4

CHAOS IN HAMILTONIAN SYSTEMS AND AREA-PRESERVING MAPPINGS

4.1 THE SURFACE OF SECTION

In the preceding chapters we have talked at length about the evolution of orbits in (multidimensional) phase space, but, apart from the phase-plane pictures of one-degree-of-freedom systems discussed in the first chapter, we have said little about how that motion can be visualized in practice. Clearly, there is the problem of dimension. For a two-degree-of-freedom (Hamiltonian) system, the phase space is four-dimensional and, if conservative, the energy shell is three-dimensional. Even following the motion on this 3-D energy shell is difficult—especially when faced with a 2-D piece of paper on which to plot it! To this end, a most valuable technique, due to Poincaré (1892) and Birkhoff (1932), termed the *surface of section*, has been developed. Although applicable to higher-dimensional systems, it is particularly well suited for conservative Hamiltonian systems of two degrees of freedom. In terms of the development of the subject of nonlinear dynamics, it is worth noting that some of the first—and still highly regarded—surface of section computations for nonintegrable Hamiltonians started to appear in the literature at about the same time as the KAM theorem. Computer

studies such as these have played a major role in both complementing and prompting theoretical developments of the subject.

4.1.a Surface of Section for Two-Degree-of-Freedom Hamiltonians

Consider a two-degree-of-freedom conservative Hamiltonian of the form

$$E = H = \frac{1}{2m}(p_x^2 + p_y^2) + V(x, y) \tag{4.1.1}$$

The study of the motion of the orbits of this system can be reduced to a two-dimensional problem in the following way. On a given energy shell, take a "slice" of the phase space at some given point, say $y = 0$. Now, follow a given orbit (obtained by numerically integrating Hamilton's equations on a computer); every time it passes through the point $y = 0$, note the corresponding values of p_x and x. If the potential $V(x, y)$ supports bounded motion, the orbit will repeatedly pass through this phase-space slice and in this way one can build up a "map" of successive (p_x, x) values as illustrated in Figure 4.1. This is the surface of section, and a point on it defines the state of the system to within a sign. This is easily seen since, given E and $y = 0$, one has

$$p_y = \pm \sqrt{2m\left(E - \frac{1}{2m}p_x^2 - V(x, 0)\right)} \tag{4.1.2}$$

The surface of section is normally constructed by just keeping one sign of p_y, say $p_y > 0$.

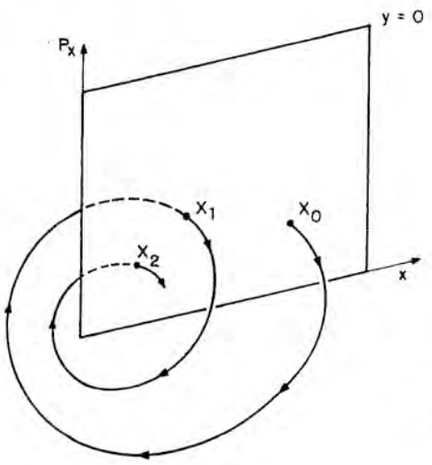


Figure 4.1 Construction of a surface of section.

If we denote the initial conditions of a given orbit (i.e., some p_x, x, E, $y = 0$) on the section as the point X_0, the successive intersections $X_1, X_2, \ldots, X_n$ provide a type of phase-plane "mapping" of the motion. This is an important concept that we shall discuss in detail later on. For now, though, we note that the times between the successive intersections with the surface of section (i.e., the points $X_0, X_1, \ldots$) are not necessarily equal. If we choose an initial condition corresponding to an orbit lying on a torus, the sequence of points $X_0, X_1, X_2, \ldots$ will lie on some smooth curve corresponding to the intersection of that torus with the surface of section (see Figure 4.2). If the chosen torus is one on which the frequency ratio ω_1/ω_2 is irrational, we know that a single orbit covers the torus ergodically. This will be manifested in the surface of section by the (gradual) "filling up" of the smooth curve by the successive iterates X_i. On the other hand, if the frequency ratio is rational, the orbit is closed and there will only be a finite number of intersections X_i $(i = 0, \ldots, n)$, such that $X_0 = X_n$, where n is determined by the rationality of ω_1/ω_2.

As discussed in the previous chapter, the KAM theorem tells us that for a weakly (nonintegrably) perturbed Hamiltonian, most tori are preserved. We have loosely talked about the other tori being, in some sense, "destroyed." Trajectories in these regions are now free to wander over larger regions of the phase space, and this is manifested on the surface of section as a random-looking "splatter" of points through which a smooth curve cannot be drawn. Of course, by eye it is difficult to be objective about what a "random splatter" is, but if a trajectory is run long enough, a pattern clearly different from smooth curves is usually apparent. Eventually, one can even hope to see (small) *areas* of the surface of section being filled up. Furthermore, there are some important computational tests (the computation of power spectra and Lyapunov exponents) to be described in Section 4.5 that enable one to distinguish objectively between the "regular" orbits, which lie on smooth curves in the surface of section, and the "irregular" (or chaotic) ones, which give rise to the random-looking patterns. Nonetheless, the surface of section is an enormously valuable tool and, when computed for a large number of initial conditions on the same

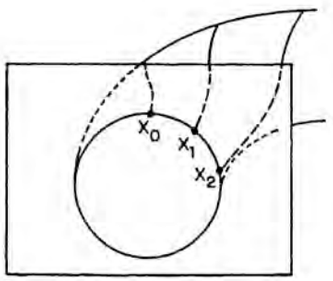


Figure 4.2 Successive intersections of a trajectory on a torus with surface of section.

energy shell, it is able to give an immediate picture of the enormously complicated phase-space structure of nonintegrable systems.

4.1.b The Henon–Heiles Hamiltonian

One of the most famous and enduring surface of section studies, carried out over 20 years ago, is due to Henon and Heiles (1964). Their paper is a model of lucid scientific writing and is a "must" on any reading list. Their Hamiltonian takes the form

$$H = \tfrac{1}{2}(p_x^2 + p_y^2 + x^2 + y^2) + x^2 y - \tfrac{1}{3}y^3 \tag{4.1.3}$$

and was chosen as a simple model for the motion of a star in a cylindrically symmetric, gravitationally smoothed galactic potential. It can also provide a simple model for a pair of nonlinearly coupled molecular bonds. The potential energy function $V(x, y) = (x^2 + y^2)/2 + x^2 y - y^3/3$, sketched in Figure 4.3, supports bounded motion up to an energy $E = \frac{1}{6}$. For small displacements, the motion is nearly linear; however, as the energy increases, the particle "samples" more and more of the nonlinearity in the potential with correspondingly interesting consequences. In Figure 4.4 we show the now-famous surfaces of section computed (for a variety of initial conditions at each energy) at $E = \frac{1}{12}$, $\frac{1}{8}$, and $\frac{1}{6}$. At $E = \frac{1}{12}$, the motion is predominantly integrable and (virtually) all the initial conditions studied lead to orbits lying on smooth curves. The self-intersecting curve, where the intersections (points A, B, C in Figure 4.4*a*) are apparently (to the eye only!) smooth, is a type of separatrix which we shall discuss in more detail later. At $E = \frac{1}{8}$, the surface of section has clearly changed. Some of the smooth curves still remain, whereas others have "broken up" in various ways. In the right-hand family of curves, there now appears a "chain" of five "islands." This is generated by a single trajectory which jumps, successively, from island to island, gradually filling up the set of five small curves. On the other hand,

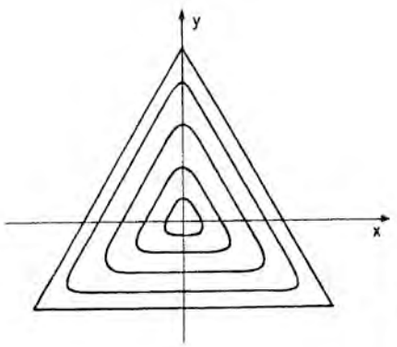


Figure 4.3 Potential energy contours for Henon–Heiles system. Beyond outermost triangle (at energy $E = \frac{1}{6}$) motion can become unbounded.

