

PERIODIC, NONPERIODIC AND IRREGULAR MOTIONS IN A HAMILTONIAN SYSTEM

PHILIP HOLMES

ABSTRACT. We study orbit structures in the neighborhoods of homoclinic orbits connecting saddlepoints with real eigenvalues in certain four-dimensional Hamiltonian systems. Under suitable hypotheses we prove the existence of dense non-periodic orbits and of periodic orbits of infinitely many periods.

1. **Introduction.** Recently Markus and Meyer [12] have proved the generic existence of solenoids in Hamiltonian systems of dimension ≥ 4 . The existence of solenoids implies that the set of orbits of the system is extremely complicated and in particular, that it contains orbits of arbitrary periods in addition to nonperiodic recurrent motions. The results of [12] are general and somewhat abstract. Devaney [4-7], has obtained more specific results on Hamiltonian systems possessing homoclinic orbits connecting hyperbolic saddle points, cf. Silnikov's studies of homoclinic orbits in general (non-Hamiltonian) system [15, 16].

Here we discuss a specific situation applicable to certain mechanical problems, e.g., the erratic behavior of a spherical pendulum swinging over arrays of two or three magnets. We shall study the structure of orbits in the neighborhood of solutions doubly asymptotic to a saddle point. Such doubly asymptotic solutions are called homoclinic orbits and occur generically in Hamiltonian systems [4, 5]. Their occurrence in non-Hamiltonian differential equations is, of course, non-generic, cf., [2].

In a four dimensional system there are essentially three structurally stable types of saddle point possible: the saddle-center, with eigenvalues $\{\pm\alpha, \pm i\beta|\alpha, \beta > 0\}$; the "saddle-saddle", with eigenvalues $\{\pm l, \pm k|l, k > 0\}$ and the saddle-focus, with eigenvalues $\{\pm(\alpha \pm i\beta) | \alpha, \beta > 0\}$. The saddle focus was treated in [4]. Here we consider the saddle-saddle and briefly mention the saddle-center. Only the saddle-focus and saddle-saddle are hyperbolic fixed points, since the saddle-center has a pair of purely imaginary eigenvalues.

Since the saddle-saddle is a hyperbolic critical point with stable and unstable manifolds M^s, M^u each of dimension 2, transverse homoclinic orbits can occur generically at the (transverse) intersection of $M^s(o)$ and $M^u(o)$ within the three dimensional energy surface $H^{-1}(0)$, where o is

Received by the editors on March 15, 1978, and in revised form on October 20, 1978.

Copyright © 1980 Rocky Mountain Mathematics Consortium

the saddle-saddle at energy level 0, cf. [4]. However, in the case of the saddle-center, $M^s(o)$ and $M^u(o)$ are each of dimension one and would not appear to intersect generically within $H^{-1}(0)$. For a homoclinic orbit to exist, $M^s(o)$ and $M^u(o)$ would have to be identified. We return to this case in §4 after discussing the saddle-saddle case.

For general background on Hamiltonian system in this context, see [3, 13].

2. The main result: the saddle-saddle. We consider the case of a saddle point o with purely real eigenvalues $\pm l, \pm k; l > k$. For comments on the case $l = k$, see §5. Our main assumptions here imply the existence of an inherent symmetry.

ASSUMPTION 1. *There are two transverse homoclinic orbits γ_a, γ_b at o ; γ_a leaves o in the first quadrant of $M_{loc}^u(o)$ and reenters in the second quadrant of $M_{loc}^s(o)$; γ_b leaves o in the third quadrant of $M_{loc}^u(o)$ and reenters in the fourth quadrant of $M_{loc}^s(o)$. (See Fig. 1).*

We can equally well replace ‘first . . . second; . . . third . . . fourth’ by ‘second . . . first; . . . fourth . . . third’ and there are other possible combinations. However, the cases ‘first . . . first; . . . third . . . third’ etc. do not appear amenable to the present methods unless we assume the presence of further orbits. We now develop our notation and can then state our assumption more precisely.

The Hamiltonian function for the system X_H at o may be written in the form

$$(2.1) \quad H(\mathbf{x}, \mathbf{y}) = lx_1y_1 + kx_2y_2 + \text{higher order terms},$$

and the equation of motion, $\dot{\mathbf{x}} = \partial H/\partial \mathbf{y}, \dot{\mathbf{y}} = -\partial H/\partial \mathbf{x}$ becomes

$$(2.2) \quad (\dot{x}_1 \ \dot{x}_2, \dot{y}_1, \dot{y}_2) = (lx_1, kx_2, -ly_1, -ky_2) + O(|\mathbf{x}|^2, |\mathbf{y}|^2).$$

With suitable assumptions on the higher order terms (cf. [6]), we can

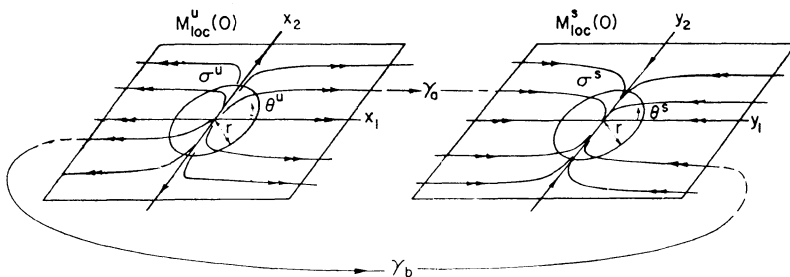


FIG. 1. Local stable and unstable manifolds.

assume that the $x_1 - x_2$ plane ($\mathbf{y} = 0$) is the local unstable manifold $M_{loc}^u(\mathbf{o})$ and the $y_1 - y_2$ plane ($\mathbf{x} = 0$) the local stable manifold $M_{loc}^s(\mathbf{o})$.

