

Generalized Hamiltonians and Optimal Control: A Geometric Study of the Extremals

IVAN A. K. KUPKA

I. Introduction to geometric optimal control.

Introduction. Optimal control problems are generalizations of classical problems in the calculus of variations. A typical one can be stated as follows: given a smooth (C^∞ or real analytic C^ω) manifold M , a compact smooth manifold U (possibly with a boundary), a smooth vector-field $E: M \times U \rightarrow TM$ (tangent space of M) on M , parametrized by U , a smooth function $c: M \times U \rightarrow R$, and two points A, B in M , let $\text{Tr}(A, B)$ be the set of all pairs $(x, u): [a, b] \rightarrow M \times U$, such that: (1) x is absolutely continuous and u measurable; (2) $dx(t)/dt = E(x(t), u(t))$ a.e.; (3) $x(a) = A, x(b) = B$.

The problem is to find a pair $(\bar{x}, \bar{u}): [\bar{a}, \bar{b}] \rightarrow M \times U$ such that

$$\int_{\bar{a}}^{\bar{b}} c(\bar{x}(t), \bar{u}(t)) dt = \inf \left[\int_a^b c(x(t), u(t)) dt \mid (x, u): \right. \\ \left. [a, b] \rightarrow M \times U, \text{ belonging to } \text{Tr}(A, B) \right].$$

A pair in $\text{Tr}(A, B)$ is called a trajectory of the system; the pair (x, u) is called an optimal trajectory.

It is well known that such a pair (x, u) is the projection on $M \times U$ of an "extremal." An extremal is the generalization of its namesake of classical calculus of variations. In the present situation, there are two families \mathcal{E}_λ , $\lambda = 0$ or 1 , of extremals: a couple $(z, u): [a, b] \rightarrow T^*M \times U$ belongs to \mathcal{E}_λ if it satisfies the following conditions:

(1) $dz/dt = \overrightarrow{H^\lambda(z(t), u(t))}$ for almost all t in $[a, b]$, $H^\lambda(z, u) = \langle z, E(x, u) \rangle - \lambda c(x, u)$, x being the projection of z onto M ; $\langle \cdot, \cdot \rangle$ denotes the canonical pairing $TM \times_M T^*M \rightarrow R$, and $\overrightarrow{H^\lambda}$ is the hamiltonian field associated to H^λ considered as a function on T^*M parametrized by u .

(2) For almost all $t \in [a, b]$, $H^\lambda(z(t), u(t)) = K^\lambda(z(t))$, where $K^\lambda(z) = \sup\{H^\lambda(z, v) \mid v \in U\}$.

The family \mathcal{E}_1 is called ordinary, the family \mathcal{E}_0 exceptional.

As in the classical calculus of variations, one tries to solve the optimal control problems using the extremals. Two methods have been exploited up to now. The first one, which could be called the direct method, is being developed by H. Sussmann and his collaborators. It has yielded some important results in the case of $\dim M$ being two or three.

The second method, the singularity method, was introduced by I. Ekeland [E] in the special case when

$$M = U = R \quad \text{and} \quad E(x, u) = u.$$

More recently F. Klok pushed this analysis further in the same case [K1]. Our approach belongs to this last line of thought.

1. *Preliminary considerations on extremals.* The main difference between the classical calculus of variations and our case is that in condition (2) of the definition of extremals the maximum K^λ can be attained for several distinct u 's. This allows for the phenomenon of "switching," that is, the extremal changing its policy $u(t)$ abruptly at some time T . Mathematically this translates into the fact that z is not differentiable at t . Let us formalize this.

DEFINITION 0. A point $z(s)$ (resp. s) on an extremal $z: [a, b] \rightarrow T^*M$ is called a switching point (resp. a switching time) if s belongs to the closure of the set of all t 's where z is not differentiable.

NOTATION. The set of all possible switching points is a subset of T^*M , called the switching surface.

The notion of switching points is crucial in the study of extremals. They determine the structure of these curves. What can we say about this structure? H. Sussmann has noticed that: (a) in the C^∞ case, any absolutely continuous curve in M is optimal for some appropriate system (E, c) ; (b) in the C^ω case, given a system (E, c) , if there exists an optimal trajectory joining two given points A, B in M , then there exists another optimal trajectory joining A to B , which is analytic on an open dense subset of times. Since any optimal trajectory is the projection on M of an extremal, this shows that in order to get any reasonable theory, we have to put some restrictions on the system (E, c) .

Now, even for a generic system (E, c) , the extremals are not smooth in general. A consequence of our results is the fact that for an extremal to have an infinite number of switching points is a very stable property. Let us note here, that the structure of the general extremal in the generic case is not known.

Finally, the extremals would be the trajectories of the hamiltonian field associated to K^λ , if K^λ were smooth, which it is not, in general. Let us mention that generalizations of the concept of hamiltonian field to include this case have been put forward.

2. *Regular points of finite multiplicity.* From now on we drop the superscript λ in H^λ . Let us denote by S the subset in $T^*M \times U$ of all couples (z, \bar{u}) such that \bar{u} is a local maximum point of the function $H: v \in U \rightarrow H(z, v)$. It is clear that if $(z, u): [a, b] \rightarrow T^*M \times U$ is an extremal, then for almost every $t \in [a, b]$,

$(z(t), u(t))$ belongs to S . Let $p: S \rightarrow T^*M$ denote the restriction to S of the canonical projection: $T^*M \times U \rightarrow T^*M$.

Without making a formal statement, it is clear that, for a generic pair, there exist stratifications of T^*M and S such that: (1) p is stratified; (2) for any stratum A , $p: p^{-1}(A) \rightarrow A$ is a finite covering; (3) for any open stratum A of T^*M , for any z in A , all points in $p^{-1}(z)$ are either nondegenerate quadratic singular points of H_z , or are regular points of H_z belonging to the boundary of U , which are nondegenerate quadratic singular points for the restriction of H_z to the boundary of U .