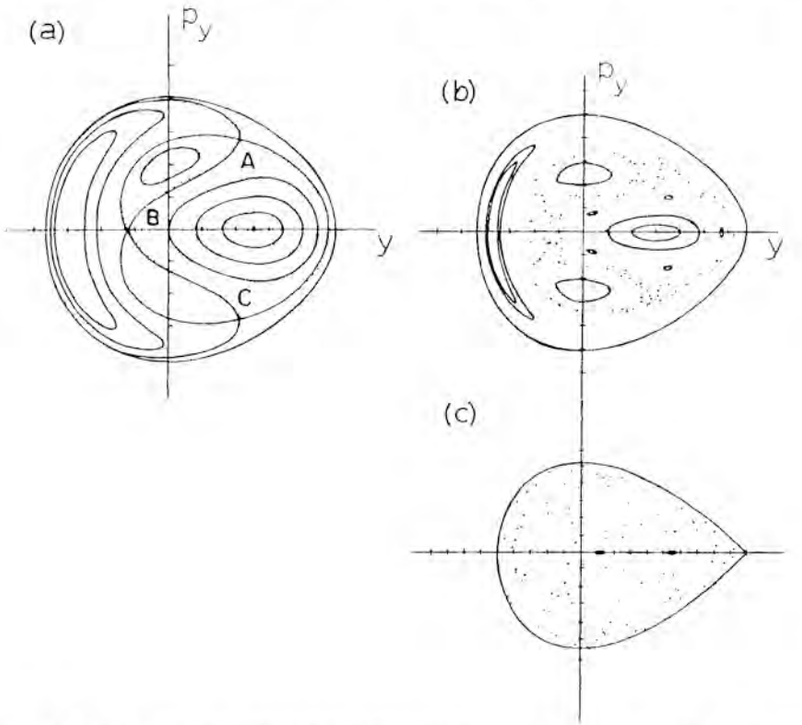


Figure 4.4 Surfaces of section for Henon–Heiles system at (a) $E = \frac{1}{12}$, (b) $E = \frac{1}{8}$, and (c) $E = \frac{1}{6}$. (Reproduced, by permission, from Ford (1975).)

the self-intersecting structure at $E = \frac{1}{12}$ has now disappeared and instead we see random splatter of points (generated by a single orbit) through which a smooth curve cannot be drawn. By $E = \frac{1}{6}$, virtually all smooth curves have disappeared except for a few tiny islands. The set of points, which fills up most of the accessible energy shell, is generated by a single orbit. This remarkable sequence of pictures gives a very clear illustration of the way in which the motion of a nonintegrable Hamiltonian (there is no other first integral apart from the energy) can change from predominantly regular to predominantly chaotic behavior.

4.1.c The Toda Lattice

Before we discuss the nature of this transition in detail, we describe the results of another rather instructive surface of section study. This is the investigation by Ford and co-workers of the three-particle Toda lattice, which consists of three particles on a ring with exponential interactions. The

Hamiltonian takes the form

$$H = \frac{p_1^2}{2m_1} + \frac{p_2^2}{2m_2} + \frac{p_3^2}{2m_3} + e^{-\nu_1(q_1-q_3)} + e^{-\nu_2(q_2-q_1)} + e^{-(q_3-q_2)} - 3 \quad (4.1.4)$$

which, owing to the fact that $p_1 + p_2 + p_3 = 0$ (check this), can be reduced to an equivalent two-dimensional form

$$H = \frac{p_x^2}{2m_x} + \frac{p_y^2}{2m_y} + \frac{1}{24}\{e^{2y+2\sqrt{3}x} + e^{2y-2\sqrt{3}x} + e^{-4y}\} - \frac{1}{8} \quad (4.1.5)$$

For small displacements, the motion is again almost linear, and, in fact, if the exponentials are expanded to third order, the potential-energy term is just the same as that of the Henon–Heiles system.† However, unlike the latter, the motion is bounded for all energies. Working with the *equal* mass case $m_x = m_y$, the surfaces of section were found to be made up entirely of smooth curves at $E = 1$, $E = 256$, and (up to) $E = 56{,}000$—which was their computer limit! There were absolutely no signs of chaos, and these results strongly suggested that the system (4.1.5) is, in fact, integrable. Motivated by these numerical results, Henon (1974) found the other first integral to be

$$\begin{aligned} F = {} & 8p_x(p_x^2 - 3p_y^2) + (p_x + \sqrt{3}p_y)\, e^{2y-2\sqrt{3}x} \\ & + (p_x - \sqrt{3}p_y)\, e^{2y+2\sqrt{3}x} - 2p_x\, e^{-4y} \end{aligned} \quad (4.1.6)$$

which, in the limit of small displacements, tends to

$$F \to 12(yp_x - xp_y) \quad (4.1.7)$$

which is just the angular momentum of the system. By contrast, a subsequent study by Casati and Ford (1975) for the case of *unequal* masses (i.e., $m_x/m_y \neq 1$) revealed chaotic behavior in the surface of section. Apart from illustrating the nice interplay between numerical experiment and theory, these results again raise the fundamental question of how the integrability of the system (4.1.5) might have been predicted without all that computational effort. A detailed discussion of this problem is postponed to Chapter 8.

4.1.d Surface of Section as a Symplectic Mapping

A fundamental property of the surface of section for Hamiltonian systems is that it corresponds to an *area-preserving* or, to be more precise, *symplectic* mapping. To see what is meant by this, it is first instructive to recall some earlier ideas. Firstly, there is Liouville's theorem (Sections 2.2 and 2.3),

†A systematic study of successive truncations of the Toda lattice has been carried out by Contopoulos and Polymilis (1987).

which tells us that phase volume is preserved under the Hamiltonian flow. For a one-degree of freedom system, this is just preservation of area in the (p, q)-phase plane. Thus, for some area A, enclosed by a closed curve $\mathscr{C}$, we can write (using the Stokes' theorem result $\oint_{\mathscr{C}} p\,dq = \int\int_A dp\,dq$)

$$\oint_{\mathscr{C}} p\,dq = \oint_{\mathscr{C}_t} p\,dq \tag{4.1.8}$$

where $\mathscr{C}_t$ is the shape of the curve after it has evolved under the Hamiltonian flow for a time t. This notion can be extended to many degrees of freedom. Thus, for a tube of trajectories in $2n$-dimensional phase space, encircled by some closed curve $\mathscr{C}$, one again has

$$\oint_{\mathscr{C}} \mathbf{p} \cdot \mathbf{dq} = \oint_{\mathscr{C}_t} \mathbf{p} \cdot \mathbf{dq} \tag{4.1.9}$$

where $\mathbf{p} = p_1, \ldots, p_n$, $\mathbf{q} = q_1, \ldots, q_n$. Of course, the integrals in (4.1.9) no longer correspond to a simple area as is the case for $n = 1$. Instead, it corresponds to the sum of areas projected onto the set of (p_i, q_i)-planes (Figure 4.5). It is in this sense that we call the preservation *symplectic* (rather than area-preserving), and, in fact, this is just another statement of the property that the Hamiltonian flow itself is a canonical transformation. For conservative *two*-degree-of-freedom systems of the type discussed above one may show (see Appendix 4.1) that an enclosed area on the

