

GEOMETRIC THEORY OF EXTREMALS IN OPTIMAL CONTROL PROBLEMS: I THE FOLD AND MAXWELL CASE

I. KUPKA

ABSTRACT. The behavior of the extremal curves in optimal control theory is much more complex than that of their namesakes in the classical calculus of variations. Here we analyze the simplest instances of singular behavior of these extremals. Among others, in sharp contrast to the classical case, a C^0 -limit of extremals may not be an extremal. In the simplest cases (elliptic fold and Maxwell points) of this occurrence, the limits are trajectories of a new vector field. A special case of this field showed up in some work of V. I. Arnold. Results related to ours have been obtained in low dimension by I. Ekeland.

(0) Let us consider the following classical optimal control problem. Its state space is an open subset M of \mathbf{R}^d and its control space is a closed subset U in \mathbf{R}^c . A control system is given on M : $dx/dt = F(x, u)$, where $F: M \times U \rightarrow \mathbf{R}^d$ is a smooth (infinitely differentiable C^∞ or real analytic C^ω) vector field on M parametrized by U . Finally we are given a smooth cost function $c: M \times U \rightarrow \mathbf{R}$.

We consider the following problem: Given two points $A, B \in M$, find a pair of functions $(\bar{x}, \bar{u}): [0, \bar{T}] \rightarrow M \times U$ such that:

(1) \bar{x} is absolutely continuous, \bar{u} is measurable and bounded, and a.e. on $[0, \bar{T}]$:

$$\frac{d\bar{x}}{dt}(t) = F(\bar{x}(t), \bar{u}(t)).$$

(2) $\bar{x}(0) = A$, $\bar{x}(\bar{T}) = B$.

(3) $\int_0^{\bar{T}} c(\bar{x}(t), \bar{u}(t)) dt = \inf \{ \int_0^T c(x(t), u(t)) dt \}$ for all pairs $\{(x, u): [0, T] \rightarrow M \times U$ satisfying conditions (1) and (2) $\}$.

Pontrjagin's maximum principle tells us that such a pair (\bar{x}, \bar{u}) , if it exists, is the projection on $M \times U$ of a triple $(\bar{x}, \bar{p}, \bar{u}): [0, \bar{T}] \rightarrow T^*M \times U$ (where T^*M is the cotangent space of M that is $M \times (\mathbf{R}^d)^*$, with $(\mathbf{R}^d)^*$ the dual of \mathbf{R}^d) satisfying the following conditions:

(1) \bar{x} , \bar{p} are absolutely continuous, \bar{u} is bounded measurable, and a.e. on $[0, \bar{T}]$:

$$\frac{d\bar{x}}{dt}(t) = \frac{H^\lambda}{\partial p}(\bar{x}(t), \bar{p}(t), \bar{u}(t)), \quad \frac{d\bar{p}}{dt} = -\frac{\partial H^\lambda}{\partial x}(\bar{x}(t), \bar{p}(t), \bar{u}(t)),$$

where λ is 0 or 1 and $H^\lambda: T^*M \times U \rightarrow \mathbf{R}$ is the function $H^\lambda(x, p, u) = \langle p, F(x, u) \rangle - \lambda c(x, u)$.

Received by the editors August 16, 1985.

1980 *Mathematics Subject Classification*. Primary 49A10, 49B10; Secondary 57R45.

©1987 American Mathematical Society
0002-9947/87 \$1.00 + \$.25 per page

(2) $H^\lambda(\bar{x}(t), \bar{p}(t), \bar{u}(t)) = \sup\{H^\lambda(\bar{x}(t), \bar{p}(t), v) \mid v \in U\}$ almost everywhere on $[0, \bar{T}]$ and $H^\lambda(\bar{x}(t), \bar{p}(t), \bar{u}(t))$ is constant on $[0, \bar{T}]$.

(3) $\bar{x}(0) = A$, $\bar{x}(\bar{T}) = B$.

REMARK 1. Since we have not assumed in the control problem that the duration of the process is fixed, it is known that $H^\lambda(\bar{x}(t), \bar{p}(t), \bar{u}(t))$ is equal to zero a.e. on $[0, \bar{T}]$. If we had assumed that the duration is fixed, this would no longer be true.

DEFINITION 0. The projection $(x, p): [0, T] \rightarrow T^*M$ on T^*M of a triple $(x, p, u): [0, T] \rightarrow T^*M \times U$ satisfying conditions (1) and (2) will be called an extremal of the problem in accordance with the classical calculus of variations. The corresponding triple will be called an extremal triple.

REMARK 2. There are two families of extremals: the normal one corresponding to $\lambda = 1$ and the exceptional one for $\lambda = 0$. This last family is independent of the cost function. Since our study of extremals will not distinguish between these two cases we shall drop the index λ in H^λ for simplicity's sake.

The extremals which we have just defined are the generalizations of the extremals of the classical calculus of variations. But their behavior is much more complex and the "flow" they define has singularities. Here we shall study the simplest cases of the latter.

1. The geometric theory of extremals.

(A) *General background.* We shall make the following assumption on U :

Basic assumption: U is the closure of its interior and its boundary is a smooth (C^∞ or C^ω) hypersurface in \mathbf{R}^c (not necessarily connected!).

Notation 1. S will denote the subset of all points (x, p, u) in $T^*M \times U$ such that u is a local maximum point of the function $v \in U \rightarrow H(x, p, v)$. $\pi: S \rightarrow T^*M$ will denote the restriction to S of the canonical projection $T^*M \times U \rightarrow T^*M$.

DEFINITION 1. A point $z_0 \in T^*M$ is called simple if there exists a nbd V_0 of z_0 in T^*M such that $(\pi^{-1}(V_0), \pi, V_0)$ is a finite smooth covering.

It is clear that the set, Sim, of all simple points is an open subset of T^*M .

Notation 2. (a) SS will denote the set $\pi^{-1}(\text{Sim})$.

(b) $H_S: \text{SS} \rightarrow \mathbf{R}$ will denote the restriction of H to SS.

(c) $\mathcal{H}: \text{Sim} \rightarrow \mathbf{R}$ will be defined as follows:

$$\mathcal{H}(z) = \sup\{H(\sigma) \mid \sigma \in \pi^{-1}(z)\}.$$

(IMPORTANT) REMARK 3. It is clear that π induces a symplectic structure on SS by pull-back. Hence we can speak of the hamiltonian field of a function on SS.

Comment. H. Sussmann has given interesting examples showing that:

(1) in the C^∞ -case, any absolutely continuous curve can be an extremal,

(2) in the real analytic case, there are restrictions on the curves that can show up as extremals.

On the other hand, to our knowledge, the structure of generic extremals is not known.

(B) *Partial classification of the simple points.*

DEFINITION 2. (a) A simple point z_0 will be called regular if $\pi^{-1}(z_0)$ contains a unique point σ such that $H_S(\sigma) = \mathcal{H}(z_0)$. A nonregular simple point will be called a switching point. The set of these will be denoted by Σ .

(b) A switching point z_0 will be called a normal switching point if:

(1) $\pi^{-1}(z_0)$ contains exactly two points σ_1, σ_2 such that $H_S(\sigma_1) = H_S(\sigma_2) = \mathcal{H}(z_0)$.

(2) In a nbd of z_0 , Σ is a smooth manifold.

(3) The hamiltonian of H_S is transverse to $\pi^{-1}(\Sigma)$ at σ_1 and σ_2 .¹

(c) A switching point z_0 will be called a fold point if it satisfies conditions (1) and (2) in (b) and

(3) The hamiltonian of H_S has a contact of order exactly two with $\pi^{-1}(\Sigma)$ at σ_1 and σ_2 .

(d) A switching point z_0 will be called a Maxwell point if:

(1) There are exactly three points $\sigma_1, \sigma_2, \sigma_3$ in $\pi^{-1}(z_0)$ such that $H_S(\sigma_1) = H_S(\sigma_2) = H_S(\sigma_3) = \mathcal{H}(z_0)$.

(2) The projections in $T_{z_0}(T^*M)$ (the tangent space to T^*M at z_0) of the values of the hamiltonian of H_S at $\sigma_1, \sigma_2, \sigma_3$ are linearly independent.

Condition (2) implies that in a sufficiently small nbd of z_0 , Σ has the classical Y form: it is the union of three submanifolds of codimension 1, having a common boundary (edge) and meeting transversally along this common edge. Then the next, and last, condition makes sense:

(3) The hamiltonian of H_S is not tangent at any strata of Σ in a nbd of z_0 .

Next we shall give conditions expressing that a point is of (a), (b), (c) or (d) type. For this choose an open nbd V_0 of z_0 such that the covering $(\pi^{-1}(V_0), \pi, V_0)$ is trivial. Then $\pi^{-1}(V_0)$ is a union of disjoint open subsets W_1, \dots, W_m ($m = \text{card } \pi^{-1}(z_0)$), each of which is mapped diffeomorphically onto V_0 by π . For each j , $i \leq j \leq m$, let us denote by $H_j: V_0 \rightarrow \mathbf{R}$ the projection of $H_S|_{W_j}$ by π . Then we have:

(a') A simple point z_0 is regular if there is an integer k , $1 \leq k \leq m$, such that $H_k(z_0) > H_l(z_0)$ for all $l \neq k$.

(b') z_0 will be a normal switching point if there exist two integers j, k such that

(1) $1 \leq j, k \leq m$, and $H_j(z_0) = H_k(z_0) > H_l(z_0)$ for all $l \neq j, k$.

(2) $\{H_j, H_k\}(z_0) \neq 0$, where $\{ \}$ denotes the Poisson bracket on T^*M .

(c') z_0 will be a fold point if there are two integers j, k satisfying condition (1) of (b') and

(2) $\{H_j, H_k\}(z_0) = 0$.

(3) $\{H_j\{H_j, H_k\}\}(z_0) \neq 0$ and $\{H_k\{H_k, H_j\}\}(z_0) \neq 0$.

(d') z_0 will be a Maxwell point if there exist three integers i, j, k such that

(1) $1 \leq i, j, k \leq m$ and $H_i(z_0) = H_j(z_0) = H_k(z_0) > H_l(z_0)$ for all $l \neq i, j, k$.