On the lower-dimensional strata, a branching of singularities takes place. Since we are mainly interested in the switching phenomena, we shall not go into branching but concentrate on the open strata. This motivates the following definition. In it, we do not assume that S and p satisfy the conditions (1)–(3) above.

DEFINITION 1. A point q in T^*M is called a regular point of multiplicity m if there exists a neighborhood V of q in T^*M , such that:

- (1) The restriction $p: p^{-1}(V) \rightarrow V$ is a trivial finite covering.
- (2) Let J be the set of all sections $\varphi: V \rightarrow S$ of this covering such that q belongs to the closure $\Gamma(\varphi)$ of the set

$$\{z \in V, z \neq q, H(\varphi(z)) = K(z)\}.$$

Then for any two sections $\varphi, \psi \in J$ the germs at q of the restrictions of $\text{Ho } \varphi$ and $\text{Ho } \psi$ to $\Gamma(\varphi)$ and $\Gamma(\psi)$ respectively, are not equal. m is the cardinal of J .

The case $m = 1$ corresponds to the classical theory of the calculus of variations. Near q , the extremals are the trajectories of a hamiltonian vector-field. We have studied the cases when m is 2 or 3. The structure of the extremals near q depends essentially on the structure of the contacts of the hamiltonian vector-fields $\text{Ho } \varphi$, $\varphi \in J$, with the switching surface and certain subsets of it defined by these contacts. This vague statement can be given a precise formulation using the Lie algebra, generated by the set of functions $[\text{Ho } \varphi / \varphi \in J]$, under the Poisson bracket. The complexity of the contact structure at a point q is measured by the minimum of the length of the brackets not zero at q .

In the remainder of this paper, we shall discuss two of our results. Both deal with the case $m = 2$.

II. Statement of the results.

1. *Notations and auxiliary concepts.* Let q be a regular point of multiplicity 2. J contains two sections φ_+, φ_- . The associated functions $\text{Ho } \varphi_+$ and $\text{Ho } \varphi_-$ will be denoted by H_+ and H_- respectively. In a neighborhood of q , the switching surface Σ is defined by $H_+ - H_- = 0$. We shall make the following assumption for the remainder of this paper:

$$0 = dH_+(q) \wedge dH_-(q) \wedge d\{H_+, H_-\}(q), \text{ where } \{H_+, H_-\} \text{ denotes the Poisson bracket of } H_+ \text{ and } H_-, \text{ that is, the Lie derivate of } H_+ \text{ in the direction of } \overrightarrow{H_-} \text{ (see [A-M]).} \quad (*)$$

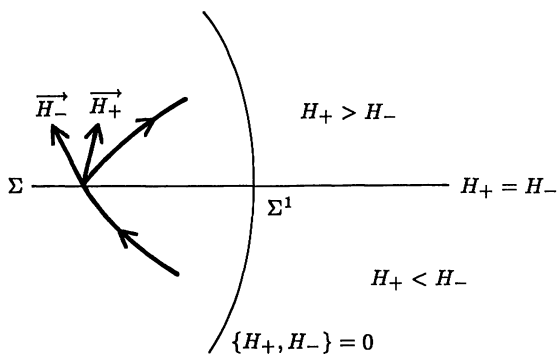


FIGURE 1

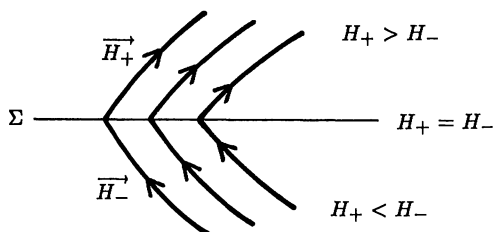


FIGURE 2

(*) implies that the switching surface is a smooth manifold in a neighborhood of q and that the same is true for the set $\Sigma^1 = \Sigma \cap [\{H_+, H_-\} = 0]$, where $\overrightarrow{H_+}$ and $\overrightarrow{H_-}$ are tangent to Σ (see Figure 1).

If q does not belong to Σ^1 then, in a neighborhood of it, the extremals are the trajectories of a piecewise smooth flow, having a tangential discontinuity along Σ (see Figure 2). Hence the only interesting cases are when q belongs to Σ^1 .

NOTATION. (1) For simplicity, let us denote by f, g the brackets

$$\{H_-, \{H_+, H_-\}\} \quad \text{and} \quad \{H_+, \{H_-, H_+\}\}$$

respectively.

(2) Given an open subset W in T^*M , a subset N of W will be called invariant in W if any extremal contained in W and meeting N is contained in N .

2. First theorem—Fold points.

DEFINITION 3. A regular point of multiplicity 2, q , satisfying the assumption (*), is called a fold point if $f(q)$ and $g(q)$ are both nonzero. If they are both positive (resp. negative) q is called hyperbolic (resp. elliptic). If they have opposite signs, q is parabolic.

NOTATION. (3) In the elliptic case, the following vector-field R , defined on Σ^1 , plays an important role:

$$R = \frac{f\overrightarrow{H_+} + g\overrightarrow{H_-}}{f + g}.$$

It is clear that R is the unique convex combination of $\overrightarrow{H_+}$ and $\overrightarrow{H_-}$ tangent to Σ^1 at the points of Σ^1 .