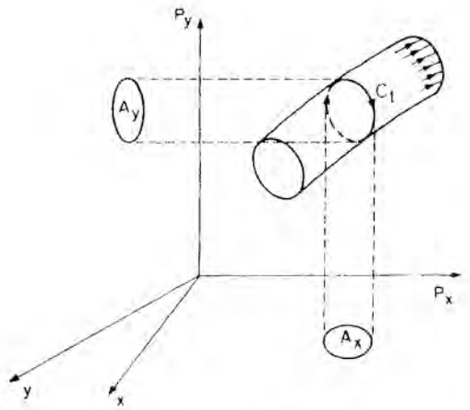


Figure 4.5 Symplectic area preservation. The area enclosed in evolving contour $\mathscr{C}_t$ is the sum of the projections A_x and A_y. This sum is conserved under the Hamiltonian flow.

surface of section (i.e., $\oint_{\mathcal{C}} p_x(x, y = 0, E)\, dx$), will be preserved under the flow.

The surface of section technique can also be used for time-dependent Hamiltonian systems. In the case of periodically driven one-degree-of-freedom systems it is particularly easy to define. For such a Hamiltonian, that is,

$$H(p, q, t + T) = H(p, q, t) \tag{4.1.10}$$

where T is the period of the time-dependent part, the phase space is just the three-dimensional space of the p, q, and t variables. An area-preserving surface of section is simply obtained by taking stroboscopic "snapshots" of the (p, q)-plane at times nT, $n = 0, 1, 2, \ldots$. The set of points $X_i = (p(t + iT), q(t + iT))$ provides the desired surface of section (Figure 4.6). Obviously in this case, as opposed to the autonomous systems discussed above, the time between successive intersections is precisely equal.

The reader will have noticed that in our discussions of the surface of section technique, the word *mapping* has appeared quite frequently. So far, though, this term has only been used in the rather vague sense of some transformation, generated by the Hamiltonian flow, that takes a phase point X_i to a new phase-space location X_{i+1}, that is, for some mapping T

$$X_{i+1} = TX_i \tag{4.1.11}$$

Furthermore, this transformation is, in some sense, "area-preserving" or

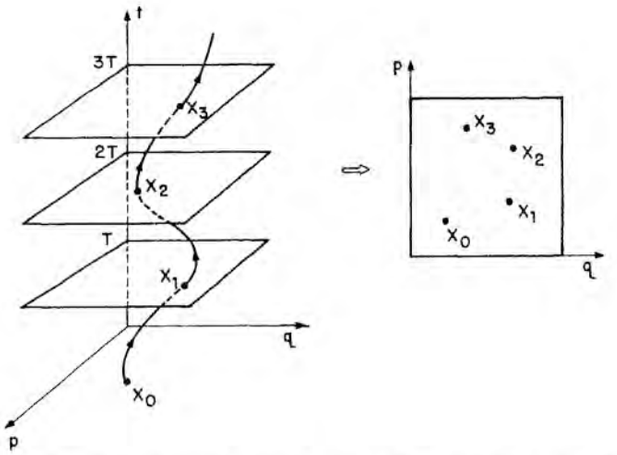


Figure 4.6 Three-dimensional phase space of periodically perturbed system. Surface of section is constructed from stroboscopic slices at $t = nT$.

symplectic. In fact, the notion of an area-preserving mapping is an enormously valuable tool for studying Hamiltonian systems. As will be described, such mappings—even very simple ones—can display all the generic properties of nonintegrable Hamiltonian systems, and eventually we shall discuss these two classes of systems on the same footing. Owing to their relative simplicity, many theorems are more easily proven for mappings than for general Hamiltonians—as well as being much easier to study numerically. So, in order to be able to provide the most detailed discussion of all the phenomena observed in, for example, the Henon–Heiles surfaces of section—such as the break up of tori and the appearance of island chains and chaotic trajectories—we first of all investigate the properties of area-preserving mappings.

4.2 AREA-PRESERVING MAPPINGS

4.2.a Twist Maps

An important class of area-preserving map is the *twist map*. A convenient way to introduce them—and to show their connection with Hamiltonian systems—is to reconsider our previous discussion of surfaces of section. Recall that, for a two-degree-of-freedom system, the surface of section for a trajectory lying on a torus is a sequence of points $X_0, X_1, \ldots$ lying on a smooth curve, which corresponds to the intersection of that torus with the surface of section. Furthermore, if the frequency ratio ω_1/ω_2 is irrational, the sequence X_i fills up the curve ergodically, whereas if ω_1/ω_2 is rational, only a finite sequence of iterates, corresponding to a closed orbit, appears. Now, assuming an integrable system, consider a family of nested tori which, for an isoenergetic nondegenerate system (cf. Eq. (3.5.2)) will have a frequency ratio which varies smoothly, say increases, from torus to torus. Considering just one of these tori, with actions I_1, I_2 (on the energy shell $E = H(I_1, I_2)$), the linear flow on the torus is, of course, just

$$\theta_1(t) = \omega_1 t + \theta_1(0) \tag{4.2.1a}$$

$$\theta_2(t) = \omega_2 t + \theta_2(0) \tag{4.2.1b}$$

where $\omega_1 = \omega_1(I_1, I_2) = \partial H/\partial I_1$ and $\omega_2 = \omega_2(I_1, I_2) = \partial H/\partial I_2$. The time, t_2, it takes for θ_2 to complete a 2π cycle is simply $t_2 = 2\pi/\omega_2$. In this amount of time, the change in θ_1 is thus

$$\begin{aligned} \theta_1(t + t_2) &= \theta_1(t) + \omega_1 t_2 \\ &= \theta_1(t) + 2\pi\omega_1/\omega_2 \\ &= \theta_1(t) + 2\pi\alpha(I_1) \end{aligned} \tag{4.2.2}$$

where $\alpha = \omega_1/\omega_2$, which is called the *rotation number*, is written as a function of just I_1 since, on a given energy shell $E = H(I_1, I_2)$, I_2 can always be expressed in terms of I_1, that is, $I_2 = I_2(E, I_1)$. If we now consider the (I_1, θ_1)-plane as a surface of section (see Figure 4.7), the successive intersections of a trajectory (on this torus) with this plane are just the points $X_i = (\theta_1(t + it_2), I_1)$. Changing to the notation $\theta_i = \theta_1(t + it_2)$ and $r = I_1$ the sequence of points $X_i = X_i(r, \theta)$ associated with the flow on a given torus (with a given "radius" I_1) can be represented by the mapping

$$T: \quad \theta_{i+1} = \theta_i + 2\pi\alpha(r_i) \tag{4.2.3a}$$

$$r_{i+1} = r_i \tag{4.2.3b}$$

where α is taken to be a smoothly varying function of r. Such a mapping is called a *twist map*. As it stands, it is rather simple in that all it does is map points, albeit uniformly or discretely, around a given circle—uniformly for irrational α and discretely for rational α. Clearly, one can think of (4.2.3) as an integrable mapping. A given *circle* $\mathscr{C}$ of points will clearly be mapped into itself. We thus say that a twist map maps circles into circles, and we

