Differential and Integral Equations Volume 14, Number 12, December 2001, Pages 1441–1468

SOLUTIONS NEAR SINGULAR POINTS TO THE EIKONAL AND RELATED FIRST-ORDER NONLINEAR PARTIAL DIFFERENTIAL EQUATIONS IN TWO INDEPENDENT VARIABLES

Emil Cornea[1]
Department of Mathematical Sciences, Northern Illinois University
DeKalb, IL 60115

Ralph Howard[2]
Department of Mathematics, University of South Carolina, Columbia, SC 29208

Per-Gunnar Martinsson[3]
Texas Institute for Computational and Applied Mathematics, MC C0200,
University of Texas-Austin, Austin TX 78712

(Submitted by: J.L. Bona)

Abstract. A detailed study of solutions to the first-order partial differential equation $H(x, y, z_x, z_y) = 0$, with special emphasis on the eikonal equation $z_x^2 + z_y^2 = h(x, y)$, is made near points where the equation becomes singular in the sense that $dH = 0$, in which case the method of characteristics does not apply. The main results are that there is a strong lack of uniqueness of solutions near such points and that solutions can be less regular than both the function H and the initial data of the problem, but that this loss of regularity only occurs along a pair of curves through the singular point. The main tools are symplectic geometry and the Sternberg normal form for Hamiltonian vector fields.

1. Introduction

The eikonal equation in two independent variables for $z = z(x, y)$ is

$$z_x^2 + z_y^2 = h(x, y) \tag{1.1}$$

where h is a nonnegative smooth function on the plane, and subscripts denote partial derivatives. Near points where $h \neq 0$ all local solutions to this

[1]Supported in part by ONR Grant No. DAAH04-96-1-0326 and N00014-97-1-0806.
[2]Supported in part by DoD Grant No. N00014-97-1-0806.
[3]Supported in part by ONR Grant Nos. N00014-97-11059 and N00014-94-11163.
Accepted for publication August 2000.
AMS Subject Classifications: 35F20.

equation can be constructed by the method of characteristics (*cf.* [2, Chapter 2], [17, Chapter 10]) and questions of local existence, uniqueness, and regularity are fully understood. However, near points where h vanishes the picture is much less complete. Our main goal is to study these questions in detail and show that near a point where h has zero as a nondegenerate local minimum value there are generally infinitely many local solutions. These results are extended to solutions of certain Hamilton–Jacobi equations $H(x, y, z_x, z_y) = 0$ near a rest point of the corresponding characteristic vector field.

As a model problem we will use

$$z_x^2 + z_y^2 = a^2x^2 + b^2y^2 \tag{1.2}$$

where a and b are positive. For ease of statement we discuss our general results in terms of this special case. Normalizing the solution so that $z(0,0) = 0$, the functions $z_{\pm,\pm} = \frac{1}{2}(\pm ax^2 \pm by^2)$ clearly solve (1.2). Moreover, it can be shown (see Proposition 4.1 below) that any C^3 solution to (1.2) near $(0,0)$ is of the form $z = z_{\pm,\pm} + O((|x| + |y|)^3)$.

The solution $z_{+,+}$ is unique in the sense that any C^2 solution $\tilde{z} = z_{+,+} + O((|x|+|y|)^3)$ agrees with $z_{+,+}$ in a neighborhood of $(0,0)$ (and on all of $\mathbf{R}^2$ if $\tilde{z}$ is globally defined), with a similar statement holding for $z_{-,-}$. (See [4] or Theorem 2.7 below for the local result and Oliensis [12] for the global version). However, the two saddle-point solutions $z_{+,-}$ and $z_{-,+}$ are far from unique. We demonstrate (Theorem 3.2) that there is an infinite-dimensional family of C^∞ solutions of the form $\tilde{z} = z_{+,-} + O((|x|+|y|)^3)$ that do not agree with $z_{+,-}$ in any neighborhood of the origin, with a similar statement holding for $z_{-,+}$. (This gives counterexamples to some claims in the literature; *cf.* [15, p. 1094], which would imply that $z_{\pm,\pm}$ are the only solutions.) We also show that for each $k \geq 2$ there are solutions that are C^k, but not C^{k+1}. The solutions, z, are obtained by looking for the graph of the differential of z, that is, the manifold of one jet of z,

$$N^2 = \{(x, y, z_x(x, y), z_y(x, y)) : (x, y) \in U\} \subset \mathbf{R}^4,$$

over an open set $U \subset \mathbf{R}^2$. We first note that the manifold, N^2, for $z_{+,-}$ is obtained by letting the two curves $\gamma_\pm \colon s \mapsto [s, \pm s, as, \mp bs]^t$ flow under the characteristic vector field of (1.2) and taking the closure of the result. Now an infinite-dimensional family of saddle-point solutions can be constructed by using as initial data for the flow, not the curves $\gamma_\pm$, but curves $\tilde{\gamma}_\pm$ obtained by perturbing $\gamma_\pm$ in certain directions. If $\tilde{\gamma}_\pm$ are C^k and agree with $\gamma_\pm$ to order $l \leq k$ at the origin, then the solution will be C^{k+1} on $\mathbf{R}^2 \setminus \{xy = 0\}$, as expected. But, surprisingly, when considered as a function on all of $\mathbf{R}^2$

the regularity is only C^n with $n := \lceil \min\{\frac{(l+1)a}{a+b}, \frac{(l+1)b}{a+b}\} \rceil - 1$ (where $\lceil \cdot \rceil$ is the ceiling function), and so there is a drop of regularity along the coordinate axes.

This existence, lack of uniqueness, and jump in regularity along a pair of distinguished curves is not specific to (1.2) but generalizes to the eikonal equation (1.1) with

$$h(x, y) = a^2x^2 + b^2y^2 + O((|x| + |y|)^3), \qquad a, b > 0, \tag{1.3}$$

under the extra assumption that the numbers a and b are linearly independent over the rational numbers. More generally, the results are stated and proven in Section 3.2 for certain Hamilton–Jacobi equations $H(x, y, z_x, z_y) = 0$ near a rest point of the characteristic vector field ξ_H of H (see (2.3) below). In the latter case it is assumed that the eigenvalues of the linearization of ξ_H at the rest point are a, b, $-a$, and $-b$, with $a, b > 0$. The proof is based on using the Sternberg normal form for such an equation (and the existence of this normal form is guaranteed only when a and b are linearly independent over the rationals) to reduce the calculations to manageable proportions.

The somewhat unexpected condition that a and b be linearly independent over the rational numbers is supported by a formal power series argument. We show that equation (1.1), with h as in (1.3), has a unique formal power series solution of the form

$$z = \tfrac{1}{2}(ax^2 - by^2) + O((|x| + |y|)^3) \tag{1.4}$$

precisely when this condition holds (see Theorems 4.2 and 4.4 below). Moreover, we show that if a and b are linearly dependent over the rationals and m and n are integers such that $ma - nb = 0$ and $m + n \geq 4$, then the equation $z_x^2 + z_y^2 = ax^2 + by^2 + x^m y^n$ has no C^k solution of the form (1.4), for $k \geq m+n$. Thus the independence condition on a and b is sufficient and very close to being necessary. The formal power series approach has previously been used for analyzing the convex and concave solutions; see Bruss [4] and the references therein.

A secondary goal of this paper is to advocate the use of differential geometric methods, in particular differential forms, symplectic geometry, and normal forms such as the Sternberg normal form, in working with first-order differential equations. To this end we have included some expository material in Section 2 about symplectic geometry and its application to the method of characteristics for first order differential equations.

The article can be summarized as follows. In Section 2, we present the basic methods of symplectic geometry as applied to the method of characteristics for first-order differential equations of the form $H(x, y, z_x, z_y) = 0$.

For several reasons the theory is easier in two dimensions, and our hope is that this will be useful in understanding the method in a concrete setting. This section also gives a proof of a slight generalization of a theorem of Bruss [4, p. 892] on the existence of concave and convex solutions to the eikonal equation (1.1). In Section 3, we first discuss in detail the eikonal model equation (1.2). We prove the existence, nonuniqueness and lack of regularity results mentioned above. We then look at more general equations of the form $H(x, y, z_x, z_y) = 0$. The same type of results hold, but we were not able to give quite as precise statements for the regularity. The proofs make essential use of both symplectic geometry and the Sternberg normal form. In Section 4, we discuss the solutions to the eikonal equation (1.1) in a formal power series framework. At this level we prove the existence and uniqueness of saddle-point solutions. This leads to examples of functions h for which the eikonal equation does not have any smooth solution of saddle type.

The article ends with Appendix A, which includes precise statements of two of the geometric tools we use: the stable submanifold theorem and the Sternberg normal form. Finally, we mention that the eikonal equation and certain related Hamilton–Jacobi equations have been of interest in computer vision because of the so-called "shape from shading problem." This problem corresponds, roughly speaking, to the reconstruction of a shape (a surface) $z = z(x, y)$ from a gray-scale image of it. That is, the function z is a solution of the Horn image irradiance equation (see [7], Chapter 10), which under some assumptions reduces to the eikonal equation [10]. Note that in applications the data function h typically has quite low regularity. In this case viscosity solutions have been considered as an alternative to classical solutions; see [10] and [14] and the references therein.

2. Application of symplectic geometry to the eikonal equation

2.1. Review of the method of characteristics and construction of concave and convex solutions. Let $\mathbf{R}^4$ have coordinates x, y, p, q. Then the *symplectic form* on $\mathbf{R}^4$ is $\omega := dp \wedge dx + dq \wedge dy$. Let $N^2 \subset \mathbf{R}^4$ be an imbedded surface. Then N^2 *projects* on an open set $U \subset \mathbf{R}^2$ if and only if there are continuous functions $p(x, y)$ and $q(x, y)$ so that

$$N^2 = \{(x, y, p(x, y), q(x, y)) : (x, y) \in U\}. \tag{2.1}$$

The submanifold N^2 is a *jet of a function* if and only if there is an open set $U \subseteq \mathbf{R}^2$ and a function $z \in C^1(U)$ such that

$$N^2 = \{(x, y, z_x(x, y), z_y(x, y)) : (x, y) \in U\}. \tag{2.2}$$

It is clear that if N^2 is a jet of a function, then it projects, but the converse is not true. The following is a standard result, but we include the short proof for those not familiar with the differential geometric set up.

Proposition 2.1. *Assume that $N^2 \subset \mathbf{R}^4$ is a simply connected two-dimensional submanifold of $\mathbf{R}^4$ of smoothness class C^k for some $k \geq 1$. Then $N^2 \subset \mathbf{R}^4$ is the jet of some function (which will be unique up to an additive constant) if and only if it projects over some open set $U \subseteq \mathbf{R}^2$ and the restriction of the symplectic form to N^2 vanishes. (That is, $\omega(X,Y) = 0$ for all vectors tangent to N^2.) Moreover, if N^2 is of class C^k and is the jet of a function z, then z is of class C^{k+1}.*

Proof. Assume that N^2 projects over U so that $N^2 = \{(x, y, p(x,y), q(x,y)) : (x,y) \in U\}$. As N^2 is simply connected, the same is true of U. Using x, y as coordinates on N^2, we see that the restriction of the symplectic form to N^2 is

$$\begin{aligned}\omega &= dp \wedge dx + dq \wedge dy = (p_x\, dx + p_y\, dy) \wedge dx + (q_x\, dx + q_y\, dy) \wedge dy \\ &= (q_x - p_y)\, dx \wedge dy.\end{aligned}$$

Therefore, the restriction of ω to N^2 vanishes if and only if $p_y = q_x$. But as U is simply connected this is exactly the condition that there is a function z (unique up to an additive constant) so that $p = z_x$ and $q = z_y$. If N^2 is of class C^k then $p(x,y)$ and $q(x,y)$ are C^k functions of (x,y). Thus z is of class C^{k+1}. $\square$

A two-dimensional surface $N^2 \subset \mathbf{R}^4$ is a *Lagrangian surface* if and only if the restriction of ω to N^2 vanishes.

