in Keen's paper, the Julia set is also the closure of the set of repelling periodic points. Thus there are repelling periodic points arbitrarily close to any escaping point, and vice versa.

The exact structure of the Julia set of transcendental functions like the exponential is known: when the Julia set is not the whole plane, it is a "Cantor bouquet". By this we mean that the Julia set is a Cantor set of curves, each of which is homeomorphic to $[0,\infty)$ and each of which extends to ∞ in the right half-plane. It is known that, in this case, the Lebesgue measure of the Julia set is zero but its Hausdorff dimension is two! See **[Mc]**. Bifurcations such as the one above occur in a variety of entire transcendental functions. For example, in Color Plates 3 and 4 we display the Julia sets of $.66i\cos z$ and $.68i\cos z$. Again one sees that the Julia set explodes as the parameter $i\lambda$ increases through $.67\ldots$, and again it is a saddle-node bifurcation that leads to this explosion. We remark that, unlike the exponential function, the Julia set for any member of the family $\lambda\cos z$ has infinite Lebesgue measure. This accounts for the fact that the colored region in Color Plate 3 seems to occupy a larger area than that of Color Plate 1, even though both Julia sets are Cantor bouquets.

REFERENCES

[A] V. I. Arnol'd, *Ordinary differential equations*, M. I. T. Press, Cambridge, Mass., 1973.

[AS] R. Abraham and C. Shaw, *Dynamics: The geometry of behavior*, Aerial Press, Santa Cruz, Calif., 1982.

[CE] P. Collet and J.-P. Eckmann, *Iterated maps of the interval as dynamical systems*, Birkhäuser, Boston, 1980.

[D] R. L. Devaney, *An introduction to chaotic dynamical systems*, Addison-Wesley, Menlo Park, 1985.

[D1] ——, *Bursts into chaos*, Phys. Lett. **104** (1984), 385-387.

[D2] ——, *Chaotic bursts in nonlinear dynamical systems*, Science **235** (1987), 342-345.

[F] M. Feigenbaum, *Quantitative universality for a class of nonlinear transformations*, J. Statist. Phys. **19** (1978), 25-52.

[GH] J. Guckenheimer and P. Holmes, *Nonlinear oscillations, dynamical systems, and bifurcation of vector fields*, Springer-Verlag, New York, 1983.

[HS] M. Hirsch and S. Smale, *Differential equations, dynamical systems, and linear algebra*, Academic Press, New York, 1974.

[M] B. Mandelbrot, *The fractal geometry of nature*, Freeman, San Francisco, 1982.

[Mc] C. McMullen, *Area and Hausdorff dimension of Julia sets of entire functions*, Trans. Amer. Math. Soc. **300** (1987), 329-342.

[PdM] J. Palis and W. de Melo, *Geometric theory of dynamical systems*, Springer-Verlag, New York, 1982.

[PR] H.-O. Peitgen and P. Richter, *The beauty of fractals*, Springer-Verlag, New York, 1986.

[S] M. Shub, *Global stability of dynamical systems*, Springer-Verlag, New York, 1986.

DEPARTMENT OF MATHEMATICS, BOSTON UNIVERSITY, BOSTON, MASSACHUSETTS 02215

Proceedings of Symposia in Applied Mathematics
Volume 39, 1989

Nonlinear Oscillations and the Smale Horseshoe Map

PHILIP HOLMES

ABSTRACT. This paper introduces, via an example, some basic ideas in the global analysis of dynamical systems. In particular, we indicate how it may be proved that Smale's horseshoe map is contained in the Poincaré map of the simple pendulum subject to periodically varying torque. We indicate the remarkable physical consequences which result.

0. Introduction. In this article we introduce some important ideas in the global theory of dynamical systems by means of a simple example: the pendulum subject to a small oscillating torque and weak dissipation. We use the second order ordinary differential equation describing this system to introduce such ideas as the Poincaré map, Smale's horseshoe map and the chaos that accompanies it, and the Melnikov perturbation method, with which one goes hunting for horseshoes. It would be imprudent to attempt, and impossible to succeed in, a comprehensive tutorial on dynamical systems within either the confines of this article or the hour of speech and gesture granted us by the organizers. For those with a year or two of interrupted leisure, the books by Arnold **[1973]**, Andronov et. al. **[1966]** or Hirsch-Smale **[1974]** provide good introductory material, while those of Arnold **[1982]**, Palis-de Melo **[1982]**, Irwin **[1980]** and (succumbing to chauvinism) Guckenheimer-Holmes **[1983]** contain more advanced material. Some aspects of the present treatment are adapted from the latter book.

In the following sections we introduce our model problem and describe the Poincaré map §1. We then discuss some general features of iterated (invertible) maps on the plane §2, before introducing and describing Smale's horseshoe map §3. We return to our example and describe the Melnikov perturbation calculation in §4, and finally summarize the fruits of our labours in §5. It is worth noting that Smale invented the horseshoe map while attempting to understand the papers of Cartwright-Littlewood **[1945]** and Levinson **[1949]** on the periodically forced Van der Pol oscillator (cf. Smale **[1963]**,

1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 58F99.

1967]: the story is told nicely in Smale [1980]). As in that case, specific examples have often led the way to general ideas in dynamical systems theory. It is therefore entirely appropriate that we should start with a (deceptively) simple model problem.