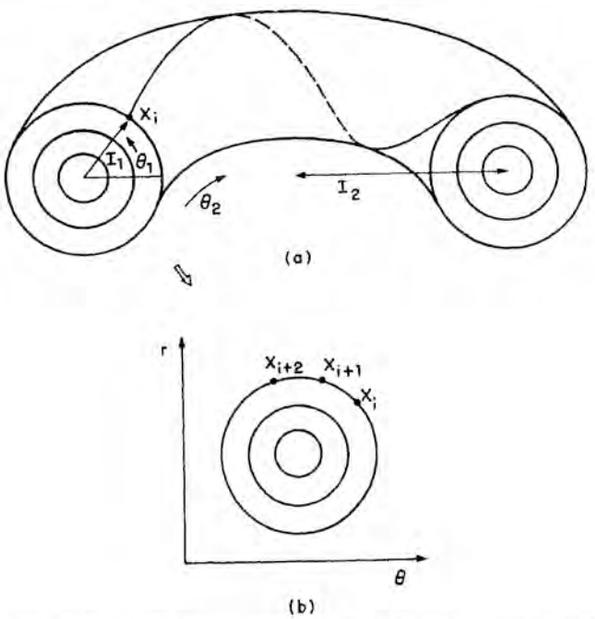


Figure 4.7 (*a*) Point X_i on a torus coordinatized by action-angle variables I_1, θ_1. (*b*) Successive points X_i, X_{i+1}, X_{i+2} of corresponding twist map in $(r(= I_1), \theta(= \theta_1))$-plane.

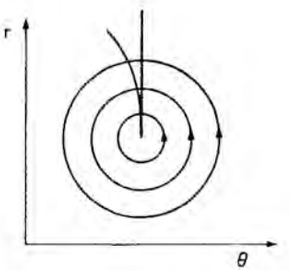


Figure 4.8 Radial line of points twisted under twist-map action.

represent this symbolically as

$$T(\mathscr{C}) = \mathscr{C} \tag{4.2.4}$$

However, since the rotation number $\alpha(r)$ increases with r, a radial line of points will clearly be twisted under T (Figure 4.8). Hence the term *twist* map. The mapping (4.2.3) is obviously area preserving since

$$\frac{\partial(\theta_{i+1}, r_{i+1})}{\partial(\theta_i, r_i)} = 1 \tag{4.2.5}$$

We also comment that, as it stands, it does not really matter whether in Eq. (4.2.3a) we write α as a function of r_i or r_{i+1}. We shall return to this point shortly.

For a nonintegrable system, the KAM theorem tells that tori with rational frequency ratios do not "survive." In terms of the twist map, we can think of the addition of some "nonintegrable" perturbation, that is,

$$T_\epsilon: \quad \theta_{i+1} = \theta_i + 2\pi\alpha(r_i) + \epsilon f(r_i, \theta_i) \tag{4.2.6a}$$

$$r_{i+1} = r_i + \epsilon g(r_i, \theta_i) \tag{4.2.6b}$$

where f and g are chosen to ensure that the area-preserving property (4.2.5) still holds. The natural question to ask about is the preservation of circles under the perturbation. This was Moser's (Moser, 1962) famous contribution to the KAM theorem in which he proved that for sufficiently small perturbation, circles with sufficiently irrational winding numbers are preserved. (See Section 3.5.)

4.2.b Mappings on the Plane

We can also write mappings in cartesian coordinates, that is,

$$T: \quad x_{i+1} = f(x_i, y_i) \tag{4.2.7a}$$

$$y_{i+1} = g(x_i, y_i) \tag{4.2.7b}$$

which will, of course, be area preserving if

$$\frac{\partial(x_{i+1}, y_{i+1})}{\partial(x_i, y_i)} = 1 \tag{4.2.8}$$

If f and g are polynomials, the mapping is termed an *entire Cremona transformation*. The properties of the mapping depend on the form of f and g. If they are only linear functions, for example,

$$x_{i+1} = x_i \cos \alpha - y_i \sin \alpha \tag{4.2.9a}$$

$$y_{i+1} = x_i \sin \alpha + y_i \cos \alpha \tag{4.2.9b}$$

the mapping is just a simple rotation through the angle α. Another simple linear transformation is

$$x_{i+1} = x_i + y_i \tag{4.2.10a}$$

$$y_{i+1} = y_i \tag{4.2.10b}$$

which corresponds to a linear shear parallel to the x-axis.

A nonlinearly perturbed version of (4.2.9), that is,

$$T: \quad \begin{aligned} x_{i+1} &= x_i \cos \alpha - (y_i - x_i^2) \sin \alpha \qquad (4.2.11a) \\ y_{i+1} &= x_i \sin \alpha + (y_i - x_i^2) \cos \alpha \qquad (4.2.11b) \end{aligned}$$

has been the subject of a beautiful study by Henon (1969). (Another piece of obligatory reading!) It is valuable to note that this mapping can be written as the "composition" of two simpler mappings corresponding to a nonlinear shear and a simple rotation (Figure 4.9). That is, one can write

$$T = T_1 T_2 \tag{4.2.12}$$

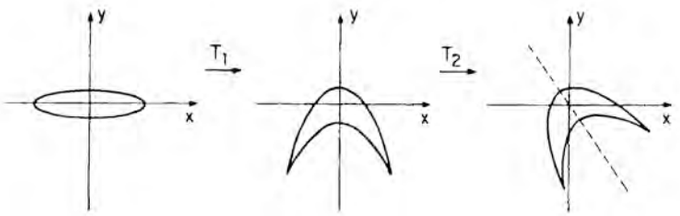


Figure 4.9 Effect on an area element of nonlinear shear T_1 followed by rotation T_2 corresponding to the Henon map.

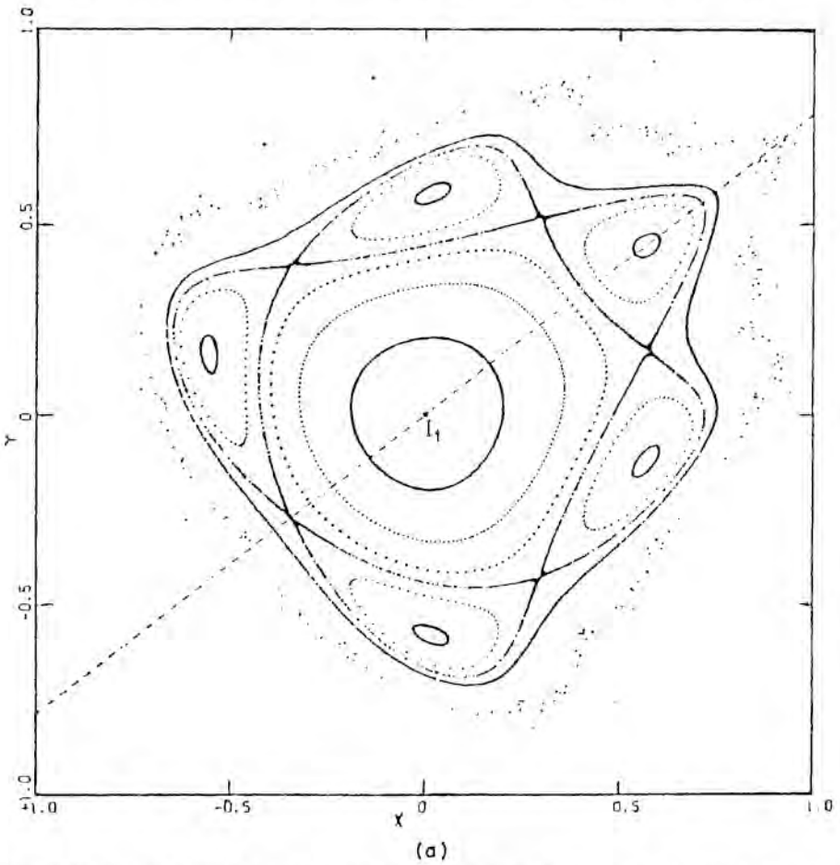


Figure 4.10 (*a*) Typical phase plane of Henon map with $\alpha = 0.2114$. (*b*) Blowup of region around right-hand-most hyperbolic point. (Reproduced, by permission, from Henon (1969).)