(2) The differentials at z_0 , $dH_i(z_0)$, $dH_j(z_0)$, and $dH_k(z_0)$, are linearly independent.

(3) $\{H_i, H_j\}(z_0) \cdot \{H_j, H_k\}(z_0) \cdot \{H_k, H_i\}(z_0) \neq 0$.

Classification of fold and Maxwell points. Fold points: A fold point will be called

Hyperbolic: if $\{H_j\{H_k, H_j\}\}(z_0) > 0$ and $\{H_k\{H_j, H_k\}\}(z_0) > 0$,

elliptic: if the above inequality signs are reversed,

parabolic: if $\{H_j\{H_k, H_j\}\}(z_0) \cdot \{H_k\{H_j, H_k\}\}(z_0) < 0$.

¹For all definitions related to the symplectic structure of T^*M we follow [3].

Maxwell points: A Maxwell point will be called

elliptic: if $\{H_j, H_j\}(z_0)$, $\{H_j, H_k\}(z_0)$, and $\{H_k, H_i\}(z_0)$ have the same sign,

parabolic: if the signs are not the same.

2. Statement of the main results. In a sufficiently small nbd of a regular point, the extremals are the trajectories of a hamiltonian field (in the above notation of (a'), the hamiltonian of H_k).

In a sufficiently small nbd of a normal switching point, the extremals are the trajectories of a continuous flow defined by a hamiltonian having a tangential discontinuity along Σ : with the notation of (b') if \vec{H}_j, \vec{H}_k denote the hamiltonians of H_j, H_k respectively, then at any $z \in \Sigma$, $\vec{H}_j(z) - \vec{H}_k(z)$ is tangent to Σ and the flow is the flow of \vec{H}_j in the half space $\{(H_j - H_k) < 0\}$.

The study of extremals in the nbd of a fold or Maxwell point is more complicated.

Notation 3. Σ^1 will denote the subset, in Σ , of all nonnormal switching points, and Σ_0^1 the subset of all fold points.

The characterization (c') in §1 of the fold points shows that Σ_0^1 is a submanifold of codimension 2 in T^*M . Locally, using the notation of (c'), Σ_0^1 is defined by the equations $H_j - H_k = 0$ and $\{H_j, H_k\} = 0$.

Along Σ_0^1 , Σ is a submanifold of T^*M and Σ_0^1 is the locus of points of Σ where both hamiltonians \vec{H}_j, \vec{H}_k of H_j, H_k are tangent to Σ . Since Σ_0^1 has codimension 1, the plane generated by \vec{H}_j, \vec{H}_k at any point z_0 in Σ_0^1 will cut the tangent hyperplane to Σ_0^1 at z_0 in a one-dimensional space at least. In fact, the intersection is exactly one dimensional since otherwise $\vec{H}_j(z_0), \vec{H}_k(z_0)$ would both be tangent to Σ_0^1 . But this would violate condition (3) in (c'). It is easy to find the expression of a generator Z of the intersection:

$$Z = \{H_j\{H_k H_j\}\}\vec{H}_k + \{H_k\{H_j H_k\}\}\vec{H}_j.$$

DEFINITION 3. The vector field $Z = \{H_j\{H_k, H_j\}\}\vec{H}_k + \{H_k\{H_j, H_k\}\}\vec{H}_j$ on Σ_0^1 , is called the residual field on Σ_0^1 and its flow is called the residual flow.

REMARK 4. In [1] Arnold has introduced the field Z under the name "geodesic direction" (see [1, p. 237, definition and remark just above Proposition 2]).

We can now state the main result about the structure at fold points. We shall use the notations of §1, (c').

THEOREM 1. Let z_0 be a fold point. For any hypersurface D passing through z_0 and transversal to Σ and Σ^1 and the hamiltonians of H_j and H_k , there exists a nbd D_0 of z_0 in D and a number $a > 0$ such that through every $\zeta \in D_0$ passes a unique extremal $t \in [-a, a] \rightarrow z(t, \zeta)$, $z(0, \zeta) = \zeta$.

(1) If z_0 is parabolic, then this extremal has one switching point if it meets Σ^1 , two if it meets the half space $\{H_j > H_k\}$ (resp. $H_j < H_k$) when $\{H_j\{H_k, H_j\}\}(z_0) < 0$ (resp. $\{H_k\{H_j, H_k\}\}(z_0) < 0$), and zero switching points otherwise. The union of the extremals meeting Σ^1 forms a smooth hypersurface separating the extremals switching twice from those that do not. Finally the mapping $]-a, a[\times D_0 \rightarrow T^*M, (t, \zeta) \rightarrow z(t, \zeta)$ is a piecewise smooth homeomorphism.

(2) If z_0 is a hyperbolic point, there exists a nbd W of z_0 and two hypersurfaces Sep_+ and Sep_- passing through Σ^1 , tangent to Σ along Σ^1 to the first order, and having the properties: for a set $A \subset W$, \bar{A} denotes the closure in W .

(a) Sep_+ (resp. Sep_-) is contained in $W \cap \{H_k > H_j\}$ (resp. $W \cap \{H_j \geq H_k\}$). $W - (\text{Sep}_+ \cup \text{Sep}_-)$ has exactly four connected components: W_+ , W_- , W_r , W_l . W_+ (resp. W_-) is contained in $\{H_k > H_j\}$ (resp. $\{H_k < H_j\}$) and W_r (resp. W_l) is contained in $\{\{H_k, H_j\} > 0\}$ (resp. $\{\{H_k, H_j\} < 0\}$). We shall set $D_+ = D_0 \cap W_+$, $D_r = D_0 \cap W_r$, and so on.

(b) Through each point $\zeta \in \bar{D}_+$ (resp. \bar{D}_-) there passes at time $t = 0$ a unique extremal $t \in [-a, a] \rightarrow z_+(t, \zeta)$ (resp. $z_-(t, \zeta)$) contained in \bar{W}_+ (resp. \bar{W}_-). If $\zeta \in \text{Bdry } D_+$ (resp. $\text{Bdry } D_-$), z_+ (resp. z_-) is contained in Sep_+ (resp. Sep_-). These extremals are all smooth and the mappings $[-a, a] \times \bar{D}_+ \rightarrow \bar{W}_+$, $(t, \zeta) \rightarrow z_+(t, \zeta)$ and $[-a, a] \times \bar{D}_- \rightarrow \bar{W}_-$, $(t, \zeta) \rightarrow z_-(t, \zeta)$ are diffeomorphisms onto.

(c) Through each point $\zeta \in \bar{D}_r$ (resp. \bar{D}_l) there passes at time $t = 0$ a unique extremal $t \in [-a, a] \rightarrow z_r(t, \zeta)$ (resp. $z_l(t, \zeta)$) contained in \bar{W}_r (resp. \bar{W}_l). If $\zeta \in \text{Bdry } D_r$ (resp. $\text{Bdry } D_l$), z_r is contained in $(\text{Sep}_+ \cup \text{Sep}_-) \cap \{\{H_k, H_j\} \geq 0\}$ (resp. $(\text{Sep}_+ \cup \text{Sep}_-) \cap \{\{H_k, H_j\} \leq 0\}$). Each of these extremals switches exactly once and the mappings $[-a, a] \times \bar{D}_r \rightarrow \bar{W}_r$, $(t, \zeta) \rightarrow z_r(t, \zeta)$ and $[-a, a] \times \bar{D}_l \rightarrow \bar{W}_l$, $(t, \zeta) \rightarrow z_l(t, \zeta)$ are piecewise smooth homeomorphisms onto. Note that through each point in $\text{Sep}_+ \cup \text{Sep}_-$ there pass two extremals, one smooth and the other switching once, both contained in $\text{Sep}_+ \cup \text{Sep}_-$.

(3)(a) If z_0 is elliptic, then through each point $\zeta \in D_0 - \Sigma^1$ there passes at time $t = 0$ a unique extremal $t \in [-a, a] \rightarrow z(t, \zeta)$. For any subinterval $J \subset [-a, a]$, the nbd of switching points of this extremal contained in J , $v(\zeta, J)$, is finite. For any $\delta > 0$, if ζ tends to $D_0 \cap \Sigma^1$, $v(\zeta, J)$ tends to infinity, uniformly in J provided that length $J \geq \delta$.

(b) For any $\zeta \in D_0 \cap \Sigma^1$, the trajectory $\tilde{z}(t, \zeta)$ of the normalized residual field $(\{H_k\{H_j, H_k\}\} + \{H_j\{H_k, H_j\}\})^{-1}Z$, passing through ζ at time $t = 0$, is defined on $[-a, a]$ and contained in Σ^1 . If we define the mapping $\rho: [-a, a] \times D_0 \rightarrow T^*M$ as

$$\rho(t, \zeta) = \begin{cases} z(t, \zeta) & \text{if } \zeta \in D_0 - \Sigma^1, \\ \tilde{z}(t, \zeta) & \text{if } \zeta \in D_0 \cap \Sigma^1, \end{cases}$$

then ρ is a homeomorphism onto a nbd of z_0 in T^*M .

Let us pass now to the Maxwell points. As in the case of the folds, there is a residual field associated to the elliptic Maxwell points. Denote by $\Sigma_{3,e}$ the set of all elliptic Maxwell points. It is a submanifold of codimension 2 in T^*M , defined locally by the equations $H_i = H_j = H_k$ using the notations of §1, (d'). Since $\{H_i, H_j\} + \{H_j, H_k\} + \{H_k, H_i\} \neq 0$, the form induced on $\Sigma_{3,e}$ by the canonical symplectic form of T^*M is nondegenerate. Hence $\Sigma_{3,e}$ has an induced symplectic structure.

At any $z_0 \in \Sigma_{3,e}$ we can consider the restriction H_{ijk} of the function

$$\frac{\{H_i, H_j\}H_k + \{H_j, H_k\}H_i + \{H_k, H_i\}H_j}{\{H_i, H_j\} + \{H_j, H_k\} + \{H_k, H_i\}}$$

to $\Sigma_{3,e}$. The hamiltonians of these functions match together and define a field E on $\Sigma_{3,e}$.