1. The pendulum equation. Consider the simple pendulum of Figure 1. A point mass m is suspended by a rigid, massless rod of length l pivoted freely at 0 to swing in a plane. Three forces act on the bob: gravitation ($-mg$, vertically), friction or dissipation due to air resistance ($-cv$, tangentially), and the external time varying torque $\delta T(t)$ applied at the pivot. The minus signs are conventional, reflecting that the forces oppose motion as indicated, and friction is modelled by the simplest possible law: resistance is linearly proportional to speed. The state of the system is uniquely specified by the pair $(\theta, \frac{d\theta}{dt})$, angular position and velocity. Resolving the forces in the tangential direction, and appealing to Newton's famous second law (force = mass×acceleration), we obtain the second order ordinary differential equation

$$(1.1)\qquad \delta l T(t) - cl\frac{d\theta}{dt} - mg\sin\theta = m\frac{d}{dt}\left(l\frac{d\theta}{dt}\right).$$

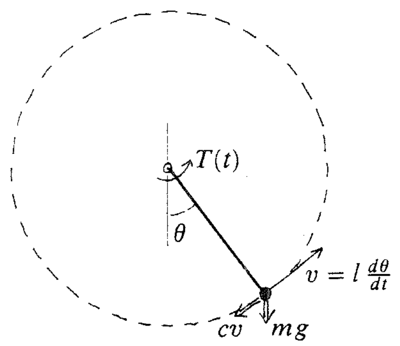

FIGURE 1. The simple pendulum subject to torque and friction.

A slight rearrangement of terms and a change of time scale yields the system we shall study:

$$(1.2)\qquad \ddot{\theta} + \sin\theta = \delta S(t) - \gamma\dot{\theta},$$

where $(\dot{\ }) = \frac{d}{dt}(\ )$ (t is the new time), $S(t) = \frac{l}{gm}T(t)$ and $\gamma = \frac{c}{m}\sqrt{\frac{g}{l}}$.

When the parameters δ, γ are equal to zero we have the classical pendulum: this equation can be solved in closed form using elliptic functions, since solutions simply run around on level sets of the hamiltonian energy function

$$(1.3)\qquad H(\theta, \dot{\theta}) = \frac{\dot{\theta}^2}{2} + (1 - \cos\theta)$$

(= kinetic + potential energy). In §4 we shall exploit this to approximate solutions of the *perturbed* problem for small δ, $\gamma \neq 0$.

Before introducing the Poincaré map, we rewrite (1.2) as a system of first order differential equations. We let $\dot{\theta} = v$ and treat time as a (trivially evolving) third dependent variable:

$$(1.4)\qquad \begin{aligned} \dot{\theta} &= v, \\ \dot{v} &= -\sin\theta + \delta S(t) - \gamma v, \\ \dot{t} &= 1. \end{aligned}$$

At this point we make the additional assumption that $S(t)$ is periodic of period T (e.g. $S(t) = \cos wt$, $T = 2\pi/w$): the *phase-space* or *state-space* of (1.4) is then $(\theta, v, t) \in S^1 \times \mathbf{R} \times S^1 \stackrel{\text{def}}{=} M$, since the state of the system depends only on the angle θ (not the total number of turns the pendulum has executed) and the phase $t \bmod T$ of the forcing function $S(t)$ (as well as the velocity $v = \dot{\theta}$). There is already some nontrivial topology in this simple example!

We next define a *cross section* $\Sigma = \{(\theta, s, t) | t = 0\} \subset M$ which solutions pierce transversely, in view of the third component $\dot{t} = 1$ of (1.4). The Poincaré map $P \colon \Sigma \to \Sigma$ is defined by picking a point $(\theta_0, v_0) \in \Sigma$ and integrating the equation (1.4) to find the point at which the solution based at (θ_0, v_0) next intersects Σ after time T has elapsed. Thus we have

$$(1.5)\qquad P(\theta_0, v_0) = (\theta(T; \theta_0, v_0), v(T; \theta_0, v_0)),$$

where $\theta(t; \theta_0, v_0)$, $v(t; \theta_0, v_0)$ $(t = t)$ is the solution to (1.4) based at (θ_0, v_0). In Figure 2 we sketch the construction for the unperturbed pendulum equation ($\delta = \gamma = 0$): note that the level *curves* of the Hamiltonian (1.3) become *sheets* in the three-dimensional (suspended) phase space M. Also, the *periodic orbits* $(\theta, v) = (0, 0)$ and $(\pi, 0)$ correspond to *fixed points* for the map P (the latter are marked γ on Figure 2(b)). In the unperturbed case the periodic orbits are trivial, so θ and v do not change with time. For $\delta \neq 0$, on a T-periodic orbit θ and v do vary as t moves from 0 to T, but it should be intuitively clear that P will still have a fixed point. Similarly, a kt-periodic orbit or *subharmonic* corresponds to a k-periodic *cycle* for P.

We end this section by remarking that it is a nice exercise to show that the existence-uniqueness theorem for ODEs implies that P is a diffeomorphism and that, for systems with linear damping like (1.4), the Jacobian derivative DP of P satisfies

$$(1.6)\qquad \det(DP) = e^{-\gamma T}.$$

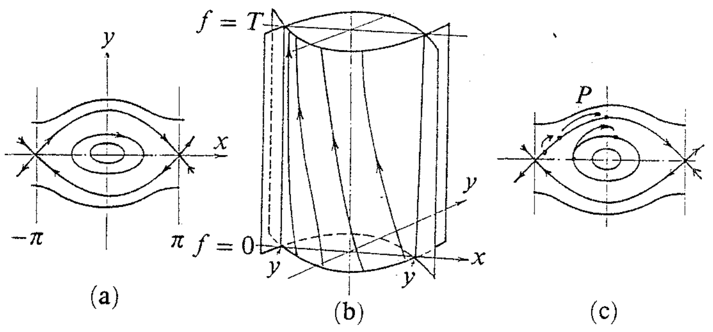

FIGURE 2. Phase plane (a), flow in extended phase space (b), and Poincaré map (c) for the unperturbed pendulum (identify $x = -\pi$ and $x = \pi$ so that (x, y) space is the cylinder (or annulus)).

2. Some basic facts about maps. We now turn to a brief review of maps, concentrating on the two-dimensional case, although everything generalizes to n dimensions. Devaney's contribution to this volume contains a good discussion of the one-dimensional case and his book (Devaney **[1986]** has additional two-dimensional information.