where

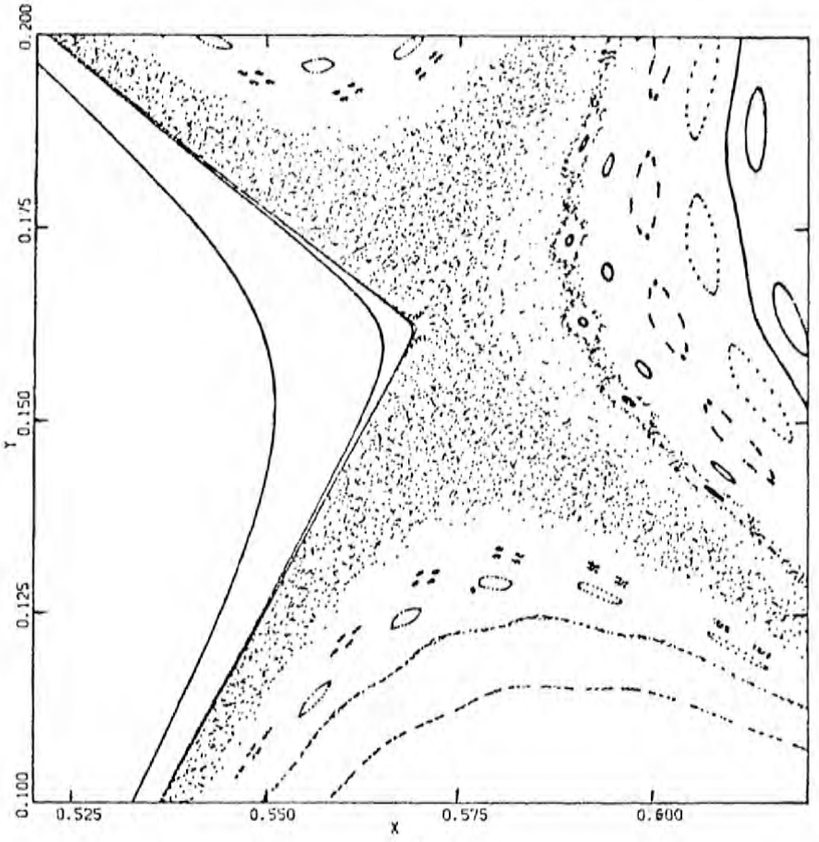
$$T_1: \quad x_{i+1/2} = x_i \tag{4.2.13a}$$

$$y_{i+1/2} = y_i - x_i^2 \tag{4.2.13b}$$

$$T_2: \quad x_{i+1} = x_{i+1/2} \cos \alpha - y_{i+1/2} \sin \alpha \tag{4.2.14a}$$

$$y_{i+1} = x_{i+1/2} \sin \alpha + y_{i+1/2} \cos \alpha \tag{4.2.14b}$$

It is important to note that the mapping (4.2.11) is *invertible*. The inverse transformation T^{-1} takes the form



(b)

Figure 4.10 *(Continued)*

$$T^{-1}: \quad \begin{aligned} x_i &= x_{i+1}\cos\alpha + y_{i+1}\sin\alpha \\ y_i &= -x_{i+1}\sin\alpha + y_{i+1}\cos\alpha + (x_{i+1}\cos\alpha + y_{i+1}\sin\alpha)^2 \end{aligned}$$

Thus some "final" point on an orbit can be uniquely "time-reversed" back to the initial point (x_0, y_0).

A numerical study of Henon's map (4.2.11) is easily performed. For a given rotation angle α, one can study—even on a pocket calculator—the evolving iterates for a variety of different initial conditions (x_0, y_0). Some typical results, as obtained by Henon (1969), are shown in Figure 4.10. They clearly show all the typical features of a Henon–Heiles-like surface of section with families of smooth curves, island chains, and chaotic trajectories. The separatrixlike structure of Figure 4.10*a* is particularly interesting. On this scale, the intersections look almost smooth; however, when

the scale is enlarged, on sees (as shown in the now famous Figure 4.10*b*) an incredibly rich, fine structure of island chains interspersed in a "sea" of chaos. In the sections that follow, we will (attempt to) explain the origins of this wonderful complexity.

4.2.c Connection between Area-Preserving Maps and Hamiltonians

Although Henon's map displays all the generic features of a nonintegrable Hamiltonian, it is not obviously derivable from one. It is therefore worth asking if area-preserving maps can be derived explicitly from Hamiltonian systems. Consider a simple one-degree-of-freedom Hamiltonian of the form

$$H(p, q) = \tfrac{1}{2}p^2 + V(q) \tag{4.2.15}$$

for which Hamilton's equations are simply

$$\dot{q} = p \tag{4.2.16a}$$

$$\dot{p} = -\frac{\partial V}{\partial q} \tag{4.2.16b}$$

One could try writing the time derivatives on the left-hand sides of (4.2.16) as first-order differences, that is, $\dot{q} = (q_{i+1} - q_i)/\Delta t$, where $q_{i+1} = q(t + \Delta t)$ and $q_i = q(t)$. The discretized version of (4.2.6) would then take the form

$$q_{i+1} = q_i + p_i \Delta t \tag{4.2.17a}$$

$$p_{i+1} = p_i - \Delta t \left(\frac{\partial V}{\partial q_i}\right)_{q = q_i} \tag{4.2.17b}$$

However, this is *not* an area-preserving transformation since

$$\frac{\partial(q_{i+1}, p_{i+1})}{\partial(q_i, p_i)} = \begin{vmatrix} 1 & -\Delta t \dfrac{\partial^2 V}{\partial q_i^2} \\ \Delta t & 1 \end{vmatrix} = 1 + (\Delta t)^2 \frac{\partial^2 V}{\partial q_i^2}$$

and we are assuming that Δt is *finite* (cf. discussion of infinitesimal canonical transformations in Chapter 2). However, if the q dependence of (4.2.17b) is changed from q_i to q_{i+1}, that is,

$$q_{i+1} = q_i + p_i \Delta t \tag{4.2.18a}$$

$$p_{i+1} = p_i - \Delta t \left(\frac{\partial V}{\partial q}\right)_{q = q_{i+1}} \tag{4.2.18b}$$

the mapping is easily seen to become area preserving.

