{aligned} \tag{5.7}$$

In case (5.7) holds we observe that $w_0 = z_0 \notin I$ by (3.4) (this is the reason for taking the closure of the interior in that formula). Therefore, it suffices that $u(z_0) - \psi(z_0) \geq 0$ at each z_0 given by (5.7) for (3) to hold. The calculation of $u(z_0) - \psi(z_0)$ is similar to the calculation immediately preceding Theorem 5.2.

In fact, let $z_1 < z_2$ be real zeros of p such that $p \leq 0$ in (z_1, z_2); then as in (5.2) $h^{-1}(z_j) = z_j$ and

$$(u - \psi)|_{z_1}^{z_2} = 2i \int_{z_1}^{z_2} p(z)^{1/2} dz = 2 \int_{z_1}^{z_2} (-p(z))^{1/2} dz.$$

Moreover, the choice of sign for $(-p(z))^{1/2}$ is as follows: if z_1 a zero of order 1 (mod 4) then $(-p(z))^{1/2} > 0$ immediately to the right of z_1; and if z_2 is such a zero, $(-p(z))^{1/2} < 0$ immediately to the left of z_2. This is as in the proof of Theorem 5.3, and gives a consistent choice of sign for $(-p)^{1/2}$ as long as (1) and (2) hold and p has no zeros of order 3 (mod 4). If z_1 or z_2 is an odd-order zero of p then it lies on $\mathscr{C}$, so $u - \psi = 0$ at that point. In order for $u - \psi \geq 0$ at the other point,

$$\int_{z_1}^{z_2} (-p(z))^{1/2}\, dz$$

must have the correct sign. This gives necessary and sufficient conditions for (3) to hold.

THEOREM 5.4. Let (1) and (2) hold as described in Theorems 4.4 and 5.2. Then (3) holds also if and only if p has the following properties:

(a) p has no zeros of order 3 (mod 4);
(b) on a maximal open interval of R where $p \leq 0$, define $(-p(z))^{1/2}$ to be holomorphic, positive immediately to the right of an odd-order zero, and negative immediately to the left of such a zero; then whenever z_0, $z_1 \in R$ are even-order and odd-order zeros respectively with $p \leq 0$ in between, we have

$$\int_{z_0}^{z_1} (-p(z))^{1/2} dz \leq 0.$$

We have, therefore, described in Theorems 4.4, 5.2 and 5.4 necessary and sufficient conditions for the curve $\mathscr{C}$ to be a free boundary for the obstacle problem, defined in some uniform neighborhood of the origin. We have noted the number n (5.3) describing the unperturbed cusp must equal 0 or 1 (mod 4); the simplest non-trivial case is thus $n = 4$, and the next simplest $n = 5$.

6. The cases $n = 4$ and 5. We give a more detailed analysis of these simpler cases. Recall the situation for general n: we are considering perturbations

$$y^2 = p(x) = (x^n + \alpha_{n-1}x^{n-1} + \cdots + \alpha_0)b(x)^2 \tag{6.1}$$

of the cusp

$$y^2 = p_0(x) = x^n b_0(x)^2$$

where p and p_0 are real analytic, $b_0(0) > 0$, $\alpha_j \in R$ are near zero, and b and b_0 are uniformly near on some small disc:

$$\sup|b(x) - b_0(x)| \leq \epsilon, \qquad x \in C, \qquad |x| \leq \delta.$$

Here $n \equiv 0$ or 1 (mod 4).

For small values of n it is easy to write down normal forms for p to satisfy the conditions derived above; the parameters α, β, γ and δ are real and near zero. When $n = 4$ there are two cases:

(A_4) $[(x-\alpha)^2 + \beta^2]^2 b(x)^2$; I consists of two regions on either side of the real axis;

(B_4) $(x-\alpha)(x-\beta)^2(x-\gamma)b(x)^2$, $\alpha \leq \beta \leq \gamma$,

$$\int_\alpha^\gamma (x-\alpha)^{1/2}(x-\beta)(\gamma - x)^{1/2} b(x)\, dx = 0$$

uniquely determines $\beta = \beta(\alpha, \gamma, b)$; I consists of two regions on either side of the vertical line $\operatorname{Re} x = \beta$.

If, in addition to symmetry about the real axis, we require symmetry about the imaginary axis these formulas are simpler:

$$(A_4^{\text{sym}}) \quad (x^2 + \beta^2)^2 b(x)^2, \qquad b(x) \equiv b(-x);$$
$$(B_4^{\text{sym}}) \quad x^2(x^2 - \alpha^2) b(x)^2, \qquad b(x) \equiv b(-x).$$

In this case necessary and sufficient conditions for

$$p(x) = (x^4 + \alpha_2 x^2 + \alpha_0) b(x)^2$$

are that b is an even function and that

$$\alpha_0 = \phi(\alpha_2) = \begin{cases} 0, & \alpha_2 \leq 0 \\ \alpha_2^2/4, & \alpha_2 \geq 0. \end{cases} \tag{6.2}$$

The function ϕ describes a C^1 curve in the (α_2, α_0)-parameter space tangent to the α_2-axis at the origin.

When $n = 5$ three normal forms occur:

(A_5) $[(x-\alpha)^2 + \beta^2]^2(x-\gamma)b(x)^2$; I consists of one region opening out to the right;

(B_5) $(x-\alpha)^2(x-\beta)^2(x-\gamma)b(x)^2$, $\alpha \leq \beta \leq \gamma$;

$$\int_\alpha^\gamma (x-\alpha)(x-\beta)(\gamma - x)^{1/2} b(x)\, dx \geq 0;$$

I again forms one region opening on the right;

(C_5) $(x-\alpha)(x-\beta)(x-\gamma)^2(x-\delta)b(x)^2$, $\alpha \leq \beta \leq \gamma \leq \delta$,

$$\int_\beta^\delta (x-\alpha)^{1/2}(x-\beta)^{1/2}(x-\gamma)(\delta - x)^{1/2} b(x) dx = 0$$

uniquely determines $\gamma = \gamma(\alpha, \beta, \delta, b)$; I forms a large region on the right and a small island to the left.

The normal forms for $n = 4$ each have two independent real parameters; for $n = 5$ there are three. In general, there are $n/2$ such parameters when $n \equiv 0 \pmod 4$ and $(n + 1)/2$ when $n \equiv 1 \pmod 4$. It is interesting to examine the set of admissible parameters $\alpha_{n-1}, \cdots, \alpha_0$ in (6.1) given by the normal forms for, say, a fixed b; in the $n = 4$ symmetric case this set was a C^1 curve (6.2) parameterized by α_2. Consider the general case for $n = 4$.

For the (A_4) normal form a direct calculation gives

$$\alpha_3 = -4\alpha \qquad \alpha_2 = 6\alpha^2 + 2\beta^2, \qquad \alpha_1 = -4\alpha(\alpha^2 + \beta^2), \qquad \alpha_0 = (\alpha^2 + \beta^2)^2.$$

The range of α_2, α_3 as α and β vary is precisely given by

$$\alpha_2 \geq \tfrac{3}{8}\,\alpha_3^2 \tag{6.3}$$

and in that region α_0 and α_1 are uniquely determined:

$$\alpha_1 = \tfrac{1}{8}\,\alpha_3(4\alpha_2 - \alpha_3^2), \qquad \alpha_0 = \tfrac{1}{64}\,(4\alpha_2 - \alpha_3^2)^2. \tag{6.4}$$

For the (B_4) form let $\beta = \alpha + \lambda\sigma$ and $\gamma = \alpha + \sigma$ where $\sigma \geq 0$ and $\lambda = \lambda(\alpha, \sigma, b) \in [0, 1]$; the defining condition in (B_4) for λ becomes after a change of variables

$$\lambda = \frac{\displaystyle\int_0^1 x^{3/2}(1 - x)^{1/2} b(\sigma x + \alpha)\,dx}{\displaystyle\int_0^1 x^{1/2}(1 - x)^{1/2} b(\sigma x + \alpha)\,dx} = \tfrac{1}{2} + O(\sigma + |\alpha|), \tag{6.5}$$

so let

$$\mu = \mu(\alpha, \sigma, b) = \lambda - \tfrac{1}{2} = O(|\alpha| + \sigma).$$