Let $P: \mathbf{R}^2 \to \mathbf{R}^2$ be a (smooth) map and p a fixed point ($p = P(p)$). We call the linear system

(2.1) $$x \mapsto DP(p)x$$

the *linearization* of P at p. $DP(p)$ is a 2×2 matrix: denote its eigenvalues λ_1, λ_2. By arguments similar to those in Devaney's article, one easily sees that p is stable if both eigenvalues of $DP(p)$ lie within the unit circle ($|\lambda_j| < 1 : j = 1, 2$). If this is the case we call p a *sink*. When $|\lambda_1| < 1 < |\lambda_2|$ p is an (unstable) *saddle point* and when $|\lambda_j| > 1$, $j = 1, 2$, p is a *source*. If $|\lambda_j| \neq 1$, $j = 1, 2$, we call p *hyperbolic* and the Hartman-Grobman theorem (cf. Devaney **[1986]**, Guckenheimer-Holmes **[1983]**) guarantees that the dynamical behavior of the linearization (2.1) holds in a neighborhood U of p for the fully nonlinear map F.

For our unperturbed example the fixed point(s) $(\theta, v) = (\pm\pi, 0)$ of the map $P = P_0$ are clearly saddle points. In fact the linearized map can be obtained by integrating the linearized differential equation linearized at $(\theta, v) = (\pm\pi, 0)$:

(2.2) $$\dot{\xi}_1 = \xi_2, \qquad \dot{\xi}_2 = -\cos(\pm\pi)\xi_1 = \xi_1.$$

Elementary analysis shows that the fundamental solution matrix to this system may be written

(2.3) $$\begin{bmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{bmatrix}$$

and hence that the time T map, which gives DP_0, is

(2.4.) $$DP_0(\pm\pi, 0)\cdot\xi = \begin{bmatrix} \cosh T & \sinh T \\ \sinh T & \cosh T \end{bmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}.$$

The matrix DP_0 has eigenvalues

(2.5) $$\lambda_{1,2} = \cosh T \pm \sinh T = e^T, e^{-T},$$

and since $e^{-T} < 1 < e^T$, the point(s) $(\pm\pi, 0)$ are, as expected, saddle points. Actually, since θ is measured modulo 2π, and both equilibria correspond to the pendulum standing straight up (Figure 1), these points are identified in M.

It is reasonable to believe, and possible to prove by a simple application of the implicit function theorem, that, for small δ, $\gamma = \mathcal{O}(\varepsilon)$, P_0 perturbs to a nearby map $P_\varepsilon = P_0 + \mathcal{O}(\varepsilon)$, which has a fixed point $P_\varepsilon = (\pi, 0) + \mathcal{O}(\varepsilon)$ with eigenvalues $e^T + \mathcal{O}(\varepsilon)$, $e^{-T} + \mathcal{O}(\varepsilon)$. We use this fact in our perturbation calculations in §4.

The linear system (2.1) can be put into a convenient form by a suitable similarity transformation. In particular, if the eigenvalues are $|\lambda_1| < 1 < |\lambda_2|$, DP may be diagonalized, so that the linear map uncoupled

(2.6) $$u \mapsto \lambda_1 u, \qquad v \mapsto \lambda_2 v$$

and the two axes $v = 0$, $u = 0$ are then the invariant *stable* and *unstable subspaces*, E^s, E^u (Figure 3(a)). The stable manifold theorem (cf. Guckenheimer-Holmes **[1983]**, Devaney **[1986]**) asserts that, locally, the structure for the nonlinear system

(2.7) $$x \mapsto P(x)$$

is qualitatively similar. More precisely, in a neighborhood U of p there exist *local* stable and unstable manifolds $W^s_{\text{loc}}(p)$, $W^u_{\text{loc}}(p)$, tangent to E^s, E^u at p, and as smooth as P. By taking backward and forward images of arcs contained in these manifolds one constructs the *global* stable and unstable manifolds.

(2.8) $$W^s(p) = \bigcup_{n\geq 0} P^{-n}(W^s_{\text{loc}}(p)), \qquad W^u(p) = \bigcup_{n\geq 0} P^n(W^u_{\text{loc}}(p)),$$

which contain all points $x \in \mathbf{R}^2$ which are forward (resp. backward) asymptotic to p under iteration of P.

While the local structure is nice, the global structure need not be, and herein lies much of the reason for "chaotic motions," as we shall see. We call a point $q \in W^u(p) \cap W^s(p)$ a *homoclinic point*, following the terminology of Poincaré **[1899]**. By definition, the orbit $\{P^n(q)\}_{n=-\infty}^{\infty}$ of q is both

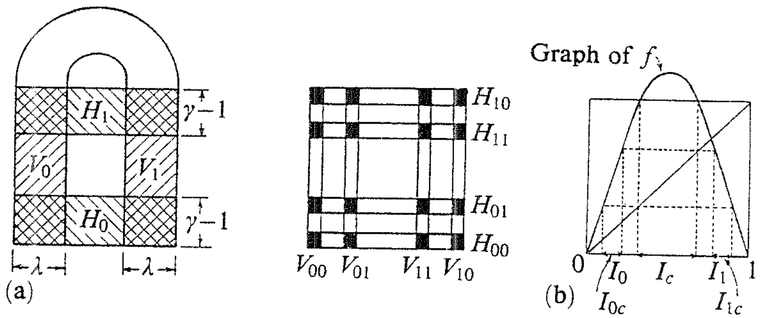

FIGURE 5. (a) The two-dimensional horseshoe and (b) its one-dimensional analogue.

the symbols 0, 1 by the rule $\phi_j(x) = i$ if $F^j(x) \in H_i$ $(i = 0, 1)$. Thus $\phi_j(F(x)) = \phi_{j+1}(x)$ and the action of F on Λ corresponds to the action of the shift σ on the space of symbol sequences Σ. Moreover, every symbol sequence corresponds to an orbit realized by F, since the images V_i lie fully across their preimages H_i. In fact that map $\phi\colon \Lambda \mapsto \Sigma$ is a homeomorphism and the diagram

$$\begin{array}{ccc} \Lambda & \xrightarrow{F} & \Lambda \\ \downarrow{\scriptstyle \phi} & & \downarrow{\scriptstyle \phi} \\ \Sigma & \xrightarrow{\sigma} & \Sigma \end{array}$$

commutes. We say that $F|\Lambda$ is *topologically conjugate to a (full) shift on two symbols*. For $x \in \Lambda'$ one does the same but using only semi-infinite (positive going) sequences since f is noninvertible. More details can be found in Devaney's article. The main advantage of this method of symbolic dynamics is that one can study the orbits of $F|_\Lambda$ (or $f|_{\Lambda'}$) combinatorially, by examining symbol sequences. For instance, the 'constant' periodic sequences $\ldots 000 \ldots \stackrel{\text{def}}{=} (0)'$ and $\ldots 111 \ldots = (1)'$ correspond to fixed points; $(01')$, $(001)'$, $(011)'$, $(0001)'$, etc. to orbits of periods 2, 3, 3, 4, etc. (here $(\ )'$ denotes periodic extension). In this way one proves the following.