It is now interesting to determine what sort of Hamiltonian would give rise to precisely such equations of motion. Instead of (4.2.15), consider a time-dependent Hamiltonian of the form

$$H(p, q, t) = \begin{cases} \dfrac{p^2}{2\gamma}, & 0 < t < \gamma T \qquad (4.2.19a) \\ \dfrac{V(q)}{1-\gamma}, & \gamma T < t < T \qquad (4.2.19b) \end{cases}$$

where $0 < \gamma < 1$. Physically, this corresponds to a situation in which a particle (of unit mass) undergoes free translation for a period γT and then experiences an impulsive force due to the potential $V(q)$, for a period $(1-\gamma)T$—a process that is then repeated periodically. Hamiltonians such as these are used to describe ray propagation in wave guides where the impulse is due to periodically spaced lenses. Integrating Hamilton's equations for (4.2.19) over any one period $t = iT$ to $t = (i+1)T$ yields precisely Eqs. (4.2.18) (with Δt replaced by T). We can also interchange the order of operations in (4.2.19), that is,

$$H(p, q, t) = \begin{cases} \dfrac{V(q)}{\gamma}, & 0 < t < \gamma T \qquad (4.2.20a) \\ \dfrac{p^2}{2(1-\gamma)}, & \gamma T < t < T \qquad (4.2.20b) \end{cases}$$

in which case integration of Hamilton's equations yields

$$q_{i+1} = q_i + T p_{i+1} \qquad (4.2.21a)$$

$$p_{i+1} = p_i - T\left(\frac{\partial V}{\partial q}\right)_{q=q_i} \qquad (4.2.21b)$$

which is again an area-preserving map on the plane.

There is, of course, really no difference between the Hamiltonians (4.2.19) and (4.2.20), although the mapping equations for the latter have a certain nice symmetry about them as follows.

4.2.d Discrete Lagrangians*

For mappings such as (4.2.21), we can introduce the discrete Lagrangian

$$L(q_{i+1}, q_i) = \frac{1}{2}\left(\frac{q_{i+1} - q_i}{T}\right)^2 - V(q_i) \qquad (4.2.22)$$

and the corresponding discrete action function

$$W(q_{i+1}, q_i) = TL(q_{i+1}, q_i) = \frac{1}{2}\frac{(q_{i+1} - q_i)^2}{T} - TV(q_i) \tag{4.2.23}$$

from which one easily deduces the symmetric pair of generating relations

$$p_i = -\frac{\partial W}{\partial q_i}(q_{i+1}, q_i), \qquad p_{i+1} = \frac{\partial W(q_{i+1}, q_i)}{\partial q_{i+1}} \tag{4.2.24}$$

Thus one can develop a rather elegant variational formalism for area-preserving maps closely analogous to the standard classical mechanical results. The discrete Lagrangian formalism will be used in our discussion of "quantum maps" in Sections 6.6 and 6.7.

4.2.e The Standard Map

A much-studied mapping is the one obtained by introducing the potential-energy function

$$V(q) = -\frac{k}{(2\pi)^2}\cos(2\pi q) \tag{4.2.25}$$

into Eqs. (4.2.21), thereby obtaining the equations of motion

$$q_{i+1} = q_i + p_{i+1} \tag{4.2.26a}$$

$$p_{i+1} = p_i + \frac{k}{2\pi}\sin(2\pi q_i) \tag{4.2.26b}$$

where we have set $T = 1$. This mapping is usually studied with both p and q taken as periodic variables with period unity; that is, the mapping is confined to the unit torus, and we rewrite (4.2.26) as

$$q_{i+1} = q_i + p_{i+1}, \qquad \text{mod } q = 1 \tag{4.2.27a}$$

$$p_{i+1} = p_i + \frac{k}{2\pi}\sin(2\pi q_i), \qquad \text{mod } p = 1 \tag{4.2.27b}$$

This is the Taylor-Chirikov or "standard" mapping, the latter name stemming from its ubiquity in a variety of theoretical and practical problems. A typical phase plane, for $k = 0.97$, is shown in Figure 4.11. Again we see a wonderfully rich structure of delicately interwoven regular and chaotic motions. The reader may also have noticed that some of the invariant curves (tori) extending across the phase plane are apparently permeated

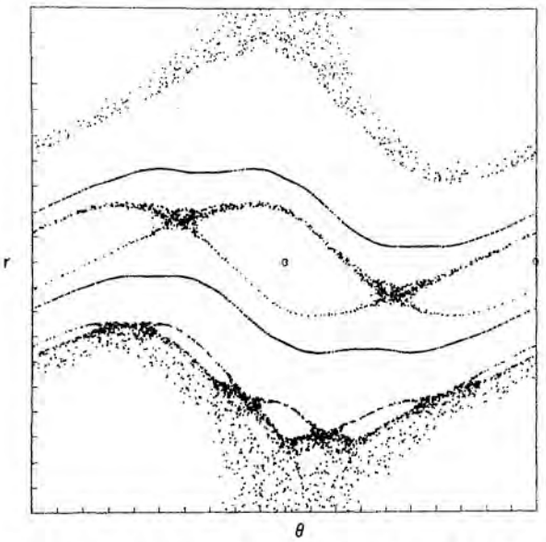


Figure 4.11 Typical phase plane of standard map with $k = 0.97$. (Reproduced, by permission, from Greene (1979).)

with small holes. This is not a numerical artifact but is, instead, a manifestation of the fact that the invariant curves are not really tori but are the so-called *cantori*. This latter term has been invoked to suggest the connection with a Cantor-set-like structure. These exciting notions are outside the scope of these lectures, although we will return to the Cantor-set concept in dynamical systems in the next chapter.

4.3 FIXED POINTS AND THE POINCARÉ–BIRKHOFF FIXED POINT THEOREM

For a mapping T, a fixed point (X^*) of T is a point for which

$$TX^* = X^* \tag{4.3.1}$$

In the case of a periodic orbit which maps out a sequence of n iterates (i.e., $X_0, X_1, \ldots, X_n$) such that $X_{n+i} = X_i$, it is clear that each X_i is a fixed point of T^n, where T^n denotes n successive applications of the mapping T, that is,

$$T^n X_i = X_i \tag{4.3.2}$$

As we found in our discussions in Chapter 1 on phase-plane dynamics, fixed points are convenient centers about which the dynamics can be

"organized." Indeed, the fixed-point analysis for mappings is virtually identical to our previous results with the important difference that the area-preserving property of T considerably restricts the allowed types of fixed points.

4.3.a The Tangent Map

Consider some mapping T whose action we denote symbolically as

$$\begin{bmatrix} x_{i+1} \\ y_{i+1} \end{bmatrix} = T \begin{bmatrix} x_i \\ y_i \end{bmatrix} \tag{4.3.3}$$

This could represent a transformation of the sort given in Eqs. (4.2.7). For simplicity, let us assume there is a fixed point of T at the phase-plane origin $(x, y) = (0, 0)$. Linearizing T in the standard way about this point, we have the (linear) mapping, often referred to as the *tangent map*:

$$\begin{bmatrix} \delta x_{i+1} \\ \delta y_{i+1} \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \begin{bmatrix} \delta x_i \\ \delta y_i \end{bmatrix} \tag{4.3.4}$$

(so, for example, for Eqs. (4.2.7), $T_{11} = (\partial f/\partial x)_{x=y=0}$, etc.). The nature of the fixed point is then determined by the eigenvalues of (4.3.4), that is,

$$\begin{vmatrix} T_{11} - \lambda & T_{12} \\ T_{21} & T_{22} - \lambda \end{vmatrix} = 0 \tag{4.3.5}$$