Given a function $H\colon \mathbf{R}^4 \to \mathbf{R}$ the *characteristic vector field* of H (also called the *symplectic gradient*) is the unique vector yield ξ_H on $\mathbf{R}^4$ such that for all vectors X, $\omega(\xi_H, X) = -dH(X)$. Then a simple calculation yields that

$$\xi_H = \frac{\partial H}{\partial p}\frac{\partial}{\partial x} + \frac{\partial H}{\partial q}\frac{\partial}{\partial y} - \frac{\partial H}{\partial x}\frac{\partial}{\partial p} - \frac{\partial H}{\partial y}\frac{\partial}{\partial q}. \tag{2.3}$$

An immediate consequence of the definition of ξ_H is that

$$dH(\xi_H) = -\omega(\xi_H, \xi_H) = 0,$$

and so H is constant on the integral curves of ξ_H.

The following is the differential-geometric justification of the method of characteristics for the Hamilton–Jacobi equation $H(x, y, z_x, z_y) = 0$. Given a smooth function $H\colon \mathbf{R}^4 \to \mathbf{R}$ and $P \in \mathbf{R}^4$, the restriction of dH to $T(\mathbf{R}^4)_P$ is denoted dH_P.

Proposition 2.2. *Let N^2 be a simply connected two-dimensional submanifold of $\mathbf{R}^4$. We also assume that the following nondegeneracy condition holds:*

(ND) *The set of points $P \in N^2$ where $dH_P \neq 0$ is dense in N^2 (here dH_P is being viewed as a linear functional on $\mathbf{R}^4$ and not just restricted to $T(N^2)$).*

Then N^2 is the jet of a solution to the equation $H(x, y, z_x, z_y) = 0$ if and only if the following three conditions hold:

1. *$N^2 \subset \{P : H(P) = 0\}$,*
2. *N^2 projects over some open set $U \subseteq \mathbf{R}^2$, and*
3. *the characteristic vector field ξ_H is tangent to N^2 at every point of N^2.*

If N^2 satisfies these conditions and is a C^k submanifold, then the solution z is of class C^{k+1}.

Remark 2.3. For our future applications it is important to realize that although H is a smooth function, the zero set $\{H = 0\}$ need not be a smooth submanifold of $\mathbf{R}^4$. For $a, b > 0$ we will be interested in $H(x, y, p, q) = p^2 + q^2 - a^2x^2 - b^2y^2$. The zero set $\{H = 0\}$ is then a cone that is singular at $(0, 0, 0, 0)$. However, if N^2 is a smooth surface in $\{H = 0\}$ that is everywhere tangent to ξ_H and which projects onto an open set in $\mathbf{R}^2$, then N^2 is the jet of a solution to $z_x^2 + z_y^2 = a^2x^2 + b^2y^2$. This is because dH vanishes only at one point, and thus the set of points on N^2 where $dH \neq 0$ is dense. $\square$

It is useful to give a name to submanifolds that satisfy two of the conditions of the proposition:

Definition 2.4. Let $H \colon \mathbf{R}^4 \to \mathbf{R}$ be a C^k function with $k \geq 2$ and characteristic vector field ξ_H. Then a connected two-dimensional submanifold N^2 of class C^1 is an *invariant Lagrangian surface* if and only if

1. N^2 is a Lagrangian surface (*i.e.,* the restriction of ω to N^2 vanishes), and
2. the characteristic vector field ξ_H is tangent to N^2 at all points of N^2.

Proof of Proposition 2.2. First assume that the three conditions hold. Then in light of Proposition 2.1 it is enough to show that the restriction of ω to N^2 is zero. If we are at a point $P \in N^2$ where $dH \neq 0$, then from the definition of ξ_H or the formula (2.3) it follows that $\xi_H \neq 0$. Let X be a tangent vector to N^2 at P linearly independent of $\xi_H(P)$. Then $\{\xi_H(P), X\}$ is a basis for the tangent space $T(N^2)_P$. Using the definition of ξ_H and that

$dH(X) = 0$ (as $H\big|_{N^2} = 0$)

$$\omega(\xi_H, X) = -dH(X) = 0.$$

Thus ω restricted to N^2 vanishes at P. But we are assuming that the set of points P where dH does not vanish is dense; thus, by continuity the restriction of ω to N^2 is zero on all of N^2.

Conversely, if N^2 is the jet of a solution, then clearly $N^2 \subset \{P : H(P) = 0\}$ and N^2 projects onto an open subset of $\mathbf{R}^2$ and also $H\big|_{N^2} = 0$. Let $P \in N^2$. Then $\omega(X, Y) = 0$ for all $X, Y \in T(N^2)_P$. But a calculation shows that if Z is any vector tangent to $\mathbf{R}^4$ at P with $\omega(Z, X) = 0$ for all $X \in T(N^2)_P$, then $Z \in T(N^2)_P$. But for $X \in T(N^2)_P$ we have $dH(X) = 0$ so (as above) $\omega(\xi_H, X) = -dH(X) = 0$. Thus $\xi_H \in T(N^2)_P$. This completes the proof. $\square$

This result makes the geometry of the method of characteristics clearer than the classical presentations. Let H be of class C^k for $k \geq 2$. Then the formula (2.3) makes it clear that the vector field ξ_H is of class C^{k-1}. Let Φ_t^H be the flow of ξ_H. That is, $\Phi_0^H(P) = P$ and $c(t) := \Phi_t^H(P)$ is an integral curve of the vector field ξ_H. Then (see [9, Theorem 1, p. 80] or [1, p. 230]) the map $(t, P) \mapsto \Phi_t^H(P)$ is C^{k-1}. Now let $c \colon (a, b) \to \mathbf{R}^4$ be a curve of class C^l ($l \geq 1$) and let $\pi \colon \mathbf{R}^4 \to \mathbf{R}^2$ be the projection $\pi(x, y, p, q) = (x, y)$. Assume

(1) $H(c(s)) \equiv 0$,
(2) c projects over an imbedded[1] curve of $\mathbf{R}^2$ (that is, $s \mapsto \pi(c(s))$ is an imbedded curve in $\mathbf{R}^2$), and
(3) at all points $c(s)$ the vectors $\pi_* c'(s)$ and $\pi_* \xi_H(c(s))$ are linearly independent.

Then let $F(s, t) := \Phi_t^H(c(s))$. By the implicit function theorem, for r small enough the submanifold $N^2 := \{F(s, t) : (s, t) \in (a, b) \times (-r, r)\}$ will project over an open subset U of $\mathbf{R}^2$. Also, as H is constant along the flow of ξ_H and $H(c(s)) = 0$, we have that $H(F(s, t)) = H(\Phi_t^H(c(s))) = 0$. Thus $H\big|_{N^2} = 0$. By construction ξ_H is tangent to N^2, and so N^2 is the jet of a solution z to $H(x, y, z_x, z_y) = 0$ by Proposition 2.2. The regularity of N^2 is $C^{\min\{k-1, l\}}$, and therefore the regularity of z is $C^{\min\{k, l+1\}}$.

[1] A curve c is imbedded if and only if it is the image of a smooth injective map $\gamma : I \to \mathbf{R}^2$, with I an interval in $\mathbf{R}$, so that the velocity vector γ' never vanishes and so that the topology of $\gamma[I]$ as a subset of $\mathbf{R}^2$ is the same as the topology induced by γ (*i.e.,* γ is a homeomorphism).

2.2. **Solutions near critical points of H.** We now look at the more interesting case of finding solutions near a critical point, P_0, of H. Assume that $dH_{P_0} = 0$, and by adding a constant to H we can assume that $H(P_0) = 0$. Then near P_0 the set $\{H = 0\}$ need not be a submanifold of $\mathbf{R}^4$. Assume that P_0 is a nondegenerate critical point so that the Hessian of H at P_0 is nonsingular. To make the notation easier we assume that $P_0 = (0,0,0,0)$. Then using the form (2.3) we see that the linearization of the characteristic system for this vector field at the origin is

$$\frac{d}{dt}\begin{bmatrix} x(t) \\ y(t) \\ p(t) \\ q(t) \end{bmatrix} = \begin{bmatrix} H_{xp} & H_{yp} & H_{pp} & H_{pq} \\ H_{xq} & H_{yq} & H_{pq} & H_{qq} \\ -H_{xx} & -H_{xy} & -H_{xp} & -H_{xq} \\ -H_{xy} & -H_{yy} & -H_{yp} & -H_{yq} \end{bmatrix} \begin{bmatrix} x(t) \\ y(t) \\ p(t) \\ q(t) \end{bmatrix} = L \begin{bmatrix} x(t) \\ y(t) \\ p(t) \\ q(t) \end{bmatrix}, \qquad (2.4)$$

where all the second partial derivatives are evaluated at $(0,0,0,0)$, and this equation defines L. Let $\det(D^2H)$ be the determinant of the Hessian at $(0,0,0,0)$, and let

$$c_2 := 2H_{xy}\,H_{pq} - 2H_{xq}\,H_{yp} + H_{yy}H_{qq} - H_{yq}^2 + H_{xx}H_{pp} - H_{xp}^2.$$

The characteristic polynomial of L is then, using that $\det(L) = \det(D^2H)$ (which can be seen by noting that $L = J\,D^2H$, where $J = \left[\begin{smallmatrix} 0 & I_2 \\ -I_2 & 0 \end{smallmatrix}\right]$, so that $\det(L) = \det(J)\det(D^2H) = \det(D^2H)$), $\det(\lambda I - L) = \lambda^4 + c_2\lambda^2 + \det(D^2H)$. Therefore, the eigenvalues are

$$\pm\sqrt{\frac{-c_2 + \sqrt{c_2^2 - 4\det(D^2H)}}{2}}, \quad \pm\sqrt{\frac{-c_2 - \sqrt{c_2^2 - 4\det(D^2H)}}{2}}. \qquad (2.5)$$

Assuming that there is no eigenvalue with zero real part we see that there are exactly two eigenvaules with positive real part and two with negative real part. Then let N_+^2 be the local stable manifold for the critical point at $(0,0,0,0)$ and N_-^2 the local unstable submanifold. (Loosely $N_\pm^2$ is the set of points $P \in \mathbf{R}^4$ so that $\lim_{t\to\pm\infty} \Phi_t^H(P) = (0,0,0,0)$. See Appendix A.1.) Because of the condition on the eigenvalues of L both N_+^2 and N_-^2 are two dimensional. If H is of class C^k for $k \geq 2$ then the flow Φ_t^H is of class C^{k-1}. Therefore N_+^2 and N_-^2 are C^{k-1} submanifolds of $\mathbf{R}^4$ (see [16, Theorem 5.20, p. 49]). Then for a point $P \in N_\pm^2$ we have, by the invariance of H under the flow, that $H(P) = \lim_{t\to\pm\infty} H(\Phi_t^H(P)) = H(0,0,0,0) = 0$. Thus $H\big|_{N_\pm^2} = 0$. Also it is clear that the characteristic vector field ξ_H is tangent to $N_\pm^2$. The condition of Proposition 2.2 that does not hold automatically for $N_\pm^2$ is that of being locally projectable. Summarizing:

Theorem 2.5. *Let $H\colon \mathbf{R}^4 \to \mathbf{R}$ be a C^k function with $k \geq 2$. Assume that H has a nondegenerate critical point at $(0,0,0,0)$ with $H(0,0,0,0) = 0$, and assume that no eigenvalue of L as defined above has zero real part. Then there are exactly two eigenvalues λ_1, λ_2 of L that have positive real part. Let e_1 and e_2 be the eigenvectors corresponding to λ_1 and λ_2, let $\mathcal{N}$ be the subspace spanned by e_1 and e_2 and let $\pi\colon \mathbf{R}^4 \to \mathbf{R}^2$ be the projection $\pi(x,y,p,q) = (x,y)$. Assume that the restriction $\pi\big|_{\mathcal{N}}\colon \mathcal{N} \to \mathbf{R}^2$ is nonsingular. Then near $(0,0,0,0)$ the unstable submanifold N_-^2 of ξ_H is the jet of a C^k solution z to $H(x,y,z_x,z_y) = 0$. An analogous statement holds for the stable submanifold N_+^2 of ξ_H.*

Proof. The tangent space to N_-^2 at $(0,0,0,0)$ is $\mathcal{N}$, so if $\pi\big|_{\mathcal{N}}\colon \mathcal{N} \to \mathbf{R}^2$ is nonsingular the implicit function theorem implies that near $(0,0,0,0)$ the submanifold N_-^2 will project onto an open subset of $\mathbf{R}^2$. As noted all the other hypotheses of Proposition 2.2 hold. This completes the proof. □

As an example of this assume that $H(x,y,p,q) = f(p,q) - h(x,y)$ where f and h have critical points at $(0,0)$ and the Hessians of f and h at $(0,0)$ are both positive definite. (This is a case that comes up in the shape from shading problem.) Then the linear map L is

$$L = \begin{bmatrix} 0 & 0 & f_{pp} & f_{pq} \\ 0 & 0 & f_{pq} & f_{qq} \\ h_{xx} & h_{xy} & 0 & 0 \\ h_{xy} & h_{yy} & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & A \\ B & 0 \end{bmatrix},$$

where A and B are positive definite 2×2 matrices. As A is positive definite it has a square root $A^{\frac{1}{2}}$. The matrix $A^{\frac{1}{2}}BA^{\frac{1}{2}}$ will also be positive definite and will consequently have positive eigenvalues. Let a^2 and b^2 be the eigenvalues of $A^{\frac{1}{2}}BA^{\frac{1}{2}}$, where $a, b > 0$, and $\mathbf{e}_1, \mathbf{e}_2$ the corresponding eigenvalues,

$$A^{\frac{1}{2}}BA^{\frac{1}{2}}\mathbf{e}_1 = a^2\mathbf{e}_1, \quad A^{\frac{1}{2}}BA^{\frac{1}{2}}\mathbf{e}_2 = b^2\mathbf{e}_2,$$

and define

$$\mathbf{v}_1 := \begin{bmatrix} A^{\frac{1}{2}}\mathbf{e}_1 \\ aA^{-\frac{1}{2}}\mathbf{e}_1 \end{bmatrix}, \ \mathbf{v}_2 := \begin{bmatrix} A^{\frac{1}{2}}\mathbf{e}_2 \\ bA^{-\frac{1}{2}}\mathbf{e}_2 \end{bmatrix}, \ \mathbf{v}_3 := \begin{bmatrix} A^{\frac{1}{2}}\mathbf{e}_1 \\ -aA^{-\frac{1}{2}}\mathbf{e}_1 \end{bmatrix}, \ \mathbf{v}_4 := \begin{bmatrix} A^{\frac{1}{2}}\mathbf{e}_2 \\ -bA^{-\frac{1}{2}}\mathbf{e}_2 \end{bmatrix}.$$

By direct calculation, we find

$$L\mathbf{v}_1 = a\mathbf{v}_1, \quad L\mathbf{v}_2 = b\mathbf{v}_2, \quad L\mathbf{v}_3 = -a\mathbf{v}_3, \quad L\mathbf{v}_4 = -b\mathbf{v}_4.$$

Thus the eigenvalues of L are $\pm a$ and $\pm b$. The eigenvectors corresponding to the two eigenvalues with positive real part are $\mathbf{v}_1$ and $\mathbf{v}_2$. For future reference we record some elementary facts about these eigenvalues and eigenvectors. The proof is left to the reader.

Lemma 2.6. *If $a \neq b$ then there are exactly six two-dimensional subspaces of the tangent space $T(\mathbf{R}^4)$ that are invariant under L, corresponding to the six ways of choosing two of the eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_4$. As $\omega(\mathbf{v}_1, \mathbf{v}_3) \neq 0$ and $\omega(\mathbf{v}_2, \mathbf{v}_4) \neq 0$ the subspaces $\operatorname{span}(\mathbf{v}_1, \mathbf{v}_3)$ and $\operatorname{span}(\mathbf{v}_2, \mathbf{v}_4)$ can never be tangent to jets of functions and therefore can be disregarded from our considerations. Each of the remaining four pairs $\{\mathbf{v}_1, \mathbf{v}_2\}$, $\{\mathbf{v}_3, \mathbf{v}_4\}$, $\{\mathbf{v}_1, \mathbf{v}_4\}$, and $\{\mathbf{v}_2, \mathbf{v}_3\}$ spans a two-dimensional subspace of the tangent space $T(\mathbf{R}^4)_0$ on which ω vanishes. If $\{\mathbf{v}_i, \mathbf{v}_j\}$ is any one of these four pairs and $\pi \colon \mathbf{R}^4 \to \mathbf{R}^2$ is the projection $\pi(x, y, p, q) = (x, y)$ then $\pi_*\mathbf{v}_i$ and $\pi_*\mathbf{v}_j$ are linearly independent. Therefore, if N^2 is any two-dimensional submanifold tangent to $\operatorname{span}(\mathbf{v}_i, \mathbf{v}_j)$ then, by the implicit function theorem, N^2 locally projects over some open neighborhood U of $(0,0)$ in $\mathbf{R}^2$. Finally, $\operatorname{span}(\mathbf{v}_1, \mathbf{v}_2)$ is the tangent space at the origin to the unstable submanifold of ξ_H and $\operatorname{span}(\mathbf{v}_3, \mathbf{v}_4)$ is the tangent space at the origin to the stable submanifold of ξ_H.*

The following gives a generalization of a theorem of Bruss [4], who considered the special case of $u_x^2 + u_y^2 = h(x, y)$. The proof uses the same geometric idea as Bruss's proof.

Theorem 2.7 (A. Bruss [4])**.** *Let $f(p, q)$ and $h(x, y)$ be functions of class C^k, with $k \geq 2$, and assume that f and h both have critical points with positive-definite Hessians at $(0,0)$. Then in the class of functions that are concave near $(0,0)$ there is a unique (up to an additive constant) solution to $f(z_x, z_y) = h(x, y)$. This solution is C^k and has as its jet near $(0,0)$ the stable manifold of the characteristic vector field ξ_H of $H(x, y, p, q) = f(p, q) - h(x, y)$. (Likewise, there is a unique convex solution; it is C^k and has as jet near $(0,0)$ the unstable submanifold of ξ_H.)*

Proof. Follows from the discussion above.

3. Construction of saddle-type solutions near a regular critical point

3.1. **Construction of all solutions to $z_x^2 + z_y^2 = a^2x^2 + b^2y^2$.** We will now construct all "saddle point" solutions of the equation

$$z_x^2 + z_y^2 = a^2x^2 + b^2y^2,$$

where $a, b > 0$. The analysis here is a model for the more general setup covered in Section 3.2. This special case is also of interest as it is possible to be somewhat more precise about the regularity of the solutions. What makes this equation especially easy to analyze is that the components of the characteristic vector field are linear so that finding the flow involves no more than linear algebra. We will assume that we have a solution defined near $(0,0)$ with $z(0,0)=0$. If $a \neq b$, then by a formal power series argument as in the proof of Proposition 4.1 below, we have that if $z \in C^3$ then near $(0,0)$ the first few terms of its Taylor series are $z = \frac{1}{2}(\pm ax^2 \pm by^2) + O((|x|+|y|)^3)$, and if $a = b$ we can bring z to this form by a rotation of the axes. If $z = \frac{1}{2}(ax^2 + by^2) + O((|x|+|y|)^3)$, then Theorem 2.7 applies, since the function is convex. Hence, $z = \frac{1}{2}(ax^2+by^2)$ is the unique solution. Likewise, if $z = -\frac{1}{2}(ax^2+by^2) + O((|x|+|y|)^3)$, then $z = -\frac{1}{2}(ax^2+by^2)$.

Next we look for a solution of the form $z = \frac{1}{2}(ax^2 - by^2) + O((|x|+|y|)^3)$. We will use the Hamiltonian $H = \frac{1}{2}(p^2+q^2-a^2x^2-b^2y^2)$. The characteristic vector field is then $\xi_H = p\frac{\partial}{\partial x} + q\frac{\partial}{\partial y} + a^2x\frac{\partial}{\partial p} + b^2y\frac{\partial}{\partial q}$. The integral curves of this vector field satisfy the differential equation (*cf.* (2.4))

$$\frac{d}{dt}\begin{bmatrix} x \\ y \\ p \\ q \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ a^2 & 0 & 0 & 0 \\ 0 & b^2 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ p \\ q \end{bmatrix} = L \begin{bmatrix} x \\ y \\ p \\ q \end{bmatrix}.$$

The eigenvectors of the matrix L are

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ a \\ 0 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ -b \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 1 \\ 0 \\ -a \\ 0 \end{bmatrix}, \quad \mathbf{v}_4 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ b \end{bmatrix}, \tag{3.1}$$

which satisfy

$$L\mathbf{v}_1 = a\mathbf{v}_1, \quad L\mathbf{v}_2 = -b\mathbf{v}_2, \quad L\mathbf{v}_3 = -a\mathbf{v}_3, \quad L\mathbf{v}_4 = b\mathbf{v}_4.$$

The jet of the function $z = \frac{1}{2}(ax^2 - by^2)$ is the span of $\mathbf{v}_1$ and $\mathbf{v}_2$.

To construct an invariant Lagrangian surface through $(0,0,0,0)$ we start with two curves $\gamma_\pm \colon \mathbf{R} \to \mathbf{R}^4$ and generate the surface by moving these curves by the flow of the characteristic vector field. Because of the nature of the singularity, two curves rather than just one are required. Toward this end let $\varphi_+, \varphi_- \colon \mathbf{R} \to \mathbf{R}$ be two functions so that

(1) $\varphi_\pm$ are C^k functions for some $k \geq 1$ and
(2) $\varphi_\pm$ vanish to order l at the origin for some $l \leq k$ (where l need not be an integer). Specifically, this means that there are functions $\overline{\varphi}_\pm \colon \mathbf{R} \to \mathbf{R}$ so that

$$\varphi_\pm(s) = \overline{\varphi}_\pm(s)s^l \tag{3.2}$$

where $\overline{\varphi}_\pm(s)$ are C^k functions on $\mathbf{R}\smallsetminus\{0\}$ and the derivatives $\frac{d^j}{ds^j}\overline{\varphi}_\pm(s)$ are bounded on $[-1,1]\smallsetminus\{0\}$ for $0\le j\le l$.