THEOREM 1. *Let q be a fold point. There exists a neighborhood W of q such that:*

(e) *If q is elliptic, we have a generalized "flow-box" result; there exist a ball B of codimension 1 in T^*M , an interval $I = [-a, a]$, and a continuous mapping $z: I \times B \rightarrow W$, with the properties:*

(i) *$z(0, b) = b$ if $b \in B$, and z is a homeomorphism onto W , piecewise smooth on any $I \times K$, K compact subset of $B - \Sigma^1$.*

(ii) *For $b \in \Sigma^1 \cap B$, the curve $z_b: I \rightarrow W$ is the trajectory of R , passing at time 0 through b .*

(iii) *For $b \in B - \Sigma^1$, z is an extremal.*

(iv) *For any subinterval d of I , any $b \in B - \Sigma^1$, let $N(d, b)$ be the number of switching times of z_b in d . When b tends to Σ^1 , $N(d, b)$ tends to ∞ and $N(d, b) \cdot h(z(t, b))$ tends to the length of d , for any $t \in d$. h is the function $2 \cdot \{H_+, H_-\} \cdot [1/f + 1/g]$.*

(h) *If q is hyperbolic, it behaves somewhat like a hyperbolic singular point of a vector-field. W contains two smooth hypersurfaces $S(+)$ and $S(-)$, having a contact of first order with Σ along Σ^1 , with the following properties:*

(i) *$W - S(+) \cup S(-)$ has four connected components W_+, W_-, W_r, W_l . W_+ and $S(+)$ (resp. W_- and $S(-)$) are located in $[H_+ \geq H_-]$ (resp. $[H_+ \leq H_-]$). W_k (resp. W_l) is located in $\{H_+, H_-\} > 0$ (resp. $\{H_+, H_-\} < 0$).*

(ii) *The sets $S(+) \cup S(-)$, W_+, W_-, W_r, W_l are invariant in W .*

(iii) *In W_+ (resp. W_-) the extremals are trajectories of $\overrightarrow{H_+}$ (resp. $\overrightarrow{H_-}$). They do not switch.*

(iv) *In W_r (resp. W_l), the extremals switch exactly once and they are the trajectories of a piecewise smooth flow.*

(v) *In $S(+) \cup S(-)$, the extremals either do not switch and then they are the trajectories of $\overrightarrow{H_+}$ in $S(+)$ or of $\overrightarrow{H_-}$ in $S(-)$, or they switch once and cross from $S(+)$ to $S(-)$ or vice versa.*

(p) *If q is parabolic, let us assume that $f(q) > 0$ and $g(q) < 0$. W contains a smooth hypersurface Sp with the following properties:*

(i) *$W - Sp$ has two connected components, W_+ and W_- .*

(ii) *W_+ is contained in $[H_+ > H_-]$, $Sp \cup W_-$ in $[H_+ \leq H_-]$.*

(iii) *Sp, W_+, W_- , are invariant in W .*

(iv) *In $Sp \cup W_-$, the extremals do not switch and they are the trajectories of $\overrightarrow{H_-}$. In W_+ they switch twice and are the trajectories of a piecewise smooth flow. The case $f(q) < 0$, $g(q) > 0$ is similar.*

(For these results see Figure 3.)

This theorem calls for some remarks:

(1) The field R is called the residual field. It also shows up in some work of Arnold (see [Ar]).

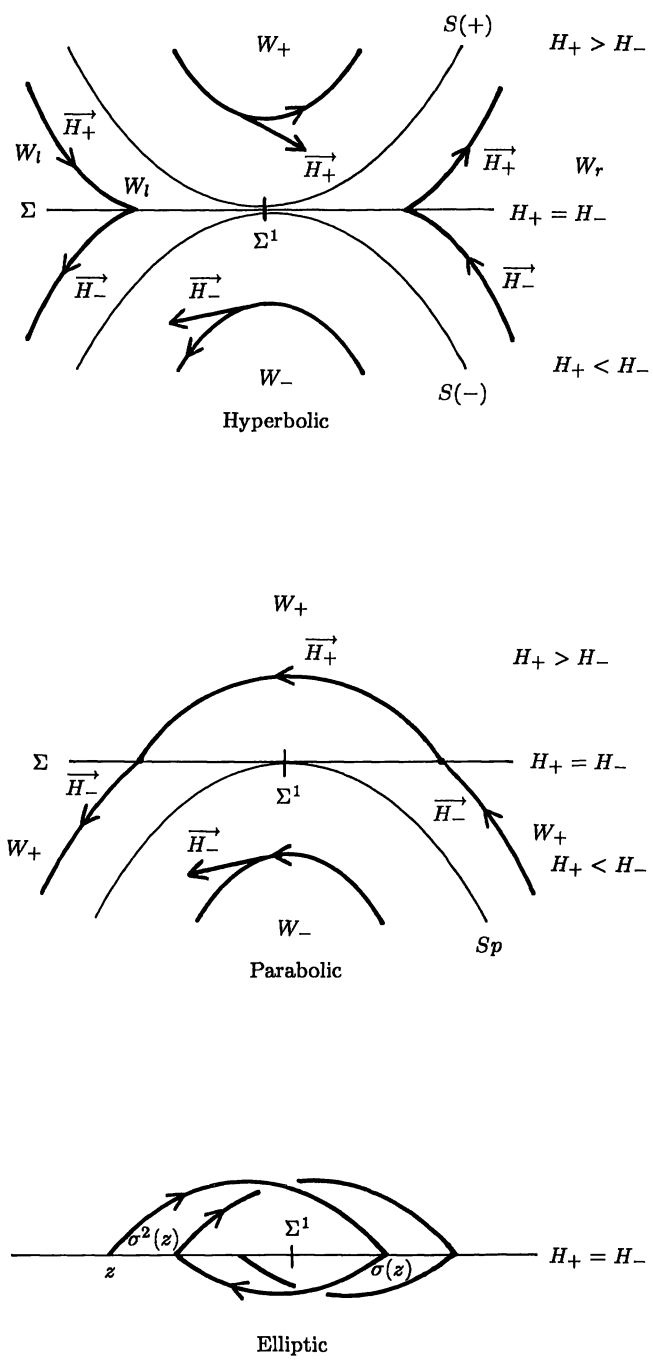


Figure 3

(2) The only really interesting and nontrivial part of the preceding theorem is the elliptic case. Near an elliptic point the number of switching points on an extremal is not bounded, and this occurrence is stable. The extremals "spiral" around Σ^1 , and as they tend to Σ^1 , they pick up more and more switching points, so that, in the limit, they become smooth.