Clearly μ is smooth in (α, σ). Again, by direct calculation we have

$$\begin{aligned} \alpha_3 &= -\alpha - 2\beta - \gamma = -4\alpha - (2 + 2\mu)\,\sigma, \\ \alpha_2 &= \alpha\gamma + 2\beta\,(\alpha + \gamma) + \beta^2 = 6\alpha^2 + (6 + 6\mu)\alpha\sigma + (\tfrac{5}{4} + 3\mu + \mu^2)\sigma^2, \\ \alpha_1 &= -2\alpha\beta\gamma - \beta^2\,(\alpha + \gamma), \qquad \alpha_0 = \alpha\beta^2\gamma, \end{aligned} \tag{6.6}$$

and so

$$\tfrac{3}{8}\alpha_3^2 - \alpha_2 = (\tfrac{1}{4} + \tfrac{1}{2}\,\mu^2)\sigma^2 \geq 0.$$

It is easy to see that this gives a one-to-one correspondence between (α, σ) with $\sigma \geq 0$ and (α_2, α_3) with $\tfrac{3}{8}\,\alpha_3^2 - \alpha_2 \geq 0$; in fact introduce variables ξ and η, and consider the system of equations

$$\xi = [\tfrac{1}{4} + \tfrac{1}{2}\,\mu\,(\alpha, \sigma)^2]^{1/2}\sigma, \qquad \eta = -4\alpha - [2 + 2\mu\,(\alpha, \sigma)]\,\sigma. \tag{6.7}$$

We see (6.7) can be solved by the implicit function theorem to obtain

$$\alpha = \alpha^*(\xi, \eta) = \frac{-1}{4}\;\eta - \xi + \cdots, \qquad \sigma = \sigma^*(\xi, \eta) = 2\xi + \cdots \tag{6.8}$$

and so

$$\alpha = \alpha^*\,((\tfrac{3}{8}\,\alpha_3^2 - \alpha_2)^{1/2}, \alpha_3), \qquad \sigma = \sigma^*\,((\tfrac{3}{8}\,\alpha_3^2 - \alpha_2)^{1/2}, \alpha_3). \tag{6.9}$$

Note here the following facts:

$$\alpha^*(0, \eta) = \frac{-1}{4}\eta,$$

$$\frac{\partial\alpha^*}{\partial\xi}(0, \eta) = \frac{-1}{2}\left[1 + \mu\left(\frac{-1}{4}\eta, 0\right)\right]\left[\frac{1}{4} + \frac{1}{2}\mu\left(\frac{-1}{4}\eta, 0\right)^2\right]^{-1/2}, \tag{6.10}$$

$$\sigma^*(0, \eta) = 0,$$

$$\frac{\partial\sigma^*}{\partial\xi}(0, \eta) = \left[\frac{1}{4} + \frac{1}{2}\mu\left(\frac{-1}{4}\eta, 0\right)^2\right]^{-1/2}.$$

From the expressions

$$\beta = \alpha + \tfrac{1}{2}\sigma + \mu(\alpha, \sigma)\sigma, \qquad \gamma = \alpha + \sigma \tag{6.11}$$

for β and γ it is now possible to consider these as functions $\beta^*(\xi, \eta)$, $\gamma^*(\xi, \eta)$ of ξ and η; and from (6.6), α_1 and α_0 are given as functions $\alpha_1^*(\xi, \eta)$ and $\alpha_0^*(\xi, \eta)$. Therefore, in the region $\frac{3}{8}\alpha_3^2 \geq \alpha_2$, α_1 and α_0 are uniquely determined from (α_2, α_3):

$$\begin{aligned} \alpha_1 &= \alpha_1^*((\tfrac{3}{8}\alpha_3^2 - \alpha_2)^{1/2}, \alpha_3) \\ \alpha_0 &= \alpha_0^*((\tfrac{3}{8}\alpha_3^2 - \alpha_2)^{1/2}, \alpha_3). \end{aligned} \tag{6.12}$$

This is as was shown for the complementary region (6.3), (6.4). In fact, the functions (6.4) and (6.12) defining α_1 and α_0 in the two regions of the (α_2, α_3)-plane fit together to form a C^1 function.

To prove this last fact, it is enough to check that the values of the functions (6.4), (6.12) and their α_2-directional derivatives are equal along the curve $\alpha_2 = \frac{3}{8}\alpha_3^2$; so it is necessary and sufficient that

$$\begin{aligned} &\alpha_1^*(0, \eta) = \frac{1}{16}\eta^3, && \frac{\partial\alpha_1^*}{\partial\xi}(0, \eta) = 0, && \frac{\partial^2\alpha_1^*}{\partial\xi^2}(0, \eta) = -\eta, \\ &\alpha_0^*(0, \eta) = \frac{1}{256}\eta^4, && \frac{\partial\alpha_0^*}{\partial\xi}(0, \eta) = 0, && \frac{\partial^2\alpha_0^*}{\partial\xi^2}(0, \eta) = \frac{-1}{8}\eta^2. \end{aligned} \tag{6.13}$$

Showing this is a direct though tedious calculation. The following identities simplify matters somewhat: from (6.6) and (6.7)

$$\begin{aligned} &2\alpha^*(\xi, \eta) + 2\beta^*(\xi, \eta) + \sigma^*(\xi, \eta) \\ &\quad = 4\alpha^*(\xi, \eta) + 2\sigma^*(\xi, \eta) + 2\mu\sigma^*(\xi, \eta) \\ &\quad = -\eta \end{aligned}$$

where $\mu = \mu(\alpha^*(\xi, \eta), \sigma^*(\xi, \eta))$. In particular,

$$\beta^*(0, \eta) = \frac{-1}{4}\eta,$$

$$\frac{\partial^m\beta^*(0, \eta)}{\partial\xi^m} = -\frac{\partial^m\alpha^*(0, \eta)}{\partial\xi^m} - \frac{1}{2}\frac{\partial^m\sigma^*(0, \eta)}{\partial\xi^m}, \qquad m \geq 1.$$

These expressions and the corresponding one

$$\frac{\partial^m \gamma^*(0, \eta)}{\partial \xi^m} = \frac{\partial^m \alpha^*(0, \eta)}{\partial \xi^m} + \frac{\partial^m \sigma^*(0, \eta)}{\partial \xi^m}, \qquad m \geq 0,$$

for γ (6.11) allow one to check (6.13) by writing everything in terms of α^*, σ^* and their first derivatives (6.10). This is now straightforward so we omit the details.

Therefore, we see that α_0 and α_1 are uniquely determined C^1 functions of α_2 and α_3 near the origin; also the first derivatives of these functions at $\alpha_2 = \alpha_3 = 0$ are zero. This is summarized in the following theorem.

THEOREM 6.1. Let $\Sigma = \Sigma_b$ denote the set of $(\alpha_0, \alpha_1, \alpha_2, \alpha_3) \in R^4$ near the origin for which the function $p(x)$ (6.1) with $n = 4$ is in one of the normal forms (A_4) or (B_4). Then Σ is a C^1 (but not C^2) manifold of dimension two, tangent to the (α_2, α_3)-plane at the origin. The normal form (A_4) occurs when $\alpha_2 \geq \frac{3}{8} \alpha_3^2$ and the form (B_4) when $\alpha_2 \leq \frac{3}{8} \alpha_3^2$.

We remark that if b is allowed to vary in the appropriate way, say in the Banach space of functions holomorphic in the disc $|z| < \delta$ in C, and continuous in $|z| \leq \delta$, with the supremum norm, then the manifold Σ_b varies in a C^1 manner. That is, the functions $\alpha_1^*(\alpha_2, \alpha_3, b)$ and $\alpha_0^*(\alpha_2, \alpha_3, b)$ describing Σ_b are C^1 in (α_2, α_3, b). For general n we have the following conjecture.

Conjecture.[1] For $n > 4$ the result analogous to Theorem 6.1 holds. When $n = 4m$, $\alpha_0, \cdots, \alpha_{2m-1}$ are C^1 functions of $(\alpha_{2m}, \cdots, \alpha_{4m-1}) \in R^{2m}$; when $n = 4m + 1$, $\alpha_0, \cdots, \alpha_{2m-1}$ are C^1 functions of $(\alpha_{2m}, \cdots, \alpha_{4m}) \in R^{2m+1}$.