Define two curves $\gamma_+,\gamma_-:\mathbf{R}\to\mathbf{R}^4$ by

$$\gamma_\pm(s)=s\mathbf{v}_1\pm s\mathbf{v}_2+\varphi_\pm(s)\mathbf{v}_3\mp\frac{a^2}{b^2}\varphi_\pm(s)\mathbf{v}_4.$$

Then $H(\gamma_\pm(s))=0$ for all s. Let e^{tL} be the exponential of L so that for $P\in\mathbf{R}^4$ the integral curve of the characteristic vector field through P is $t\mapsto e^{tL}P$. Define $F_\pm\colon\mathbf{R}^2\to\mathbf{R}^4$ by

$$F_\pm(s,t):=e^{tL}\gamma_\pm(s)=se^{at}\mathbf{v}_1\pm se^{-bt}\mathbf{v}_2+\varphi_\pm(s)e^{-at}\mathbf{v}_3\mp\frac{a^2}{b^2}\varphi_\pm(s)e^{bt}\mathbf{v}_4. \quad (3.3)$$

We now consider $F_+(s,t)$ for $s>0$. Do the change of variables $(s,t)\mapsto(u,v)$ with $u,v>0$ given by

$$\begin{cases} u=se^{at}\\ v=se^{-bt}\end{cases}\qquad \begin{cases} t=\dfrac{1}{a+b}\ln\dfrac{u}{v}\\ s=u^{b/(a+b)}v^{a/(a+b)}\end{cases}.$$

In these coordinates we have that the image of $F_+(s,t)$ with $s>0$ can be parametrized by

$$\begin{aligned}G(u,v)=&u\mathbf{v}_1+v\mathbf{v}_2\\&+u^{\frac{-a}{a+b}}v^{\frac{a}{a+b}}\varphi_+(u^{\frac{b}{a+b}}v^{\frac{a}{a+b}})\mathbf{v}_3-\frac{a^2}{b^2}u^{\frac{b}{a+b}}v^{\frac{-b}{a+b}}\varphi_+(u^{\frac{b}{a+b}}v^{\frac{a}{a+b}})\mathbf{v}_4.\end{aligned}$$

Lemma 3.1. *Let φ be a function on $[0,\infty)$ of class C^k that vanishes to order $l\le k$ at $t=0$, and let $1>\alpha,\beta>0$. Then the function*

$$E(u,v):=\begin{cases}u^{\alpha-1}v^\beta\varphi(u^\alpha v^\beta), & u,v>0\\ 0, & \textit{otherwise}\end{cases}$$

is C^n on $\mathbf{R}^2$ for all $n<\min\{(l+1)\alpha-1,(l+1)\beta\}$. Conversely, if $E(u,v)$ is of class C^n, then φ vanishes of order l at $t=0$ for all $l>\max\{(n+1)/\alpha-1,n/\beta-1\}$.

Proof. The hypothesis on φ implies that $\varphi(t)=t^l\psi(t)$ where ψ has continuous derivatives up to order k on the open interval $(0,\infty)$ and is bounded on the closed interval $[0,\infty)$. Thus E can be rewritten as

$$E(u,v)=u^{(l+1)\alpha-1}v^{(l+1)\beta}\psi(u^\alpha v^\beta)\quad\text{for } u,v>0.$$

The result now follows by direct calculation of the derivatives. □

This directly implies

Theorem 3.2. *Let $\varphi_\pm\colon [0,\infty) \to \mathbf{R}$ be C^k functions that vanish to order $l \le k$ at 0. Let $F_\pm$ be defined by equation (3.3) and let $N^2(\varphi_+, \varphi_-)$ be the closure of the union of the images of F_+ and F_-; that is,*

$$N^2(\varphi_+, \varphi_-) := \overline{\operatorname{image}(F_+) \cup \operatorname{image}(F_-)}.$$

Set $n := \lceil \min\{\frac{(l+1)a}{a+b}, \frac{(l+1)b}{a+b}\}\rceil - 1$, where $\lceil\cdot\rceil$ is the ceiling function. Then $N^2(\varphi_+, \varphi_-)$ is a C^{n-1} submanifold of $\mathbf{R}^4$ that is the jet of a C^n solution $z(x,y)$ to the equation $z_x^2 + z_y^2 = a^2x^2 + b^2y^2$. Conversely, any C^n solution of this equation is of this form for unique φ_+ and φ_-. Regardless of the value of l, the function z will be of class C^{k+1} on $\mathbf{R}^2 \smallsetminus \{xy = 0\}$. Thus, there is a decrease in regularity of solutions from C^{k+1} to C^n along the coordinate axes.

Proof. All but the statements about the decrease in regularity along the coordinate axes follow from Lemma 3.1. The two curves where the regularity of the surface $N^2(\varphi_+, \varphi_-)$ drops are along the lines $u \mapsto G(u,0)$ (which projects down onto the x-axis) and $v \mapsto G(0,v)$ (which projects down onto the y-axis). All other points of $N^2(\varphi_+, \varphi_-)$ are C^k. As in Proposition 2.2, this implies that z is C^{k+1} off of the coordinate axes.

3.2. **Solutions to $H(x,y,z_x,z_y) = 0$ near a critical point of H.** We now consider an equation $H(x,y,z_x,z_y) = 0$ near a critical point of H. We assume that $H \in C^\infty$, $H(0,0,0,0) = 0$ and $dH = 0$ at the origin. Let L be the linearization at the origin of the characteristic system defined by (2.4) and assume that the eigenvalues of L are $a, b, -a, -b$ where $a, b > 0$. (By (2.5) the eigenvalues are of this form if they are real and nonzero.) Let e_1, e_2, e_3, e_4 be the eigenvectors for $a, b, -a, -b$ respectively. We make the two assumptions

$$a \text{ and } b \text{ are linearly independent over the rational numbers,} \tag{3.4}$$

and

$$\left\{\begin{array}{l}\{\pi_*e_1, \pi_*e_2\},\ \{\pi_*e_3, \pi_*e_4\},\ \{\pi_*e_1, \pi_*e_4\},\ \{\pi_*e_2, \pi_*e_3\}\\ \text{are each linearly independent sets.}\end{array}\right. \tag{3.5}$$

(Where, as usual, $\pi\colon \mathbf{R}^4 \to \mathbf{R}^2$ is $\pi(x,y,p,q) = (x,y)$.) This and the implicit function theorem imply that, if N^2 is any C^1 surface in $\mathbf{R}^4$ that is tangent to both e_1 and e_2 (or one of the other pairs $\{e_3, e_4\}$, $\{e_1, e_4\}$ or $\{e_2, e_3\}$), then locally N^2 projects over an open neighborhood of the origin in $\mathbf{R}^2$.

The number-theoretic assumption (3.4) is what is needed to invoke the Sternberg normal form (Theorem A.2 in the Appendix) and conclude there are local coordinates $\overline{x}, \overline{y}, \overline{p}, \overline{q}$ centered at the origin of $\mathbf{R}^4$ with $\omega = d\overline{p} \wedge$

$d\overline{x} + d\overline{q} \wedge d\overline{y}$ and a function $f(u,v)$ with $f(0,0) = 0$, $f_u(0,0) = f_v(0,0) = 1$ and $H = \frac{1}{2} f(\overline{p}^2 - a^2\overline{x}^2, \overline{q}^2 - b^2\overline{y}^2)$. In the coordinates $\overline{x}, \overline{y}, \overline{p}, \overline{q}$ the eigendirections of L are given by the vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$ of (3.9) below. So in these coordinates we can take $e_1 = \mathbf{v}_1$, $e_2 = \mathbf{v}_4$, $e_3 = \mathbf{v}_3$ and $e_4 = \mathbf{v}_2$. We then can compute that $\omega(e_1, e_2) = \omega(e_3, e_4) = \omega(e_1, e_4) = \omega(e_2, e_3) = 0$, but that $\omega(e_1, e_3), \omega(e_2, e_4) \neq 0$. This leads to a general existence result.

Theorem 3.3. *Let the origin of $\mathbf{R}^4$ be a nondegenerate critical point of H such that the linearization L of the characteristic vector field ξ_H at the origin satisfies the conditions* (3.4) *and* (3.5). *Then, with the notation above, each of the pairs $\{e_1, e_2\}$, $\{e_3, e_4\}$, $\{e_1, e_4\}$ and $\{e_2, e_3\}$ is tangent to the jet of a C^∞ solution to $H(x, y, z_x, z_y) = 0$.*

Proof. Consider the four submanifolds $N^2_{\pm,\pm}$ of $\mathbf{R}^4$ defined locally near the origin in the coordinates $\overline{x}, \overline{y}, \overline{p}, \overline{q}$ by

$$N^2_{\pm,\pm} := \{(\overline{x}, \overline{y}, \pm a\overline{x}, \pm b\overline{y}) : (\overline{x}, \overline{y}) \in U\},$$

where U is a small neighborhood of $(0,0)$ in $\mathbf{R}^2$. On $N_{\pm,\pm}$ the relations $\overline{p} = \pm ax$ and $\overline{q} = \pm by$ hold. Thus on $N^2_{\pm,\pm}$ we have $H = \frac{1}{2} f(\overline{p}^2 - a^2\overline{x}^2, \overline{q}^2 - b^2\overline{y}^2) = \frac{1}{2} f(0,0) = 0$, and therefore $N^2_{\pm,\pm} \subset \{H = 0\}$. A direct calculation using the form of the characteristic vector field ξ_H in the coordinates $\overline{x}, \overline{y}, \overline{p}, \overline{q}$ (see (3.7) below) shows that ξ_H is tangent to $N^2_{\pm,\pm}$. Thus, $N^2_{\pm,\pm}$ is an invariant Lagrangian surface. We also have

$$T(N^2_{+,+})_0 = \operatorname{span}(e_1, e_2), \quad T(N^2_{-,-})_0 = \operatorname{span}(e_3, e_4),$$
$$T(N^2_{+,-})_0 = \operatorname{span}(e_1, e_4), \quad T(N^2_{-,+})_0 = \operatorname{span}(e_2, e_3).$$

Therefore the assumption (3.5) implies that locally near $(0,0,0,0)$ each of the surfaces $N^2_{\pm,\pm}$ projects onto an open neighborhood of $(0,0)$ in $\mathbf{R}^2$. Consequently Proposition 2.1 implies that $N^2_{\pm,\pm}$ is the jet of a C^∞ solution to $H(x, y, z_x, z_y) = 0$. $\square$

If $N^2 \subset \mathbf{R}^4$ is the jet of a solution to $H(x, y, z_x, z_y) = 0$ that passes through $(0,0,0,0)$, then N^2 is an invariant Lagrangian surface in $\mathbf{R}^4$. The invariance of N^2 under the flow of the characteristic vector field ξ_H implies that $T(N^2)_0$ is invariant under the linearization L of ξ_H. This, coupled with the fact that ω vanishes on $T(N^2)_0$, implies that $T(N^2)_0$ is one of the four subspaces $\operatorname{span}(e_1, e_2)$, $\operatorname{span}(e_3, e_4)$, $\operatorname{span}(e_1, e_4)$, or $\operatorname{span}(e_2, e_3)$ of $T(\mathbf{R}^4)_0$.

The subspace $\operatorname{span}(e_1, e_2)$ is tangent to $N^2_{+,+}$ at the origin. This is the unstable submanifold of ξ_H at $(0,0,0,0)$ and is unique in the sense that if $\widetilde{N}^2$ is a C^1 invariant submanifold for ξ_H with $T(\widetilde{N}^2)_0 = \operatorname{span}(e_1, e_2)$ then $\widetilde{N}^2 = N^2_{+,+}$ in a neighborhood of the origin. To see this, note that the restriction $\xi_H\big|_{\widetilde{N}^2}$ has a source at the origin (as the eigenvalues of the restriction $L\big|_{\widetilde{N}^2}$ are $a, b > 0$), and so for all points P of $\widetilde{N}^2$ sufficiently near the origin $\lim_{t\to-\infty} \Phi_t(P) = (0,0,0,0)$. But this is exactly the condition that P be on the unstable submanifold. Thus $\widetilde{N}^2 = N^2_{+,+}$ as claimed. There is a corresponding uniqueness statement for the stable submanifold $N^2_{-,-}$.