We now define certain submanifolds in the neighborhood of \mathbf{o} [4, 15]:

$$\Sigma^u = \{(\mathbf{x}, \mathbf{y}) \in \mathbf{R}^4 \mid |\mathbf{x}| = r; |\mathbf{y}| \leq \rho_u\}$$

$$\Sigma^s = \{(\mathbf{x}, \mathbf{y}) \in \mathbf{R}^4 \mid |\mathbf{x}| \leq \rho_s; |\mathbf{y}| = r\}.$$

For small enough r and $\rho_u, \rho_s < r$, the three dimensional manifolds Σ^u, Σ^s are transverse to the flow of (2.2). We shall also need the two dimensional manifolds

$$\Sigma_\varepsilon^u = \Sigma^u \cap H^{-1}(\varepsilon); \Sigma_\varepsilon^s = \Sigma^s \cap H^{-1}(\varepsilon)$$

(although we shall be particularly concerned with the case $\varepsilon = 0$), and the one dimensional circles

$$\sigma^u = \Sigma_0^u \cap M_{loc}^u(\mathbf{o}); \sigma^s = \Sigma_0^s \cap M_{loc}^s(\mathbf{o}).$$

σ^u and σ^s are the center circles of the solid tori Σ^u and Σ^s .

We now let γ_a and γ_b intersect σ^u and σ^s at the points $p_a, p_b; q_a, q_b$ respectively, and for the Poincaré map construction below, we consider neighborhoods of these points as follows

$$D_a^u, D_b^u \subset \Sigma_0^u; D_a^s, D_b^s \subset \Sigma_0^s.$$

These cross sections belong to the energy surface $H^{-1}(0)$ and are transverse to the flow; γ_a pierces D_a^u and D_a^s , γ_b pierces D_b^u and D_b^s . We can immediately define two (Poincaré) maps $\Phi_a: D_a^u \rightarrow D_a^s, \Phi_b: D_b^u \rightarrow D_b^s$ by following orbits forward in time until their first intersections with D_a^s or D_b^s . Since D_a^u, D_b^u , etc., are small neighborhoods of γ_a and γ_b , Φ_a and Φ_b are local diffeomorphisms by the implicit function theorem. The situation so far is illustrated in figure 1, where it is understood that $M^u(\mathbf{o})$ and $M^s(\mathbf{o})$ intersect transversely along γ_a and γ_b (cf. figure 4).

A little calculation shows that the submanifolds $\Sigma^u, \Sigma_0^u, \Sigma^s, \Sigma_0^s$ take the forms shown in figure 2. Here we have parametrized σ^u and σ^s by θ^u, θ^s and ‘unrolled’ the tori Σ^u, Σ^s .

We must now show that points near q_a, q_b in D_a^s, D_b^s are mapped into $D_a^u \cup D_b^u$ by the action of the flow. Since Σ_0^s and Σ_0^u are near \mathbf{o} , the flow is determined by equation (2.2) and the higher order terms may be neglected. Specifically, we have

$$(2.3) \quad x_1(t) = x_{10}e^{t}; x_2(t) = x_{20}e^{kt}; y_1(t) = y_{10}e^{-t}; y_2(t) = y_{20}e^{-kt}.$$

We shall consider what happens to thin strips $S_a \subset D_a^s \setminus \sigma^s, S_b \subset D_b^s \setminus \sigma^s$, transverse to σ^s . In particular, we shall study their images in $\Sigma_0^s \setminus \sigma^u$ under the flow (2.3). To do this we define two further maps

$$\Psi_a: D_a^s \setminus \sigma^s \rightarrow \Sigma_0^u \setminus \sigma^u; \Psi_b: D_b^s \setminus \sigma^s \rightarrow \Sigma_0^u \setminus \sigma^u,$$

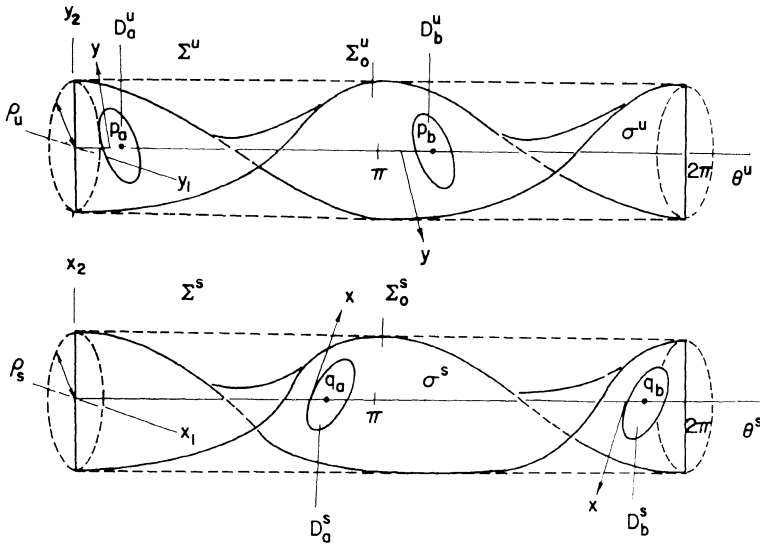


FIG. 2. The manifolds Σ^u , Σ_o^u , Σ^s and Σ_o^s .

noting that points in σ^s, σ^u approach o as $t \rightarrow +\infty$ and $-\infty$ respectively and that the maps are therefore not defined on those circles. In the subsequent analysis we shall be concerned with the maps $\Phi_a \circ \Psi_a, \Phi_b \circ \Psi_b, \Phi_b \circ \Psi_a$ and $\Phi_a \circ \Psi_b$. In order that these maps make sense, we must first show that, under suitable assumptions, Ψ_a and Ψ_b map points of $D_a^s \setminus \sigma^s$ and $D_b^s \setminus \sigma^s$ into $(D_a^u \cup D_b^u) \setminus \sigma^u$. In Devaney's study this follows from the assumption that o is a hyperbolic *saddle-focus*, since the winding actions associated with the nonzero imaginary parts of the eigenvalues guarantees that the images of strips in D^s wind infinitely often around Σ_o^u . Here, however, we must make the following assumption.