DEFINITION 4. The symplectic vector field E on $\Sigma_{3,e}$ is called the residual field on $\Sigma_{3,e}$, defined locally by the hamiltonians of the H_{ijk} .

THEOREM 2. (1) Let z_0 be a parabolic Maxwell point. For any hypersurface D through z_0 which is transversal to Σ_3 and to the hamiltonians of H_i , H_j and H_k , there exists a nbd D_0 of z_0 in D and a number $a > 0$ such that through any $\zeta \in D_0$ there passes a unique extremal $t \in [-a, a] \rightarrow z(t, \zeta) \in T^*M$. Each of these extremals has at most two switching points and the mapping $]-a, a[\times D_0 \rightarrow T^*M, (t, \zeta) \rightarrow z(t, \zeta)$ is a piecewise smooth homeomorphism.

(2)(a) Let z_0 be an elliptic Maxwell point. For any hypersurface D through z_0 which is transversal to Σ_3 , to the hamiltonians of H_i , H_j , H_k , and to the residual field E , there exist a nbd D_0 of z_0 in D and a number $a > 0$ such that through any point $\zeta \in D_0 - \Sigma^1$ there passes a unique extremal $t \in [-a, a] \rightarrow z(t, \zeta) \in T^*M$.

If $v[\zeta; [\alpha, \beta]]$ denotes the number of switching points of $z(t, \zeta)$ in the interval $[\alpha, \beta]$ for any $\zeta \in D_0 - \Sigma^1$, then for any $\delta > 0$, $v[\zeta; [\alpha, \beta]]$ tends to infinity as ζ tends to Σ^1 uniformly in $[\alpha, \beta]$ provided that $\beta - \alpha \geq \delta$.

(b) If we define the mapping $\rho: [-a, a] \times D_0 \rightarrow T^*M$ as

$$\rho(t, \zeta) = \begin{cases} z(t, \zeta) & \text{if } \zeta \in D_0 - \Sigma_3, \\ e^{tE}(\zeta) & \text{if } \zeta \in D_0 \cap \Sigma_3, \end{cases}$$

then ρ is a homeomorphism onto a nbd of z_0 in T^*M .

REMARK. To better understand the statements of Theorems 1 and 2, one should study the diagrams at the end of this paper.

3. Basic concepts needed for the proof of Theorems 1 and 2. (A) The proof of Theorem 1 is based on the following lemma. We use the notations of §1, (c').

LEMMA 1. Given a fold point z_0 there exist a nbd W_0 of z_0 and two diffeomorphisms $\chi_j, \chi_k: (W_0, z_0) \rightarrow (\mathbf{R}^{2d}, 0)$ ($d = \dim M$) such that

$$(1) \quad \chi_j^*(x_0) = -\chi_k^*(x_0) = H_k - H_j,$$

$$(2) \quad \chi_{j*}(\vec{H}_j) = \omega_j \left[\frac{\partial}{\partial x_1} - x_1 \frac{\partial}{\partial x_0} \right], \quad \chi_{k*}(\vec{H}_k) = \omega_k \left[\frac{\partial}{\partial x_1} - x_1 \frac{\partial}{\partial x_0} \right],$$

where (x_0, \dots, x_{2d-1}) denotes the canonical system of coordinates in \mathbf{R}^{2d} , \vec{H}_j and \vec{H}_k are the hamiltonians of H_j and H_k respectively, and ω_j, ω_k are two smooth functions never vanishing in W_0 and such that

$$\omega_j(0) = \{H_j\{H_k, H_j\}\}(z_0), \quad \omega_k(0) = \{H_k\{H_j, H_k\}\}(z_0).$$

REMARK 5. Doing more work one can show that if we endow \mathbf{R}^{2d} with the canonical symplectic structure $\sum_{i=0}^{d-1} dx_{2i} \wedge dx_{2i+1}$, then one can choose χ_j and χ_k to be symplectic.

PROOF OF THE LEMMA. An old theorem of Whitney states that one can find a nbd W_0 of z_0 and two diffeomorphisms $\bar{\chi}_j, \bar{\chi}_k: (W_0, z_0) \rightarrow (\mathbf{R}^{2d}, 0)$ such that:

$$(1) \quad \bar{\chi}_j^* \left(\frac{x_1^2}{2} - x_0 \right) = -\bar{\chi}_k^* \left(\frac{x_1^2}{2} - x_0 \right) = H_j - H_k,$$

$$(2) \quad \bar{\chi}_{j*}(\bar{H}_j) = \bar{\omega}_j \frac{\partial}{\partial x_1}, \quad \bar{\chi}_{k*}(\bar{H}_k) = \bar{\omega}_k \frac{\partial}{\partial x_1},$$

where $\bar{\omega}_j, \bar{\omega}_k$ are two units. Moreover if $g = x_1^2/2 - x_0$, then

$$\begin{aligned} \bar{\omega}_j(0) &= \left(\theta \left(\bar{\omega}_j \frac{\partial}{\partial x_1} \right)^2 g \right) (0) = \chi_j^* \left[\theta \left(\chi_{j*}(\bar{H}_j) \right)^2 g \right] (z_0) \\ &= \left(\theta(\bar{H}_j)^2 \chi_j^*(g) \right) (z_0) = \{ H_j \{ H_k, H_j \} \} (z_0).^2 \end{aligned}$$

One proves in a similar fashion that $\bar{\omega}_k(0) = \{ H_k \{ H_j, H_k \} \} (z_0)$. Now if we make the change of coordinates $\lambda: \mathbf{R}^{2d} \rightarrow \mathbf{R}^{2d}$, $\lambda^*(x_j) = x_j$, $j \geq 1$, $\lambda^*(x_0) = x_0 - x_1^2/2$, then the compositions $\chi_j = \lambda \circ \bar{\chi}_j$ and $\chi_k = \lambda \circ \bar{\chi}_k$ satisfy all the conditions of the lemma.

From this lemma it is easy to prove all the facts stated in Theorem 1 about parabolic and hyperbolic fold points. The proof of the statements about the elliptic points is more involved.

(B) *The Poincaré map.* Let z_0 be an elliptic fold point. From Lemma 1, it is clear that any extremal $\psi: [\alpha, \beta] \rightarrow T^*M$ contained in a sufficiently small neighborhood of z_0 has the following structure: there exists a sequence $t_0 = \alpha \leq t_1 < t_2 < t_3 < \dots < t_N \leq \beta = t_{N+1}$ such that (a) $\psi(t) \in \Sigma$ if and only if t is in the sequence; (b) on each interval $[t_i, t_{i+1}]$, ψ is the trajectory of \bar{H}_j (resp. \bar{H}_k) if i is even and of \bar{H}_k (resp. \bar{H}_j) if i is odd, the alternative depending on whether $H_j - H_k$ is positive (resp. negative) on $]\alpha_1, t_1[$.

This remark shows that the extremals in a nbd of z_0 are determined by the sequence of points of Σ , $\{\psi(t_1), \psi(t_2), \dots, \psi(t_N)\}$. Let us make these remarks more precise. In order to make computations, we have to fix the definition of the Poisson brackets since two choices, differing in sign, are possible.

Poisson bracket. If F, G are two C^∞ functions on an open subset of T^*M , then their Poisson bracket $\{F, G\} = \theta(\vec{G})F$ where \vec{G} is the hamiltonian of G and θ denotes the Lie derivative.

Distinguished nbds. Lemma 1 shows that z_0 has a basis of nbds W , called distinguished, as follows: for any $z \in W \cap \{H_j - H_k \geq 0\}$ (resp. $W \cap \{H_j - H_k \leq 0\}$), there exist two nbs $\tau_-(z)$ and $\tau_+(z)$, $\tau_-(z) \leq 0 \leq \tau_+(z)$, such that:

(1) the trajectory $h_j(t, z)$ (resp. $h_k(t, z)$) of \bar{H}_j (resp. \bar{H}_k) passing through z at time $t = 0$ is defined on $[\tau_-(z), \tau_+(z)]$ contained in W and:

$$(H_j - H_k)(h_j(t, z)) > 0 \quad \text{if } \tau_-(z) < t < \tau_+(z),$$

$$(H_j - H_k)(h_j(t, z)) = 0 \quad \text{if } t = \tau_+(z) \text{ or } \tau_-(z).$$

$$\{H_k - H_j\}(h_j(t, z)) < 0 \quad \text{if } t = \tau_+(z), \quad < 0 \quad \text{if } t = \tau_-(z).$$

Similar relations obtained by exchanging j and k are also valid.

² θ denotes the Lie derivative.

(2) If $z \in W \cap \Sigma$, then both h_j and h_k are defined for z . If $\{H_j, H_k\}(z) > 0$, then $\tau_+(z) = 0$ for h_j and $\tau_-(z) = 0$ for h_k , and if $\{H_j, H_k\}(z) < 0$, then it is the opposite. Finally if $z \in \Sigma^1$ (i.e. $\{H_j, H_k\}(z) = 0$), then $\tau_-(z) = \tau_+(z) = 0$ for both trajectories.

For any distinguished nbd W , let

$$\Sigma_+ = \{z \in W \cap \Sigma \mid \{H_k, H_j\}(z) > 0\}$$

and

$$\Sigma_- = \{z \mid z \in W \cap \Sigma, \{H_k, H_j\}(z) < 0\}.$$

DEFINITION 5 (POINCARÉ MAPPING). For any distinguished nbd W we define two diffeomorphisms

$$\Pi_+ : \Sigma_+ \rightarrow \Sigma_- \quad (\text{resp. } \Pi_- : \Sigma_- \rightarrow \Sigma_+),$$

$$\Pi_+(z) = h_k(\tau_+(z), z) \quad (\text{resp. } \Pi_-(z) = h_j(\tau_-(z), z)).$$

We set $\Pi = \Pi_- \circ \Pi_+ : \Sigma_+ \rightarrow \Sigma_+$.