Call a solution z to $H(x,y,z_x,z_y) = 0$ with $z_x(0,0) = z_y(0,0) = 0$ a *solution with unstable jet* if and only if its jet near the origin is the unstable submanifold $N^2_{+,+}$ for ξ_H. Likewise, z is a *solution with stable jet* if and only if its jet near the origin is the stable submanifold $N^2_{-,-}$ for ξ_H. Our discussion leads at once to a uniqueness result.

Theorem 3.4. *If the hypotheses of Theorem* 3.3 *hold and if z is a solution to $H(x,y,z_x,z_y) = 0$ with $z(0,0) = z_x(0,0) = z_y(0,0) = 0$ of either stable or unstable type, then z is unique in the sense that if $\tilde{z}$ is any C^2 solution with $\tilde{z}(0,0) = \tilde{z}_x(0,0) = \tilde{z}_y(0,0) = 0$, such that the jet of $\tilde{z}$ is tangent to the jet of z at $(0,0,0,0)$, then $\tilde{z} = z$ in a neighborhood of the origin. (The tangency condition of the jets is equivalent to $\tilde{z}_{xx}(0,0) = z_{xx}(0,0)$, $\tilde{z}_{xy}(0,0) = z_{xy}(0,0)$ and $\tilde{z}_{yy}(0,0) = z_{yy}(0,0)$.)*

Call a solution to $H(x,y,z_x,z_y) = 0$ with $z(0,0) = z_x(0,0) = z_y(0,0) = 0$ whose jet at the origin is not the stable or unstable submanifold of the characteristic vector field ξ_H a *saddle solution*. The motivation of this terminology is that for the eikonal equation $z_x^2 + z_y^2 = a^2x^2 + b^2y^2$ with $a, b > 0$ the solution with unstable jet is the convex solution $z = \frac{1}{2}(ax^2 + by^2)$ and the solution with stable jet is the concave solution $z = -\frac{1}{2}(ax^2 + by^2)$. The other two obvious solutions $z = \pm\frac{1}{2}(ax^2 - by^2)$ (which are not unique by Theorem 3.2) have saddle points at the origin.

We now wish to investigate in detail the lack of uniqueness of saddle solutions. To do this it is more convenient to work in the coordinates $\overline{x}, \overline{y}, \overline{p}, \overline{q}$. To simplify notation we drop the bars and just write x, y, p, q, but keep in mind that these are not the original symplectic coordinates on $\mathbf{R}^4$ and that once we have constructed the jets of solutions we have to translate these results back to statements in the original symplectic coordinates. With this in mind, we assume that H has form

$$H = \tfrac{1}{2} f(p^2 - a^2x^2, q^2 - b^2y^2), \tag{3.6}$$

where $f(0,0) = 0$, $f_u(0,0) = f_v(0,0) = 1$, and $\omega = dp \wedge dx + dq \wedge dy$. Then the characteristic vector field is

$$\xi_H = f_u p \frac{\partial}{\partial x} + f_v q \frac{\partial}{\partial y} + a^2 f_u x \frac{\partial}{\partial p} + b^2 f_v y \frac{\partial}{\partial q}. \tag{3.7}$$

The integral curves of this vector field satisfy the differential equation

$$\frac{d}{dt}\begin{bmatrix} x \\ y \\ p \\ q \end{bmatrix} = \begin{bmatrix} 0 & 0 & f_u & 0 \\ 0 & 0 & 0 & f_v \\ f_u a^2 & 0 & 0 & 0 \\ 0 & f_v b^2 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ p \\ q \end{bmatrix}. \tag{3.8}$$

This is not a linear system, as the functions f_u and f_v depend on x, y, p and q. However, we will show that because of its special structure it can be treated almost as if it were linear. Letting L be the matrix of this system we find that its eigenvectors are (*cf.* (3.1))

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ a \\ 0 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ -b \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 1 \\ 0 \\ -a \\ 0 \end{bmatrix}, \quad \mathbf{v}_4 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ b \end{bmatrix}, \tag{3.9}$$

and that

$$L\mathbf{v}_1 = f_u a \mathbf{v}_1, \quad L\mathbf{v}_2 = -f_v b \mathbf{v}_2, \quad L\mathbf{v}_3 = -f_u a \mathbf{v}_3, \quad L\mathbf{v}_4 = f_v b \mathbf{v}_4.$$

Thus L has a basis of eigenvectors that are independent of the variables x, y, p and q. For use in the statement and proof of Theorem 3.9 we define the two curves

$$\mathbf{a}_1(t) = t\mathbf{v}_1, \quad \mathbf{a}_2(t) = t\mathbf{v}_2, \tag{3.10}$$

which are just the first two coordinate axes for the basis $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$. What is special about these curves is that while we will construct many surfaces that are jets of solutions and tangent to $\operatorname{span}(\mathbf{v}_1, \mathbf{v}_2)$, these curves will lie on all of these surfaces.

Let w_1, w_2, w_3, w_4 be coordinates on $\mathbf{R}^4$ such that

$$\begin{bmatrix} x \\ y \\ p \\ q \end{bmatrix} = w_1\mathbf{v}_1 + w_2\mathbf{v}_2 + w_3\mathbf{v}_3 + w_4\mathbf{v}_4.$$

Then

$$\begin{cases} w_1 = \frac{1}{2}\left(x + p/a\right), & w_2 = \frac{1}{2}\left(y - q/b\right), \\ w_3 = \frac{1}{2}\left(x - p/a\right), & w_4 = \frac{1}{2}\left(y + q/b\right). \end{cases} \tag{3.11}$$

In these coordinates the differential equations for the characteristics become

$$\begin{aligned} w_1'(t) &= f_u a w_1(t), \quad & w_2'(t) &= -f_v b w_2(t), \\ w_3'(t) &= -f_u a w_3(t), \quad & w_4'(t) &= f_v b w_4'(t). \end{aligned}$$

Recall that $f_u(0,0) = f_v(0,0) = 1$. Therefore, given $\varepsilon > 0$ there is a $\delta > 0$ so that if $B(\delta) := \{(w_1, w_2, w_3, w_4) : \sum_{i=1}^4 w_i^2 \le \delta^2\}$ is the closed ball of radius δ at the origin, then $|f_u - 1|, |f_v - 1| < \varepsilon \le 1/2$ in $B(\delta)$. We now use the fact that if a function $w(t)$ satisfies a differential equation $w'(t) = c\,h(t, w(t))w(t)$ on an interval $[-\kappa, \kappa]$, where c is a constant and $|h(t, w(t)) - 1| \le \varepsilon$, then on $[-\kappa, \kappa]$, $w(t) = w(0)e^{c\theta(t)t}$ for a function $\theta(t)$ that satisfies $|1 - \theta| \le \varepsilon$. The differential equations for the characteristics are all of this form, so if $(w_1(0), w_2(0), w_3(0), w_4(0)) \in B(\delta)$ and t is so that $(w_1(\tau), w_2(\tau), w_3(\tau), w_4(\tau)) \in B(\delta)$ for τ between 0 and t, then

$$w_1(t) = w_1(0)e^{a\theta_1 t}, \quad w_2(t) = w_2(0)e^{-b\theta_2 t},$$
$$w_3(t) = w_3(0)e^{-a\theta_3 t}, \quad w_4(t) = w_4(0)e^{b\theta_4 t},$$

where $|1-\theta_i| \le \varepsilon \le \frac{1}{2}$ for $i = 1, 2, 3, 4$. Let $\varphi_+, \varphi_- \colon \mathbf{R} \to \mathbf{R}$ be two functions so that

(1) $\varphi_\pm$ are C^k functions for some $k \ge 1$ and
(2) $\varphi_\pm$ vanished to order l at the origin for some $l \le k$.

Lemma 3.5. *There are unique functions $\psi_\pm$ defined in a neighborhood of* 0 *so that if $\gamma_\pm$ are the curves*

$$\gamma_\pm(s) := s\mathbf{v}_1 \pm s\mathbf{v}_2 + \varphi_\pm(s)\mathbf{v}_3 + \psi_\pm(s)\mathbf{v}_4,$$

then

$$H(\gamma_\pm(s)) \equiv 0.$$

Moreover,

(1) *$\psi_\pm$ are C^k functions, and*
(2) *$\psi_\pm$ vanish to the same order $l \le k$ at the origin that $\varphi_\pm$ do. As above, this means there are $\overline{\psi}_\pm \colon \mathbf{R} \to \mathbf{R}$ so that*

$$\psi_\pm(s) = \overline{\psi}_\pm(s)s^l \tag{3.12}$$

where $\overline{\psi}_\pm(s)$ are C^k functions on $\mathbf{R}\setminus\{0\}$ and the derivatives $\frac{d^j}{ds^j}\overline{\psi}_\pm(s)$ are bounded on $[-1,1] \setminus \{0\}$ for $0 \le j \le l$.

(*In the case $H = \frac{1}{2}(p^2 + q^2 - a^2x^2 - b^2y^2)$ we have $\psi_\pm(s) = \mp(a^2/b^2)\varphi_\pm(s)$.*)

Proof. We will show how to find ψ_+, the derivation for ψ_- being similar. In terms of the coordinates w_1, w_2, w_3, w_4, we are looking for $\psi_+(s)$ so that if $w_1 = s$, $w_2 = s$, $w_3 = \varphi_+(s)$, and $w_4 = \psi_+(s)$, then $\gamma_+(s) = w_1\mathbf{v}_1 + w_2\mathbf{v}_2 + w_3\mathbf{v}_3 + w_4\mathbf{v}_4$ is the required curve. Writing H in terms of the coordinates x, y, p, q (related to the coordinates w_1, w_2, w_3, w_4 by (3.11)) we have $H = \frac{1}{2}f(p^2 - a^2x^2, q^2 - b^2y^2)$. Then we want ψ_+ so that $H(\gamma_+(s)) = \frac{1}{2}f(-4a^2s\varphi_+(s), -4b^2s\psi_+(s)) = 0$. Recalling the assumptions on f (*i.e.*, $f(0,0) = 0$ and $f_u(0,0) = f_v(0,0) = 1$) we have by the implicit function theorem that there is a smooth function $\tilde{v}(\cdot)$ defined in a neighborhood U of 0 with $\tilde{v}(0) = 0$, $\tilde{v}'(0) = -1$ so that $f(u, \tilde{v}(u)) \equiv 0$. Then near $(0,0)$ we have $\{(u,v) : f(u,v) = 0\} = \{(u, \tilde{v}(u)) : u \in U\}$. Therefore, near $s = 0$ we are trying to solve $-4b^2 s\psi_+(s) = v(-4a^2 s\varphi_+(s))$ for $\psi_+(s)$. As $\tilde{v}(u)$ vanishes at the origin it can be expressed as $\tilde{v}(u) = u\overline{v}(u)$, where $\overline{v}$ is a smooth function and $\tilde{v}'(0) = -1$ implies $\overline{v}(0) = -1$. Therefore, $\psi_+(s) = (a^2/b^2)\varphi_+(s)\overline{v}(-4a^2 s\varphi_+(s))$. Note that if $\varphi_+(s)$ vanishes to order l at $s = 0$, then this formula makes it clear that the same is true for $\psi_+(s)$. Finally, if $H = \frac{1}{2}(p^2 + q^2 - a^2x^2 - b^2y^2)$ then $f(u,v) = u + v$, and so $\tilde{v}(u) = -u$ and $\overline{v}(u) = -1$, which implies $\psi_+(s) = -(a^2/b^2)\varphi_+(s)$. $\square$