ASSUMPTION 2. $l > k$ and γ_a (resp. γ_b) leaves o in $M_{loc}^u(o)$ such that $p_a \in \sigma^u$ (resp. $p_b \in \sigma^u$) lies at an angle $\theta_a^u \in (0, \pi/2)$ (resp. $\theta_b^u \in (\pi, 3\pi/2)$) and enters o in $M_{loc}^s(o)$ such that $q_a \in \sigma^s$ (resp. $q_b \in \sigma^s$) lies at an angle $\theta_a^s \in (\pi/2, \pi)$ (resp. $\theta_b^s \in (3\pi/2, 2\pi)$); moreover,

$$|\arctan \theta^u - \arctan \theta^s| < -\delta + (l/k) \exp[(k/l - 1) \ln(\rho_s/r)], \text{ for } \delta > 0, \text{ small.}$$

We now state the main result of this section.

THEOREM A. Under the above assumptions, the Hamiltonian system X_H possesses an infinite set of hyperbolic periodic orbits of all periods and dense hyperbolic non-periodic orbits within the energy surface $H^{-1}(0)$. All the orbits are of unstable (saddle) type and each orbit \mathcal{O} can be denoted by a bi-infinite sequence of two symbols. $\dots, i_{-n}, \dots, i_{-2}, i_{-1}, i_0, i_1, i_2, \dots$

i_n, \dots where i_n takes the symbol a or b depending on whether the n 'th pass of \mathcal{O} lies in the tubular neighborhood of γ_a or γ_b .

3. Proof of theorem A. We first study Ψ_a and Ψ_b and establish the following lemma.

LEMMA 1. $\Psi_a(S_a)$ consists of two components $\Psi_a(S_a^+)$ and $\Psi_a(S_a^-)$ lying across D_a^u and D_b^u respectively, "parallel" to σ^u . $\Psi_b(S_b)$ likewise consists of two components $\Psi_b(S_b^+)$ and $\Psi_b(S_b^-)$ lying across D_a^u and D_b^u respectively, "parallel" to σ^u . (cf. Figure 3).

PROOF. To clarify matters, we choose coordinates in Σ_0^s and Σ_0^u as indicated in figure 2. In figure 3 we show these two-dimensional strips "unwound". Note that we distinguish the components $S_{a,b}^+, S_{a,b}^-$ of $S_{a,b}$ lying "above" and "below" σ_s . Here $|x| = (x_1^2 + x_2^2)^{1/2}$ and $|y| = (y_1^2 + y_2^2)^{1/2}$ and $+x$ and $+y$ are oriented on Σ_0^s, Σ_0^u as indicated in figure 2.

The proof proceeds in several stages. We first find conditions on r, ρ_u and ρ_s such that the images of S_a and S_b lie within Σ_0^u . To do this, we consider a curve $c \subset S_a$ normal to σ^s at q_a and show that $\Psi_a(c)$ lies within Σ_0^u ; that is, we require that $y_1^2(t) + y_2^2(t) < \rho_u^2$.

But

$$y_1^2(t) + y_2^2(t) = y_{10}^2 e^{-2t} + y_{20}^2 e^{-2kt} < (y_{10}^2 + y_{20}^2) e^{-2kt}.$$

Since $y_{10}^2 + y_{20}^2 = r^2 (c \subset \Sigma_0^s)$, our requirement becomes

$$(2.4) \quad r^2 e^{-2kt} < \rho_u^2 \text{ or } r/\rho_u < e^{kt}.$$

We next estimate t , or more particularly, the minimum time required for points of c to be mapped to Σ_0^u . This evidently occurs for points at the extrema of c , where $x_{10}^2 + x_{20}^2 = \rho_s^2$. Thus we obtain the equation

$$(2.5) \quad x_1^2(t) + x_2^2(t) = x_{10}^2 e^{2t} + x_{20}^2 e^{2kt} = r^2.$$

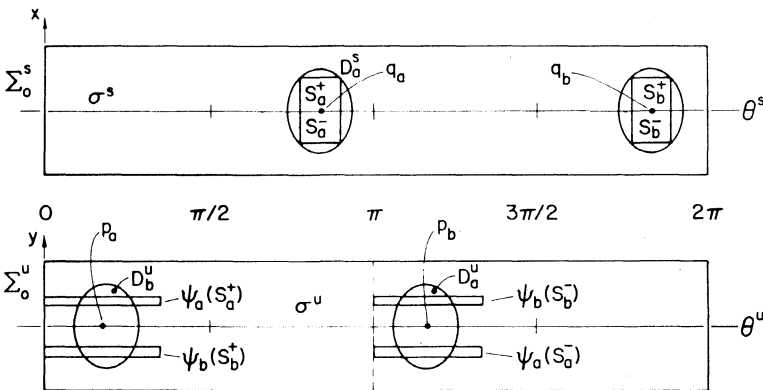


FIG. 3. The maps Ψ_a and Ψ_b .

Using the Hamiltonian (2.1) ($lx_1y_1 = -kx_2y_2$ locally) we can express x_{10}, x_{20} in terms of ρ_s and (2.5) becomes

$$r^2 = \rho_s^2 \left(\frac{e^{2lt}}{1+a^2} + \frac{e^{2kt}}{1+a^{-2}} \right); \quad a = \left(\frac{ly_1}{ky_2} \right).$$

Thus $e^{kt} < r/\rho_s < e^{lt}$, and we can estimate the minimum time for the flow of X_H to carry points of c into Σ_0^u as

$$(2.6) \quad (1/l) \ln(r/\rho_s) < t_{\min} < (1/k) \ln(r/\rho_s).$$

Using (2.6) in (2.4) we obtain, after some manipulation

$$(2.7) \quad \rho_u^l / \rho_s^k > r^{l-k}.$$

(2.7) expresses the condition for Ψ_a (and hence Ψ_b) to map c into Σ_0^u .