Coming back to our considerations at the beginning of (B) in this section, it is clear that in the sequence $\{\psi(t_1), \dots, \psi(t_N)\}$, we have

$$\psi(t_{2n+1}) = \Pi^n(\psi(t_1)), \quad \psi(t_{2n+2}) = \Pi_+ \Pi^n(\psi(t_1))$$

when $H_j - H_k$ is positive on $\psi([\alpha, t_1])$ and

$$\psi(t_{2n}) = \Pi^{n-1} \circ \Pi_-(\psi(t_1)), \quad \psi(t_{2n+1}) = \Pi_+(\psi(t_{2n}))$$

when $H_j - H_k$ is negative on $\psi([\alpha, t_1])$. Hence to determine the structure of extremals in the elliptic case, it is convenient to study the Poincaré mapping Π .

LEMMA 2. (a) Π can be extended as a smooth diffeomorphism $\tilde{\Pi}$ to a nbd Σ' of z_0 in Σ .

(b) For any $z \in \Sigma'$,

$$\begin{aligned} \tilde{\Pi}(z) = z + 2\{H_j, H_k\}(z) & \left[\frac{\vec{H}_j(z)}{\{H_j\{H_k, H_j\}\}(z)} + \frac{\vec{H}_k(z)}{\{H_k\{H_j, H_k\}\}(z)} \right] \\ & + \{H_j, H_k\}^2(z) S(z), \end{aligned}$$

where S is a smooth mapping.

PROOF. $\Pi = \Pi_- \circ \Pi_+$. $\Pi_+(z) = h_k(\tau_1(z), z)$ (resp. $\Pi_-(z) = h_j(\tau_2(z), z)$), where τ_1 (resp. τ_2) is the smallest positive solution of $(H_j - H_k)[h_k(t, z)] = 0$ in Σ_+ (resp. $(H_j - H_k)[h_j(t, z)] = 0$ in Σ_-). It follows from Lemma 1 that these solutions exist, are smooth functions of z , and extend smoothly to a nbd Σ' of z_0 . (The extensions are not everywhere positive anymore!) This proves (a). In particular Π_+ and Π_- have smooth extensions $\tilde{\Pi}_+$, $\tilde{\Pi}_-$ to Σ' .

In the following computations $f = g + O(h^n)$ means that there exists a smooth function or mapping k such that $f = g + h^n k$. We have

$$(H_j - H_k)[h_k(t, z)] = t \left[f(z) - \frac{t}{2} g_k(z) \right] + O(t^3),$$

where $f = \{H_j, H_k\}$ and $g_k = \{H_k\{H_j, H_k\}\}$. Hence $\tau_1 = 2f/g_k + O(f^2)$.

In the same way, $\tau_2 = -2f/g_j + O(f^2)$, where $g_j = \{H_j\{H_k, H_j\}\}$.

$$\begin{aligned}\tilde{\Pi}_+ &= \text{Id} + \frac{2f}{g_k} \vec{H}_k + O(f^2), \\ \tilde{\Pi}_- &= \text{Id} - \frac{2f}{g_j} \vec{H}_j + O(f^2), \\ \tilde{\Pi} &= \tilde{\Pi}_+ - \left(\frac{2f}{g_j} \vec{H}_j \right) \circ \tilde{\Pi}_+ + O((f \circ \Pi_+)^2).\end{aligned}$$

Id = identity mapping,

Now if φ is any smooth function or mapping we have $\varphi \circ \Pi_+ = \varphi + O(f)$. Also

$$f \circ \Pi_+ = f + \frac{2f}{g_k} (\theta(\vec{H}_k)f) + O(f^2).$$

Since $\theta(\vec{H}_k)f = \{\{H_j, H_k\}H_k\} = -g_k$, $f \circ \Pi_+ = -f + O(f^2)$. Hence

$$\tilde{\Pi} = \text{Id} + 2f \left[\frac{\vec{H}_k}{g_k} + \frac{\vec{H}_j}{g_j} \right] + O(f^2).$$

LEMMA 3. For any $z \in \Sigma_+$, the time τ spent by the extremal between z and $\Pi(z)$, as a function of z , is roughly

$$\tau = 2f \left[\frac{1}{g_k} + \frac{1}{g_j} \right] + O(f^2).$$

PROOF. Using the notation and the facts of Lemma 2,

$$\tau = \tau_1 + \tau_2 \circ \Pi_+ = 2f \left[\frac{1}{g_k} + \frac{1}{g_j} \right] + O(f^2).$$

The next lemma gives us the normal form of Π .

BASIC LEMMA 4. Let $\Phi: (\mathbf{R}^n, 0) \hookrightarrow$ be a germ of a diffeomorphism at 0 having the form:

$$\Phi(z) = z + \alpha(z)A(z),$$

where α and A are respectively the germ of a smooth function and a smooth mapping satisfying the conditions:

(a) $\alpha(0) \neq 0, A(0) \neq 0$;

(b) $d\alpha(z)A(z) = 0$

for all z such that $\alpha(z) = 0$.

Then Φ is C^∞ -conjugate to its linear part Φ_0 :

$$\Phi_0(z) = z + (d\alpha(0) \cdot z)A(0).$$

IMPORTANT REMARK. In the case where Φ is analytic, it is not true, in general, that it is analytically conjugate to Φ_0 .

The basic lemma is probably well known but the author could not find any reference for it. In the appendix a proof of the equivalent "coordinate version" of the lemma is supplied.

BASIC LEMMA 4—COORDINATE VERSION. Let $\Phi: (\mathbf{R}^n, 0) \hookrightarrow$ be a germ of a diffeomorphism at 0 such that in the canonical coordinates of \mathbf{R}^n , Φ has the form

$$\Phi_1(x) = x_1 + x_1^2 A_1(x), \quad \Phi_j(x) = x_j + x_1 A_j(x), \quad j \geq 2,$$

where the A_j are smooth and the $A_j(0)$ with $j \geq 2$ are not all 0.

Then Φ is conjugate to its linear part $\Phi_0: x \rightarrow x + x_1 v$, where $v = (0, A_2(0), \dots, A_n(0))$. Let A denote the field $(0, A_2(x), \dots, A_n(x))$.

A restricted to the hypersurface $F = \{\alpha = 0\}$ defines a vector field tangent to F , which we denote by Z .

COROLLARY. There exist a nbd Q of 0 and a constant $C > 0$ such that for all $z \in Q$, all $z_\infty \in Q \cap F$, and for any $\Phi^n(z)$ belonging to the orbit of z in Q ,

$$\|\Phi^n(z) - \exp[S_n(z)Z](z_\infty)\| \leq e^{C\|S_n(z)\|} \|z - z_\infty\|,$$

where $S_n(z) = \sum_{k=0}^{n-1} \alpha(\Phi^k(z))$.

PROOF. Lemma 4 implies that there exists a C^∞ -diffeomorphism $\psi: (Q, 0) \rightarrow (N, 0)$, Q and N open nbds of 0, $\psi(Q) = N$, N convex, such that $\psi \circ \Phi \circ \psi^{-1} = \Phi_0$, where (x_1, \dots, x_n) is the canonical system of coordinates on \mathbf{R}^n , and $v = (0, v_2, \dots, v_n)$ is a constant vector field. Now $\psi \circ \Phi \circ \psi^{-1} = x + \alpha \circ \psi^{-1}(x)V(x)$, where V is a field on N :

$$V(x) = \int_0^1 d\psi[(1-s)\psi^{-1}(x) + s\Phi(\psi^{-1}(x))] ds A(\psi^{-1}(x)).$$

It follows that there exists a unit ω on N such that $\alpha \circ \psi^{-1} = \omega x_1$, $v = \omega V$. Also: (1) ψ_1 is constant on the orbits of Φ in Q , (2) $e^{tv}(x) = \exp[\sigma(t, x)V](x)$ for any $x \in N$ and any $t \in \mathbf{R}$ such that $e^{tv}(x) \in N$, where $\sigma(t, x) = \int_0^t \omega(x + sv) ds$, (3) V restricted to F is $\psi_*(Z)$.

Let $z \in Q$ and let $\Phi^n(z)$ be a point of the orbit of z in Q . Then $\Phi^n(z) = \psi^{-1}[e^{(nx_1)v}(x)]$, where $x = \psi(z)$. By property (2) above,

$$\Phi^n(z) = \exp[\sigma_n(z)\psi_*^{-1}(V)](z),$$

where $\sigma_n(z) = \sigma(n\psi_1(z), \psi(z))$.

By property (1) above:

$$S_n(z) = \sum_{k=0}^{n-1} \psi_1(z) \omega(\psi \Phi^k(z)),$$

$$S_n(z) = \sum_{k=1}^{n-1} x_1 \omega(x + kx_1 v).$$

Hence

$$S_n(z) - \sigma_n(z) = \sum_{k=0}^{n-1} \left[x_1 \omega(x + kx_1 v) - \int_{kx_1}^{(k+1)x_1} \omega(x + sv) ds \right].$$

It follows that

$$|S_n - \sigma_n| \leq \frac{n\psi_1^2}{2} \sup_N |\theta(v)\omega|.$$

Since $n|\psi_1| \leq S_n/\inf|\omega|$, we have $|S_n - \sigma_n| \leq C_0|S_n||\psi_1|$, where C_0 is some constant:

$$\Delta = \|\Phi''(z) - \exp[S_n(z)\psi_*^{-1}(V)](z)\| \leq C_0|S_n(z)\psi_1(z)| \sup_Q \|\psi_*^{-1}(V)\|.$$

Hence with some constant C_1 :

$$\Delta \leq C_1|S_n(z)| \|z - z_\infty\|.$$

With this estimate and property (3) above, it is easy to finish the proof.

Let us go back to the proof of Theorem 1(3). We can apply Lemma 4 and its corollary to $\Phi = \tilde{\Pi}$, taking $\alpha = \tau$ (see Lemma 2) and $A = (\tilde{\Pi} - \text{Id})/\alpha$. Choose a nbd Q' of z_0 such that $z(t, \xi) \in Q$ for all $\xi \in Q'$ and all $t \in [-a, a]$. For any $\xi \in Q' \cap \Sigma$, let $-a \leq T_{n_-(\xi)}(\xi) < \dots < T_0(\xi) = 0 < T_1(\xi) < \dots < T_{n_+(\xi)}(\xi) < a$ be the switching times of $z(\cdot, \xi)$ in $] -a, a[$. It will be convenient to set $a = T_{n_+(\xi)+1}(\xi)$, $-a = T_{n_-(\xi)-1}(\xi)$.