(3) The residual field R is also defined in the hyperbolic domain and it is important in the study of relaxed trajectories. The trajectories of R are relaxed trajectories of the system (E, c) . In the elliptic case, these curves are more expensive than the nearby extremals. But in the hyperbolic case, they are cheaper and together with the extremals in $S(+) \cup S(-)$, they can be used to construct local optimal control synthesis of the "turnpike" type.

(4) The proof of the above result and some other ones will appear in the Transactions of the American Mathematical Society.

In agreement with our remarks at the end of paragraph 2 of §1 the preceding result dealt with the case where the bracket of length 2 of the Lie algebra L generated by (H_+, H_-) is zero at q , but those of length 3 are not. The situation when these latter are zero but those of length 4 are not, we shall not examine here and instead pass to the next stage where new phenomena appear. In our preceding considerations, a single extremal switched a finite number of times only. The question is: is it possible to have a general system (E, c) such that above each point from an open subset of M , there passes an extremal that switches an infinite number of times in a finite time-interval? We shall answer this question next.

3. *The Fuller curves.* Let q be a point in T^*M as in Definition 3.

DEFINITION 4. A pair of smooth arcs of curves, $C(+)$ and $C(-)$, contained in $\Sigma - \Sigma^1$, having both q as extremity and no other point in common, is called a Fuller pair if it has the following properties: (1) there is a continuous function

$$T: C(+) \cup C(-) \rightarrow \mathbf{R},$$

such that any extremal z starting at a point s in $C(+) \cup C(-)$, is defined on the interval $[0, T(s)]$ and $z(t)$ tends to q when t tends to $T(s)$.

(2) Let $s \in C(u)$, $u = +$ or $-$. The switching times of z form an increasing sequence $[0, t_1, \dots, t_n, \dots]$ such that $z(t_n)$ belongs to $C(u)$ (resp. $C(-u)$) if n is even (resp. odd).

(3) There exist a constant $k > 1$, depending only on the pair $(C(+), C(-))$ and a continuous function $D: C(+) \cup C(-) \rightarrow \mathbf{R}$, such that $k^n(t_{n+1} - t_n)$ tends to $D(s)$ as n tends to infinity (see Figure 4).

To state our second theorem we need one more notation.

NOTATION 4. Let Γ denote the vector space of all smooth functions on an unspecified open subset of T^*M . $\alpha: \Gamma \times \Gamma \rightarrow \Gamma \times \Gamma \times \dots \times \Gamma$ (6 times) is the mapping defined as follows: $\alpha(f, g) = (\alpha_1(f, g), \dots, \alpha_6(f, g))$ where $\alpha_1, \dots, \alpha_6$ is the ordered set of all elements of length 5 from a Hall basis built on the set (f, g) ordered by $f \prec g$, $\text{ad } f(g) = \{f, g\}$.

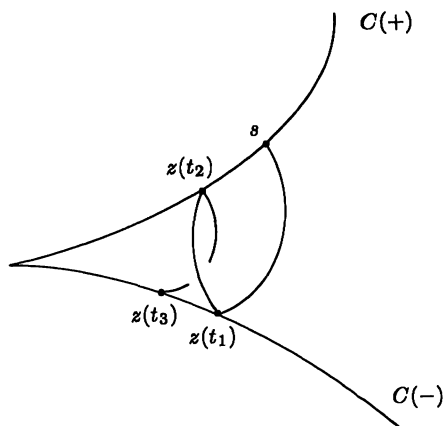


FIGURE 4

EXAMPLE.

$$\begin{aligned}\alpha_1(f, g) &= \text{ad}^4 f(g), & \alpha_2(f, g) &= \text{ad } g \text{ad}^3 f(g), & \alpha_3(f, g) &= \text{ad}^2 g \text{ad}^2 f(g), \\ \alpha_4(f, g) &= -\text{ad}^4 g(f), & \alpha_5(f, g) &= \{\text{ad } f(g), \text{ad}^2 f(g)\}, \\ \alpha_6(f, g) &= \{\text{ad } f(g), -\text{ad}^2 g(f)\}.\end{aligned}$$

THEOREM 2. *There exists a semialgebraic set \mathcal{F} (explicit) in R^6 , with non-empty interior having the property: let q be a regular point of multiplicity 2, satisfying the assumption (*). If the couple (H_+, H_-) has the properties (a)–(b) below, then there is a Fuller pair passing through q .*

(a) *All the Poisson brackets of length 2, 3, and 4, built on H_+, H_- , are zero at q .*

(b) *$\alpha(H_+, H_-)$ belongs to \mathcal{F} .*

COMMENT. This result shows that the presence of a Fuller pair is a remarkably stable phenomenon. Using it, one can show that on any smooth manifold M , there is an open (C^∞ -topology) set of systems (E, c) , such that, for any one of these systems, there exists an open subset O of M with the property that above any point x in O , there is a point q with a Fuller pair passing through it.

III. Short review of the techniques used. Essentially, three types of techniques are used: (i) discrete dynamical systems, (ii) partial normal forms, and (iii) blowing up procedures.

1. *Discrete dynamical systems.* To each point q regular of multiplicity 2, satisfying the assumption (*), we associate a discrete dynamical system (DDS for short), σ , as follows: the domain of σ , $\text{dom}(\sigma)$, is the set of all z in Σ , such that:

- (1) $\{H_+, H_-\}(z) \neq 0$. Let u be the sign of this number.
 - (2) The trajectory of $\overrightarrow{H_u}$, starting at z , meets Σ again, at z_u for the first time.
 - (3) $u\{H_+, H_-\}(z_u) < 0$.
- (2) implies that: $u\{H_+, H_-\}(z) \leq 0$. If z belongs to $\text{dom}(\sigma)$, we set $\sigma(z) = z_u$.