We next find how far around σ^u the image of c extends. Clearly, points of c arbitrarily close to σ^s take arbitrarily long to arrive at Σ_0^u and hence, since $x_{10}e^{lt} \gg x_{20}e^{kt}$ for x_{10}, x_{20} fixed $\neq 0$ and t large, such points arrive in Σ_0^u with angles θ^u in the first and third quadrants close to 0 and π . We now estimate the angles to which the extrema of c are mapped. Using the initial values x_{10}, x_{20} derived above, we have

$$(2.8) \quad \frac{x_2}{x_1} = \sqrt{\frac{1+a^2}{1+a^{-2}}} e^{(k-l)t} = \frac{ly_1}{ky_2} e^{(k-l)t}.$$

Hence, taking $t_{\min} < (l/k)\ln(r/\rho_s)$, we have,

$$\left| \frac{x_2}{x_1} \frac{y_2}{y_1} \right| > (l/k) \exp[(1 - l/k) \ln(r/\rho_s)],$$

or

$$(2.9) \quad |\arctan \theta^u \arctan \theta^s| > (l/k) \exp[(1 - l/k) \ln(r/\rho_s)].$$

Thus, since the images of points of c close to σ^s lie at $\theta^u \approx 0$ or π , by continuity each component of the image of c will extend around Σ_0^u from 0 (resp π) to an angle $\theta^u(\theta^u + \pi)$ given by (2.9) and will hence intersect transversely a foliation of Σ_0^u "normal" to σ^u . To ensure that the images of c intersect curves t_a, t_b "normal" to σ^u at p_a and p_b , we must therefore ensure that $p_a, p_b; q_a, q_b$ lie in σ^u, σ^s at angles θ^u, θ^s such that

$$(2.10) \quad |\arctan \theta^u \arctan \theta^s| < -\delta + (l/k) \exp[(1 - l/k) \ln(r/\rho_s)],$$

for some small $\delta > 1$. The presence of δ ensures that we can widen the curves c, t into strips S, T in σ^s and σ^u and preserve the intersections. If (2.7) and (2.10) are satisfied, we have the situation of figure 3, where S_a and S_b are sufficiently narrow strips and D_a^u, D_b^u are chosen sufficiently narrow in angular extent. This completes the proof of Lemma 1.

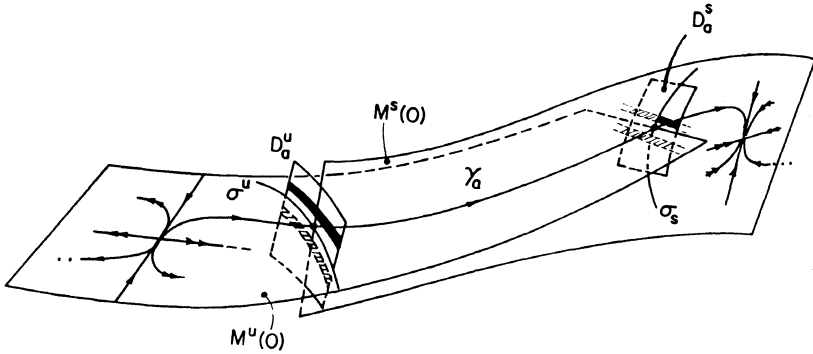


FIG. 4. The map Φ_* .

We next consider the maps $\Phi_a: D_a^u \rightarrow D_a^s$ and $\Phi_b: D_b^u \rightarrow D_b^s$ in more detail. We first recall that, for transverse homoclinic orbits such as γ_a , and γ_b , $M^s(o)$ and $M^u(o)$ intersect transversely and only along γ_a and γ_b in their tubular neighborhoods. The global structure near γ_a , for instance, thus appears as in figure 4. The intersection lies within the energy surface $H^{-1}(0)$, which locally has the structure of \mathbf{R}^3 . Since $\Psi_a(S_a^+)$ and $\Psi_b(S_b^+)$ lie "parallel" to σ^u their images under Φ_a must lie transverse to σ^s and must intersect and overlap D_a^s . A similar argument holds for $\Psi_a(S_a^-)$ and $\Psi_b(S_b^-)$.

We thus establish the following lemma.

LEMMA 2. $\Phi_a: D_a^u \rightarrow D_a^s$ and $\Phi_b: D_b^u \rightarrow D_b^s$ map strips parallel to σ^u into strips transverse to σ^s .

Note that the contraction in $M^s(o)$ and expansion in $M^u(o)$ ensures that the strips become even narrower and closer to $M^u(o)$, so that the images of $\Psi_a(S_a)$ and $\Psi_b(S_b)$ intersect S_a and S_b . We summarize the situation in figure 5, where the map is now the full Poincaré map P given by $P = \Phi_a \circ \Psi_a$ (for points in S_a^+); $\Phi_a \circ \Psi_b(S_b^+)$; $\Phi_b \circ \Psi_a(S_a^-)$ and $\Phi_b \circ \Psi_b(S_b^-)$. The exact orientations of the images $P(S_a^\pm)$ and $P(S_b^\pm)$ are unimportant, but note that they extend "vertically" across S_a and S_b and lie within them.

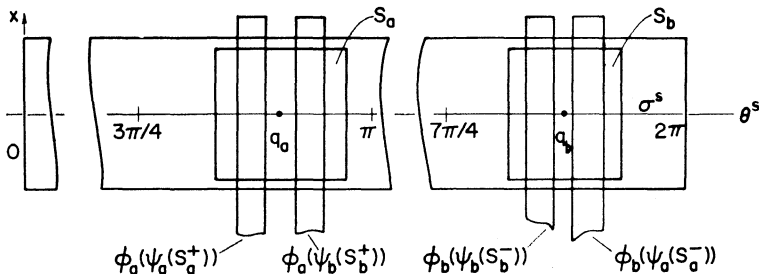


FIG. 5. The Poincaré map.

The next images of S_a and S_b , under P^2 , will consist of eight thin strips lying in pairs within the four strips of $(P(S_a) \cup P(S_b)) \cap (S_a \cup S_b)$. Due to this splitting and doubling, after n iterates the images of S_a and S_b each consist of 2^n vertical strips lying across D_a^s and D_b^s . Half of the images of S_a intersect D_a^s and half intersect D_b^s ; similarly for the images of S_b . The mapping clearly has much in common with the Horseshoe map of Smale [17, 18]. To complete our analysis we study it in some detail, following the symbolic dynamic methods originally used by Smale (see Chillingworth [2] for a good introduction and a summary).