For any $t \in [-a, a[$ and any $\xi \in Q' \cap \Sigma$, there is a unique n such that $T_n(\xi) \leq t < T_{n+1}(\xi)$. If $t = a$, set $n = n_+(\xi) + 1$. Then setting $Z = Z_{\text{norm}}$

$$\|z(t, \xi) - \exp(tZ)[\xi_\infty]\| \leq G_1 + G_2 + G_3,$$

where

$$G_1 = \|z(t, \xi) - z(T_n(\xi), \xi)\|, \quad G_2 = \|z(T_n(\xi), \xi) - \exp|T_n(\xi)Z|(\xi_\infty)\|,$$

$$G_3 = \|\exp[T_n(\xi)Z](\xi_\infty) - \exp[tZ](\xi_\infty)\|.$$

By the corollary, since $S_n = T_n$ and $z(T_n(\xi), \xi) = \Pi^n(\xi)$, we have

$$G_2 \leq e^{C|T_n(\xi)|} \|\xi - \xi_\infty\|, \quad G_1 + G_3 \leq C|t - T_n(\xi)| \leq C|T_{n+1}(\xi) - T_n(\xi)|,$$

where $C = \sup_Q \|Z\| + \sup_Q \|\vec{H}_j\| + \sup_Q \|\vec{H}_k\|$. Also

$$|T_{n+1}(\xi) - T_n(\xi)| = |\alpha[\Pi^n(\xi)]| \leq \left(\sup_N |\omega| \right) \psi_1(\xi),$$

$|T_{n+1}(\xi) - T_n(\xi)| \leq C_4 \|\xi - \xi_\infty\|$ with some constant C_4 .

It follows from the preceding inequalities that, with some constant C_5 , for all $\xi \in Q' \cap \Sigma$ and all $t \in [-a, a]$,

$$\|z(t, \xi) - \exp(tZ)[\xi_\infty]\| \leq e^{C_5|t|} \|\xi - \xi_\infty\|.$$

This ends the proof of Theorem 1.

4. Proof of Theorem 2. (A) We shall use the notations of §1, (d'). Let z_0 be a Maxwell point. In a sufficiently small neighborhood W , Σ will be a Y-shaped stratified set, the union of the three manifolds $\Sigma(ij)$, $\Sigma(jk)$, $\Sigma(ki)$, having $\Sigma_3 \cap W$ as a common boundary. $\Sigma(ij)$ (resp. $\Sigma(jk)$, $\Sigma(ki)$) is the set of all $z \in W$ such that $H_i(z) = H_j(z) \geq H_k(z)$ (resp. $H_j(z) = H_k(z) \geq H_i(z)$, $H_k(z) = H_i(z) \geq H_j(z)$).

If W is sufficiently small, then $W - \Sigma$ will have exactly three connected components W_i , W_j , W_k , where W_i (resp. W_j , W_k) is the set of all z where $H_i(z) > H_j(z)$, $H_k(z)$ (resp. $H_j(z) > H_k(z)$, $H_i(z)$, $H_k(z) > H_i(z)$, $H_j(z)$). W_i (resp. W_j , W_k) has the shape of a wedge with $\Sigma_3 \cap W$ as edge and $(\Sigma(ij), \Sigma(ik))$ (resp. $(\Sigma(jk), \Sigma(ji))$, $(\Sigma(ki), \Sigma(kj))$) as sides.

In what follows it will be convenient to set $\Sigma(ji) = \Sigma(ij)$ and so on...

LEMMA 5. (a) *If $\{H_k, H_i\}(z_0)$ and $\{H_j, H_i\}(z_0)$ are both of the same sign, then, for a sufficiently small W , the positive (resp. negative) flow of \vec{H}_i leaves W_i invariant if $\{H_k, H_i\}(z_0) < 0$ (resp. > 0).*

(b) *If the above quantities are of opposite signs, then for a sufficiently small W the flow of \vec{H}_i crosses W_i from $\Sigma(ij)$ to $\Sigma(ik)$ if $\{H_j, H_i\}(z_0) < 0$, and from $\Sigma(ik)$ to $\Sigma(ij)$ in the opposite case.*

This lemma is trivial. It enables us to settle the parabolic case immediately: without restricting the generality we can always assume that $\{H_i, H_j\}(z_0)$ and $\{H_k, H_i\}(z_0)$ are of the same sign ε and $\{H_j, H_k\}(z_0)$ of the opposite sign $-\varepsilon$. From case (b) of Lemma 5, it follows that for a sufficiently small W , the flow of H_i crosses W_i from $\Sigma(ij)$ to $\Sigma(ik)$ if $\varepsilon > 0$, and from $\Sigma(ik)$ to $\Sigma(ij)$ if $\varepsilon < 0$. By case (a) the flow of \vec{H}_j (resp. \vec{H}_k) flows out of W_j (resp. W_k) if $\varepsilon > 0$ (resp. $\varepsilon < 0$) and into W_j (resp. W_k) in the contrary case.

We are left with the elliptic case which is more involved.

(B) *The elliptic case and the Poincaré mapping.* Case (b) of Lemma 5 allows us to state the following definition. For simplicity let us denote by Σ_{ij} the hypersurface $H_i - H_j = 0$, and similarly for Σ_{ik} , Σ_{kj} .

DEFINITION 6. Assume that $\{H_j, H_i\}(z_0) < 0$ and $\{H_k, H_i\}(z_0) > 0$. Then there exists a sufficiently small nbd $W_1 \subset W$ of z_0 , such that the positive trajectory of \vec{H}_i , in W_1 , starting at a point $z \in \Sigma_{ij} \cap W_1$, cuts Σ_{ik} for the first time at $\tilde{\phi}_i(z)$ and $\tilde{\phi}_i(z)$ is contained in $\Sigma_{ik} \cap W_1$. We shall denote by ϕ_i the restriction of $\tilde{\phi}_i$ to $\Sigma(ij) \cap W_1$. Then $\tilde{\phi}_i$ and ϕ_i are diffeomorphisms from $\Sigma_{ij} \cap W_1$ onto $\Sigma_{ik} \cap W_1$ and $\Sigma(ij) \cap W_1$ onto $\Sigma(ik) \cap W_1$ respectively.

In case $\{H_j, H_i\}(z_0) > 0$ and $\{H_k, H_i\}(z_0) < 0$, we have the same definitions with the roles of j and k interchanged.

Assume now that z_0 is an elliptic point. Without restricting the generality, we can assume that $\{H_i, H_j\}(z_0) > 0$. Then for an appropriately chosen W_1 , all six diffeomorphisms

$$\begin{aligned} \tilde{\phi}_i: \Sigma_{ij} \cap W_1 &\rightarrow \Sigma_{ik} \cap W_1, & \phi_i: \Sigma(ij) \cap W_1 &\rightarrow \Sigma(ik) \cap W_1, \\ \tilde{\phi}_k: \Sigma_{ki} \cap W_1 &\rightarrow \Sigma_{kj} \cap W_1, & \phi_k: \Sigma(ki) \cap W_1 &\rightarrow \Sigma(kj) \cap W_1, \\ \tilde{\phi}_j: \Sigma_{jk} \cap W_1 &\rightarrow \Sigma_{ji} \cap W_1 & \phi_j: \Sigma(jk) \cap W_1 &\rightarrow \Sigma(ji) \cap W_1 \end{aligned}$$

are defined. (Remember that $\Sigma_{ij} = \Sigma_{ji}$, $\Sigma(ij) = \Sigma(ji)$, and so on.)

Now if $\psi: [\alpha, \beta] \rightarrow W_1$ is an extremal and $\alpha \leq t_0 < t_2 < \dots < t_N \leq \beta$ are its switching times, assuming for simplicity that $\psi(t_0) \in \Sigma(ij)$, it is clear that $\psi(t_{3n+1}) = \phi_i(\psi(t_{3n}))$, $\psi(t_{3n+2}) = \phi_k(\psi(t_{3n+1}))$, $\psi(t_{3n+3}) = \phi_j(\psi(t_{3n+2}))$ for all n , $0 \leq n < N/3$, n an integer.

In particular $\psi(t_{3n+3}) = \phi_j \circ \phi_k \circ \phi_i(\psi(t_{3n}))$. This justifies the following

DEFINITION 7. We denote by Π the composition

$$\phi_j \circ \phi_k \circ \phi_i: \Sigma(ij) \cap W_1 \rightarrow \Sigma(ij) \cap W_1.$$

Π has a smooth extension

$$\tilde{\Pi} = \tilde{\phi}_j \circ \tilde{\phi}_k \circ \tilde{\phi}_i: \Sigma_{ij} \cap W_1 \rightarrow \Sigma_{ij} \cap W_1.$$

Both Π and $\tilde{\Pi}$ are diffeomorphisms.

(C) *The structure of Π and the end of the proof.* We shall consider a linear model of our situation: in \mathbf{R}^{2d} endowed with the canonical symplectic structure $\sum_{n=0}^{d-1} dx_{2n} \wedge dx_{2n+1}$, let L_i, L_j, L_k be three linearly independent linear forms, such that $L_j - L_k = x_1$ and $L_i - L_k = \sigma x_0$, where σ is a constant, $\sigma = \{L_i, L_j\} + \{L_j, L_k\} + \{L_k, L_i\}$.

If we set $H_i = L_i, H_j = L_j$, and $H_k = L_k$, then the corresponding Π, Π_L , will be the linear mapping

$$\Pi_L(x) = x + \frac{\sigma(L_i(x) - L_k(x))}{\{L_j, L_k\}\{L_k, L_i\}} Z_{\text{norm}},^3$$

where $Z_{\text{norm}} = \sigma^{-1}(\{L_j, L_k\}\tilde{L}_i + \{L_k, L_i\}\tilde{L}_j + \{L_i, L_j\}\tilde{L}_k)$.