The DDS σ is useful in keeping track of the switching points, and it determines the behavior of the extremals: if $z: [a, b] \rightarrow W - \Sigma^1$ is such a curve, its set of switching points is discrete. Let y be the first (timewise) of these points. Then the set of them, ordered by increasing time, is a partial orbit of $y: [y, \sigma(y), \sigma^2(y), \dots, \sigma^N(y)]$.

As an example, the main ingredient in the proof of the elliptic case of Theorem 1 is a very convenient normal form for the associated DDS σ .

2. *Partial normal forms.* Using symplectic coordinate transformations, we determine partial normal forms for the pair (H_+, H_-) . Such a form is the sum of a normal form for some set of (H_+, H_-) and a remainder term. More precisely: let

$$F = \frac{1}{2}(H_+ + H_-), \quad G = \frac{1}{2}(H_+ - H_-).$$

We determine a symplectic system of coordinates, centered at q , $(x_1, \dots, x_d, p_1, \dots, p_d)$, such that $G = p_1$, $F = F_0 + F_1$, where F_0 is the normal form and F_1 the remainder. There is a gradation on the coordinates such that F_0 is a homogeneous polynomial and the order of F_1 at q (degree of the lowest degree terms in the Taylor series of F_1 at q) is greater than the degree of F_0 . This gradation is intimately linked with the structure at q of the Lie algebra generated by H_+ and H_- . It defines a local group action of the multiplicative group of all positive reals on a neighborhood of q , for which F_0 and G are semi-invariants. In the next proposition, $w(P)$ will denote the degree of the homogeneous polynomial P , and $\text{ord}(h)$, the order of h at q .

PROPOSITION 1. (i) If $\{H_+, H_-\}(q) = 0$, but not all brackets of length 3 are zero at q , then

$$\begin{aligned} F_0 &= p_2 + x_1(\tfrac{1}{2}ax_1 + bx_2), \\ w(x_n) &= 1 \text{ and } w(p_n) = 2 \quad \text{if } n = 1 \text{ or } 2, \\ w(x_n) &= w(p_n) = 3 \quad \text{if } n > 3, \\ w(F_0) &= 2, \quad \text{ord}(F_1) \geq 3, \end{aligned}$$

$a = \{G\{G, F\}\}(q)$, $b = \{F\{G, F\}\}(q)$. This is the fold case.

(ii) If all the Poisson brackets of length 2 and 3 of H_+, H_- are 0 at q but not all brackets of length 4:

$$\begin{aligned} F_0 &= p_2 + x_1(p_3 + \tfrac{1}{6}ax_1^2 + \tfrac{1}{2}bx_1x_2 + \tfrac{1}{2}cx_2^2), \\ w(x_n) &= 1 \text{ and } w(p_n) = 2 \quad \text{if } n = 1 \text{ or } 2, \\ w(x_n) &= w(p_n) = 3 \quad \text{if } n \geq 3, \\ W(F_0) &= 3, \quad \text{ord}(F_1) \geq 4. \end{aligned}$$

$a = -\text{ad}^3 G(F)(q)$, $c = \text{ad}^3 F(G)(q)$, $b = -\text{ad } F \text{ad}^2 G(F)(q)$. This is the cusp case.

(iii) If all the brackets of length 2, 3, and 4 of H_+ , H_- at q are 0 but not all of length 5:

$$F = p_2 + x_1(p_3 + \frac{1}{2}bx_1^2x_3 + ax_2x_3 - \frac{1}{6}c_4x_2^3 + \frac{1}{4}c_3x_1x_2^2 + \frac{1}{6}c_2x_1^2x_2 + \frac{1}{4}c_1x_1^3),$$

$$w(x_n) = 1 \text{ and } w(p_n) = 4 \quad \text{if } n = 1 \text{ or } 2,$$

$$w(x_3) = 2, \quad w(p_3) = 3, \quad w(x_n) = w(p_n) = 3 \quad \text{if } n \geq 4,$$

$$w(F_0) = 4, \quad \text{ord}(F_1) \geq 5, \quad c_n = \alpha_1(G, F)(q) \quad \text{if } i \leq n \leq 4,$$

$$b = \alpha_5(G, F)(q), \quad a = \alpha_6(G, F)(q).$$

3. *Blowing up technique.* Using the action of the multiplicative group of the positive reals, \mathbf{R}_+^* , we can blow up the point q on the manifold T^*M . This is not the classical blowing up procedure but a “weighted” version of it, q is replaced by the quotient, Q , of $V - q$ under the action of \mathbf{R}_+^* , V being a suitable neighborhood of q .

4. *Sketch of the proof of Theorem 2.* Let σ and σ_0 denote the associated DDS to the couples (H_+, H_-) and $(F_0 + G, F_0 - G)$. Using the blowing up technique, the stability theorem of hyperbolic manifolds we reduce the problem of finding a Fuller pair for the couple (H_+, H_-) to that same problem for the second couple. A suitable neighborhood of Q in the blown up space is fibered by the orbits of the \mathbf{R}_+^* -action. Since the couple $(F_0 + G, F_0 - G)$ is semi-invariant, the lifting $\tilde{\sigma}_0$ of σ_0 to the blown-up space preserves Q and this fibration. Assume we can find a fixed point h of the restriction of $\tilde{\sigma}_0^2$ to Q . The fiber $O(h)$ above h is then invariant under $\tilde{\sigma}_0^2$. If we can choose h in such a way that $O(h)$ is a contracting curve for $\tilde{\sigma}_0^2$, then the projection on T^*M of the pair $(O(h), \sigma_0(O(h)))$ is a Fuller pair for the couple $(F_0 + G, F_0 - G)$. This ends the “proof” of Theorem 2.