The Poincaré map P. We wish to establish that, under P , there is an hyperbolic invariant set $A \subset D_a^s \cup D_b^s$ and then to study the structure of this set. First consider the sets

$$\begin{aligned} S_1 &= (P(S_a) \cup P(S_b)) \cap (S_a \cup S_b) \\ &= (P(S_a) \cap S_a) \cup (P(S_a) \cap S_b) \cup (P(S_b) \cap S_a) \cup (P(S_b) \cap S_b) \end{aligned}$$

and

$$\begin{aligned} S_{-1} &= (P^{-1}(S_a) \cup P^{-1}(S_b)) \cap (S_a \cup S_b) \\ &= (P^{-1}(S_a) \cap S_a) \cup (P^{-1}(S_b) \cap S_a) \cup (P^{-1}(S_a) \cap S_b) \\ &\quad \cup (P^{-1}(S_b) \cap S_b). \end{aligned}$$

S_1 consists of four vertical and S_{-1} of four horizontal strips. We extend this in the obvious way to considering the sets S_{+n} and S_{-n} consisting of 2^{n+1} vertical and horizontal strips each, and then to the sets $S_{+\infty}$ and $S_{-\infty}$ formed under infinitely many iterates. Evidently the invariant set $A \subset S_a \cup S_b$ is given by $A = S_{+\infty} \cap S_{-\infty}$, is divided into two clearly distinct components in S_a and S_b , and each component has the structure of the product of two Cantor sets, or, equivalently, of a single Cantor set (cf [2], p. 233).

We now prove that there are orbits of all periods in S_a and S_b . Since $S_a \subset D_a^s$ and $S_b \subset D_b^s$ are disjoint, we cannot use Smale's methods directly, but must proceed in stages.

Orbits of period one. Consider S_a only, and the horizontal and vertical strips $S_a \cap P^{-1}(S_a)$ and $P(S_a) \cap S_a$ lying across it. Continue to iterate infinitely many times backward and forward to obtain the strips denoted by the semi-infinite sequences

$$S_a \cap P^{-1}(S_a) \cap P^{-2}(S_a) \cap \dots \cap P^{-m}(S_a) \cap \dots$$

and

$$\dots \cap P^n(S_a) \cap \dots \cap P^2(S_a) \cap P(S_a) \cap S_a.$$

These infinitesimally thin strips clearly intersect at just one point in

$P(S_a) \cap S_a \cap P^{-1}(S_a)$. We thus have a point of period one in $S_a^+ \subset S_a$. By similar arguments, there is a second period one point in $S_b^- \subset S_b$. Both fixed points are clearly hyperbolic, since we have exponential expansion in the vertical direction and contraction in the horizontal direction in their neighborhoods. For more details on hyperbolicity, see [2, 17, 18].

Orbits of all even periods. Consider S_a and its images under two iterates forward and backward of P . There are two vertical and two horizontal strips; note that one each of these has been mapped from S_a to S_b and then back to S_a . The map P^2 restricted to S_a evidently behaves exactly as does the Smale Horseshoe map (or, depending on the orientations of $P^{-2}(S_a) \cap P^{-1}(S_b) \cap S_a$ and $P^2(S_a) \cap P(S_b) \cap S_a$, as a variant of it, cf [18], p. 772). A similar argument applies to the map P^2 restricted to S_b . We can thus conclude that P^2 has points of all periods plus nonperiodic orbits dense in its invariant set \mathcal{A}' . Of course, $\mathcal{A}' \subset \mathcal{A}$, the invariant set of P .

Orbits of period three and other odd periods. Consider P^3 restricted to S_a . The image and inverse images of S_a under P^3 each consist of four strips lying across S_a . If we continue the iteration process as in the case for orbits of period one above, it is clear that there will be four fixed points of P^3 lying in S_a . (One of these, the top left, is the fixed point of P found above). We thus have points of period 3: again a similar argument applies to P^3 restricted to S_b . One can continue by induction in this manner to show the existence of points of all periods. P^3 restricted to S_a is, in fact, a map of Horseshoe type corresponding to a shift on a four symbol sequence, since $P^3(S_a) \cap S_a$ and $P^{-3}(S_a) \cap S_a$ each have four components.

Note that the labelling of the strips denotes the histories of the orbits: for example, the period 3 point lying in

$$P^3(S_a) \cap P^2(S_b) \cap P(S_b) \cap S_a \cap P^{-1}(S_b) \cap P^{-2}(S_b) \cap P^{-3}(S_a)$$

corresponds to a periodic orbit of the differential equation which passes once close to γ_a , then twice close to γ_b , then once close to γ_a , etc. We can denote this orbit as . . . abbabb There is a second period 3 orbit . . . baabaaba In this manner one can denumerate all periodic and nonperiodic orbits by doubly infinite sequences of two symbols. This completes the proof of Theorem A.

REMARK 1. With appropriate modifications to the assumptions and the proof, one can extend the above results to the case of (multiple) heteroclinic orbits connecting two or more saddle points. (cf. Devaney [4]). The symbol sequence would need to include as many symbols as there were heteroclinic orbits.

REMARK 2. The period of the orbits referred to above is a topological

concept in the sense that period n denotes n circuits or passes in the neighborhoods of γ_a and γ_b before returning to the starting point on D_a^s or D_b^s . The temporal period, which is of interest in applications, can clearly vary very much for orbits close to homoclinic orbits. In particular, we have the following theorem.

THEOREM B. [5, 6]. *Let γ be a non-degenerate homoclinic orbit to a hyperbolic equilibrium point. Then there exists a one-parameter family of closed orbits γ_τ which converges to γ as the parameter τ approaches infinity. Moreover, τ may be chosen to be the period of the closed orbit.*

We therefore can expect some of the periodic motions to have long temporal periods, even though their topological periods might be relatively low.