Given any elliptic point z_0 it is clear that one can find a symplectic diffeomorphism χ of a nbd N_1 of 0 in \mathbf{R}^{2d} onto the nbd W_1 of z_0 (see Definition 6) such that $\chi(0) = z_0$ and L_i, L_j, L_k are the differentials at 0 of $\chi^*(H_i), \chi^*(H_j), \chi^*(H_k)$ respectively. It is then easy to see that $\tilde{\Pi}_L$ is the differential at 0 of the mapping $\chi^{-1} \circ \Pi \circ \chi$. Applying the Basic Lemma 4 ends the proof of Theorem 2.

5. Proof of the Basic Lemma 4. (A) The proof will be in two steps: First we reduce the diffeomorphism Φ to Φ_0 formally in x along the hypersurface $\mathbf{R}^{n-1}: x_1 = 0$.

Let $P: \mathbf{R}^n \rightarrow \mathbf{R}^n$ be the projection having \mathbf{R}^{n-1} as its image and the x_1 -axis as its kernel. Call $\hat{\mathbf{R}}^n$ the space \mathbf{R}^n completed formally along \mathbf{R}^{n-1} . A germ $\hat{\Phi}: (\hat{\mathbf{R}}^n, 0) \rightarrow (\hat{\mathbf{R}}^n, 0)$ of a formal diffeomorphism at 0 is a formal series $\sum_{i=0}^{\infty} x_1^i \alpha_i(Px)$ where (a) the α_i are smooth mappings from a nbd N of 0 in \mathbf{R}^{n-1} (depending on $\hat{\Phi}$) into \mathbf{R}^n , (b) the linear part of $\hat{\Phi}$ at 0 is invertible. These formal diffeomorphisms form a group $\hat{\mathcal{G}}$.

We shall denote by $J(\hat{\Phi})$ the mapping $x \in N \rightarrow (\alpha_{0,1}(Px) + x_1 \alpha_{1,1}(Px), P\alpha_0(Px) + x_1 P\alpha_1(0))$, where $\alpha_0 = (\alpha_{0,1}, P\alpha_1), \alpha_1 = (\alpha_{1,1}, P\alpha_1)$, and so on.

Let $\text{Tr}(1)$ be the subgroup of $\text{GL}(n)$ of all linear endomorphisms of \mathbf{R}^n whose restriction to \mathbf{R}^{n-1} is the identity and who induce the identity on the quotient $\mathbf{R}^n/\mathbf{R}^{n-1}$. $\hat{\mathcal{G}}_m$ for $m \geq 1$ will denote the subgroup of $\hat{\mathcal{G}}$ of all $\hat{\Phi}$ such that: (1) $J(\hat{\Phi}) \in \text{Tr}(1)$, (2) $\hat{\Phi}_1 - (J(\hat{\Phi}))_1 \equiv 0 \pmod{x_1^{m+1}}, P\hat{\Phi} - PJ(\hat{\Phi}) \equiv 0 \pmod{x_1^m}$.

The following lemma settles the formal part. We convene that $\hat{\mathcal{G}}_0 = \hat{\mathcal{G}}$.

LEMMA 6. *Let $\hat{\Phi} \in \hat{\mathcal{G}}_m$ ($m \geq 1$) be such that $J(\hat{\Phi})$ is not the identity. Then there exists a $\hat{\psi} \in \hat{\mathcal{G}}_{m-1}$ such that $\hat{\psi} \circ \hat{\Phi} \circ \hat{\psi}^{-1}$ belongs to $\hat{\mathcal{G}}_{m+1}$.*

If $m \geq 2$, $J(\hat{\psi})$ is the identity and if $m = 1$, $\hat{\psi}(x) = [x_1(1 + \lambda(Px)), Px + \mu(Px)]$.

PROOF OF THE LEMMA. We have to find a $\hat{\psi} \in \hat{\mathcal{G}}_{m-1}$ satisfying the following two conditions:

$$(A) (\hat{\psi} \circ \hat{\Phi})_1 \equiv \hat{\psi} \pmod{x_1^{m+2}}.$$

$$(B) P \circ \hat{\psi} \circ \hat{\Phi} \equiv P\hat{\psi} + \hat{\psi}_1 v \pmod{x_1^{m+1}}, v = P\alpha_1(0).$$

³ The coefficient of Z_{norm} is the time needed by the extremal to go from x to $\Pi_L(x)$.

Now

$$\begin{aligned}\hat{\Phi}_1 &\equiv x_1 + x_1^{m+1}\alpha_{m+1,1} \pmod{x_1^{m+2}}, \\ P\hat{\Phi} &\equiv Px + x_1v + x_1^m P\alpha_m \pmod{x_1^{m+1}}, \quad \alpha_{m+1} = (\alpha_{m+1,1}, P\alpha_{m+1}).\end{aligned}$$

We look for a $\hat{\psi}$ in the form: $\hat{\psi}_1 = x_1 + x_1^m \lambda(Px)$, $P\hat{\psi} = Px + x_1^{m-1} \mu(P(x))$, where λ, μ are germs at 0 of mappings from \mathbf{R}^{n-1} to \mathbf{R} and \mathbf{R}^{n-1} respectively.

It is easy to see that conditions (A) and (B) are equivalent to the following (A'), (B'):

- (A') $\alpha_{m+1,1} + d\lambda(v) = 0$ if $m \geq 2$,
- (B') $P\alpha_m - \lambda v + d\mu(v) = 0$ if $m \geq 2$,
- (A') $(1 + \lambda)\alpha_{2,1} + d\lambda(P\alpha_1) = 0$ if $m = 1$,
- (B') $P\alpha_1 - (1 + \lambda)v + d\mu(P\alpha_1) = 0$ if $m = 1$.

These equations can be solved as follows: Let $S \subset \mathbf{R}^{n-1}$ be a hypersurface in \mathbf{R}^{n-1} through 0, transversal to v if $m \geq 2$ and to $P\alpha_1$ if $m = 1$. Then given any sufficiently small nbd N of 0 in \mathbf{R}^{n-1} , there exists a unique solution (λ, μ) of equations (A')–(B') such that the restrictions of λ and μ to $N \cap S$ are given.

Going back to the proof of the basic lemma, one can associate to Φ its formal counterpart $\hat{\Phi}$, $\hat{\Phi} = \sum_{k=0}^{\infty} x_1^k \alpha_k$, where

$$\alpha_k = \frac{1}{k!} \left. \frac{\partial^k \Phi}{\partial x_1^k} \right|_{x_1=0}.$$

Then $\hat{\Phi} \in \hat{\mathcal{D}}_1$. Applying Lemma 6 inductively one can construct a sequence $\hat{\psi}_1, \hat{\psi}_2, \dots, \hat{\psi}_m, \dots$ such that (a) $\hat{\psi}_m \in \hat{\mathcal{D}}_{m-1}$, (b) $J^1 \hat{\psi} = \text{Id}$ if $m \geq 2$, (c) for any integer $m \geq 1$:

$$\hat{\psi}_m \circ \dots \circ \hat{\psi}_1 \circ \hat{\Phi} \circ \hat{\psi}_1^{-1} \circ \hat{\psi}_1^{-1} \circ \dots \circ \hat{\psi}_m^{-1} \in \hat{\mathcal{D}}_{m+1}.$$

Now the compositions $\{\hat{\psi}_m \circ \dots \circ \hat{\psi}_1 \mid m \geq 1\}$ converge in the x_1 -adic topology to an element $\hat{\psi} \in \hat{\mathcal{D}}$, which has the property that $\hat{\psi} \circ \hat{\Phi} \circ \hat{\psi}^{-1} = \hat{\Phi}_0$.

Using Borel's theorem, we can find a germ of a C^∞ -diffeomorphism ψ at 0 such that its x_1 -adic expansion is the $\hat{\psi}$ constructed above. If we set $\Phi^1 = \psi \circ \Phi \circ \psi^{-1}$, then $\Phi^1 \equiv \Phi_0 \pmod{x_1^\infty}$ which means that for any integer $N > 0$ the mapping $(\Phi^1 - \Phi_0)/x_1^N$, which is defined on $\mathbf{R}^n - \mathbf{R}^{n-1}$ in a nbd of 0, has a smooth extension to a nbd of 0.

To end the proof of the basic lemma it will be sufficient to prove the following proposition:

PROPOSITION 1. *Let Φ be a germ of a C^∞ -diffeomorphism $(\mathbf{R}^n, 0)$ such that $\Phi \equiv \Phi_0 \pmod{x_1^\infty}$. Then there exists a germ of a C^∞ -diffeomorphism $\phi: (\mathbf{R}^n, 0) \rightrightarrows$ such that $\phi \circ \Phi \circ \phi^{-1} = \Phi_0$ and $\phi \equiv \text{identity} \pmod{x_1^\infty}$.*

PROOF OF THE PROPOSITION. Let V be a compact nbd of 0 in \mathbf{R}^n such that Φ has a representant defined on V . The latter will also be denoted by Φ . Let $l: \mathbf{R}^n \rightarrow \mathbf{R}$ be a linear form such that $l(v) = 1$.

It is easy to find a C^∞ function $a: \mathbf{R}^n \rightarrow [0, 1]$ having the following properties: (1) a has its support in V , (2) a is equal to 1 in a nbd of 0, (3) the mapping $x \in \mathbf{R}^n$,

$$x \in \mathbf{R}^n \rightarrow \Phi'(x) = \begin{cases} \Phi_0(x) & \text{if } x \notin V, \\ \Phi_0(x) + a(x)(\Phi(x) - \Phi_0(x)), & \text{if } x \in V, \end{cases}$$

is a C^∞ -diffeomorphism of \mathbf{R}^n onto \mathbf{R}^n , (4) for all $x \in \mathbf{R}^n$,

$$l(\Phi'(x)) > l(x) + \frac{1}{2}x_1, \quad l(\Phi'^{-1}(x)) < l(x) - \frac{1}{2}x_1.$$

To prove Proposition 1 it is sufficient to construct a C^∞ -diffeomorphism ϕ of \mathbf{R}^n onto itself such that $\Phi' \circ \phi = \phi \circ \Phi_0$ and $\phi \equiv \text{identity} \bmod x_1^\infty$. It is quite easy to define such a ϕ but it is technically harder to prove it has the right properties. For simplicity, from now on we denote Φ' simply by Φ . Hence Φ is a diffeomorphism of \mathbf{R}^n onto \mathbf{R}^n such that: (1) $\Phi - \Phi_0$ has a compact support, (2) $\Phi \equiv \Phi_0 \bmod x_1^\infty$, (3) for all $x \in \mathbf{R}^n$, $l(\Phi(x)) > l(x) + \frac{1}{2}x_1$, $l(\Phi^{-1}(x)) < l(x) - \frac{1}{2}x_1$.