In the next section we shall return to the obstacle problem; some formal calculations there strongly support this conjecture.

7. **Generic bifurcation—a formal basis.** Below we give a formal perturbation analysis of the obstacle problem of Sec. 2. We state again that our interest and motivation goes beyond a formal analysis, however—such calculations are a necessary first step before a rigorous treatment (such as from the restricted unfolding viewpoint [4, 5, 6, 7, 16]) can be made. In particular, the formal analysis is necessary to obtain candidates u^- and y^+, as in [6], giving *a priori* estimates for the true solution u and location of the free boundary ∂I.

Recall the situation as in Sec. 2. On the smooth bounded domain $\Omega \subseteq R^2$ is given the boundary condition

$$g: \partial\Omega \to R$$

so that the membrane there is above the obstacle:

$$\psi(w) < g(w) \quad \text{on} \quad \partial\Omega$$

$$\psi(w) = \frac{-1}{2} |w|^2.$$

The solution $u(w)$ satisfies

$$u: \bar{\Omega} \to R,$$

$$u = g \quad \text{on} \quad \partial\Omega,$$

[1] We have recently proved this conjecture and obtained some bifurcation results for general n [17].

$$u \geq \psi \quad \text{in} \quad \Omega,$$
$$\Delta u = 0 \quad \text{in} \quad \Omega - I,$$
$$I = \{w \in \Omega |\, u(w) = \psi(w)\},$$
$$u \quad \text{is} \quad C^1 \quad \text{in} \quad \bar{\Omega}.$$

The fact that u is C^1 across the free boundary ∂I means, in particular, that u solves the Dirichlet problem

$$\Delta u = 0 \quad \text{in} \quad \Omega - I,$$
$$u = g \quad \text{on} \quad \partial\Omega,$$
$$u = \psi \quad \text{on} \quad \partial I$$

and, in addition

$$\partial u/\partial \nu = \partial\psi/\partial\nu \quad \text{on} \quad \partial I$$

where ν is the outward unit normal on ∂I.

Now consider a particular problem with datum g_0 and solution u_0; let g_0 be perturbed so there is a resulting perturbation of u_0:

$$g = g_0 + \epsilon g_1 + o(\epsilon), \qquad u = u_0 + \epsilon u_1 + o(\epsilon);$$

the free boundary ∂I_0 is also perturbed to ∂I:

$$w \in \partial I_0 \quad \text{moves to} \quad w + \epsilon N(w)\nu + o(\epsilon) \in \partial I$$
$$N: \partial I_0 \to R.$$

The differential equation describing u_1 and N was given by Schaeffer [13]; one first has

$$\Delta u_1 = 0 \quad \text{in} \quad \Omega - I_0, \qquad u_1 = g_1 \quad \text{on} \quad \partial\Omega. \tag{7.1}$$

Then differentiating the formulas

$$u_0(w + \epsilon N(w)\nu) + \epsilon u_1(w) = \psi(w + \epsilon N(w)\nu) + o(\epsilon),$$

$$\frac{\partial u_0}{\partial \nu}(w + \epsilon N(w)\nu) + \epsilon \frac{\partial u_1}{\partial \nu}(w) = \frac{\partial \psi}{\partial \nu}(w + \epsilon N(w)\nu) + o(\epsilon)$$

and noting $u_0 - \psi$ vanishes to second order on ∂I_0 gives

$$u_1 = 0 \quad \text{on} \quad \partial I_0 \tag{7.2}$$

$$\left.\begin{aligned} \frac{\partial u_1}{\partial \nu} &= \Delta(\psi - u_0)N \\ &= -2N \end{aligned}\right. \quad \text{on} \quad \partial I_0. \tag{7.3}$$

Thus, (7.1), (7.2) is a Dirichlet problem for u_1, and (7.3) uniquely determines N.

We wish to relate the function $N(w)$ describing the perturbation of the free boundary to a perturbation of the function $p(z)$ describing a cusp on the boundary. We assume henceforth, as was the case in Secs. 3 through 6, symmetry with respect to the $x = \text{Re } w$

axis; if Ω and g are symmetric, then u and I will also be. Let the unperturbed boundary ∂I_0 contain a cusp, given locally as in Sec. 3:

$$y^2 = p_0(x) = x^n b_0(x)^2.$$

The perturbed boundary is given to first order as

$$\begin{aligned} y^2 &= p_0(x) + \epsilon p_1(x) \\ &= x^n b_0(x)^2 + \epsilon(\beta_{n-1}x^{n-1} + \cdots + \beta_0)b_0(x)^2 \\ &+ 2\epsilon x^n b_0(x)b_1(x), \end{aligned} \tag{7.4}$$

where ϵb_1 is the perturbation of b_0, and $\epsilon\beta_j$ represents the coefficient α_j. At $w_0 = x_0 + iy_0 \in \partial I_0$ the unit outward normal ν to I_0 is

$$\nu = [(p_0'(x_0))^2 + 4p_0(x_0)]^{-1/2}(-p_0'(x_0), 2p_0(x_0)^{1/2}).$$

(Suppose here $p(x_0) > 0$ and choose $p(x_0)^{1/2} = y_0 > 0$.) Let the point w_0 move under the perturbation to

$$(x_0 + \epsilon x_1, y_0 + \epsilon y_1) = (x_0, y_0) + \epsilon N(w_0)\nu;$$

then to first order

$$(y_0 + \epsilon y_1)^2 = p_0(x_0 + \epsilon x_1) + \epsilon p_1(x_0).$$

Equating terms in ϵ gives

$$x_1 = -N(w_0)[(p_0'(x_0))^2 + 4p_0(x_0)]^{-1/2}p_0'(x_0),$$

$$y_1 = 2N(w_0)[(p_0'(x_0))^2 + 4p_0(x_0)]^{-1/2}p_0(x_0)^{1/2},$$

$$2y_0y_1 = p_0'(x_0)x_1 + p_1(x_0),$$

so solving for p_1 gives

$$p_1(x_0) = N(w_0)[(p_0'(x_0))^2 + 4p_0(x_0)]^{1/2}. \tag{7.5}$$

There remains now the question of analyzing the function $N(w_0)$ near the cusp. Two cases occur:

$n = 4m$. Here ∂I_0 forms two analytic curves tangent at the origin, and not crossing. The harmonic function u_1 obtained by solving the Dirichlet problem (7.1), (7.2) is analytic up to ∂I_0 so, in particular, $N = -1/2(\partial u_1/\partial\nu)$ is an analytic function of the parameter $x = \operatorname{Re} w$:

$$N(w) = \frac{-1}{2}\frac{\partial u_1}{\partial \nu} = \frac{-1}{2}(A_0 + A_1x + A_2x^2 + \cdots).$$

From (7.4) and (7.5) $p_1(x)$ is determined:

$$\begin{aligned} p_1(x) &= (\beta_{4m-1}x^{4m-1} + \cdots + \beta_0)b_0(x)^2 + 2x^{4m}b_0(x)b_1(x) \\ &= -(A_0 + A_1x + A_2x^2 + \cdots)(x^{2m}b_0(0) + O(x^{2m+1})). \end{aligned}$$

We see that this forces

$$\beta_0 = \beta_1 = \cdots = \beta_{2m-1} = 0$$

while there is no essential restriction on the values of $\beta_{2m}, \cdots, \beta_{4m-1}$. This fact is the basis for the conjecture of Sec. 6, as β_j represents the first-order variation in α_j from $\alpha_j = 0$.