REMARK 3. Since the hyperbolic invariant set $A \subset D_a^s \cup D_b^s$ is homeomorphic to a Cantor set, orbits of P near γ_a and γ_b will depend on initial conditions in a sensitive manner: that is—it is impossible to predict the long term behaviour of orbits of P starting near points of A . Moreover, the divergent fates of orbits starting nearby manifest themselves in a dramatic manner in that they alternately leave the neighborhood of o in distinct directions: either close to γ_a or close to γ_b . Such behaviour would be clearly recognizable as an irregular motion in physical configuration space. Note that the periodic orbits in A are all of saddle type, since there is a vertical expansion and a horizontal contraction associated with each fixed point of P^n .

REMARK 4. To see that the conditions on the homoclinic orbits of Lemma 1 are not excessively strict, consider the case $l = 4, k = 1$. We set $r = 1$ and $\rho_s = 1 - \varepsilon; 0 < \varepsilon < 1$. Condition (2. 7) then becomes $\rho_u > (1 - \varepsilon)^{k/l}$. Since $1 - \varepsilon < 1$ we can choose $(1 - \varepsilon)^{k/l} < \rho_u < 1$. The conditions on θ^u, θ^s are then

$$|\arctan \theta^u \arctan \theta^s| < -\delta + (l/k) \exp[(k/l - 1) \ln(1 - \varepsilon)].$$

Setting $l = 4, k = 1, \theta^s = \pi/4$ (resp $5\pi/4$), $\varepsilon = 0.1, \delta = 0.2$, we have

$$|\arctan \theta^u| < 4 \exp(-0.0792) - 0.2 \approx 3.28,$$

or $\theta^u < 73.1^\circ$ (resp 253.1°). Note, however, that our assumptions on the existence of two homoclinic orbits leaving and arriving at appropriate angles are essential. Devaney [7] has an example of a completely integrable system with homoclinic orbits which do not satisfy such a condition. His example has no chaotic motions such as those detected here.

4. The Saddle-center. We now briefly consider the case of a saddle point

with real and imaginary eigenvalues $\pm\alpha, \pm i\beta$. Locally the Hamiltonian function may be written

$$(4.1) \quad H(\mathbf{x}, \mathbf{y}) = \alpha x_1 y_1 + \beta x_2^2/2 + \beta y_2^2/2 + \text{higher order terms},$$

and the equation of the system as

$$(4.2) \quad (\dot{x}_1, \dot{x}_2, \dot{y}_1, \dot{y}_2) = (\alpha x_1, \beta y_2, -\alpha y_1, -\beta x_2) + O(|\mathbf{x}|^2, |\mathbf{y}|^2),$$

The stable and unstable manifolds $M^s(\mathbf{o})$ and $M^u(\mathbf{o})$ are each one dimensional and tangent to the y_1 and x_1 axes respectively, and there is also locally a center manifold $M_{loc}^c(\mathbf{o})$, tangent to the $x_2 - y_2$ plane.

Unlike the cases of purely real and strictly complex eigenvalues, homoclinic orbits to center-saddles do not appear to be generic, since the one dimensional manifolds $M^u(\mathbf{o})$ and $M^s(\mathbf{o})$ would not normally intersect within $H^{-1}(\mathbf{0})$ (they would, in fact, have to be identical). However, if we assume rather special symmetries, we can see that there is a situation in which $M^u(\mathbf{o}) \equiv M^s(\mathbf{o})$. To picture this, we suppress one dimension of $M^c(\mathbf{o})$ and represent it by a single (polar) coordinate $z = (x_2^2 + y_2^2)^{1/2}$. In x_1, y_1, z space the energy surface $H^{-1}(\mathbf{0})$ is thus two dimensional and, under suitable assumptions on the existence of a pair of homoclinic orbits as in §2, takes the form shown in figure 6.

Clearly $H^{-1}(\mathbf{0})$ falls into two distinct components which meet only at \mathbf{o} . There can thus be no orbits passing alternately between these components and thus no chaotic behaviour of the type discussed in §3. Within the separate components of $H^{-1}(\mathbf{0})$, however, one probably does have sets of periodic and non-periodic motions close to γ_a and γ_b .

For a definite example, consider the two degree of freedom system with Hamiltonian

$$(4.3) \quad H(\mathbf{x}, \mathbf{y}) = u^2/2 + v^2/2 - w^2/2 + z^2/2 + w^4/4 + \delta \tilde{H}(w, z).$$

For $\delta = 0$, since there are two independent first integrals $H_1 = u^2/2 -$

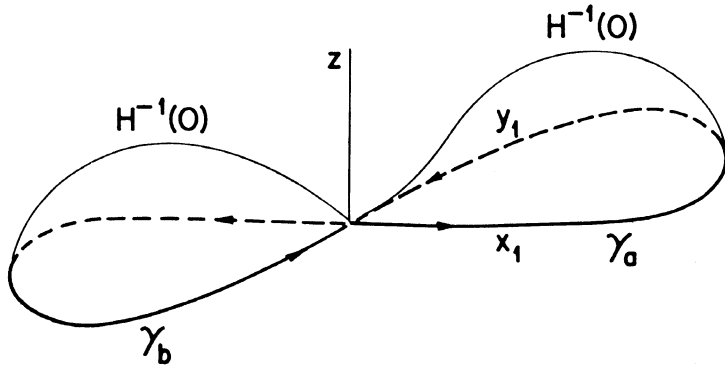


FIG. 6. The energy surface $H^{-1}(\mathbf{0})$ for the the saddle-center.

$w^2/2 + w^4/4$ and $H_2 = v^2/2 + z^2/2$, the behaviour of (4.3) is considerably simpler than that of the systems discussed above. In particular, since the system decouples into two independent two dimensional (single degree of freedom) oscillators and is completely integrable, there can be no orbits corresponding to shift automorphisms of the type described in §3 (cf. Devaney [4–6]). However, for $\delta \neq 0$ the oscillators are coupled and more complex behaviour can arise within the two components of $H^{-1}(0)$. The forthcoming survey by Churchill, et al., [3] contains details of the saddle center case.

5. Passage from saddle-saddle to saddle-focus: bifurcations. In a number of physical problems, such as the flutter studies of [9, 14] the eigenvalues of a fixed point evolve under the action of an external parameter in such a way that it is transformed from a saddle-saddle to a saddle-focus. At the transition there are two pairs of eigenvalues equal in magnitude but opposite in sign ($-\alpha, -\alpha; +\alpha, +\alpha$).