(B) *Construction of ϕ .* Let us define the two functions $\nu, \nu_1: \mathbf{R}^n - \mathbf{R}^{n-1} \rightarrow \mathbf{N}$ ($\mathbf{N} = \text{integers}$):

$$\nu(x) = \inf \{ n | n \in \mathbf{N} \text{ for all } k \geq n, \Phi_0^{-k}(x) \notin \text{support of } \Phi - \Phi_0 \},$$

$$\nu_1(x) = \inf \{ n | n \in \mathbf{N} \text{ for all } k \geq n, \Phi^{-k}(x) \notin \text{support of } \Phi - \Phi_0 \}.$$

Property (3) above insures that $\nu_1(x) < +\infty$ for all $x \in \mathbf{R}^n - \mathbf{R}^{n-1}$.

Now we can define two mappings $\phi: \mathbf{R}^n \rightarrow \mathbf{R}^n$ and $\phi_1: \mathbf{R}^n \rightarrow \mathbf{R}^n$:

$$\begin{aligned} \phi(x) &= \begin{cases} x & \text{if } x \in \mathbf{R}^{n-1}, \\ \Phi^{\nu(x)} \circ \Phi_0^{-\nu(x)}(x) & \text{if } x \in \mathbf{R}^n - \mathbf{R}^{n-1}. \end{cases} \\ \phi_1(x) &= \begin{cases} x & \text{if } x \in \mathbf{R}^{n-1}, \\ \Phi_0^{\nu_1(x)} \circ \Phi^{-\nu_1(x)}(x) & \text{if } x \in \mathbf{R}^n - \mathbf{R}^{n-1}. \end{cases} \end{aligned}$$

LEMMA 7. (a) $\phi \circ \phi_1 = \phi_1 \circ \phi = \text{Id}$. (b) *The restrictions of ϕ and ϕ_1 to $\mathbf{R}^n - \mathbf{R}^{n-1}$ are C^∞ -diffeomorphisms of this open set onto itself.*

PROOF. Notice first that if k is an integer and $k > \nu(x)$ (resp. $\nu_1(x)$), then $\phi(x) = \Phi^k \circ \Phi_0^{-k}(x)$ (resp. $\Phi_0^k \circ \Phi^{-k}(x)$). This implies (a). It is obvious that ν is locally bounded and for ν_1 the same property follows from property (3) of Φ . The local boundedness of ν and ν_1 implies (b).

We have to show that ϕ is C^∞ everywhere and that $\phi \equiv \text{identity} \bmod x_1^\infty$. This easily implies that ϕ is a diffeomorphism of \mathbf{R}^n onto itself.

Let us introduce the following notations: For any integer k , $f_k = \Phi_0^k \circ \Phi \circ \Phi_0^{-k-1}$, $\Psi_k = f_k - \text{Id} = \Phi_0^k \circ \Psi \circ \Phi_0^{-k}$, where $\Psi = (\Phi - \Phi_0) \circ \Phi_0^{-1}$. It is easy to see that we have the relation:

$$(I) \quad \phi = f_0 \circ f_1 \circ \cdots \circ f_{\nu-1}.$$

Since Ψ is flat on \mathbf{R}^{n-1} and has a compact support, we have the following estimate:

(E1) For any integers $m, N > 0$ there exists a constant $K(m, N)$ such that for all $x \in \mathbf{R}^n$

$$\|d^m \Psi(x)\| \leq K(m, N)|x|^N.$$

$d^m \Psi(x)$ is a symmetric m -linear mapping: $\mathbf{R}^n \times \cdots \times \mathbf{R}^n \rightarrow \mathbf{R}^n$ (m times) and $\| \cdot \|$ denotes the classical norm on these mappings.

It is easy to see that estimate (E1) implies the following two basic estimates:

$$(E2) \quad \|\Psi_k(x)\| \leq [K(0, N) + k\|v\||x_1|K(0, N+1)]|x_1|^N.$$

$$(E3) \quad \|d^m \Psi_k(x)\| \leq K(m, N)(1 + k\|v\|)^{m+1}|x_1|^N.$$

The next estimate is trivial but important.

(E4) For all $x \in \mathbf{R}^n$, $\nu(x) \leq (C + \|Px\|)/|x_1|$, where C is the diameter of the support of $\Phi - \Phi_0$.

The following lemmas will clinch the proof.

LEMMA 8. Let $R > 0$ and let

$$\delta = \delta(R) = \frac{1}{2(C + R)[K(0, 3) + K(0, 4)\|v\|(C + R)]}.$$

For all $x \in \mathbf{R}^n - \mathbf{R}^{n-1}$ such that $|x_1| \leq \delta$, $\|Px\| \leq R$, and for all integers j , $0 \leq j \leq \nu(x) - 1$,

$$(E5) \|f_j(x) - x\| \leq \delta[K(0, 3) + K(0, 4)\|v\|(C + R)]|x_1|^2.$$

For all $x \in \mathbf{R}^n - \mathbf{R}^{n-1}$ such that $|x_1| \leq \delta(R)/2$, $\|Px\| \leq R/2$, and for all integers j , $0 \leq j \leq \nu(x) - 1$,

$$(E6) |x_1 \circ f_j \circ \cdots \circ f_{\nu(x)-1}(x)| \leq 2|x_1|.$$

$$(E7) \|f_j \circ \cdots \circ f_{\nu(x)-1}(x) - x\| \leq 2|x_1|.$$

LEMMA 9. Under the conditions of Lemma 8

$$(E8) \|df_j(f_{j+1} \circ \cdots \circ f_{\nu(x)-1}(x)) - \text{Id}\| \leq C(R, N)|x_1|^N.$$

$$(E9) \text{ For all } m \geq 2, \|d^m f_j(f_{j+1} \circ \cdots \circ f_{\nu(x)-1}(x))\| \leq C_1(R, m, N)|x_1|^N.$$

C and C_1 are functions of R , N and R , m , N respectively.

Assume the lemmas for the moment. Introduce the pseudo-norms: $g \in C^\infty(\mathbf{R}^n; \mathbf{R}^n)$

$$\|g\|_{m, N, R} = \sup \left\{ \frac{1}{|x_1|^N} \|d^k g(x)\| \mid 0 \leq k \leq m, |x_1| \leq \delta(R), \|Px\| \leq R \right\}.$$

Then we claim that $\|\phi - \text{Id}\|_{m, N, R} < +\infty$. In fact:

$$\phi(x) - x = \sum_{j=0}^{\nu(x)-1} (f_j \circ \cdots \circ f_{\nu(x)-1}(x) - f_{j+1} \circ f_{j+2} \circ \cdots \circ f_{\nu(x)-1}(x)).$$

Using (E2–4), if $x \in \mathbf{R}^n$, $|x_1| \leq \delta(R)/2$, and $\|Px\| \leq R/2$, then:

$$\|\phi(x) - x\| \leq 3^{N+2} [K(0, N+1) + 3(C + R)\|v\|K(0, N+2)]|x_1|^N.$$

Hence

$$\|\phi - \text{Id}\|_{0, N, R/2} \leq 3^{N+2} [K(0, N+1) + 3(C + R)\|v\|K(0, N+2)].$$

To prove that ϕ is C^∞ , it is sufficient to show that for any real positive number R and any integers m, N ,

$$\sup \left\{ \frac{1}{|x_1|^N} \|d^m \phi(x)\| \mid x_1 \neq 0, |x_1| \leq \frac{\delta(R)}{2}, \|Px\| \leq \frac{R}{2} \right\} < +\infty.$$

Let us look at the structure of $d^m \phi$ first.

For all integers $m \geq 1$, $d^m \phi$ is a sum $\sum \sigma$ of “monomials.” We shall define these now. A monomial σ of degree k and height $\nu - 1$ is just the differential $d^k f_{\nu-1}$. Assume we have defined the monomials of degree k and height $j + 1$. A monomial σ of degree n and height j is a composition $d^l f_j(f_{j+1} \circ \cdots \circ f_{\nu-1})(\sigma_1 \times \cdots \times \sigma_l)$,

where each σ_r is a monomial of degree k_r and height $j + 1$ and the k_r satisfy the relations $k_1 + \dots + k_l = n$, $k_1 \geq 1, \dots, k_l \geq 1$. The factors of σ are $d^l f_j(f_{j+1} \circ \dots \circ f_{v-1})$ and the factors of the $\sigma_1, \dots, \sigma_l$.

Let us denote by $s(\sigma)$ the number of factors in σ and by $h(\sigma)$ the number of factors $d^p f_k(f_{k+1} \circ \dots \circ f_{v-1})$ such that $p > 1$. If $m > 1$, $h(\sigma) > 0$. Let $S(m)$ be the maximum of $s(\sigma)$ over all "monomials" σ appearing in the sum $d^m \phi = \sum \sigma$ and by $N(m)$ the number of σ appearing in this sum, each computed with its multiplicity: in the sum $\sum \sigma$, many σ 's are repeated several times.

Let us go from m to $m + 1$: $d^{m+1} \phi = \sum d\sigma = \sum \sigma'$. A "monomial" σ in $d^m \phi$ gives rise to $s(\sigma)$ "monomials" σ' in $d^{m+1} \phi$. Moreover each of these monomials σ' contains at most $s(\sigma) + \nu$ factors. Hence:

$$S(m+1) \leq S(m) + \nu, \quad N(m+1) \leq S(m)N(m).$$

This shows that $S(m) \leq m\nu$ since $S(1) = \nu$, and $N(m) \leq (M-1)! \nu^{m-1}$ since $N(1) = 1$.