n = 4m + 1. Since ∂I_0 forms a sharp cusp at the origin in this case u_1 may not be analytic up to this cusp—$(\partial u_1/\partial \nu)$ is unbounded in general. It is necessary to unfold the cusp with the conformal mapping

$$w = k(\zeta) = \zeta^2 + ip(\zeta^2)^{1/2}$$

which locally maps the real ζ-axis onto ∂I_0, taking $\operatorname{Im} \zeta > 0$ to the exterior of I_0. Then

$$v(\zeta) = u_1(k(\zeta))$$

is harmonic for $\operatorname{Im} \zeta > 0$, even in $\operatorname{Re} \zeta$, and vanishes for ζ real. Letting ν_ζ and ν_w denote unit normals to $\operatorname{Im} \zeta = 0$ and I_0 respectively, we have

$$\begin{aligned} N(w) &= \frac{-1}{2} \frac{\partial u_1}{\partial \nu_w} = \frac{-1}{2} \frac{\partial v}{\partial \nu_\zeta} |k'(\zeta)|^{-1} \\ &= \frac{-1}{2} \frac{\partial v}{\partial \nu_\zeta} [4 + p(\zeta^2)^{-1}(p'(\zeta^2))^2]^{-1/2} \zeta^{-1} \\ &= \frac{-1}{2} \frac{\partial v}{\partial \nu_\zeta} [4 + p(x)^{-1}(p'(x))^2]^{-1/2} \zeta^{-1} \end{aligned} \tag{7.6}$$

where $\zeta^2 = x = \operatorname{Re} w$ in the last formula. Also, because $v(\zeta)$ is even in $\operatorname{Re} \zeta$ and analytic, for ζ real

$$\begin{aligned} \frac{\partial v}{\partial \nu_\zeta} \zeta^{-1} &= (B_0 + B_1 \zeta^2 + B_2 \zeta^4 + \cdots)\zeta^{-1} \\ &= (B_0 + B_1 x + B_2 x^2 + \cdots)x^{-1/2}. \end{aligned}$$

Therefore, from (7.5) and (7.6)

$$\begin{aligned} p_1(x) &= (\beta_{4m} x^{4m} + \cdots + \beta_0) b_0(x)^2 + 2x^{4m+1} b_0(x) b_1(x) \\ &= \frac{-1}{2} (B_0 + B_1 x + B_2 x^2 + \cdots)(x^{2m} b_0(0) + O(x^{2m+1})). \end{aligned}$$

Again,

$$\beta_0 = \beta_1 = \cdots = \beta_{2m-1} = 0$$

while $\beta_{2m}, \cdots, \beta_{4m}$ are essentially arbitrary, in accordance with the conjecture.

In summary, the formal perturbation scheme above associates to a perturbation $g_0 + \epsilon g_1$ of the boundary condition a corresponding perturbation $p_0 + \epsilon p_1$ of the function describing the free boundary near a cusp. Moreover, p_1 depends linearly on g_1 and has the form

$$p_1(x) = (\beta_{2m} x^{2m} + \cdots + \beta_{4m-1} x^{4m-1}) b_0(x)^2 + O(x^{4m})$$

when

$$p_0(x) = x^{4m} b_0(x)^2$$

and the form

$$p_1(x) = (\beta_{2m}x^{2m} + \cdots + \beta_{4m}x^{4m})b_0(x)^2 + O(x^{4m+1})$$

when

$$p_0(x) = x^{4m+1}b_0(x)^2.$$

In particular, the coefficients β_j may be regarded as derivatives $\beta_j = \partial\alpha_j/\partial\epsilon$ of α_j for a variation of g in the direction g_1.

8. Generic bifurcation—one-parameter unfoldings. We continue here the formal analysis of Sec. 7. Given a perturbation $g(w,\epsilon) = g_0(w) + \epsilon g_1(w) + \cdots$ of the boundary condition, quantities $\beta_j = \partial\alpha_j/\partial\epsilon$ depending linearly on g_1 are calculated. How does this knowledge, coupled with the knowledge of the admissible normal forms for p, indicate how the singularity unfolds? In particular, what normal forms occur for generic perturbations, and how does one characterize such perturbations? The analysis in this section concerns the cases $n = 4$ and 5, and describes the generic unfoldings for perturbations depending on one scalar parameter ϵ. In Sec. 9, the case of a two-parameter perturbation is considered where the parameters ϵ_1 and ϵ_2 vary independently in a neighborhood of the origin.

Theorem 6.1 makes the case $n = 4$ very clear; since the boundary between the cases (A_4) and (B_4) is the curve $\alpha_2 = \frac{3}{8}\alpha_3^2$, the natural generic condition is that (α_2, α_3) move across the curve transversally, that is, $\partial\alpha_2/\partial\epsilon \neq 0$. For ϵ on one side of zero (A_4) should be encountered, and (B_4) for ϵ on the other side. In fact, set

$$\alpha_2 = \epsilon, \qquad \alpha_3 = -4c\epsilon + o(\epsilon)$$

$$(\alpha_0, \alpha_1, \alpha_2, \alpha_3) \in \Sigma, \quad \text{so} \quad \alpha_0, \alpha_1 = o(\epsilon)$$

for some real constant c. For $\epsilon > 0$ p has the normal form (A_4); since $\alpha_3 = -4\alpha$ and $\alpha_2 = 6\alpha^2 + 2\beta^2$ it follows $\alpha \sim c\epsilon$ and $\beta^2 \sim \frac{1}{2}\epsilon$. Therefore,

$$p(x) = [(x - c\epsilon + \cdots)^2 + \tfrac{1}{2}\epsilon + \cdots]^2 b(x,\epsilon)^2, \qquad \epsilon > 0.$$

(We let b depend on ϵ, presumably in a C^1 manner.) When $\epsilon < 0$ (B_4) occurs; in that case from (6.8) and (6.9) $\alpha \sim -(-\epsilon)^{1/2}$ and $\gamma = \alpha + \sigma \sim (-\epsilon)^{1/2}$ and from (6.6) and (6.11)

$$\beta = \frac{-1}{4}\alpha_3 + \frac{1}{2}\mu\sigma = c\epsilon + \frac{1}{2}\mu\sigma + o(\epsilon).$$

Expanding $\mu = \lambda - \frac{1}{2}$ (6.5) to first order about $\alpha = \sigma = 0$ and using the asymptotic forms for α and σ immediately above then shows $\beta \sim c'\epsilon$ for some constant c'. Thus, p has the form

$$p(x) = [(x - c'\epsilon + \cdots)^2 + \epsilon + \cdots](x - c'\epsilon + \cdots)^2 b(x,\epsilon)^2, \qquad \epsilon < 0.$$

Roughly, the parameter α_3 causes a translation of the free boundary to the left or right, as seen by the term $x - c\epsilon + \cdots$ above. The parameter α_2 describes, near the cusp, whether the membrane moves down (A_4) or up (B_4) relative to the obstacle. A particular case when the generic hypothesis $\partial\alpha_2/\partial\epsilon \neq 0$ holds is when $\partial g/\partial\epsilon = g_1$ is of one sign; if $g_1 < 0$, so the membrane is pushed down, then applying the maximum principle to u_1 shows $\partial\alpha_2/\partial\epsilon = \beta_2 > 0$.

Now let $n = 5$ and $g = g_0 + \epsilon g_1 + \cdots$ again depend on the scalar parameter ϵ, so quantities β_2, β_3 and β_4 are defined, and $\beta_0 = \beta_1 = 0$. The generic hypothesis we consider here is $\beta_2 \neq 0$; in this case the normal form (A_5) appears for both $\epsilon < 0$ and $\epsilon > 0$, as the following theorem indicates.

THEOREM 8.1. Let

$$p(x) = (x^5 + \alpha_4 x^4 + \cdots + \alpha_0) b(x)^2$$

satisfy one of the normal forms (A_5), (B_5), (C_5) and be such that

$$\max \{|\alpha_3|, |\alpha_4|, \| b - b_0 \|_\delta\} \leq C|\alpha_2| \tag{8.1}$$

for some constants C and δ, where $\| \cdot \|_\delta$ represents the supremum norm in the Banach space of functions holomorphic in some complex disc $|x| \leq \delta$. Then for $|\alpha_0|$, $|\alpha_1|$ and $|\alpha_2|$ sufficiently small p cannot be in the normal forms (B_5) or (C_5).

Proof. First translate the variable x by replacing it with $x - \alpha_4/5$. This has the effect of eliminating α_4, transforming α_3 and α_2 to $\tilde{\alpha}_3 = \alpha_3 + O(\alpha_4^2)$ and $\tilde{\alpha}_2 = \alpha_2 + O(\alpha_3^2 + \alpha_4^2)$, and transforming b to $\tilde{b}(x) = b(x - \alpha_4/5)$. The hypothesis (8.1) still holds, possibly with larger C and smaller δ. Also, p is still in the same normal form after such a transformation.