These eigenvalues are found by solving the quadratic equation

$$\lambda^2 - \lambda(T_{11} + T_{22}) + \lambda(T_{11}T_{22} - T_{12}T_{21}) = 0$$

which we write as

$$\lambda^2 - \lambda(\text{trace}(T)) + \lambda(\det(T)) = 0$$

Since T is an area-preserving transformation (i.e., $\det(T) = 1$), the roots are simply

$$\lambda_{1,2} = \tfrac{1}{2}(\text{trace}(T)) \pm \tfrac{1}{2}\sqrt{(\text{trace}(T))^2 - 4} \tag{4.3.6}$$

There are three possibilities, depending on the value of $\text{trace}(T)$:

1. $|\text{trace}(T)| < 2$; λ_1, λ_2 are a complex conjugate pair lying on the unit circle, that is,

$$\lambda_1 = e^{+i\alpha}, \qquad \lambda_2 = e^{-i\alpha} \tag{4.3.7}$$

2. $|\text{trace}(T)| > 2$; the eigenvalues are real numbers satisfying

$$\lambda_2 = 1/\lambda_1 \tag{4.3.8}$$

3. $|\text{trace}(T)| = 2$; the eigenvalues take on the value

$$\lambda_1 = \lambda_2 = \pm 1 \tag{4.3.9}$$

Using the standard techniques of linear algebra, it is always possible to transform to a representation that diagonalizes (4.3.4), that is,

$$\begin{bmatrix} \xi_{i+1} \\ \eta_{i+1} \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} \xi_i \\ \eta_i \end{bmatrix}$$

where, given the transformation matrix A,

$$A \begin{bmatrix} \delta x_i \\ \delta y_i \end{bmatrix} = \begin{bmatrix} \xi_i \\ \eta_i \end{bmatrix} \quad \text{and} \quad ATA^{-1} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

The three different eigenvalue cases are now easily interpreted.

4.3.b Classification of Fixed Points

The first case (i.e., $\lambda_{1,2} = e^{\pm i\alpha}$) is simply the rotation

$$\begin{bmatrix} \xi_{i+1} \\ \eta_{i+1} \end{bmatrix} = \begin{bmatrix} e^{i\alpha} & 0 \\ 0 & e^{-i\alpha} \end{bmatrix} \begin{bmatrix} \xi_i \\ \eta_i \end{bmatrix} \tag{4.3.10}$$

in the neighborhood of the fixed point (0, 0). This obviously corresponds to a stable or elliptic fixed point. Thus in the immediate neighborhood of (0, 0) we expect to find invariant curves (see Figure 1.10c). (Also see the discussion of the KAM theorem for equilibrium points in Section 3.5.)

For the second case (i.e., $\lambda_2 = 1/\lambda_1$), the linearized transformation takes the form

$$\begin{bmatrix} \xi_{i+1} \\ \eta_{i+1} \end{bmatrix} = \begin{bmatrix} \lambda & 0 \\ 0 & 1/\lambda \end{bmatrix} \begin{bmatrix} \xi_i \\ \eta_i \end{bmatrix} \tag{4.3.11}$$

which leads to hyperbolic motion in the neighborhood of (0, 0). The precise behavior of (4.3.11) will be determined by the sign of λ:

Case (a) $\lambda > 0$ gives a regular *hyperbolic* fixed point in which successive iterates of (4.3.11) stay on the same branch of the hyperbola (see Figure 4.12*a*),

Case (b) $\lambda > 0$ gives a *hyperbolic-with-reflection* fixed point in which successive iterates of (4.3.11) jump backwards and forwards between opposite branches of the hyperbola (see Figure 4.12*b*). This is easy to see

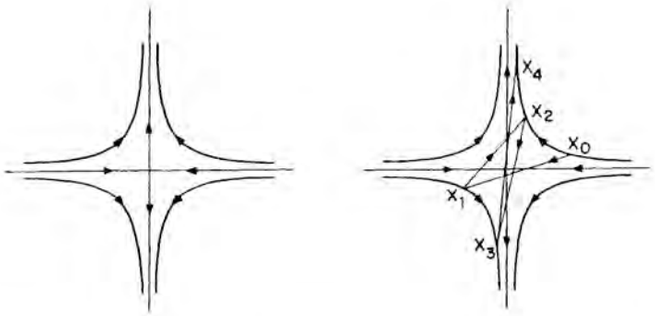


Figure 4.12 (*a*) Hyperbolic fixed point. (*b*) Hyperbolic-with-reflection fixed point.

since

$$\begin{bmatrix}\delta\xi_1\\ \delta\eta_1\end{bmatrix}=\begin{bmatrix}-|\lambda|\delta\xi_0\\ -\dfrac{1}{|\lambda|}\delta\eta_0\end{bmatrix},\quad \begin{bmatrix}\delta\xi_2\\ \delta\eta_2\end{bmatrix}=\begin{bmatrix}|\lambda|^2\delta\xi_0\\ \dfrac{1}{|\lambda|^2}\delta\eta_0\end{bmatrix},\quad \begin{bmatrix}\delta x_3\\ \delta y_3\end{bmatrix}=\begin{bmatrix}-|\lambda|^3\delta\xi_0\\ \dfrac{1}{|\lambda|^3}\delta\eta_0\end{bmatrix},\quad \text{etc.}$$

The third case, which is a special case corresponding to $\lambda_1=\lambda_2=\pm 1$, is best understood by recognizing that in the original variables $(\delta x_i, \delta y_i)$ the (linearized) transformation (4.3.4) can always be written as (choosing $\lambda_1 = +1$)

$$\begin{bmatrix}\delta x_{i+1}\\ \delta y_{i+1}\end{bmatrix}=\begin{bmatrix}1 & c\\ 0 & 1\end{bmatrix}\begin{bmatrix}\delta x_i\\ \delta y_i\end{bmatrix} \tag{4.3.12}$$

where c is any constant, which corresponds to a translation parallel to the x-axis. This is known as a *parabolic* fixed point (see Figure 4.13). For the particular choice $\delta y_0 = 0$, we note that every point on the x-axis is a fixed point of (4.3.12). This is the situation that arises for invariant tori, or

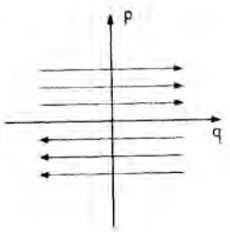


Figure 4.13 Parabolic fixed point.

curves, covered with closed orbits, with every point on the curve (in the phase plane or on the surface of section) being a fixed point of the flow.

4.3.c Poincaré–Birkhoff Fixed-Point Theorem

We are now in the position to tackle, in some detail, the fundamental question of the fate of tori with rational frequency ratios, or curves with rational rotation number, under (small) perturbation. It is convenient to approach this problem from the point of view of twist maps, which we write in the form (see Section 4.3) that makes the connection with two-degree-of-freedom Hamiltonians. Again thinking of the (transversal) intersection of family of tori, we can write the unperturbed twist map as

$$\varphi' = \varphi + \frac{\partial}{\partial I'} S_0(I') \tag{4.3.13a}$$

$$I' = I \tag{4.3.13b}$$

where, in the case of 2-D tori, we note that $\partial S_0/\partial I = 2\pi\omega_1/\omega_2$. (Here the primed variables indicate the $(i+1)$th iterate of the mapping and unprimed variables the ith iterate.) The KAM theorem tells us that for sufficiently small perturbation $\epsilon S_1(I, \varphi)$, that is, the mapping

$$\varphi' = \varphi + \frac{\partial}{\partial I'} S_0(I') + \epsilon \frac{\partial}{\partial I'} S_1(I', \varphi) \tag{4.3.14a}$$

$$I' = I + \epsilon \frac{\partial}{\partial \varphi} S_1(I, \varphi) \tag{4.3.14b}$$

"most" invariant curves will be preserved, provided that the nondegeneracy condition

$$\det\left|\frac{\partial^2 S_0}{\partial I_i \, \partial I_j}\right| \neq 0 \tag{4.3.15}$$

is satisfied, with the term "most" excluding those curves with rational rotation number $\alpha = \omega_1/\omega_2 = r/s$.