PROPOSITION. *The invariant set Λ of the horseshoe contains*: (1) *a countable infinity of periodic orbits, including orbits of arbitrarily high period* ($\simeq 2^k/k$ *orbits of each period* k): (2) *an uncountable infinity of nonperiodic orbits, including countably many homoclinic and heteroclinic orbits, and* (3) *a dense orbit.*

Since $F|_{H_1 \cup H_2}$ contracts uniformly by λ horizontally and expands by γ vertically, the eigenvalues $\mu_{1,2}$ of DF^k for any k-periodic orbit satisfy $|\mu_1| = \lambda^k < 1 < |\mu_2| = \gamma^k$ and thus all such orbits are (unstable) saddles. In fact all

orbits in Λ have associated with them exponentially strong unstable manifolds and thus almost all pair of points Λ separate exponentially fast under F^n. This *sensitive dependence on initial conditions* leads to what we popularly call "chaos". More strikingly, since *every* bi-infinite sequence in Σ corresponds to an orbit of $F|_\Lambda$, there are uncountably many orbits which behave in a manner indistinguishable from the outcome of repeated tossing of a coin: a quintessentially random process.

Perhaps most important is the fact that Λ is a *structurally stable* set; small perturbations $\tilde{F}$ of F possess a topologically equivalent set $\tilde{\Lambda} \sim \Lambda$. In fact to prove the existence of such sets one does not need linearity of F or f, as in Smale's example; it is sufficient to establish uniform bounds on contraction and expansion. See Moser **[1973]** and Guckenheimer and Holmes **[1983]** for more details.

The constructions we have sketched above and in Figure 4 lead one to the fundamental

SMALE-BIRKHOFF HOMOCLINIC THEOREM. *Let $P\colon \mathbf{R}^2 \mapsto \mathbf{R}^2$ be a diffeomorphism possessing a transversal homoclinic point q to a hyperbolic saddle point p. Then, for some $N < \infty$, P has a hyperbolic invariant set Λ on which the Nth iterate P^N is topologically conjugate to a shift on two symbols.*

Birkhoff **[1927]** had already proved the existence of countably many periodic points in any neighborhood of a homoclinic point, but Smale's construction provided a more complete picture and he extended it to $\mathbf{R}^n$. Infinite dimensional versions of the theorem are also available.

4. Melnikov's perturbation method. Although Smale constructed the horseshoe in connection with a periodically forced oscillator problem it was not until the work of Melnikov **[1963]** that a general method existed for proving that horseshoes exist in specific Poincaré maps. Tantalizing hints of this technique can be found in Poincaré's **[1890]** paper on the three body problem and Arnold **[1964]** applied the idea to Hamiltonian systems around the same time as Melnikov. Thus, as Jerry Marsden has remarked, the method should probably be called the Poincaré-Arnold-Melnikov method. What one actually does is prove that a suitably perturbed, almost Hamiltonian system has a transversal homoclinic orbit and then apply the Smale-Birkhoff homoclinic theorem.

We only outline the simplest version of the method here. See Holmes-Marsden **[1981, 1982a, b, 1983]** and Wiggins **[1988]** for extensions to many (even infinitely many) dimensions. Consider a planar ordinary differential equation subject to a small time-periodic perturbation:

$$(4.1) \qquad \dot{x} = f(x) + \varepsilon g(x, t), \qquad g(x, t) = g(x, t + T), \quad x \in \mathbf{R}^2.$$

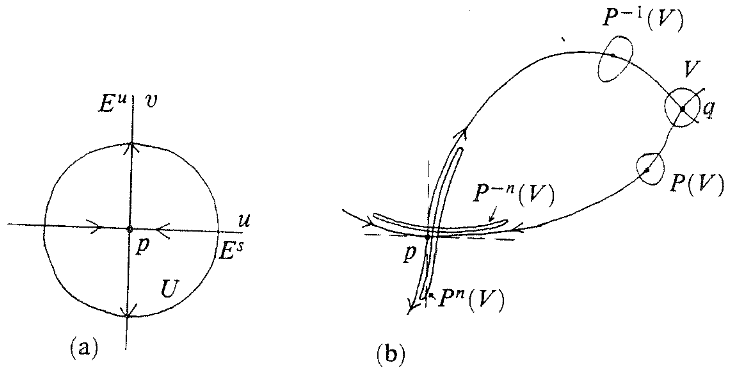

FIGURE 3. (a) Invariant subspaces for linear map; (b) Invariant manifolds for nonlinear map, showing a homoclinic point, q.

forward and backward asymptotic to p. If the manifolds $W^s(p)$, $W^u(p)$ intersect transversely at q, then iteration of a small region V containing q causes $P^n(V)$ and $P^{-n}(V)$ to "pile up" on $W^u(p)$, $W^s(p)$ respectively as $n \mapsto \infty$ (Figure 3(b)). (That this occurs in the controlled fashion of C^1-convergence of transversals to W^u, W^s at q is the content of the Lambda Lemma; Newhouse **[1980]**, Guckenheimer-Holmes **[1983]**.) In such a situation the Smale-Birkhoff homoclinic theorem, described in the next section, shows that V and its images contain a very complicated invariant set for P.