Assume that $\varphi_\pm$ is defined on the interval $(-s_0, s_0)$, and let Φ_t be the flow of ξ_H. Then define $F_\pm \colon (-s_0, s_0) \times \mathbf{R} \to \mathbf{R}^4$ by

$$F_\pm(s,t) := \Phi_t(\gamma_\pm(s)). \tag{3.13}$$

Writing this as

$$F_\pm(s,t) = w_1(s,t)\mathbf{v}_1 + w_2(s,t)\mathbf{v}_2 + w_3(s,t)\mathbf{v}_3 + w_4(s,t)\mathbf{v}_4$$

from the discussion of differential equations for the characteristics we have

$$\begin{cases} w_1(s,t) &= se^{a\theta_1 t}, \qquad\qquad w_2(s,t) = \pm se^{-b\theta_2 t}, \\ \\ w_3(s,t) &= \varphi_\pm(s)e^{-a\theta_3 t}, \quad w_4(s,t) = \psi_\pm(s)e^{b\theta_4 t}, \end{cases} \tag{3.14}$$

where $|\theta_i - 1| \le \varepsilon \le 1/2$. We are now going use $u = w_1(s,t)$ and $v = w_2(s,t)$ as new variables and express $w_3(s,t)$ and $w_4(s,t)$ in terms of these variables. As $w_2(s,t) = \pm se^{-b\theta_2 t}$ this calculation breaks up corresponding to the choice of signs. To simplify notation we will do the calculation in the first quadrant of the uv-plane, which corresponds to choosing $+$ in the definition of w_2, using the functions φ_+, ψ_+, and restricting to $s > 0$. The calculations in the other quadrants are identical. In the first quadrant

$$u = w_1(s,t) = se^{a\theta_1 t}, \quad v = w_2(s,t) = se^{-b\theta_2 t},$$

with s, u, and v positive. Solving for s and t we get

$$s = u^{b/(a+b)}v^{a/(a+b)}e^{ab(\theta_2-\theta_1)t}, \quad t = \frac{1}{a\theta_1 + b\theta_2}\ln\left(\frac{u}{v}\right).$$

Thus,

$$e^t = \left(\frac{u}{v}\right)^{1/(a\theta_1+b\theta_2)} = \left(\frac{u}{v}\right)^{(1/(a+b))+\rho_1}$$

where

$$\rho_1 = \frac{1}{a\theta_1 + b\theta_2} - \frac{1}{a+b} = \frac{(1-\theta_1)a + (1-\theta_2)b}{(a+b)(a\theta_1 + b\theta_2)}.$$

As we are assuming that $|1-\theta_i| \le \varepsilon \le 1/2$, this leads to

$$|\rho_1| \le \frac{2\varepsilon}{a+b} =: C_1\varepsilon.$$

This implies

$$\begin{aligned} s &= u^{\frac{b}{a+b}} v^{\frac{a}{a+b}} (e^t)^{ab(\theta_2-\theta_1)} \\ &= u^{\frac{b}{a+b}} v^{\frac{a}{a+b}} \left(\left(uv^{-1}\right)^{\frac{1}{a+b}+\rho_1} \right)^{ab(\theta_2-\theta_1)} = u^{\frac{b}{a+b}} v^{\frac{a}{a+b}} \left(uv^{-1}\right)^{\rho_2}, \end{aligned}$$

where $\rho_2 := (\frac{1}{a+b}+\rho_1)ab(\theta_2-\theta_1)$. As $|\theta_1-\theta_2| \le 2\varepsilon$, $|\rho_1| \le 2\varepsilon/(a+b) \le 1/(a+b)$,

$$|\rho_2| \le \frac{4ab\varepsilon}{a+b} =: C_2\varepsilon.$$

Now

$$\begin{aligned} w_3(s,t) &= \varphi_\pm(s)e^{-a\theta_3 t} = \varphi_\pm\left(u^{\frac{b}{a+b}} v^{\frac{a}{a+b}} \left(uv^{-1}\right)^{\rho_2}\right) \left(uv^{-1}\right)^{\left(\frac{1}{a+b}+\rho_1\right)(-a\theta_3)} \\ &= \varphi_\pm\left(u^{\frac{b}{a+b}} v^{\frac{a}{a+b}} \left(uv^{-1}\right)^{\rho_2}\right) \left(uv^{-1}\right)^{\left(\frac{-a}{a+b}+\rho_3\right)}, \end{aligned}$$

where $\rho_3 := \frac{(1-\theta_3)a}{a+b} - \rho_1 a\theta_3$. Using the estimates $|\rho_1| \le 2\varepsilon/(a+b)$ and $|1-\theta_3| \le \varepsilon \le 1/2$ (so that $1/2 \le \theta_3 \le 3/2$) this gives

$$|\rho_3| \le \frac{\varepsilon a}{a+b} + \frac{2\varepsilon}{a+b}\frac{3}{2}a = \frac{4a\varepsilon}{a+b} =: C_3\varepsilon.$$

Using the form of φ_+ given by equation (3.2) we have

$$\begin{aligned} w_3 &= \overline{\varphi}_+\left(u^{\frac{b}{a+b}} v^{\frac{a}{a+b}} \left(uv^{-1}\right)^{\rho_2}\right) \left(u^{\frac{b}{a+b}} v^{\frac{a}{a+b}} \left(uv^{-1}\right)^{\rho_2}\right)^l \left(uv^{-1}\right)^{\left(\frac{-a}{a+b}+\rho_3\right)} \\ &= \overline{\varphi}_+\left(u^{\frac{b}{a+b}} v^{\frac{a}{a+b}} \left(uv^{-1}\right)^{\rho_2}\right) u^{\frac{bl-a}{a+b}} v^{\frac{a(l+1)}{a+b}} (uv^{-1})^{l\rho_2+\rho_3}. \end{aligned} \tag{3.15}$$

Likewise

$$\begin{aligned} w_4(s,t) &= \psi_+(s)e^{b\theta_4 t} = \psi_+\left(u^{\frac{b}{a+b}} v^{\frac{a}{a+b}} \left(uv^{-1}\right)^{\rho_2}\right) \left(uv^{-1}\right)^{\left(\frac{1}{a+b}+\rho_1\right)b\theta_4} \\ &= \psi_+\left(u^{\frac{b}{a+b}} v^{\frac{a}{a+b}} \left(uv^{-1}\right)^{\rho_2}\right) \left(uv^{-1}\right)^{\left(\frac{b}{a+b}+\rho_4\right)}, \end{aligned}$$

where $\rho_4 = \frac{(\theta_4-1)b}{a+b} + \rho_1 b\theta_4$. Thus,

$$|\rho_4| \le \frac{\varepsilon b}{a+b} + \frac{2\varepsilon}{a+b}\frac{3}{2}b = \frac{4b\varepsilon}{a+b} =: C_4\varepsilon.$$

Using the form of ψ_+ given by equation (3.12) we have

$$\begin{aligned} w_4 &= \overline{\psi}_+\big(u^{\frac{b}{a+b}} v^{\frac{a}{a+b}} \left(uv^{-1}\right)^{\rho_2}\big) \left(u^{\frac{b}{a+b}} v^{\frac{a}{a+b}} \left(uv^{-1}\right)^{\rho_2}\right)^l \left(uv^{-1}\right)^{\left(\frac{b}{a+b}+\rho_4\right)} \\ &= \overline{\psi}_+\big(u^{\frac{b}{a+b}} v^{\frac{a}{a+b}} \left(uv^{-1}\right)^{\rho_2}\big) u^{\frac{b(l+1)}{a+b}} v^{\frac{al-b}{a+b}} (uv^{-1})^{l\rho_2+\rho_4}. \end{aligned} \tag{3.16}$$

Definition 3.6. Let k be a positive integer, m a positive real number, and U a connected neighborhood of the origin in $\mathbf{R}^2$. Then we denote the class of functions defined on U that are C^k off of the coordinate axes and which vanish to order m along the axes by $C^{(k,m)}_{\text{axes}} = C^{(k,m)}_{\text{axes}}(U)$. Explicitly $h \in C^{(k,m)}_{\text{axes}}(U)$ if and only if h is C^k on $U \smallsetminus \{uv = 0\}$ and for any compact subset V of U there is a constant $C = C_V$ so that

$$|h(u,v)| \le C|u|^m \quad \text{and} \quad |h(u,v)| \le C|v|^m$$

hold for all $(u,v) \in V$. $\square$

Using this definition we now summarize and extend the calculations above.

Proposition 3.7. *Let $H\colon \mathbf{R}^4 \to \mathbf{R}$ of class C^∞ be given by* (3.6), *let $\varphi_\pm\colon \mathbf{R} \to \mathbf{R}$ be C^k (with $1 \le k \le \infty$) functions that vanish to order l at the origin* (*as in* (3.2)), *and let $\psi_\pm\colon \mathbf{R} \to \mathbf{R}$ and the curves $\gamma_\pm\colon \mathbf{R} \to \mathbf{R}^4$ be given by Lemma* 3.5. *Let $w_1(s,t), \dots, w_4(s,t)$ be the solutions of the characteristic equations given by* (3.14) *and define $u(s,t) = w_1(s,t)$, $v(s,t) = w_2(s,t)$. As above, we can then solve for s and t in terms of u and v. Then define functions on a neighborhood of the origin in $\mathbf{R}^2$ by*

$$g(u,v) := w_3(s,t), \quad h(u,v) = w_4(s,t), \quad \textit{for } u, v \ne 0,$$

and extend this to the coordinate axes by

$$g(u,0) = g(0,v) = h(u,0) = h(0,v) = 0.$$

Let m be any real number with

$$m < \min\left\{\frac{bl-a}{a+b}, \frac{al-b}{a+b}\right\}.$$

Then there is a connected neighborhood U of the origin such that $g, h \in C^{(k,m)}_{\text{axes}}(U)$. If $\varphi_\pm(s) = 0$ in a small neighborhood of $s = 0$, then g and h vanish in a neighborhood of the coordinate axes, and are C^k. The surface

$$N^2 = N^2(\varphi_+, \varphi_-) := \{u\mathbf{v}_1 + v\mathbf{v}_2 + g(u,v)\mathbf{v}_3 + h(u,v)\mathbf{v}_4 : (u,v) \in U\}$$

is an invariant Lagrangian submanifold of $\mathbf{R}^4$ for some sufficiently small neighborhood U of the origin in $\mathbf{R}^2$.