REFERENCES

- [A-M] R. Abraham and J. Marsden, *Foundations of mechanics*, Benjamin, 1967.
- [Ar] V. Arnold, *Lagrangian manifolds with singularities asymptotic rays and the open swallowtail*, Functional Anal. Appl. **15** (1981), 235–246.
- [E] I. Ekeland, *Discontinuite des champs hamiltoniens et existence de solutions optimale en calcul des variations*, Inst. Hautes Études Sci. Publ. Math. **43** (1977), 5–32.
- [Kl] F. Klok, *Broken solutions of homogeneous variational problems*, J. Differential Equations **55** (1984), 101–134.
- [Ku] I. Kupka, *Geometric theory of extremals in optimal control problems I: The fold and Maxwell case*, Trans. Amer. Math. Soc. **299** (1987), 225.
- [Ku2] —, *The ubiquity of the Fuller phenomenon*, Preprint, Inst. Fourier (1986).

INSTITUT FOURIER, 38402 SAINT MARTIN D'HÈRES, FRANCE



Dynamics of Area Preserving Maps

JOHN N. MATHER

Poincaré initiated the study of the dynamics of area preserving mappings, in his studies of celestial mechanics [20]. He showed that the study of the dynamics of the restricted three body problem (two positive masses, one zero mass) could be reduced to the study of the dynamics of an area preserving mapping. He showed, moreover, that even in this case, which is nearly the simplest nontrivial case of Hamiltonian mechanics, the dynamics is so complicated that there is no hope of “solving” the n -body problem (or even the restricted three body problem), in the sense of finding exact expressions of the trajectories as a function of time.

As a consequence of Poincaré’s pioneering work, the focus of mathematical studies related to celestial mechanics has shifted to the more topological and analytical approach which Poincaré dubbed “dynamical systems.” The books of Arnold and Avez [1] and Moser [15] and the articles of Kolmogorov [10] and Smale [21] present overviews of modern developments in the theory of dynamical systems.

One of the main questions of dynamical systems is the extent to which they display randomness or stability. Many studies in the past century have dealt with these questions. The KAM (Kolmogorov, Arnold, Moser) theory shows that small Hamiltonian perturbations of integrable Hamiltonian systems display a great deal of stability. Invariant tori on which the flow is conjugate to a linear flow exist and fill up most of phase space in the sense of Lebesgue measure. (See, e.g., Moser [15].) In contrast, hyperbolic systems exhibit a great deal of randomness, as is discussed, for example, in Hadamard [8], Anosov [2], Smale [21], Bowen [6], and Pesin [19]. But, even small Hamiltonian perturbations of integrable Hamiltonian systems have regions of instability or randomness alongside the regions of stability. This instability was discovered by Poincaré, further explored by Birkhoff, and given a very transparent form by Smale [21] in terms of “horseshoes.”

All this work shows that, typically, one finds a pattern of stability and instability mixed together in a complicated way. But there are many unresolved questions. In the Newtonian n -body problem are the unbounded trajectories

dense in phase space (where the center of mass is fixed at the origin)? Newton integrated the 2-body problem and thereby showed that for $n = 2$, the answer to this question is no. The restricted 3-body problem is a Hamiltonian system in 2 degrees of freedom and it is possible to deduce from KAM theory that the answer is no in that case, too (Moser [16]). But all other cases are unsolved.

Is it generically the case that Hamiltonian systems on a smooth compact symplectic manifold are topologically transitive? Here, one must specify what one means by “generically.” A popular notion of genericity is that a property of Hamiltonian systems is C^r generic if the set of C^r Hamiltonian systems on the given manifold having that property is a residual set in the C^r topology, in the sense of Baire category. We recall that a dynamical system is said to be topologically transitive if it has a dense orbit. Here again, KAM theory shows that the answer is no for systems in 2 degrees of freedom (i.e., on a 4-manifold) and r sufficiently large, but other cases are unsolved. The KAM theory resolved the analogous problem for “topologically transitive” replaced by “ergodic” and r sufficiently large, the answer being no, contrary to what was expected.

These problems are very difficult and no solution is in sight. In this article, I will report on some recent progress on Hamiltonian systems in two degrees of freedom and the closely related subject of area preserving mappings. Even for such an apparently simple case, there are many difficult unresolved questions, and these questions have attracted engineers, who have recently done a great deal of numerical work on them (surveyed in [11]), as well as inspired mathematicians to obtain deep results (e.g., [9]).

In this article, I will report on one aspect of recent work on dynamics of area preserving mappings, based on variational methods. Although these methods do not apply to all area preserving homeomorphisms, they apply to a large class of such homeomorphisms, the monotone tilt homeomorphisms. This work is an extension of earlier work of Aubry [3] and myself [12]. Bangert [5], Chenciner [7], and Moser [17, 18] have given very complete expositions of this earlier work and related matters, so I will use this opportunity to announce extensions of this earlier work, which are not yet published.

For simplicity, I will confine the discussion to C^1 monotone twist (area preserving) diffeomorphisms, of an annulus. There is no loss of generality in considering only positive twist diffeomorphisms, since the inverse of such a diffeomorphism is a negative twist diffeomorphism. This is the class of mappings considered, for example, in [12].

A mapping in this class is a C^1 diffeomorphism \bar{f} of the annulus

$$\bar{A} = S^1 \times [0, 1]$$

onto itself which maps each boundary component to itself, preserves area and orientation, and has the “positive twist” property, i.e., for each $\theta \in S^1$, the mapping $y \mapsto \text{pr}_1 \bar{f}(\theta, y)$ has positive derivative at each point, where pr_1 denotes the projection of $S^1 \times [0, 1]$ on its first factor. We let f be a lift of \bar{f} to the universal cover $A = \mathbf{R} \times [0, 1]$ of \bar{A} . Then the *rotation interval* $(\rho(f_0), \rho(f_1))$ of

f is defined, where $f_i = f|_{\mathbf{R} \times i}$, $i = 0, 1$, and $\rho(f_i)$ is the Poincaré rotation number of f , i.e., $\rho(f_i) = \lim_{n \rightarrow \pm\infty} f_i^n(x)/n$ for any $x \in \mathbf{R}$.