Therefore, without loss of generality $\alpha_4 = 0$. Suppose first p is in the form (B_5); so

$$\begin{aligned} \alpha_4 &= -2\alpha - 2\beta - \gamma = 0, \\ \alpha_3 &= -2\alpha^2 - \alpha\gamma - \tfrac{3}{4}\gamma^2, \\ \alpha_2 &= \alpha^2\gamma + \tfrac{1}{2}\alpha\gamma^2 - \tfrac{1}{4}\gamma^3. \end{aligned} \tag{8.2}$$

The conditions $\alpha \leq \beta = -\alpha - \gamma/2 \leq \gamma$ restrict α and γ:

$$\rho = -\frac{\gamma}{\alpha} \in \left[\frac{2}{3}, 4\right], \qquad \alpha \leq 0. \tag{8.3}$$

A further restriction is the inequality

$$\int_\alpha^\gamma (x - \alpha)(x - \beta)(\gamma - x)^{1/2} b(x)\, dx \geq 0$$

which implies

$$\begin{aligned} \frac{\beta}{\gamma - \alpha} &\leq \frac{\int_\alpha^\gamma (x - \alpha) x (\gamma - x)^{1/2} b(x)\, dx}{(\gamma - \alpha) \int_\alpha^\gamma (x - \alpha)(\gamma - x)^{1/2} b(x)\, dx} \\ &= \frac{\int_0^1 y^2 (1 - y)^{1/2} dy + \left(\frac{\alpha}{\gamma - \alpha}\right) \int_0^1 y(1 - y)^{1/2} dy}{\int_0^1 y(1 - y)^{1/2} dy} + O(|\alpha| + |\alpha - \gamma| + |\alpha_2|) \\ &= \frac{4}{7} + \frac{\alpha}{\gamma - \alpha} + O(|\alpha| + |\alpha - \gamma| + |\alpha_2|). \end{aligned} \tag{8.4}$$

The error term of order $|\alpha| + |\alpha - \gamma| + |\alpha_2|$ arises from the charge of variable in the integration and from $\|b - b_0\|_\delta = O(|\alpha_2|)$. In terms of ρ (8.4) is

$$\rho \geq \frac{4}{3} + O(|\alpha| + |\alpha - \gamma| + |\alpha_2|). \tag{8.5}$$

Now Eq. (8.2) for α_2 gives

$$\alpha_2 = \alpha^3\left(-\rho + \frac{1}{2}\rho^2 + \frac{1}{4}\rho^3\right) \tag{8.6}$$

and so for some $C_1 > 0$, $|\alpha_2| \leq C_1|\alpha|^3$. But then

$$\alpha_3 = \alpha^2\left(-2 + \rho - \frac{3}{4}\rho^2\right), \qquad -2 + \rho - \frac{3}{4}\rho^2 < 0 \quad \text{in} \quad \left[\frac{4}{3}, 4\right] \tag{8.7}$$

implies $|\alpha_3| \geq C_2|\alpha|^2 \geq C_3|\alpha_2|^{2/3}$, contradicting the hypothesis (8.1).

The case for p in the form (C_5) proceeds similarly. Again assume $\alpha_4 = 0$ so

$$\alpha_4 = -\alpha - \beta - 2\gamma - \delta = 0. \tag{8.8}$$

From this fact it follows that

$$\begin{aligned} \alpha_3 &= -10\beta^2 + (\beta - \alpha)(\beta - \delta) + (\beta - \gamma)^2 + 2(\beta - \alpha)(\beta - \gamma) + 2(\beta - \gamma)(\beta - \delta), \\ \alpha_2 &= 20\beta^3 - 3\beta(\alpha_3 + 10\beta^2) + 2(\beta - \alpha)(\beta - \gamma)(\beta - \delta) \\ &\qquad + (\beta - \alpha)(\beta - \gamma)^2 + (\beta - \gamma)^2(\beta - \delta). \end{aligned} \tag{8.9}$$

The integral condition in (C_5) determining γ is

$$\begin{aligned} \frac{-1}{2}(\alpha + \beta + \gamma) &= \frac{\displaystyle\int_\beta^\delta (x - \alpha)^{1/2}(x - \beta)^{1/2}x(\delta - x)^{1/2}b(x)\,dx}{\displaystyle\int_\beta^\delta (x - \alpha)^{1/2}(x - \beta)^{1/2}(\delta - x)^{1/2}b(x)\,dx} \\ &= (\delta - \beta)\frac{\displaystyle\int_0^1 \left(y + \frac{\beta - \alpha}{\delta - \beta}\right)^{1/2} y^{3/2}(1 - y)^{1/2}b((\delta - \beta)y + \beta)\,dy}{\displaystyle\int_0^1 \left(y + \frac{\beta - \alpha}{\delta - \beta}\right)^{1/2} y^{1/2}(1 - y)^{1/2}b((\delta - \beta)y + \beta)\,dy} + \beta \\ &= (\delta - \beta)\left[\Phi\left(\frac{\beta - \alpha}{\delta - \beta}\right) + O(|\beta| + |\delta - \beta| + |\alpha_2|)\right] + \beta \end{aligned} \tag{8.10}$$

where Φ: $[0, \infty] \to R$ is the function

$$\Phi(\sigma) = \frac{\displaystyle\int_0^1 (y + \sigma)^{1/2}y^{3/2}(1 - y)^{1/2}\,dy}{\displaystyle\int_0^1 (y + \sigma)^{1/2}y^{1/2}(1 - y)^{1/2}\,dy}.$$

Observe that for positive σ

$$0 < \Phi(\sigma) < 1, \qquad \Phi(0) = \frac{4}{7}, \qquad \Phi(\infty) = \frac{1}{2}\cdot \tag{8.11}$$

In terms of new variables

$$\sigma = \frac{\beta - \alpha}{\delta - \beta} \in [0, \infty], \qquad \tau = \delta - \beta, \qquad 0 \le \tau \ll 1, \tag{8.12}$$

(8.10) becomes

$$\beta = \frac{\tau}{5}[-1 + \sigma - 2\Phi(\sigma) + O(|\beta| + |\tau| + |\alpha_2|)];$$

also note

$$\gamma - \beta = \tau[\Phi(\sigma) + O(|\beta| + |\tau| + |\alpha_2|)]. \tag{8.13}$$

Therefore from (8.9) α_3 and α_2 can be expressed in terms of τ and σ:

$$5\tau^{-2}\alpha_3 = -2\sigma^2 - [1 + 2\Phi(\sigma)]\sigma - [2 - 2\Phi(\sigma) + 3\Phi(\sigma)^2] + O((1 + \sigma)\eta), \tag{8.14}$$

$$\begin{aligned} 25\tau^{-3}\alpha_2 = {} & 4\sigma^3 + [3 + 6\Phi(\sigma)]\sigma^2 + [-3 + 8\Phi(\sigma) - 2\Phi(\sigma)^2]\sigma \\ & + [-4 + 6\Phi(\sigma) + 2\Phi(\sigma)^2 - 2\Phi(\sigma)^3] + O((1 + \sigma)\eta^2 + (1 + \sigma)^2\eta) \end{aligned} \tag{8.15}$$

where η is the error term of order $|\beta| + |\tau| + |\alpha_2|$. From the hypothesis $|\alpha_3| \le C|\alpha_2|$ of the theorem it follows σ has an upper bound $\sigma \le C_4$; for otherwise from (8.14), (8.15) we have $\alpha_3 \sim -\frac{2}{5}\tau^2\sigma^2$ and $\alpha_2 \sim \frac{4}{25}\tau^3\sigma^3$, a contradiction. With this bound, therefore, $|\alpha_2| \le C_5|\tau|^3$ holds by (8.15); and in (8.14)

$$-2\sigma^2 - [1 + 2\Phi(\sigma)]\sigma - [2 - 2\Phi(\sigma) + 3\Phi(\sigma)^2] < 0 \quad \text{on} \quad [0, \infty)$$

(which follows from (8.11)) leads to the estimates $|\alpha_3| \ge C_6|\tau|^2 \ge C_7|\alpha_2|^{2/3}$. This contradiction completes the proof of the theorem.