Proof. Most of this follows at once from the calculation above. The flow Φ_t of the characteristic vector field is C^∞ and the curves $\gamma_\pm$ are C^k. Therefore $w_i(s,t)$ is a C^k function of (s,t) for $s \neq 0$. So by the inverse function theorem s and t are C^k functions of (u,v), and g and h are C^k functions off of the coordinate axes $u=0$ and $v=0$. Along the $u=0$ axis the formula (3.15) for w_3 in terms of u and v implies that

$$|g(u,v)| = |w_3| \leq C|u|^{\frac{bl-a}{a+b}}(uv^{-1})^{|l\rho_2+\rho_3|}$$

on any compact set. We have

$$|l\rho_2 + \rho_3| \leq (lC_2 + C_3)\varepsilon.$$

Since $\frac{bl-a}{a+b} > m$ we can, by choosing δ small enough, make ε so small that $\frac{bl-a}{a+b} - (lC_2 + C_3)\varepsilon > m$. But then (for a possibly larger value of C)

$$|g(u,v)| = |w_3| \leq C|u|^m.$$

Similar considerations, using (3.15) and (3.16), show

$$|h(u,v)| \leq C|u|^m, \quad |g(u,v)| \leq C|v|^m, \quad |h(u,v)| \leq C|v|^m.$$

While this only shows that this estimate holds in a small neighborhood of $(0,0)$, we can extend it along the $u=0$ axis as follows. From the construction the subset $N^2 := \{u\mathbf{v}_1 + v\mathbf{v}_2 + g(u,v)\mathbf{v}_3 + h(u,v)\mathbf{v}_4 : (u,v) \in V\}$ is invariant under the local flow of Φ_t. Likewise the uv-plane, realized as $\mathbf{R}^2_{u,v} := \{u\mathbf{v}_1 + v\mathbf{v}_2 : (u,v) \in \mathbf{R}^2\}$, is invariant under the flow of Φ_t, as is the line defined by $u=0$, realized as $\mathbf{R}_v := \{v\mathbf{v}_2 : v \in \mathbf{R}\}$. The estimates above imply that in some small neighborhood of the origin N^2 and $\mathbf{R}^2_{u,v}$ have contact to order m along $\mathbf{R}_v$. But the flow Φ_t is smooth and therefore preserves the order of contact. Thus, invariance of N^2, $\mathbf{R}^2_{u,v}$ and $\mathbf{R}_v$ under Φ_t shows that N^2 and $\mathbf{R}^2_{u,v}$ have order m contact along $\mathbf{R}_v$ at all points that can be moved close to the origin by Φ_t. Similarly, we can extend the estimates along the $v=0$ axis.

Finally, from the construction we have that $N^2 \subset \{H=0\}$ and that the characteristic vector field ξ_H is tangent to N^2. Let X be a vector tangent to N^2; then $dH(X)=0$, because $N^2 \subset \{H=0\}$ and therefore H is constant on N^2. From the definition of ξ_H we have $\omega(\xi_H, X) = -dH(X) = 0$. As ξ_H is tangent to N^2 and $T(N^2)$ is two dimensional, this implies that $\omega = 0$ at points of N^2 where $\xi_H \neq 0$. But ξ_H has an isolated zero at the origin, and therefore by continuity $\omega\big|_{N^2} = 0$ in some neighborhood of the origin. This shows that N^2 is an invariant Lagrangian surface near the origin and completes the proof. $\square$

Recall that the coordinates x, y, p, q we have been working with are not the standard coordinates on $\mathbf{R}^4$, but rather the coordinates $\overline{x}, \overline{y}, \overline{p}, \overline{q}$ that reduced H to the Sternberg normal form. We now need to translate Proposition 3.7 back into a result about solutions to the original equation $H(x, y, z_x, z_y) = 0$. However, as the notion of being an invariant Lagrangian surface is invariant under a symplectic change of coordinates, this translation does not involve much calculation.

Let N^2 be the invariant Lagrangian surface from the conclusion of Proposition 3.7. Then from the construction of N^2 the two curves $\mathbf{a}_1$ and $\mathbf{a_2}$ in $\mathbf{R}^4$ defined by (3.10) lie in N^2. Thus $\mathbf{a}_1'(0) = \mathbf{v}_1$ and $\mathbf{a}_2'(0) = \mathbf{v}_2$ is a basis for $T(N^2)_0$. Then (recalling that $\mathbf{v}_1 = e_1$ and $\mathbf{v}_2 = e_4$ at the origin) the condition (3.5) and the implicit function theorem imply that by restricting the size of N^2 we can assume that it projects over a neighborhood of $(0, 0)$. Again letting $\pi\colon \mathbf{R}^4 \to \mathbf{R}^2$ be the natural projection, we then have that the two curves $\pi \circ \mathbf{a}_1$ and $\pi \circ \mathbf{a}_2$ cross transversely at the origin. Let $\tilde{a}_1(t) = \pi(\mathbf{a}_1(t))$ and $\tilde{a}_2(t) = \pi(\mathbf{a}_2(t))$. Then these curves depend only on the curves $\mathbf{a}_1$ and $\mathbf{a}_2$ and thus only depend on the Sternberg normal coordinates $\overline{x}, \overline{y}, \overline{p}, \overline{q}$ and the function H, but are independent of the invariant Lagrangian surface N^2. As $\tilde{a}_1$ and $\tilde{a}_2$ are C^∞ and cross transversely at the origin, there are C^∞ local coordinates $\tilde{u}, \tilde{v}$ centered at $(0, 0)$ so that for t near zero

$$\tilde{u}(\tilde{a}_2(t)) \equiv 0, \quad \tilde{v}(\tilde{a}_1(t)) \equiv 0.$$

Definition 3.8. Let k be a positive integer, m a positive real number, and U a connected neighborhood of the origin in $\mathbf{R}^2$ contained in the domain of the coordinates $\tilde{u}, \tilde{v}$. Then denote by $C^{(k,m)}_{\tilde{u}\text{-}\tilde{v}\ \text{axes}}(U)$ the class of functions defined exactly as in Definition 3.6 but with the coordinates $\tilde{u}$ and $\tilde{v}$ replacing the coordinates u and v.

With this definition we can now translate Proposition 3.7 into a statement about solutions to the differential equation $H(x, y, z_x, z_y) = 0$.

Theorem 3.9. *Use the notation of Proposition 3.7 and assume that (3.4) and (3.5) hold. Then given the C^k (with $1 \le k \le \infty$) functions $\varphi_\pm\colon \mathbf{R} \to \mathbf{R}$ there is a solution z to $H(x, y, z_x, z_y) = 0$ defined in a neighborhood U of the origin that has $N^2(\varphi_+, \varphi_-)$ as its jet near the origin. Moreover, for any real number m with*

$$m < \min\left\{\frac{bl - a}{a + b}, \frac{al - b}{a + b}\right\},$$

we have $z \in C^{(k+1,m+1)}_{\tilde{u}\text{-}\tilde{v}\ \text{axes}}(U)$.

Proof. That there is a solution z with $N^2 = N^2(\varphi_+, \varphi_-)$ as jet follows from Proposition 2.2. As $N^2 = \{(x, y, z_x, z_y) : (x, y) \in U\}$ there is a gain in regularity of one derivative in going from the jet N^2 to the function z. As N^2 has regularity $C^{(k,m)}_{\text{axes}}$, we see that z has regularity $C^{(k+1,m+1)}_{\tilde{u}\text{-}\tilde{v}\text{ axes}}$. This completes the proof. □

Remark 3.10. Choosing φ_+ and φ_- to vanish to infinite order at the origin we see that under assumptions (3.4) and (3.5) any saddle-type solution to $H(x, y, z_x, z_y) = 0$ will have an infinite-dimensional family of smooth deformations.

4. Formal power series and nonexistence of smooth solutions

In this section we will use power series methods to investigate the local existence and regularity of solutions to the eikonal equation,

$$z_x^2 + z_y^2 = h, \tag{4.1}$$

near a point where $h = 0$. We assume that h is smooth and that it has a nondegenerate Hessian at the zero we are considering. We make, if necessary, an affine change of coordinates of translation type followed by a rotation, to bring the series expansion at this zero into the form

$$h(x, y) = a^2x^2 + b^2y^2 + \sum_{m+n\geq 3} h_{m,n}x^m y^n \tag{4.2}$$

for some $a, b > 0$.

Proposition 4.1. *If z is a C^3 solution to* (4.1) *such that $z(0, 0) = 0$, and if $a \neq b$, then*

$$z(x, y) = \tfrac{1}{2}(\pm ax^2 \pm by^2) + O((|x| + |y|)^3). \tag{4.3}$$

If $a = b$, then (4.3) *still holds after making a change of coordinates*

$$\begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} \cos\xi & \sin\xi \\ -\sin\xi & \cos\xi \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \tag{4.4}$$

for some angle of rotation ξ.

Proof. Suppose that z has a series expansion $z(x, y) = \alpha x + \beta y + \frac{1}{2}(Ax^2 + 2Bxy + Cy^2) + O((|x| + |y|)^3)$. Then $z_x^2 + z_y^2 = \alpha^2 + \beta^2 + O(|x| + |y|)$. But h vanishes at $(0, 0)$ which, along with (4.1), implies that $\alpha = \beta = 0$. Evaluating the second-order terms of $z_x^2 + z_y^2$ and equating them with the second-order terms of (4.2), we get the equations

$$A^2 + B^2 = a^2, \qquad B(A + C) = 0, \qquad B^2 + C^2 = b^2. \tag{4.5}$$

Suppose first that $a \neq b$; then we must have $B = 0$ since otherwise the middle equation would imply that $A = -C$ which, together with the first and the last equations, contradict the assumption that $a \neq b$. It then follows immediately that $A = \pm a$ and $C = \pm b$, and so (4.3) holds. This holds also if $a = b$ and $B = 0$, *i.e.*, $A = \pm a$ and $C = \pm a$ with any combination of signs.

If $a = b$ and $B \neq 0$, then any solution of (4.5) is of the form

$$A_\theta = \pm a\sqrt{1-\theta^2}, \quad B_\theta = a\theta, \quad C_\theta = \mp a\sqrt{1-\theta^2},$$

for some $\theta \in [-1, 1]$, $\theta \neq 0$. For a given θ, the function $z = \frac{1}{2}(A_\theta x^2 + 2B_\theta xy + C_\theta y^2) + \cdots$ is brought to the form (4.3) by the change of variables (4.4), if ξ is chosen such that $\tan\xi = \theta/(1+\sqrt{1-\theta^2})$. $\square$

As the convex and concave solutions are well understood (these are the ones corresponding to leading terms $\pm(ax^2+by^2)/2$), we will focus on the saddle solutions. Thus, assume that z has Taylor expansion

$$z(x,y) = \frac{1}{2}(ax^2 - by^2) + \sum_{m+n\geq 3} z_{m,n}x^m y^n. \tag{4.6}$$

(Note that, if $a = b$, we work with the rotated coordinates.) Then

$$z_x = ax + \sum_{m+n\geq 3} m z_{m,n} x^{m-1} y^n$$

so that

$$z_x^2 = a^2x^2 + \sum_{m+n\geq 3} 2ma z_{m,n} x^m y^n + \Big(\sum_{m+n\geq 3} m z_{m,n} x^{m-1} y^n\Big)^2.$$

In $\left(\sum_{m+n\geq 3} m z_{m,n} x^{m-1} y^n\right)^2$ we get terms

$$m z_{m,n} x^{m-1} y^n k z_{k,l} x^{k-1} y^l = mk z_{m,n} z_{k,l} x^{m+k-2} y^{n+l} = mk z_{m,n} z_{k,l} x^p y^q,$$

where $p + q = m + n + k + l - 2 \geq k + l + 1$ as $m + n \geq 3$ and likewise $p+q \geq m+n+1$ as $k+l \geq 3$. Thus, there are polynomials $P_{m,n}^{[i+j<m+n]}(z_{i,j})$ in the variables $z_{i,j}$ with $i + j < m + n$ such that

$$\Big(\sum_{m+n\geq 3} m z_{m,n} x^{m-1} y^n\Big)^2 = \sum_{m+n\geq 4} P_{m,n}^{[i+j<m+n]}(z_{i,j}) x^m y^n.$$

Putting this all together (and setting $P_{m+n} = 0$ when $m + n = 3$) we get

$$z_x^2 = a^2x^2 + \sum_{m+n\geq 3} (2ma z_{m,n} + P_{m,n}^{[i+j<m+n]}(z_{i,j})) x^m y^n.$$