We let $h: B \rightarrow \mathbf{R}$ be a “generating function” of f , i.e., $B = \{(x, x') \in \mathbf{R}^2: \text{there exists } y \in [0, 1] \text{ with } \text{pr}_1 f(x, y) = x'\}$, and h is the function defined (up to addition of a constant) by $f(x, y) = (x', y')$ if and only if $y = -\partial_1 h(x, x')$ and $y' = \partial_2 h(x, x')$.

We let C denote the subset of $\mathbf{R}^{\mathbf{Z}}$ consisting of bi-infinite sequences $x = (x_i)_{i \in \mathbf{Z}} \in \mathbf{R}^{\mathbf{Z}}$ with $(x_i, x_{i+1}) \in B$. We let M denote the set of $x \in C$ such that for all $m, n \in \mathbf{Z}$ with $m < n$, we have that $x' \in C$, $x'_m = x_m$, and $x'_n = x_n$ imply

$$\sum_{i=m}^n h(x_i, x_{i+1}) \leq \sum h(x'_i, x'_{i+1}).$$

Elements of C are called *configurations* and elements of M are called *minimal energy configurations*. Aubry and Le Daeron [3] have developed a more or less complete theory of minimal energy configurations. See Bangert [5] for a complete exposition of this theory. If x is a minimal energy configuration, the rotation number $\rho(x) = \lim_{i \rightarrow \pm\infty} x_i/i$ exists and lies in $[\rho(f_0), \rho(f_1)]$. We define the *rotation symbol* $\tilde{\rho}(x)$ of x to be $\rho(x)$ if $\rho(x)$ is irrational or if $\rho(x) = p/q$ and $x_{i+q} = x_i + p$ for all $i \in \mathbf{Z}$. If $\rho(x) = p/q$ and $x_{i+q} > x_i + p$ for all $i \in \mathbf{Z}$, we set $\tilde{\rho}(x) = p/q+$. If $\rho(x) = p/q$ and $x_{i+q} < x_i + p$ for all $i \in \mathbf{Z}$, we set $\tilde{\rho}(x) = p/q-$. According to the theory of Aubry and Le Daeron, one of these three possibilities always holds for $x \in M$. If ω is a rotation symbol, we let M_ω denote the set of minimal energy configurations of rotation symbol ω . We let $\Phi_\omega = p_0(M_\omega)$, where $p_0(x) = x_0$. If $\omega \in \mathbf{R}$, then Φ_ω is a closed subset of \mathbf{R} . Also $\text{cl } \Phi_{p/q+} = \Phi_{p/q+} \cup \Phi_{p/q}$, $\text{cl } \Phi_{p/q-} = \Phi_{p/q-} \cup \Phi_{p/q}$, where cl means “closure.” These results are due to Aubry and Le Daeron [3]. Bangert [5] has explained them clearly.

This machinery permits us to define “Peierls’s energy barrier” $P_\omega(\xi)$ for a real number ξ and a rotation symbol ω , whose underlying number is in the rotation interval of f . If $\xi \in \text{cl } \Phi_\omega$, we set $P_\omega(\xi) = 0$. Otherwise, we let (a, b) be the complementary interval of $\text{cl } \Phi_\omega$ which contains ξ . By the theory of Aubry and Le Daeron, there exist $x, y \in M_\omega$ such that $x_0 = a$ and $y_0 = b$. Moreover, $y_i > x_i$ for all $i \in \mathbf{Z}$, and $\sum_{i \in A} y_i - x_i \leq 1$, where $A = \mathbf{Z}$ if ω is an irrational number or of the form $p/q+$ or $p/q-$ and $A = \{0, \dots, q-1\}$ if $\omega = p/q$. We set

$$P_\omega(\xi) = \min \left\{ \sum_{i \in A} h(z_i, z_{i+1}) - h(x_i, x_{i+1}) \right\},$$

where z ranges over all configurations such that $x_i \leq z_i \leq y_i$ and $z_0 = \xi$. This was defined and called Peierls’s energy barrier in Aubry, Le Daeron, and André [4]. See also Mather [13], where the basic properties of $P_\omega(\xi)$ are developed. I defined a closely related quantity ΔW_ω in [14], where I showed that as a function of the number ω , this quantity is continuous at irrationals, although it is discontinuous at rational ω , for generic f . The definition of ΔW_ω may be extended to rotation symbols ω , and then the functions $\omega \mapsto \Delta W_\omega$ and

$\omega \mapsto P_\omega(\xi)$ are continuous on the space of rotation symbols. We provide the set of rotation symbols with the topology associated to the obvious order. In this topology, rational numbers are isolated points. The intervals $[p/q+, p/q + \varepsilon)$ in the set of rotation symbols form a basis of neighborhoods of $p/q+$, where ε ranges over all positive numbers. The continuity of these functions follows from [14] or slight extensions of the results of [14]. Its importance derives from the fact I proved in [14] that there is an invariant circle for f of rotation number ω (where ω is irrational) if and only if $\Delta W_\omega = 0$, or equivalently, P_ω vanishes identically.