Therefore a generic one-parameter perturbation gives rise to the normal form (A_5); in fact the parameters α, β, γ in this form can be given quite explicitly, as the following theorem shows.

THEOREM 8.2. Let

$$p(x, \epsilon) = (x^5 + \alpha_4(\epsilon)x^4 + \cdots + \alpha_0(\epsilon))b(x, \epsilon)^2$$

satisfy one of the normal forms (A_5), (B_5), (C_5) for scalar ϵ near zero (and on either side of zero). Suppose as $\epsilon \to 0$

$$\alpha_2(\epsilon) \sim \epsilon, \qquad \max\{|\alpha_3(\epsilon)|, |\alpha_4(\epsilon)|, \|b - b_0\|_\delta\} \le C|\epsilon|, \qquad \alpha_0(\epsilon), \alpha_1(\epsilon) \to 0.$$

Then, in fact, p has the normal form (A_5), and

$$p(x, \epsilon) = [(x - \alpha)^2 + \beta^2]^2(x - \gamma)b(x, \epsilon)^2, \qquad \alpha \sim \left(\frac{\epsilon}{40}\right)^{1/3},$$

$$\beta^2 \sim 5\left(\frac{\epsilon}{40}\right)^{2/3}, \qquad \gamma \sim -4\left(\frac{\epsilon}{40}\right)^{1/3}\cdot$$

Proof. By Theorem 8.1 p must be in the form (A_5). By direct calculation

$$\alpha_4 = -4\alpha - \gamma, \qquad \alpha_3 = \frac{2}{5}\alpha_4^2 + 2\beta^2 - \frac{2}{5}(\alpha - \gamma)^2,$$

$$\alpha_2 = \frac{-4}{25}\alpha_4^3 + \frac{3}{5}\alpha_3\alpha_4 + \frac{4}{25}(\alpha - \gamma)^3 + \frac{4}{5}(\alpha - \gamma)\beta^2,$$

so by the hypotheses on $\alpha_j(\epsilon)$ it follows that

$$-4\alpha - \gamma = O(\epsilon), \tag{8.16}$$

$$2\beta^2 - \frac{2}{5}(\alpha - \gamma)^2 = O(\epsilon), \tag{8.17}$$

$$\frac{4}{25}(\alpha - \gamma)^3 + \frac{4}{5}(\alpha - \gamma)\beta^2 \sim \epsilon. \tag{8.18}$$

Substituting for β^2 in (8.18) from (8.17) gives

$$\alpha - \gamma \sim \left(\frac{25\epsilon}{8}\right)^{1/3} \tag{8.19}$$

so from (8.16) we have

$$\alpha \sim \frac{1}{5}\left(\frac{25\epsilon}{8}\right)^{1/3} = \left(\frac{\epsilon}{40}\right)^{1/3}$$

and $\gamma \sim -4\alpha$. Finally, (8.17) and (8.19) give

$$\beta^2 \sim \frac{1}{5}\left(\frac{25\epsilon}{8}\right)^{2/3} = 5\left(\frac{\epsilon}{40}\right)^{2/3},$$

which completes the proof.

9. Two-parameter unfoldings. Our final topic is a formal analysis of the case $n = 5$ under a generic two-parameter perturbation $g(w, \epsilon) = g_0(w) + \epsilon_1 g_{11}(w) + \epsilon_2 g_{12}(w) + o(|\epsilon_1| + |\epsilon_2|)$ of the boundary conditions. The parameters ϵ_1 and ϵ_2 vary independently near zero, that is, $\epsilon = (\epsilon_1, \epsilon_2)$ varies in a full neighborhood of the origin in R^2. The object is to describe the form of the singularity in terms of the parameter ϵ; in particular, to decide which regions of the ϵ-plane correspond to the normal forms (A_5), (B_5) or (C_5). We shall see that each of these forms will occur. In particular, the contact set I has one large connected component for (A_5) or (B_5) and has a large component (the mainland) and a small island nearby for (C_5). The transition from (A_5) to (C_5) occurs as the island breaks away from the mainland; from (B_5) to (C_5) it occurs as the island rises out of the sea.

The generic condition imposed is that α_2 and α_3 vary independently with ϵ_1 and ϵ_2, that is, $|\partial(\alpha_2, \alpha_3)/\partial(\epsilon_1, \epsilon_2)| \neq 0$. This can be normalized to

$$\frac{\partial(\alpha_2, \alpha_3)}{\partial(\epsilon_1, \epsilon_2)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{9.1}$$

A further normalization is to assume

$$\alpha_4 = 0.$$

This can be accomplished by a translation, i.e., by replacing x with $x - \alpha_4/5$; then α_2 and α_3 change by orders $O(\alpha_3^2 + \alpha_4^2)$ and $O(\alpha_4^2)$ so (9.1) still holds and the description of the bifurcation set in the ϵ-plane remains the same.

The analysis below is roughly the analogue for $n = 5$ of the case $n = 4$ considered in Sec. 6, culminating in Theorem 6.1. In effect we are taking the cross-section

$$\Sigma_b{}^0 = \Sigma_b \cap \{\alpha | \alpha_4 = 0\}$$

of the set Σ of all $(\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4) \in R^5$ for which $p(x)$ is in some normal form (A_5), (B_5) or (C_5). The question of interest is to characterize those $(\alpha_2, \alpha_3) \in R^2$ which occur in each of these forms. To simplify matters we neglect the higher-order terms arising from the non-zero factor $b(x)^2$ in $p(x)$ by assuming $b(x) \equiv b(0) > 0$ is constant. We shall show there is a continuous one-to-one correspondence between $(\alpha_2, \alpha_3) \in R^2$ near the origin, and coefficients $(\alpha_0, \alpha_1, \alpha_2, \alpha_3, 0) \in R^5$ corresponding to some normal form; that is, $\Sigma_b{}^0$, for $b(x)$ constant, is the graph of a continuous function expressing α_0 and α_1 in terms of α_2 and α_3. Moreover, the regions in the (α_2, α_3)-plane corresponding to the various normal forms are given as follows:

$$(A_5): \quad \alpha_3 \geq -\left(\frac{5}{2}\right)^{1/3} \alpha_2^{2/3}$$

$$(B_5): \quad -\frac{9}{2^{1/3}} \alpha_2^{2/3} \leq \alpha_3 \leq -\left(\frac{5}{2}\right)^{1/3} \alpha_2^{2/3} \quad \text{and} \quad \alpha_2 \leq 0$$

$$(C_5): \quad \begin{cases} \alpha_3 \leq -\dfrac{9}{2^{1/3}} \alpha_2^{2/3} & \text{when} \quad \alpha_2 \leq 0 \\ \alpha_3 \leq -\left(\dfrac{5}{2}\right)^{1/3} \alpha_2^{2/3} & \text{when} \quad \alpha_2 \geq 0. \end{cases}$$

The transition from (A_5) to (C_5) occurs along the curve

$$\alpha_3^3 = -\frac{5}{2} \alpha_2^2, \qquad \alpha_2 \geq 0 \tag{9.2}$$

and from (B_5) to (C_5) along

$$\alpha_3^3 = -\frac{729}{2} \alpha_2^2, \qquad \alpha_2 \leq 0. \tag{9.3}$$

This is illustrated as a bifurcation diagram in Fig. 1. Presumably this question could be analyzed without assuming $b(x)$ to be constant, that is, by studying a two-dimensional embedding $(\epsilon_2, \epsilon_3) \rightarrow \Sigma_b$ with possibly $b(x) = b(x, \epsilon)$ also depending on ϵ in a C^1 manner. This should lead to regions bounded by transition curves with asymptotic forms (9.2) and