To see how the homoclinic orbit structure of the saddle-saddle changes to the saddle-focus structure studied in [4, 6] consider the system with Hamiltonian

$$(5.1) \quad H(\mathbf{x}, \mathbf{y}) = x_2 y_1 - x_1 y_2 + 2\delta x_2 y_2 + \text{higher order terms},$$

where $\delta > 0$ is a parameter. The equations of motion are

$$(5.2) \quad (\dot{x}_1, \dot{x}_2, \dot{y}_1, \dot{y}_2) = (x_2, -x_1 + 2\delta x_2, y_2, -y_1 - 2\delta y_2) \\ + O(|\mathbf{x}|^2, |\mathbf{y}|^2)$$

and the eigenvalues of the saddle-point \mathbf{o} at the origin are $\delta \pm (\delta^2 - 1)^{1/2}$, $-\delta \pm (\delta^2 - 1)^{1/2}$. For $\delta > 1$ we have the saddle-saddle case, for $\delta = 1$ the critical case and for $\delta < 1$ the saddle-focus studied by Devaney. The system studied here is not in normal form, but it has the advantage that we can consider the effect of changing the parameter δ without an accompanying change of coordinates at the critical case. Note that, as in §2–3, the \mathbf{x} -plane is locally the unstable manifold $M_{\text{loc}}^u(\mathbf{o})$ and the \mathbf{y} -plane the stable manifold $M_{\text{loc}}^s(\mathbf{o})$. In the critical case the flow restricted to either $M_{\text{loc}}^u(\mathbf{o})$ or $M_{\text{loc}}^s(\mathbf{o})$ takes the form of an improper node (cf [8], p. 94); by analogy we call the four dimensional saddle occurring in this case an improper saddle.

In [4] (Lemma, p. 434) it is proven that a smooth curve $S \subset \Sigma_0^s$ intersecting σ^s transversely at the point q (at which the homoclinic orbit γ enters) is mapped under the flow such that the image of $S - q$ spirals infinitely often around Σ_0^u . The image of a small vertical strip containing $S \subset \Sigma_0^s$ must therefore intersect itself a countable number of times, since the map $\Phi: D^u \rightarrow D^s$, where $D^u \subset \Sigma_0^u$, $D^s \subset \Sigma_0^s$ are neighborhoods of the points where γ pierces Σ_0^u and Σ_0^s , is a local diffeomorphism. In the saddle focus

case, then, there is no need for the restrictive assumption (2) necessary in §2 above. It is interesting to see how the rigour of this assumption can be relaxed as the saddle-saddle approaches the improper saddle.

We define the submanifolds $\Sigma^u, \Sigma^s, \Sigma_\delta^u, \Sigma_\delta^s, \sigma^u, \sigma^s$ and the maps $\Phi_a, \Phi_b, \Psi_a, \Psi_b$ in a manner similar to that of §2, and such that $\Sigma_\delta^u, \Sigma_\delta^s$ are always transverse to the flow. We similarly postulate the existence of two homoclinic orbits γ_a and γ_b . We will see how the hypothesis on the dispositions of $\gamma_{a,b} \cap \sigma^u$ and $\gamma_{a,b} \cap \sigma^s$ can be relaxed as $\delta \geq 1$ approaches 1 such that images of points near D_a^s and D_b^s still lie in $D_a^s \cup D_b^s$.

For $\delta \gg 1$ the situation is essentially as in §2-3, with the x_1 and x_2 and y_1 and y_2 axes interchanged. As δ decreases the eigenvectors in both $M_{loc}^s(o)$ and $M_{loc}^u(o)$ rotate so that two "quadrants" increase in size and two decrease until at $\delta = 1$ the improper saddle situation occurs; cf. figure 7.

Some straightforward calculations using (5.1) and (5.2) show that, if γ_a leaves o in quadrant $A^u \subset M_{loc}^u(o)$ (resp. C^u) and enters in quadrant $A^s \subset M_{loc}^s(o)$ (resp. C^s), then the angular extent in Σ_δ^u of the images of strips in Σ_δ^s increases as δ approaches 1. Thus we can successively relax our conditions on the dispositions of γ_a and γ_b in $M_{loc}^u(o)$ and $M_{loc}^s(o)$.

As δ passes through 1, a dramatic change takes place. At $\delta = 1$ the sections B^u, D^u, B^s and D^s have shrunk to the positive and negative 45° halflines and every strip transverse to $\sigma^s \subset \Sigma_\delta^s$ is mapped under the flow in such a way that its images extend almost all the way around Σ . Each strip is, of course, split into two components just as in §3: one component lies over A^u and one over C^u . For $\delta < 1$, however, each of these components now winds repeatedly around Σ_δ^u as described in [4].

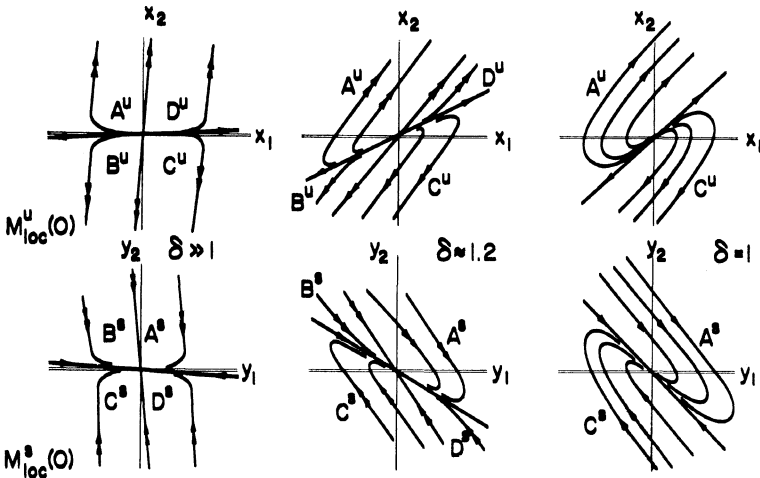


FIG. 7. Transition from saddle-saddle to improper saddle.