We end by noting that, for the unperturbed pendulum Poincaré map of Figure 2(c), all points on the level sets

$$H(\theta, v) = \frac{v^2}{2} + (1 - \cos\theta) = 0 \tag{2.9}$$

are homoclinic to the point $(-\pi, 0) = (\pi, 0)$. However, these points are all nontransversal, since in this very special case the two manifolds are identical. Perturbation of this degenerate structure to produce transversal homoclinic points is treated in §4.

3. Smale's horseshoe map. As Poincaré **[1890]** realized, the presence of homoclinic points can vastly complicate dynamical behavior. However, the very fact that their existence implies recurrent motions makes the situation amenable to at least a partial analysis. Consider the effect of the map P of Figure 3(b), containing a transverse homoclinic point q to a hyperbolic saddle p, on a "rectangular" strip S containing p and q in its boundary. As n increases, $P^n(S)$ is contracted horizontally and expanded vertically until the image $P^N(S)$ loops around and intersects S and P in a 'horseshoe' shape

(Figure 4). To prove that the rates of contraction and expansion are uniformly bounded, one shrinks the width of S until many iterates occur for which $P^j(S)$ lies in a neighborhood U of p and the dynamics is therefore dominated by the linear map $DP(p)$ (cf. (2.6)).

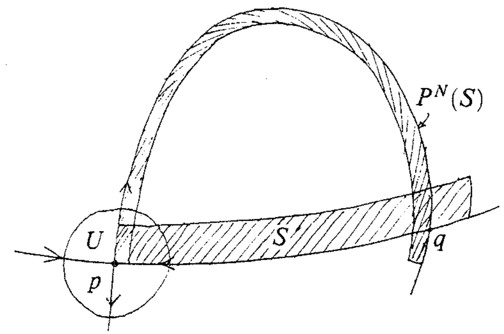

FIGURE 4. P^N has a horseshoe.

A simple model for this situation was provided by Smale **[1963]**, who introduced the map $F: S \mapsto \mathbf{R}^2$ of the square $[0,1] \times [0,1] \subset \mathbf{R}^2$ sketched in Figure 5(a). The map is linear on the two horizontal strips H_i whose images are the vertical strips V_i, $i = 0, 1$; the linearizations being

$$DF(x)|_{x \in H_1} = \begin{bmatrix} \lambda & 0 \\ 0 & \gamma \end{bmatrix}, \qquad DF(x)|_{x \in H_2} = \begin{bmatrix} -\lambda & 0 \\ 0 & -\gamma \end{bmatrix}, \tag{3.1}$$

with $0 < \lambda < 1 < \gamma$. Thus $F|_{H_i}$ contracts horizontally and expands vertically in a uniform manner. Smale studied the structure of the set of points Λ which never leave S under iteration of F. By definition $\Lambda = \bigcap_{n=-\infty}^{\infty} F^n(S)$: the intersection of all images and preimages of S. Now $F^{-1}(S) \cap S = H_1 \cup H_2$ and $S \cap F(S) = V_1 \cup V_2$ so $F^n(S)$ is the union of four rectangles of height γ^{-1} and width λ (Figure 5(a)). Similarly $\bigcap_{n=-2}^{2} F^n(S)$ is the union of 16 rectangles of height γ^{-2} and width λ^2, $\bigcap_{n=-k}^{k} F^n(S)$ is the union of 2^{2k} rectangles and, passing to the limit, Λ turns out to be a *Cantor set*: an uncountable point set, every member of which is a limit point.

To see this more easily, consider the set of points which never leave $I = [0,1] \subset R$ under iteration of the one-dimensional map f of Figure 5(b). After one iterate the 'middle' interval I_c is lost, after two iterates its preimages I_{0c}, I_{1c} are lost, etc. Removing middle intervals of fixed proportional size (α, say) produces the classic 'middle α' Cantor set Λ', the one-dimensional analogue of Λ. We remark that the map f is qualitatively like the famous quadratic map $x \mapsto ax(1-x)$ or $x \mapsto c - x^2$ (cf. the article by Devaney in this volume).

The sets Λ and Λ' can be coded in a way which describes their dynamics. To each $x \in \Lambda$ we assign a bi-infinite sequence $\phi(x) = \{\phi_j(x)\}_{j=-\infty}^{\infty}$ of

We suppose that f and g are sufficiently smooth and bounded on bounded sets and that the unperturbed system is Hamiltonian, i.e. there exists a function $H(x)\colon \mathbf{R}^2 \mapsto \mathbf{R}$ such that

$$\dot x_1 = f_1(x_1,x_2) = \frac{\partial H}{\partial x_2}(x_1,x_2), \qquad \dot x_2 = f_1(x_1,x_2) = \frac{-\partial H}{\partial x_1}(x_1,x_2). \tag{4.2}$$

We assume that this unperturbed vector field contains a hyperbolic saddle point p_0 lying in a closed level set of H: thus there is a (degenerate, nontransversal) loop of homoclinic points: Figure 6(a). The orbits on this loop are denoted $x = x_0(t - t_0)$, where t_0 denotes a shift in the initial condition or base point. For precise technical hypothesis see Guckenheimer-Holmes [**1983**, §4.5].

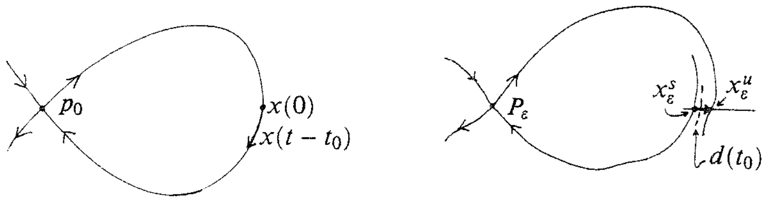

FIGURE 6. (a) The unperturbed loop; (b) The perturbed Poincaré map.