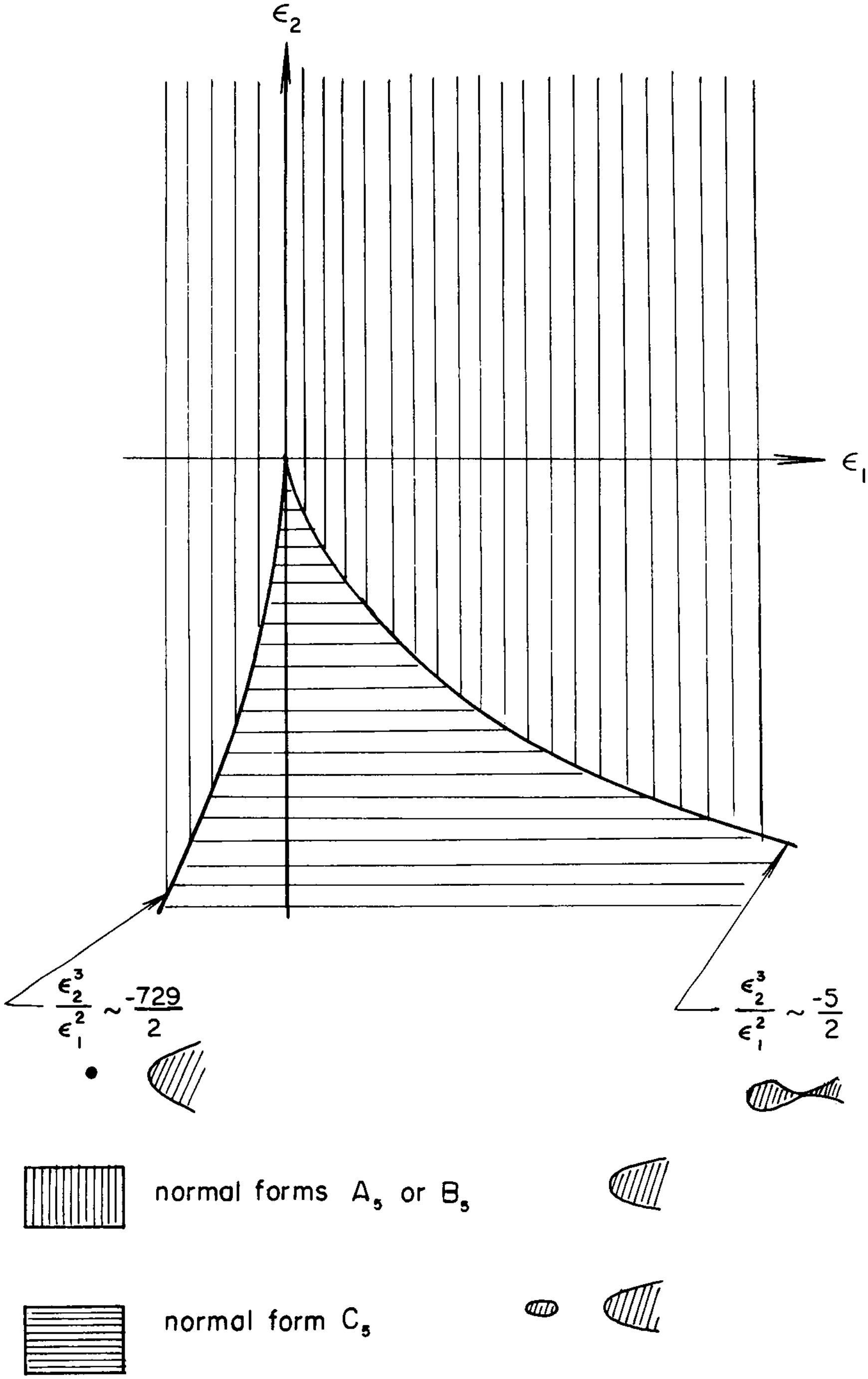

FIG. 1. Generic two-parameter bifurcation for $n = 5$.

(9.3). Our interest at this point, however, is in the derivation of these formulas rather than their generalization, so we do not do this.

Consider first the normal form (A_5); here $0 = \alpha_4 = -4\alpha - \gamma$ so $\gamma = -4\alpha$. Then

$$\alpha_3 = -10\alpha^2 + 2\beta^2, \qquad \alpha_2 = 4\alpha(5\alpha^2 + \beta^2). \tag{9.4}$$

A continuous one-to-one parameterization of the part of Σ^0 corresponding to (A_5) is given by letting

$$|\alpha| \ll 1, \qquad 0 \le \beta \ll 1;$$

for such α and β, (9.4) gives a local homeomorphism onto the region

$$\alpha_3 \ge -\left(\frac{5}{2}\right)^{1/3} \alpha_2^{2/3},$$

mapping $\beta = 0$ onto the boundary of this region.

For the case of (B_5) the corresponding formulas were derived in the proof of Theorem 8.1:

$$\alpha_3 = \alpha^2 K_1(\rho), \qquad \alpha_2 = \alpha^3 K_2(\rho), \qquad 0 \le -\alpha \ll 1, \qquad \frac{4}{3} \le \rho \le 4$$

where

$$K_1(\rho) = -2 + \rho - \frac{3}{4}\rho^2, \qquad K_2(\rho) = -\rho + \frac{1}{2}\rho^2 + \frac{1}{4}\rho^3$$

as in (8.3), (8.5), (8.6) and (8.7). For $\rho = 4/3$ and 4 this describes curves $\alpha_3 = -(9/2^{1/3})\alpha_2^{2/3}$ and $\alpha_3 = -(5/2)^{1/3}\alpha_2^{2/3}$ respectively, where $\alpha_2 \le 0$. To show the region between these curves is covered in a one-to-one manner it suffices to show that the function

$$\Lambda(\rho) = \alpha_3^3/\alpha_2^2 = K_1(\rho)^3/K_2(\rho)^2$$

satisfies $d\Lambda/d\rho \le 0$ on [4/3,4]; and since $K_2(\rho) < 0$ and $K_1(\rho) \ne 0$ there, this is equivalent to showing

$$3K_2\frac{dK_1}{d\rho} - 2K_1\frac{dK_2}{d\rho} = -\frac{3}{2}\rho^3 + \frac{11}{2}\rho^2 + 3\rho - 4 \ge 0.$$

But this is clear since

$$\frac{-3}{2}\rho^3 + \frac{11}{2}\rho^2 + 3\rho - 4 = -(\rho - 4)\left(\frac{3}{2}\rho - 1\right)(\rho + 1).$$

For the (C_5) normal form we have again from (8.12), (8.14) and (8.15) in the proof of Theorem 8.1

$$\alpha_3 = \tau^2 K_3(\sigma), \qquad \alpha_2 = \tau^3 K_4(\sigma) \tag{9.5}$$

where

$$5K_3(\sigma) = -2\sigma^2 - [1 + 2\Phi(\sigma)]\sigma - [2 - 2\Phi(\sigma) + 3\Phi(\sigma)^2],$$

$$25K_4(\sigma) = 4\sigma^3 + [3 + 6\Phi(\sigma)]\sigma^2 + [-3 + 8\Phi(\sigma) - 2\Phi(\sigma)^2]\sigma + [-4 + 6\Phi(\sigma) + 2\Phi(\sigma)^2 - 2\Phi(\sigma)^3]. \tag{9.6}$$

At $\sigma = 0$ and ∞ the curves $\alpha_3 = -(9/2^{1/3})\,\alpha_2^{2/3}$, $\alpha_2 \le 0$ and $\alpha_3 = -(5/2)^{1/3}\alpha_2^{2/3}$, $\alpha_2 \ge 0$ are described. (To handle the situation near $\sigma = \infty$ introduce variables $\sigma_1 = 1/\sigma$ and $\tau_1 = \sigma\tau$ where $0 \le \sigma_1$ and $0 \le \tau_1 \ll 1$.) Because $K_3(\sigma) < 0$ on $[0,\infty]$ it would be sufficient to have

$$3K_4 \frac{dK_3}{d\sigma} - 2K_3 \frac{dK_4}{d\sigma} \ge 0 \tag{9.7}$$

to show we have a one-to-one mapping onto the region between these curves. We could prove this directly as in the (B_5) case, but this is not feasible as it entails rather difficult calculations involving the non-elementary function Φ and its derivative. An indirect method of proving (9.7) is better.