Since orbits in the neighborhood of a homoclinic orbit to a saddle-focus are related to a shift on symbol sequence S_∞ with a countable (infinite) alphabet, while the saddle-saddle case corresponds to a finite alphabet, as δ passes through 1 the Poincaré map P undergoes a bifurcation in which infinitely many new points of all periods, in addition to infinitely many dense nonperiodic orbits are created. The situation is evidently quite complicated and would repay further study.

Double and single homoclinic orbits. In the saddle-saddle and improper saddle cases it is necessary to postulate the existence of (at least) two homoclinic orbits at o , while only one is necessary for the saddle-focus. However, if there is only one such orbit, γ , then, since the recurrent behaviour is confined to a (small) tubular neighborhood of γ , its physical manifestations will not be of great interest. All orbits, whatever their symbol sequences, will follow much the same paths and, within the limits of physical measurement, may even appear to be identical. However, if we again have two distinct homoclinic orbits γ_a and γ_b , then by methods analogous to those of Devaney and those used in §3 above we can prove the existence of a countable number of orbits of each period. Of the orbits of period 3, for example, half will have the sequence . . . aabaabaab . . . and half . . . bbabbabba . . . , in the terms of §3. (Recall that in the saddle-saddle case there is precisely one orbit with sequence . . . aabaab . . . and one with sequence . . . bbabba . . .). If γ_a and γ_b are well separated in configuration space then these two families of orbits will be quite distinct, as will the various sets of orbits of higher periods and the non-periodic orbits.

6. Physical Implications. The invariant set A described in §3 is hyperbolic and therefore will persist under small perturbations; in particular, a set topologically equivalent to A will be found at nearby energy levels $H^{-1}(\varepsilon)$, $\varepsilon \ll 1$; cf. [2]. The existence of a set such as A is therefore significant if it occurs in models of physical situations such as those of [9, 14].

Since all the orbits in A are of saddle type, orbits originating near points of A will diverge from them exponentially fast, while displaying a sensitive dependence on initial conditions. In physical problems this will manifest itself as an erratic wandering: the orbit being thrown back and forth for a while near A and finally ejected from its neighborhood.

Horseshoe mappings have also been detected in a study of a single degree of freedom forced oscillator, which is another dynamical system in a three dimensional manifold ($\mathbf{R}^2 \times S^1$) [10]. They also occur in certain 'perturbulent' parameter ranges in the Lorenz equations [11]. In both of these cases, while the attractors are relatively simple (closed orbits or fixed points), a 'metastable' hyperbolic set of horseshoe type causes typical orbits to move erratically for some time before approaching an attractor.

In the Hamiltonian case, since there are no attractors, one expects orbits to move towards regions containing families of elliptic orbits and therefore ultimately to tend toward periodic or almost periodic motions.

ACKNOWLEDGEMENT. The author would like to thank Robert Devaney for a number of helpful comments.

REFERENCES

1. G.D. Birkhoff, *Dynamical systems with two degrees of freedom*, Trans. Amer. Math. Soc. **18** (1917), 199–300.
2. D.R.J. Chillingworth, *Differential Topology with a View to Applications*, Pitman, London, 1976.
3. R.C. Churchill, G. Pecelli and D.L. Rod, *A survey of the Hénon-Heiles Hamiltonian with applications to related examples*, Como Conference Proc. on Stochastic Behavior in Classical and Quantum mechanical systems, Springer Lecture Notes in Physics. (to appear).
4. R. Devaney, *Homoclinic orbits in Hamiltonian systems*, J. Differential Equations **21** (2) (1976), 431–438.
5. ———, *Blue sky catastrophes in reversible and Hamiltonian systems*, Indiana University Math. J. **26** (2) (1977), 247–263.
6. ———, *Homoclinic orbits to hyperbolic equilibria*, Ann. N.Y. Acad. Sci. **316** (1979), Bifurcation theory and applications in the scientific disciplines, O. Gurel and O. Rössler ed., 108–117.
7. ———, *Transverse homoclinic orbits in an integrable system*, Amer. J. Math., (to appear).
8. M.W. Hirsch and S. Smale, *Differential Equations, Dynamical Systems and Linear Algebra*, Academic Press, New York, 1974.
9. P.J. Holmes, *Bifurcations to divergence and flutter in flow-induced oscillations: a finite dimensional analysis*, J. Sound and Vibration **53** (4) (1977), 471–503.
10. ———, *A nonlinear oscillator with a strange attractor*, Phil. Trans. Roy. Soc. A292 No. 1394 (1979), 419–448.
11. J.L. Kaplan and J.A. Yorke, *The onset of chaos in a fluid flow model of Lorenz*, Ann. N.Y. Acad. Sci. **316** (1979), 400–407.
12. L. Markus and K. Meyer, *Solenoids in generic Hamiltonian dynamics*, (to appear).
13. J. Moser, *Stable and Random Motions in Dynamical Systems* Annal. of Math Studies **71**, Princeton University Press, Princeton, NJ, 1973.
14. M.P. Paidoussis and N.Y. Issid, *Dynamic Stability of pipes conveying fluid*, J. Sound and Vibration **33** (3) (1974), 267–294.
15. L.P. Silnikov, *The existence of a denumerable set of periodic motions in four-dimensional space in an extended neighborhood of a saddle-focus*, Soviet. Math. Dokl. **8** (1967), 54–58.
16. ———, *A contribution to the problem of the structure of an extended neighborhood of a rough equilibrium state of saddle-focus type*, Math. USSR Sb. **10** (1970), 91–102.
17. S. Smale, *Diffeomorphisms with many periodic points*, Differential and Combinational Topology, ed. S.S. Cairns, Princeton University Press, Princeton, NJ, 1965.
18. ———, *Differentiable dynamical systems*, Bull. Amer. Math. Soc. **73** (1967), 747–817

DEPARTMENT OF THEORETICAL AND APPLIED MECHANICS, CORNELL UNIVERSITY, ITHACA, NY 14853.