Now we can estimate $d^m \phi$. First assume $m = 1$.

$$d\phi(x) - \text{Id} = \sum_{j=0}^{r-1} (A_j - \text{Id}) \circ A_{j+1} \circ \dots \circ A_{v-1}$$

if $A_j = df_j(f_{j+1} \circ f_{j+2} \circ \dots \circ f_{v-1})$. Hence if $x \in \mathbf{R}^n$, $|x_1| \leq \delta(R)/2$, and $\|Px\| \leq R/2$, then using (E8) we have

$$\begin{aligned} \|d\phi(x) - \text{Id}\| &\leq |x_1|^N C_1(R, N) \sum_{j=0}^{\nu(x)-1} (1 + C_1(R, 1)|x_1|)^{\nu(x)-1-j}, \\ \|d\phi(x) - \text{Id}\| &\leq |x_1|^{N-1} \frac{C_1(R, N)}{C_1(R, 1)} (1 + C_1(R, 1)|x_1|)^{\nu(x)}. \end{aligned}$$

Obviously

$$(1 + C_1(R, 1)|x_1|)^{\nu(x)-1} \leq \exp(C_1(R, 1)|x_1|\nu(x)).$$

Hence

$$\|d\phi(x) - \text{Id}\| \leq \frac{C_1(R, N)}{C_1(R, 1)} \exp(C_1(R, 1)(C + R)) |x_1|^{N-1}.$$

This shows that $\|d\phi - \text{Id}\|_{N, R/2} < +\infty$.

In the case $m \geq 2$; $d^m \phi = \sum \sigma$. As we have seen before, for each σ , $h(\sigma) \geq 1$. Let

$$\begin{aligned} M(x) &= \sup \left\{ \|df_k(f_{k+1} \circ \dots \circ f_{\nu(x)-1}(x))\| \mid a \leq k \leq \nu(x) - 1 \right\}, \\ M_1(x) &= \sup \left\{ \|d^l f_k(f_{k+1} \circ \dots \circ f_{\nu(x)-1}(x))\| \mid 0 \leq k \leq \nu(x) - 1, 2 \leq l \leq m \right\}. \end{aligned}$$

Then $\|\sigma(x)\| \leq M(x)^{s(\sigma)-h(\sigma)} M_1(x)^{h(\sigma)}$.

If $x \in \mathbf{R}^n$, $|x_1| \leq \delta(R)/2$, and $\|Px\| \leq R/2$, then using (E8)–(E9) we have

$$M(x) \leq (1 + C(R, 1)|x_1|), \quad M_1(x) \leq C_1(R, m, N)|x_1|^N.$$

Hence

$$\|\sigma(x)\| \leq (1 + C(R, 1)|x_1|)^{m\nu(x)} C_1(R, m, N)|x_1|^N$$

if $|x_1| \leq \delta(R)/2$, $1/C_1(R, m, N)$ and $\|Px\| \leq R/2$, since $s(\sigma) \leq m\nu$ and $h(\sigma) \geq 1$. Under these conditions:

$$\begin{aligned} \|d^m \phi(x)\| &\leq N(m)(1 + C(R, 1)|x_1|)^{m\nu(x)} C_1(R, m, N)|x_1|^N, \\ \|d^m \phi(x)\| &\leq (m-1)!(1 + C(R, 1)|x_1|)^{m\nu(x)} C_1(R, m, N)|x_1|^N \nu(x)^{m-1}, \\ \|d^m \phi(x)\| &\leq (m-1)!e^{C(R, 1)(C+R)}(C+R)^{m-1} C_1(R, m, N)|x_1|^{N-m}. \end{aligned}$$

This last inequality completes the proof. It remains to prove the lemmas.

(C) *Proof of Lemmas 8 and 9.* (E5) is an immediate consequence of (E2) and (E4).

(E6)–(E7). Let $\varepsilon = \delta[K(0, 3) + K(0, 4)]\nu/(C+R)$. Using a going down induction on j it is easy to see that if $x \in \mathbf{R}^n$, $\|Px\| \leq R/2$, and $\varepsilon|x_1| < 1$, then for all j , $0 \leq j \leq \nu(x) - 1$:

$$|x_1 \circ f_j \circ \cdots \circ f_{\nu(x)-1}(x)| \leq g^{\nu(x)-1-j}(|x_1|)$$

and

$$\begin{aligned} \|f_j \circ \cdots \circ f_{\nu(x)-1}(x) - x\| &\leq \sum_{s=j}^{\nu(x)-1} \|f_s \circ \cdots \circ f_{\nu(x)-1}(x) - f_{s+1} \circ \cdots \circ f_{\nu(x)-1}(x)\| \\ &\leq \varepsilon \sum_{s=j}^{\nu(x)-1} (g^{\nu(x)-1-s}(|x_1|))^2, \end{aligned}$$

where $g(t) = t(1 + \varepsilon t)$ and $g^k = g \circ \cdots \circ g$ (k times). If $0 < q\varepsilon t < 1$, then for all integers $k \leq q$, $g^k(t) \leq g_1^k(t)$ where $g_1^k = g_1 \circ \cdots \circ g_1$ (k times) and $g_1 = t/(1 - \varepsilon t)$. Hence

$$|x_1 \circ f_0 \circ \cdots \circ f_{\nu(x)-1}(x)| \leq \frac{|x_1|}{1 - \varepsilon(C+R)}$$

and

$$\|f_0 \circ \cdots \circ f_{\nu(x)-1}(x) - x\| \leq \frac{\varepsilon|x_1|^2 \nu(x)}{(1 - \varepsilon(C+R))^2}$$

provided that $\varepsilon(C+R) < 1$. Hence

$$\|f_0 \circ \cdots \circ f_{\nu(x)-1}(x) - x\| \leq \frac{\varepsilon(C+R)}{(1 - \varepsilon(C+R))^2} |x_1|.$$

Let us choose δ such that $\varepsilon(C+R) = \frac{1}{2}$. Then:

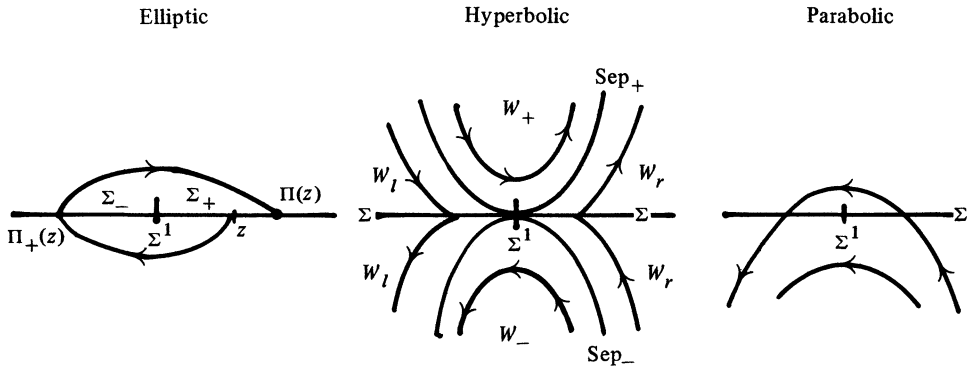
$$\begin{aligned} |x_1 \circ f_0 \circ f_1 \circ \cdots \circ f_{\nu(x)-1}(x)| &\leq 2|x_1|, \\ |f_0 \circ \cdots \circ f_{\nu(x)-1}(x) - x| &\leq 2|x_1|, \end{aligned}$$

and

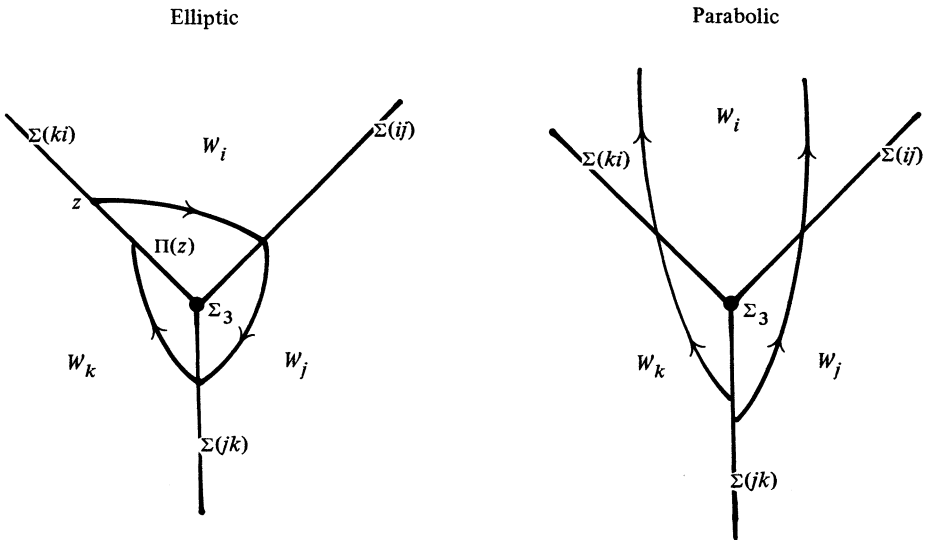
$$\delta(R) = \frac{1}{2(C+R)[K(0, 3) + K(0, 4)]\nu/(C+R)}.$$

(E8)–(E9) follow immediately from (E2–3–6–7).

Fold



Maxwell



REFERENCES

1. V. I. Arnold, *Lagrangian manifolds with singularities, asymptotic rays, and the open swallow-tail*, Functional Anal. Appl. **15** (1981), 235–246.
2. I. Ekeland, *Discontinuité des champs hamiltoniens et existence de solutions optimales en calcul des variations*, Inst. Hautes Études Sci. Publ. Math. no. 47, 1977, pp. 1–32.
3. R. Abraham and J. E. Marsden, *Foundations of mechanics*, Benjamin, New York, 1966.

INSTITUT FOURIER, UNIVERSITÉ DE GRENOBLE I, B. P. 74, 38402 SAINT MARTIN D'HERES CEDEX, FRANCE