Express α_1 and α_0 in terms of τ and σ using (8.12); this gives

$$\alpha_1 = \tau^4 K_5(\sigma), \qquad \alpha_0 = \tau^5 K_6(\sigma) \tag{9.8}$$

where K_5 and K_6 are certain polynomials in σ and $\Phi(\sigma)$. We may consider the formulas (9.5), (9.8) for α_j as a composition

$$(\sigma, \tau) \to (\xi, \eta, \zeta) \to (\alpha_3, \alpha_2, \alpha_1, \alpha_0)$$

where ((8.12), (8.13))

$$\xi = \beta - \alpha = \tau\sigma, \qquad \eta = \beta - \gamma = -\tau\Phi(\sigma), \qquad \zeta = \beta - \delta = -\tau \tag{9.9}$$

and from (8.8)

$$\beta = \frac{1}{5}(\xi + 2\eta + \zeta). \tag{9.10}$$

Let $M(\sigma, \tau)$ and $N(\sigma, \tau)$ be two by two matrices defined by

$$\frac{\partial(\alpha_3, \alpha_2, \alpha_1, \alpha_0)}{\partial(\sigma, \tau)} = \begin{pmatrix} \tau^2 K_3' & 2\tau K_3 \\ \tau^3 K_4' & 3\tau^2 K_4 \\ \tau^4 K_5' & 4\tau^3 K_5 \\ \tau^5 K_6' & 5\tau^4 K_6 \end{pmatrix} = \begin{pmatrix} M \\ \\ N \end{pmatrix} = \frac{\partial(\alpha_3, \alpha_2, \alpha_1, \alpha_0)}{\partial(\xi, \eta, \zeta)} \frac{\partial(\xi, \eta, \zeta)}{\partial(\sigma, \tau)}.$$

Thus (9.7) is equivalent to showing $\det M \ge 0$ for $\tau \ne 0$; we show in fact, $\det M \ne 0$ for $\tau \ne 0$ and $\sigma \in (0,\infty)$. To see that this determinant is positive note that (8.11), (9.6) give $K_4(0) = -4/343$ but $K_4(\sigma) > 0$ for large σ, so that for some σ_0, $K_4(\sigma_0) = 0$, $K_4'(\sigma_0) \ge 0$; but also $K_3(\sigma_0) < 0$.

The proof that M is non-singular proceeds in three steps. First note $\partial(\xi,\eta,\zeta)/\partial(\sigma,\tau)$ has zero kernel; in fact,

$$\partial(\xi,\zeta)/\partial(\sigma,\tau) \text{ is non-singular.} \tag{9.11}$$

Secondly, we show from (9.11) that

$$\ker \begin{pmatrix} M \\ N \end{pmatrix} = \{0\}; \tag{9.12}$$

and finally, we shall show

$$\ker M \subseteq \ker N \tag{9.13}$$

from which it follows $\det M \ne 0$.

Suppose for some

$$\tau_0 \neq 0, \sigma_0 > 0 \quad \text{that} \quad \begin{pmatrix} \sigma_1 \\ \tau_1 \end{pmatrix} \neq 0$$

lies in the kernel of

$$\begin{pmatrix} M \\ N \end{pmatrix}$$

Letting

$$p(x,\sigma,\tau) = x^5 + \alpha_3 x^3 + \alpha_2 x^2 + \alpha_1 x + \alpha_0$$

gives identically in x

$$\frac{\partial p(x, \sigma_0, \tau_0)}{\partial \sigma} \sigma_1 + \frac{\partial p(x, \sigma_0, \tau_0)}{\partial \tau} \tau_1 = 0. \tag{9.14}$$

Regard the simple zeros α, β and δ of p as functions of σ and τ given by (9.9) and (9.10) and, using (9.14), differentiate $p = 0$ at these points in the (σ_1, τ_1) direction. It follows that

$$\begin{pmatrix} \sigma_1 \\ \tau_1 \end{pmatrix} \text{ is in the kernel of } \partial(\alpha,\beta,\delta)/\partial(\sigma,\tau),$$

and therefore of $\partial(\xi,\zeta)/\partial(\sigma,\tau) = \partial(\beta-\alpha,\beta-\delta)/\partial(\sigma-\tau)$, a contradiction. Thus (9.12) holds.

Now, let

$$\begin{pmatrix} \sigma_1 \\ \tau_1 \end{pmatrix} \text{ lie in the kernel of } M = \partial(\alpha_3, \alpha_2)/\partial(\sigma,\tau),$$

so at (σ_0, τ_0), $(\partial p/\partial\sigma)\,\sigma_1 + (\partial p/\partial\tau)\,\tau_1$ is linear in x. Differentiating $p = 0$ at the double zero γ shows this linear function vanishes at γ, so

$$\frac{\partial p(x, \sigma_0, \tau_0)}{\partial \sigma} \sigma_1 + \frac{\partial p(x, \sigma_0, \tau_0)}{\partial \tau} \tau_1 = c(x - \gamma) \tag{9.15}$$

for some c. And differentiating the integral condition

$$\int_\beta^\delta (-p)^{1/2} dx = 0$$

in the (σ_1, τ_1) direction gives

$$0 = \frac{c}{2} \int_\beta^\delta (-p(x,\sigma_0,\tau_0))^{-1/2}(x-\gamma)dx$$

$$= \frac{c}{2} \int_\beta^\delta (x-\alpha)^{-1/2}(x-\beta)^{-1/2}(\delta - x)^{-1/2} dx$$

implying $c = 0$. Therefore, (9.15) vanishes, so

$$\begin{pmatrix} \sigma_1 \\ \tau_1 \end{pmatrix}$$

also lies in the kernel of N, proving (9.13). Therefore, (9.7) holds.

REFERENCES

[1] H. Brézis and D. Kinderlehrer, *The smoothness of solutions to non-linear variational inequalities*, Indiana Univ. Math. J. **23**, 831–844 (1974)

[2] L. A. Caffarelli, *The regularity of free boundaries in higher dimensions*, Acta Math. **139**, 155–184 (1977)

[3] L. A. Caffarelli and N. M. Rivière, *Smoothness and analyticity of free boundaries in variational inequalities*, Ann. Scuola Norm. Sup. Pisa, Ser. **IV 3**, 289–310 (1976)

[4] S.-N. Chow, J. K. Hale and J. Mallet-Paret, *Applications of generic bifurcation.* I, Arch. Rat. Mech. Anal. **59**, 159–188 (1975)

[5] S.-N. Chow, J. K. Hale and J. Mallet-Paret, *Applications of generic bifurcation. II*, Arch. Rat. Mech. Anal. **62**, 209–235 (1976)

[6] S.-N. Chow and J. Mallet-Paret, *The parameterized obstacle problem*, submitted for publication.

[7] J. K. Hale, *Bifurcation with several parameters*, in *VII. Internationale Konferenz über nichtlineare Schwingungen*, **I**, 1, pp. 321–332, Akademie-Verlag, Berlin 1977

[8] J. K. Hale, *Restricted generic bifurcation*, in *Nonlinear analysis*, pp. 83–98, Academic Press, 1978

[9] D. Kinderlehrer, *The free boundary determined by the solution to a differential equation*, Indiana Univ. Math. J. **25**, 195–208 (1976)

[10] D. Kinderlehrer and L. Nirenberg, *Regularity in free boundary problems*, Ann Scuola Norm. Sup. Pisa, Ser. **IV 4**, 373–391 (1977)

[11] H. Lewy and G. Stampacchia, *On the regularity of the solution of a variational inequality*, Comm. Pure Appl. Math. **22**, 153–188 (1969)

[12] J. Mallet-Paret, *Buckling of cylindrical shells with small curvature*, Quart. Appl. Math. **35**, 383–400 (1977)

[13] D. G. Schaeffer, *A stability theorem for the obstacle problem*, Adv. Math. 17, 34–47 (1975)

[14] D. G. Schaeffer, *Some examples of singularities in a free boundary*, Ann. Scuola Norm. Sup. Pisa, Ser. **IV**, **4**, 133–144 (1977)

[15] D. G. Schaeffer, *One-sided estimates for the curvature of the free boundary in the obstacle problem*, Adv. Math. **24**, 78–98 (1977)

[16] M. Golubitsky and D. Schaeffer, *A theory for imperfect bifurcation via singularity theory*, Comm. Pure Appl. Math. **23**, 21–98 (1979)

[17] J. Mallet-Paret, *Generic unfoldings and normal forms of some singularities arising in the obstacle problem*, Duke Math, J. **46**, 645–683 (1979)