As in §2, we consider the unperturbed and perturbed Poincaré maps P_0, P_ε corresponding to (4.1) with $\varepsilon = 0$ and $\varepsilon \neq 0$. Implicit function arguments show that the hyperbolic fixed point p_0 of p_0 perturbs to a nearby hyperbolic fixed point $p_\varepsilon = p_0 + \mathscr{O}(\varepsilon)$ for P_ε and its stable and unstable manifolds remain close, as indicated in the sketch of Figure 6(b). In fact one proves that the power series representations of solutions $x_\varepsilon^{s,u}$ lying in the perturbed stable and unstable manifolds of the small periodic orbit $\gamma_\varepsilon = p_0 + \mathscr{O}(\varepsilon)$ of (4.1; $\varepsilon \neq 0$) are valid in the following *semi-infinite* time intervals:

$$\begin{aligned} x_\varepsilon^s(t,t_0) &= x_0(t-t_0) + \varepsilon x_1^s(t,t_0) + \mathscr{O}(\varepsilon^2), \qquad t \in [t_0, \infty);\\ x_\varepsilon^u(t,t_0) &= x_0(t-t_0) + \varepsilon x_1^u(t,t_0) + \mathscr{O}(\varepsilon^2), \qquad t \in (-\infty, t_0]. \end{aligned} \tag{4.3}$$

This follows from the usual finite time Gronwall estimates and the fact that these special solutions are "trapped" in the local stable and unstable manifolds and thus have well controlled asymptotic behavior as $|t| \mapsto \infty$. One can therefore seek the leading order terms $x_1^{s,u}(t,t_0)$ as solutions of the first variational equation obtained by substituting (4.3) into (4.1) and expanding in powers of ε:

$$\dot x_1^{s,u} = Df(x_0(t-t_0))x_1^{s,u} + g(x_0(t-t_0),t). \tag{4.4}$$

Now, while equation (4.4) is linear, it is usually very hard to solve, since $Df(x_0(t-t_0))$ is a time varying 2×2 matrix and is not even periodic. Here the idea of Melnikov comes to our rescue. He realized that, to estimate the distance $d(t_0)$ between the perturbed stable and unstable manifolds at a base point t_0 of the unperturbed solution, one need not solve (4.4) explicitly. His method goes as follows.

From (4.3) and Figure 6(b), we have

$$\begin{aligned} d(t_0) &= x_\varepsilon^u(t_0,t_0) - x_\varepsilon^s(t_0,t_0)\\ &= \frac{\varepsilon(x_1^u(t_0,t_0) - x_1^s(t_0,t_0)) \cdot f^\perp(x_0(0))}{\|f(x_0(0))\|} + \mathscr{O}(\varepsilon)^2, \end{aligned} \tag{4.5}$$

where $f^\perp(x_0(0))$ denotes the normal to the unperturbed solution vector $f(x_0(0))$. Since $a \cdot b^\perp = b \times a$ for vectors in $\mathbf{R}^2$, we can rewrite (4.5) as

$$\begin{aligned} d(t_0) &= \varepsilon \frac{f(x_0(0)) \times (x_1^u(t_0,t_0)) - x_1^s(t_0,t_0)}{\|f(x_0(0))\|} + \mathscr{O}(\varepsilon^2)\\ &\overset{\text{def}}{=} \varepsilon \frac{\Delta^u(t_0,t_0) - \Delta^s(t_0,t_0)}{\|f(x_0(0))\|} + \mathscr{O}(\varepsilon^2). \end{aligned} \tag{4.6}$$

If the quantity $\Delta^u - \Delta^s$ has simple zeros as t_0 varies it follows from the implicit function theorem that, for $\varepsilon \neq 0$ small enough, the distance $d(t_0)$ changes sign as to varies and consequently that the perturbed manifolds intersect transversely. To compute $\Delta^u - \Delta^s$ we introduce time varying functions

$$\Delta^{u,s}(t,t_0) = f(x_0(t-t_0)) \times x_1^{u,s}(t,t_0)$$

and compute

$$\begin{aligned} \dot\Delta^s &= Df(x_0)\dot x_0 \times x_1^s + f(x_0) \times \dot x_1^s\\ &= Df(x_0)f(x_0) \times x_1^s + f(x_0) \times [Df(x_0)x_1^s + g(x_0,t)]\\ &= \operatorname{trace} Df(x_0)f(x_0) \times x_1^s + f(x_0) \times g(x_0,t)\\ &= f(x_0(t-t_0)) \times g(x_0(t-t_0),t). \end{aligned} \tag{4.7}$$

Here we substitute for $\dot x_1^s$ from (4.4) and use $\dot x_0 = f(x_0)$, a matrix-cross product identity, and finally appeal to the fact that

$$\operatorname{trace} Df = \frac{\partial f_1}{\partial x_1} + \frac{\partial f_2}{\partial x_2} = \frac{\partial^2 H}{\partial x_1 \partial x_2} - \frac{\partial^2 H}{\partial x_2 \partial x_1} \equiv 0, \tag{4.8}$$

since f is Hamiltonian. Integrating (4.7) we have

$$\Delta^s(t,t_0) - \Delta^s(t_0,t_0) = \int_{t_0}^{t} f(x_0(s-t_0)) \times g(x_0(s-t_0),s)\,ds$$

and, taking the limit $t \to +\infty$ and using the fact that $f(x_0(t)) \to f(p_0) = 0$ as $t \to \infty$, so that $\Delta^s(t,t_0) \to 0$, we obtain

$$-\Delta^s(t_0,t_0) = \int_{t_0}^{\infty} (f \times g)(x_0(s-t_0),t)\,ds. \tag{4.9}$$

Note that we have used the validity of (4.3) on $[t_0, \infty)$ in this computation. Together with a similar computation for Δ^u, (4.9) yields

(4.10) $$\Delta^u(t_0, t_0) - \Delta^s(t_0, t_0) \overset{\text{def}}{=} M(t_0) = \int_{-\infty}^{\infty} (f \times g)(x_0(s - t_0), t)\, ds.$$

We have completed our sketch of the proof of

MELNIKOV'S THEOREM. *Under the hypotheses stated in* (4.1), *if* $M(t_0)$ *has simple zeros, then for* $\varepsilon \neq 0$ *sufficiently small the manifolds* $W^s(p_\varepsilon)$, $W^u(p_\varepsilon)$ *intersect transversely. If* $M(t_0)$ *is bounded away from zero then* $W^s(p_\varepsilon) \cap W^u(p_\varepsilon) = 0$.