Recently, I have improved these results, to give moduli of continuity for ΔW_ω or $P_\omega(\xi)$, as functions of ω . It is easy to see that there exists $C > 0$ such that $|P_\omega(\xi) - P_\omega(\xi')| \leq C|\xi - \xi'|$, for all $\xi, \xi' \in \mathbb{R}$. The dependence on ω , however, is more complicated. For $P_\omega(\xi)$, we have

$$|P_{p/q}(\xi) - P_\omega(\xi)| \leq C(q^{-1} + |q\omega - p|),$$

where C depends only on f . Moreover,

$$\begin{aligned} |P_{p/q+}(\xi) - P_\omega(\xi)| &\leq C|q\omega - p|, & \text{if } \omega > p/q, \\ |P_{p/q-}(\xi) - P_\omega(\xi)| &\leq C|q\omega - p|, & \text{if } \omega < p/q. \end{aligned}$$

There are similar estimates for ΔW_ω .

Using these estimates, I have been able to prove that if ω is a Liouville number, then there is a dense set \mathcal{D} in the space of C^∞ monotone twist diffeomorphisms such that a homotopically nontrivial invariant circle of a diffeomorphism in \mathcal{D} has rotation number ω . This is a converse of well-known results in KAM theory. The proof of this result is based on the theorem of Mather [14] that f has an invariant circle of rotation number ω if and only if $\Delta W_\omega = 0$, and this holds if and only if P_ω vanishes identically.

In another direction, I have shown that in a certain sense it is possible to “shadow” minimal energy orbits in a fixed Birkhoff region of instability by local minimal energy orbits. Recall that a minimal energy configuration x is *stationary*, in the sense that $\partial_1 h(x_{i-1}, x_i) + \partial_2 h(x_i, x_{i+1}) = 0$, and therefore if we set $y_i = -\partial_1 h(x_i, x_{i+1})$, we have $f(x_i, y_i) = (x_{i+1}, y_{i+1})$. Thus, to every minimal energy configuration, we may associate an orbit, and we call the resulting orbit a *minimal energy orbit*. Consider two homotopically nontrivial invariant circles which do not intersect, so they bound an annulus. If the annulus which they bound contains no invariant circle, then the region between the circles is called a *Birkhoff region of instability*. A *local minimal energy configuration* x minimizes in the same sense that a minimal energy configuration minimizes, but only for small perturbations of x . Local minimal energy orbits are the orbits corresponding to local minimal energy configurations. Then we have the following result: given a sequence $(\mathcal{O}_i)_{i \in \mathbb{Z}}$ of minimal energy orbits, all in the same Birkhoff region of instability, and numbers $\varepsilon_i > 0$, there is a local minimal energy orbit $\mathcal{O} = (P_j)_{j \in \mathbb{Z}}$ and an increasing sequence $(n_i)_{i \in \mathbb{Z}}$ of integers, such that

$\text{dist.}(P_{n(i)}, O_i) < \varepsilon_i$, i.e., O comes as close as we please to each orbit O_i in turn. Proofs will appear elsewhere.

REFERENCES

1. V. I. Arnold and A. Avez, *Ergodic problems of classical mechanics*, Benjamin, New York, 1968.
2. D. V. Anosov, *Geodesic flows on closed Riemann manifolds with negative curvature*, Proc. Steklov Inst. Math., vol. 90, Amer. Math. Soc., Providence, R.I., 1969.
3. S. Aubry and P. Y. Le Daeron, *The discrete Frenkel-Kontorova model and its extensions*, Phys. D **8** (1983), 381-422.
4. S. Aubry, P. Y. Le Daeron, and G. André, *Classical ground-states of a one-dimensional model for incommensurate structures*, Preprint, 1982.
5. V. Bangert, *Mather sets for twist maps and geodesics on tori*, Preprint, 1968.
6. R. Bowen, *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, Lecture Notes in Math., vol. 470, Springer-Verlag, Berlin-New York, 1975.
7. A. Chenciner, *La dynamique au voisinage d'un point fixe elliptique conservatif: de Poincaré et Birkhoff à Aubry et Mather*, Sémin. Bourbaki, Exp. 622, Astérisque, no. 121/122, Soc. Math. France, Paris, 1985, pp. 147-170.
8. J. Hadamard, *Les surfaces à courbures opposées et leurs lignes géodesiques*, J. Math. **5** (1898), 27-73.
9. M. R. Herman, *Introduction à l'étude des courbes invariantes par les difféomorphismes de l'anneau*, Astérisque, no. 103/104, Soc. Math. France, Paris, 1983.
10. A. N. Kolmogorov, *Théorie générale des systèmes dynamiques et mécanique classique*, Proc. Internat. Congr. Math. (Amsterdam, 1954), Vol. 1, North-Holland, Amsterdam, 1957, pp. 315-333.
11. A. J. Lichtenberg and M. A. Lieberman, *Regular and stochastic motion*, Springer-Verlag, New York, 1983.
12. J. N. Mather, *Existence of quasi-periodic orbits for twist homeomorphisms of the annulus*, Topology **21** (1982), 457-467.
13. —, *More Denjoy minimal sets for area preserving diffeomorphisms*, Comment. Math. Helv. **60** (1985), 508-557.
14. —, *A criterion for the non-existence of invariant circles*, Inst. Hautes Études Sci. Publ. Math. (to appear).
15. J. Moser, *Stable and random motions in dynamical systems*, Ann. of Math. Studies, no. 77, Princeton Univ. Press, Princeton, N.J., and Univ. of Tokyo Press, Tokyo, 1973.
16. —, *Quasi-periodic solutions in the three-body problem*, Bull. Astr. **3** (1968), 53-59.
17. —, *Break-down of stability*, Preprint, Forschungsinst. für Math. ETH Zürich, March, 1985.
18. —, *Recent developments in the theory of Hamiltonian systems*, Preprint, Forschungsinst. für Math. ETH Zürich, September, 1985.
19. Ya. B. Pesin, *Characteristic Lyapunov exponents and smooth ergodic theory*, Russian Math. Surveys **32** (1977), 55-114.
20. H. Poincaré, *Les méthodes nouvelles de la mécanique céleste*, Paris, 1893.
21. S. Smale, *Differentiable dynamical systems*, Bull. Amer. Math. Soc. **73** (1967), 747-817.