Likewise,

$$z_y = -bx + \sum_{m+n\geq 4} n z_{m,n} x^m y^{n-1},$$

so a similar calculation gives

$$z_y^2 = b^2y^2 + \sum_{m+n\geq 3} (-2nbz_{m,n} + Q_{m,n}^{[i+j<m+n]}(z_{i,j}))x^my^n.$$

Plugging these equations into (4.1) we get the recursion

$$2(ma - nb)z_{m,n} = h_{m,n} - \left(P_{m,n}^{[i+j<m+n]}(z_{i,j}) + Q_{m,n}^{[i+j<m+n]}(z_{i,j})\right). \tag{4.7}$$

Theorem 4.2. *Let h be a smooth function so that the Taylor expansion of $h(x,y)$ near $(0,0)$ is of the form (4.2), where a and b are positive and linearly independent over the rational numbers. Then at the level of formal power series there is a unique saddle-point solution z to (4.1) with leading terms of the form (4.6).*

Proof. If a and b are linearly independent over the rationals, then $ma-mb \neq 0$ for all m and n. Thus, the equation (4.7) uniquely determines all the coefficients of z as required. □

Corollary 4.3. *With the same hypothesis as in the last theorem, there is a smooth function $\tilde{z}$ so that $\tilde{z}_x^2 + \tilde{z}_y^2 - h$ vanishes to infinite order at the origin $(0,0)$.*

Proof. This follows from the well-known theorem of Borel (*cf.* [3] and [8, Theorem 1.2.6. p. 16]) that given any formal power series there is a smooth (*i.e.,* C^∞) function having the given formal power series as its Taylor series. Thus, if z is the formal power series solution given by the theorem, then let $\tilde{z}$ be a smooth function that has z as its Taylor expansion. □

Theorem 4.4. *Let a and b be positive real numbers that are linearly dependent over the rationals. Then there is a smooth function h of the form (4.2) such that there is no solution z of the saddle-point form (4.6) to the equation (4.1). In particular, if m and n are positive integers with $ma - nb = 0$ and $m+n \geq 4$ and $c \neq 0$, then for $h = a^2x^2+b^2y^2+cx^my^n$ the equation (4.1) will not have any saddle-point solution in formal power series (or any C^k saddle-point solution for $k \geq m+n$).*

Proof. If a and b are positive and dependent over the rationals, then there are positive integers m and n with $m+n \geq 4$ so that $ma-nb = 0$. Therefore, if there is a solution z to (4.1) of the form (4.6), the equation (4.7) implies

$$h_{m,n} = \left(P_{m,n}^{[i+j<m+n]}(z_{i,j}) + Q_{m,n}^{[i+j<m+n]}(z_{i,j})\right).$$

But then, by adding cx^my^m to h, $c \neq 0$, we get a new h of the form (4.2) so that there is no solution to (4.1) in formal power series. For a specific example, let $h_0 = a^2x^2 + b^2y^2$ and $z_0 = \frac{1}{2}(ax^2 - by^2)$. Then $(z_0)_x^2 + (z_0)_y^2 = h_0$ and the equation $z_x^2 + z_y^2 = h_0 + cx^my^n$ will have no solution in formal power series, if $ma - nb = 0$ and $c \neq 0$.

A. Appendix: The stable submanifold theorem and the Sternberg normal form

A.1. **The stable submanifold theorem.** The following is a statement of the stable submanifold theorem adapted to our applications. If $k \geq 1$ and X is a C^k vector field on a smooth n-dimensional manifold M, we denote by Φ_t the flow of X. That is, Φ_t is the locally defined one-parameter group of diffeomorphisms of M that satisfy $\frac{d}{dt}\Phi_t(P) = X(P)$. As X is C^k, the function $(P, t) \mapsto \Phi_t(P)$ is also C^k (*cf.* [1, p. 230] or [9, Theorem 5 p. 86]). A point $P_0 \in M$ is a *hyperbolic critical point* of X if and only if $X(P_0) = 0$ and the linearization of X at P_0, $L\colon T(M)_{P_0} \to T(M)_{P_0}$, has no eigenvalue with zero real part. This implies that P_0 is an isolated critical point of X. Letting, as usual, $e^{tL} = \sum_{k=0}^{\infty}(tL)/k!$, the *stable subspace* of X at P_0 is the linear subspace of $T(M)_{P_0}$ defined by

$$T_+(M)_{P_0} := \{v \in T(M)_{P_0} : \lim_{t\to\infty} e^{tL}v = 0\}.$$

The *unstable subspace* is likewise defined by

$$T_-(M)_{P_0} := \{v \in T(M)_{P_0} : \lim_{t\to-\infty} e^{tL}v = 0\}.$$

If no eigenvalue of L has zero real part, then basic linear algebra implies the direct-sum decomposition

$$T(M)_{P_0} = T_-(M)_{P_0} \oplus T_+(M)_{P_0}. \tag{A.1}$$

When all the eigenvalues of L are real and distinct (which is the case for most of the applications in this paper) then $T_+(M)_{P_0}$ is the span of the eigenvectors corresponding to the negative eigenvalues and $T_-(M)_{P_0}$ is the span of the eigenvectors corresponding to the positive eigenvalues.

Theorem A.1 (stable submanifold theorem)**.** *If $k \geq 1$ and P_0 is a hyperbolic critical point of the C^k vector field X, then there is a connected open neighborhood U of P_0 such that*

$$N_+ := \{P \in U : \lim_{t\to+\infty} \Phi_t(P) = P_0\}$$

(*the stable submanifold of X at P_0*) *and*

$$N_- := \{P \in U : \lim_{t\to-\infty} \Phi_t(P) = P_0\}$$

(the unstable submanifold of X at P_0) are both embedded C^k submanifolds of U. The tangent space to $N_\pm$ at P_0 is $T_\pm(M)_{P_0}$ so $\dim N_\pm = \dim T_\pm(M)_{P_0}$, and therefore $(A.1)$ implies that N_+ and N_- intersect transversely and that $\dim N_+ + \dim N_- = \dim M$.

This was originally proven in various degrees of generality by Hadamard [5], Liapunov [11] and Perron [13]. Modern presentations of the proof can be found in [6] and [16, Chapter 6].

A.2. **The Sternberg normal form.** Let H be a smooth function on $\mathbf{R}^4$ and ξ_H the characteristic vector field of H as given by (2.3). Let P be a critical point of H (and thus a rest point of ξ_H), and let L be the linearization of ξ_H at P. Then the matrix of L is given by equation (2.4). If the eigenvalues of L are real and nonzero, then the formulas (2.5) imply that they are of the form $a, b, -a, -b$, for some $a, b > 0$. With this in mind, we can give Sternberg's normal form for a Hamiltonian near a critical point.

Theorem A.2 (Sternberg [19]). *Let H be a smooth function on $\mathbf{R}^4$ and P a critical point of H with $H(P) = 0$. Assume that the eigenvalues of the linearization L of ξ_H at P are $a, b, -a, -b$ where $a, b > 0$ and that*

$$a \text{ and } b \text{ are linearly independent over the rational numbers.} \tag{A.2}$$

Then there are local coordinates $\overline{x}, \overline{y}, \overline{p}, \overline{q}$ centered at P and a smooth function $f(u, v)$ of two variables so that the following hold:

(1) *$\overline{x}, \overline{y}, \overline{p}, \overline{q}$ are symplectic coordinates. That is, if ω is the symplectic form, then*

$$\omega = d\overline{p} \wedge d\overline{x} + d\overline{q} \wedge d\overline{y}.$$

(2) *The function f satisfies $f(u, v) = u + v + O(u^2 + v^2)$ or more formally*

$$f(0,0) = 0, \quad f_u(0,0) = f_v(0,0) = 1.$$

(3) *In these coordinates H is given by*

$$H = \frac{1}{2} f(\overline{p}^2 - a^2\overline{x}^2, \overline{q}^2 - b^2\overline{y}^2).$$

This follows from Theorem 9 in Sternberg [19, p. 603] applied to the transformations given by the flow of a Hamiltonian vector field, using the considerations from Section 7 in Sternberg [18, pp. 818– 819].

Acknowledgments. We have benefited from numerous discussions with Björn Jawerth, and in fact this paper evolved out of Björn's seminars on the shape from shading problem.

References

[1] V.I. Arnold, "Ordinary Differential Equations," The M.I.T. Press, Cambridge, Mass.-London, 1973, translated from the Russian and edited by Richard A. Silverman.

[2] ______, "Geometrical Methods in the Theory of Ordinary Differential Equations," Springer-Verlag, New York, 1983, translated from the Russian by Joseph Szücs, translation edited by Mark Levi.

[3] L. Borel, *Sur quelques points de la théorie des fonctions*, Ann. Sci. École Norm. Sup. **12** (1895), 9–55.

[4] A. Bruss, *The eikonal equation: some results applicable to computer vision*, J. Math. Phys. **23** (1982), 890–896.

[5] J. Hadamard, *Sur l'iteration et les solutions asymptotiques des equations differentielles*, Bull. Soc. Math. France **29** (1901), 224–228.

[6] M.W. Hirsch, C.C. Pugh, and M. Shub, "Invariant Manifolds," Springer-Verlag, Berlin, 1977, Lecture Notes in Mathematics, Vol. 583.

[7] B.K.P. Horn, "Robot Vision," MIT Engineering and Computer Science Series, MIT Press, MacGraw Hill, 1986.

[8] L. Hörmander, "The Analysis of Linear Partial Differential Operators," I, second ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 256, Springer-Verlag, Berlin, 1990, Distribution Theory and Fourier Analysis.

[9] Serge Lang, "Differential Manifolds," Addison-Wesley Publishing Co., Inc., Reading, Mass.-London-Don Mills, Ont., 1972.

[10] P.L. Lions, E. Rouy, and A. Tourin, *Shape-from-shading, viscosity solutions and edges*, Numer. Math. **64** (1993), 323–353.

[11] A.M. Lyapunov, "The General Problem of the Stability of Motion," Taylor & Francis Ltd., London, 1992, translated from Edouard Davaux's French translation (1907) of the 1892 Russian original and edited by A.T. Fuller, with an introduction and preface by Fuller, a biography of Lyapunov by V.I. Smirnov, and a bibliography of Lyapunov's works compiled by J.F. Barrett, Lyapunov centenary issue, Reprint of Internat. J. Control **55** (1992), no. 3 with a foreword by Ian Stewart.

[12] J. Oliensis, *Uniqueness in shape from shading*, IJCV **6** (1991), 75–104.

[13] O. Perron, *Uber stabiltat und asymptotisches verhalten der integrale von differentialgleichungssystem*, Math. Z. **29** (1928), 129–160.

[14] E. Rouy and A. Tourin, *A viscosity solutions approach to shape-from-shading*, Siam J. Numer. Anal. **29** (1992), 867–884.

[15] B.V.H. Saxberg, *An application of dynamical systems theory to shape from shading*, DARPA89, 1989, pp. 1089–1104.

[16] M. Shub, A. Fathi, and R. Langevin, "Global Stability of Dynamical Systems," Springer-Verlag, New York, 1987, translated from the French by Joseph Christy.

[17] M. Spivak, "A Comprehensive Introduction to Differential Geometry," 2 ed., vol. 5, Publish or Perish Inc., Berkeley, 1979.

[18] Shlomo Sternberg, *Local contractions and a theorem of Poincaré*, Amer. J. Math. **79** (1957), 809–824.

[19] ______, *The structure of local homeomorphisms. III*, Amer. J. Math. **81** (1959), 578–604.