We illustrate how easy the theorem is to apply by returning to our example (1.4). Here

(4.11) $$f = \begin{pmatrix} v \\ -\sin\theta \end{pmatrix}, \qquad \varepsilon g = \begin{pmatrix} 0 \\ \delta S(t) - \gamma v \end{pmatrix}.$$

The unperturbed homoclinic solution may be written

$$v(t - t_0) = \pm 2\,\text{sech}(t - t_0)$$

(we will not need the θ component), and to make explicit calculations we shall take

(4.12) $$\delta S(t) = \varepsilon\bar{\delta}\cos\omega t, \qquad \gamma = \varepsilon\bar{\gamma},$$

so that the damping and applied torque are assumed to be small and of the same order. We then have $f \times g = \bar{\delta} v \cos\omega t - \gamma v^2$ and $M(t_0)$ may be written

(4.13) $$M(t_0) = \pm 2\bar{\delta}\int_{-\infty}^{\infty} \text{sech}(\sigma)\cos\omega(\sigma + t_0)\, d\sigma - 4\gamma\int_{-\infty}^{\infty} \text{sech}^2(\sigma)\, d\sigma$$

after a change of variables $\sigma = s - t_0$. The second integral of (4.13) is elementary and the first can be evaluated by the method of residues to give

(4.14) $$M(t_0) = \pm 2\bar{\delta}\pi\,\text{sech}\frac{\pi\omega}{2}\cos\omega t_0 - 8\bar{\gamma}.$$

Clearly, $M(t_0)$ has simple zeros iff

(4.15) $$\bar{\delta}\pi > 4\bar{\gamma}\cosh\left(\frac{\pi\omega}{2}\right),$$

which is therefore an explicit criterion for the existence of transverse homoclinic orbits in the limit $\varepsilon \to 0$. Observe that this makes good physical sense. If dissipation $\gamma = \varepsilon\bar{\gamma}$ is large compared to force amplitude $\delta = \varepsilon\bar{\delta}$ then one does not expect recurrent behavior, since the pendulum will simply settle towards the stable position $(\theta, v) = (0, 0)$ as energy is absorbed and will asymptotically approach a (small) periodic solution about that equilibrium.

5. Conclusions. We now show that a remarkable physical conclusion follows from the analysis of the preceding sections. In Figure 7 we indicate how a modest generalization of the horseshoe arises in the Poincaré map of the perturbed pendulum. The "horizontal" strips H_R, H_L are carried by P^N into "vertical" strips V_R, V_L as indicated. Since the saddle points near (θ, v): $(\pm\pi, 0)$ are identified, these images intersect H_R, H_L much as in the canonical Smale example of Figure 5 (cf. Figure 4). As in §3, one obtains a homeomorphism between the shift on the two symbols R, L and some iterate P^N of the Poincaré map restricted to a suitable (Cantor) set $\Lambda^N = \bigcap_{n=-\infty}^{\infty} P^{nN}(H_R \cup H_L)$. Note that our construction guarantees that a point lying in H_R will be mapped around near the stable and unstable manifolds with $\dot{\theta} = v > 0$ while a point lying in H_L is mapped around with $\dot{\theta} < 0$. Thus, an 'R' in the symbol sequence corresponds to a passage of the pendulum bob past $\theta = 0$ with $\dot{\theta} > 0$ and an 'L' to a passage with $\dot{\theta} < 0$. Since we have a full shift ($P^N(H_L)$ and $P^N(H_R)$ both lie across $H_L \cup H_R$) we conclude that any "random" sequence of the symbols L, R corresponds to an orbit of the pendulum, rotating "chaotically" to the left and to the right.

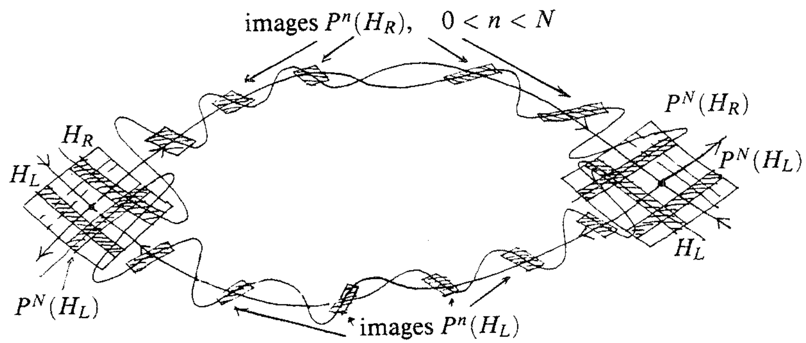

FIGURE 7. Poincaré map for the perturbed pendulum.

This conclusion is perhaps not too surprising, if we consider the effect of a small periodic perturbation on the unclamped, Hamiltonian pendulum swinging near its seperatrix orbit $H(\theta, v) = 2$. Each time the pendulum reaches the top of its swing, near the inverted, unstable state, the oscillating torque supplies a small push either to the left or right depending on the phase (time). Thus the precise time at which the bob arrives near this position is crucial and this, in turn, is determined by the time at which it left the same position after the preceding swing. Here is the physical interpretation of sensitive dependence upon initial conditions.

At this point honesty compels us to point out that all is not rosy in the study of chaotic dynamics. The analysis sketched in this paper establishes that a specific deterministic differential equation possesses chaotic orbits and provides an estimate for the parameter range(s) in which they exist. This does not necessarily imply that we have a *strange or chaotic attractor.* An *attractor* for a flow or map is an *indecomposable, closed, invariant set* for the flow or map, which attracts all orbits starting at points in some neighborhood. The maximal such neighborhood is the *domain of attraction*, or *basin.* Jim Yorke's lecture deals with this idea. In the pendulum example it is easy to see that all orbits of P remain trapped in a band $\mathscr{B} = \{(\theta, v) | |v| \leq \Gamma\}$ in the phase space; one simply observes that, if we choose $\Gamma > \frac{1+S_{\max}}{\delta}$ ($S_{\max} = \max_t |S(t)|$), then the second component of (1.4) admits the bounds

$$\begin{aligned} \dot{v} &\leq |-\sin\theta| + \delta|S(t)| - \gamma|v| \\ &\leq -\gamma|v| + 1 + \delta S_{\max} \end{aligned}$$

for $v > 0$ and

$$\dot{v} \geq \gamma|v| - 1 - \delta S_{\max}$$

for $v < 0$. Thus the vector field points into the band $\mathscr{B}$ and hence it is a forward invariant region for P ($P(\mathscr{B}) \subset \mathscr{B}$). The *attracting set* $\mathscr{A}$ is the intersection of all forward images of $\mathscr{B}$ and, since $\det DP = e^{-\gamma T} < 1$ (equation (1.6)), P contracts areas by a constant factor and

$$\mathscr{A} = \bigcap_{n=0}^{\infty} P^n(\mathscr{B})$$

has zero area. $\mathscr{A}$ certainly contains the homoclinic points and their attendant horseshoes displayed above, and any attractors are certainly contained in $\mathscr{A}$, but $\mathscr{A}$ itself need not be indecomposable. To display parameter values for which $\mathscr{A}$ *as a whole* and not just $\Lambda \subset \mathscr{A}$ has a dense orbit (or even a chain recurrent dense orbit, cf. Guckenheimer and Holmes **[1983]**) appears very difficult. In fact work of Newhouse **[1980]** on wild hyperbolic sets and the presence of infinitely many stable periodic orbits at certain parameter values for maps like P shows that there are a lot of values for which $\mathscr{A}$ *cannot* be indecomposable. Thus a "typical" solution approaching $\mathscr{A}$ might eventually settle down to stable periodic behavior, perhaps after a chaotic transient played out near Λ. In spite of the suggestive nature of numerical simulations (cf. Yorke's lecture), this issue still awaits clarification. I prefer to say that P has a *strange attracting set.*

We conclude this article by remarking that the ideas we have outlined seem to be of general relevance in the study of nonlinear differential equations arising in engineering and the sciences. It is now almost a commonplace that "chaotic solutions" are observed in numerical simulation of diverse model systems. Centers for nonlinear science are producing color graphics of fractals and strange attractors almost faster than one can look at them. However, the tools introduced in this article, and especially the perturbative analytical

method of Melnikov, show that analysis can be brought to bear on these systems, and rigorous results obtained. Let us hope that their application brings some order into chaos.

References

A. A. Andronov, E. A. Vitt, and S. E. Khaiken, [1966] *Theory of oscillators*, Pergamon Press, New York.

V. I. Arnold [1964], *Instability of dynamical systems with several degrees of freedom*, Soviet Math. Dokl. **5**, 581–585.

V. I. Arnold [1973], *Ordinary differential equations*, M. I. T. Press.

V. I. Arnold [1982], *Geometrical methods in the theory of ordinary differential equations*, Springer-Verlag.

G. D. Birkhoff [1927], *Dynamical systems*, Amer. Math. Soc.

M. L. Cartwright and J. E. Littlewood [1945], *On nonlinear differential equations of the second order*, I: *the equation* $\ddot{y} + k(1 - y^2)\dot{y} + y = b\lambda k \cos(\lambda t + a), k$ *large*, J. London Math. Soc. **20**, 180–189.

R. L. Devaney [1986], *An introduction to chaotic dynamical systems*, Benjamin Cummings.

J. Guckenheimer and P. Holmes [1983], *Nonlinear oscillations, dynamical systems and bifurcation of vectorfields*, Springer-Verlag.

P. Holmes and J. E. Marsden [1981], *A partial differential equation with infinitely many periodic orbits: chaotic oscillations of a forced beam*, Arch. Rat. Mech. Anal. **76**, 135–166.

P. Holmes and J. E. Marsden [1982a], *Horseshoes in perturbations of Hamiltonians with two degrees of freedom*, Comm. Math. Phys. **82**, 523–544.

P. Holmes and J. E. Marsden [1982b], *Melnikov's method and Arnold diffusion for perturbations of integrable Hamiltonian systems*, J. Math. Phys. **23**, 669–675.

P. Holmes and J. E. Marsden [1983], *Horseshoes and Arnold diffusion for Hamiltonian systems on Lie groups*, Indiana Univ. Math. J. **32**, 273–310.

V. K. Melnikov [1983], *On the stability of the center for time periodic perturbations*, Trans. Moscow Math. Soc. **12**, 1–57.

J. Moser [1973], *Stable and random motions in dynamical systems*, Princeton University Press.

S. E. Newhouse [1980], Lectures on dynamical systems. In *Dynamical Systems*, CIME Lectures, Bressanone, Italy, June 1978, Birkhauser, pp. 1–114.

J. Palis and W. de Melo [1982], *Geometric theory of dynamical systems: An Introduction*, Springer-Verlag.

H. Poincaré [1890], *Sur le problème des trois corps et les équations de la dynamique*, mémoire Couronné du Prix de S. M. le Roi Oscar II, Paris, Acta Math. **13**, 1–271.

S. Smale [1963], Diffeomorphisms with many periodic points. In *Differential and Combinatorial Topology*, ed. S. S. Cairns, Princeton University Press, pp. 63–80.

S. Smale [1967], *Differential dynamical systems*, Bull. Amer. Math. Soc. **73**, 747–817.

S. Smale [1980], *The mathematics of time: essays on dynamical systems, economic processes and related topics*, Springer-Verlag.

S. Wiggins [1988], *Global bifurcations and chaos: analytic methods*, Springer-Verlag.

Departments of Theoretical & Applied Mechanics and Mathematics & Center for Applied Mathematics, Cornell University, Ithaca, New York 14853