We can use these maps to study the precise fate of the rational curves under perturbation. Consider two curves $\mathscr{C}^+$ and $\mathscr{C}^-$ which lie on either side of the curve $\mathscr{C}$ with rational rotation number $\alpha = r/s$, as sketched in Figure 4.14*a*. We also assume that $\alpha = \alpha(I)$ increases smoothly with increasing I. If we denote the mapping by T, that is,

$$\begin{bmatrix} \varphi' \\ I' \end{bmatrix} = T \begin{bmatrix} \varphi \\ I \end{bmatrix} \tag{4.3.16}$$

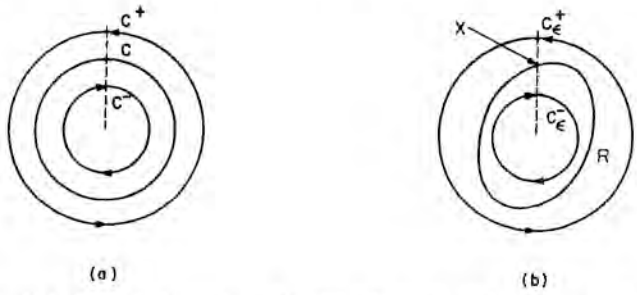


Figure 4.14 (*a*) Invariant curves of unperturbed twist map T with rotation numbers $\alpha < r/s$ for $\mathscr{C}^-$, $\alpha = r/s$ for $\mathscr{C}$, and $\alpha > r/s$ for $\mathscr{C}^+$. (*b*) Effect of perturbed twist map T_ϵ on these curves. Since relative twists of $\mathscr{C}^+$ and $\mathscr{C}^-$ are preserved, there will be one point X between them whose angular coordinate is preserved. Curve R drawn between $\mathscr{C}^+$ and $\mathscr{C}^-$ is curve of these points.

then *every* point on $\mathscr{C}$ is a fixed point of T^s since

$$T^s\begin{bmatrix}\varphi\\ I\end{bmatrix} = \begin{bmatrix}\varphi + s\left(\dfrac{\partial S_0}{\partial I}\right)\\ I\end{bmatrix} = \begin{bmatrix}\varphi + s2\pi\left(\dfrac{r}{s}\right)\\ I\end{bmatrix} = \begin{bmatrix}\varphi + 2\pi r\\ I\end{bmatrix} = \begin{bmatrix}\varphi\\ I\end{bmatrix} \tag{4.3.17}$$

Thus, *relative* to $\mathscr{C}$, $\mathscr{C}_+$ rotates anticlockwise and $\mathscr{C}_-$ rotates clockwise under the mapping T^s.

Now consider the (weakly) perturbed mapping T_ϵ. By the KAM theorem, $\mathscr{C}^+$ and $\mathscr{C}^-$ are preserved, albeit in slightly distorted form, say $\mathscr{C}^+_\epsilon$ and $\mathscr{C}^-_\epsilon$. These curves are now invariant curves of T_ϵ, that is,

$$T_\epsilon(\mathscr{C}^+_\epsilon) = \mathscr{C}^+_\epsilon \quad \text{and} \quad T_\epsilon(\mathscr{C}^-_\epsilon) = \mathscr{C}^-_\epsilon \tag{4.3.18}$$

Furthermore, we assume that ϵ is sufficiently small such that the relative twists of $\mathscr{C}^+$ and $\mathscr{C}^-$ are preserved under T^s_ϵ. If this is so, then there must be only one point between $\mathscr{C}^+_\epsilon$ and $\mathscr{C}^-_\epsilon$ whose angular coordinate φ is preserved under T^s_ϵ. In fact, along each radius (emanating from the center) there must be one such point, so we can draw a curve R of these points (Figure 4.14*b*). R is not an invariant curve of T_ϵ, but on R must lie the fixed points of T^s_ϵ. (A fixed point has preserved "angle" and "radius"—so far we just have the curve of preserved angles.) These may be found by subjecting R to the mapping T^s_ϵ, that is,

$$R' = T^s_\epsilon R \tag{4.3.19}$$

The new curve R' will intersect R at an *even* number of points (this follows

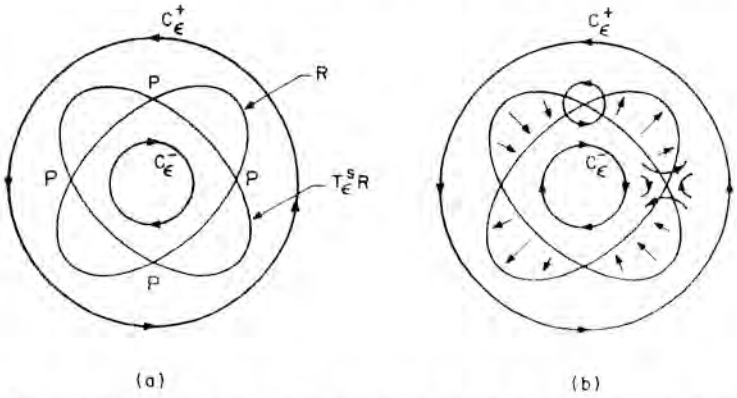


Figure 4.15 (a) Mapping of curve of points R into new curve $R' = T_\epsilon^s R$. Points P are the intersection points of these two curves. (*b*) By following "flow lines," these points are seen to be alternatively hyperbolic and elliptic fixed points.

from simple geometry)—these are the fixed points of T_ϵ^s (see Figure 4.15*a*). (We exclude from this argument any nongeneric points of tangency between R and R'.) This is the famous *Poincaré–Birkhoff fixed-point theorem*, which states that for the rational curve of an unperturbed system, with rotation number r/s (for which *every* point is a fixed point of T^s), only an even number of fixed points $2ks$ ($k = 1, 2, 3, \ldots$) will remain under perturbation. (We shall soon see that these fixed points are alternately stable and unstable.) That the number of fixed points is an even multiple of s is easy to see. Consider one of the (even number of) fixed points found from the intersection of R and R'. By definition, it is a fixed point of T_ϵ^s. Under T_ϵ the orbit generated is X, $T_\epsilon X$, $T_\epsilon^2 X, \ldots, T_\epsilon^{s-1} X$. However, each point of this closed orbit is also a fixed point of T_ϵ^s. Hence there are s fixed points associated with each intersection point of R and R'—hence $2ks$ fixed points overall.

Returning to Figure 4.15, all we have to do is follow the "flow lines" to see that the fixed points are *alternately* elliptic and hyperbolic. Thus, under perturbation of the rational curve with $\alpha = r/s$, $2ks$ fixed points of T_ϵ^s remain, of which ks are elliptic and ks are hyperbolic in alternating sequence. Now, around each elliptic fixed-point we will find a family of invariant curves. This family is itself subject to the KAM theorem (see Section 3.5.d), so its rational members will break up according to the Poincaré–Birkhoff fixed point theorem. The same structure must then be repeated about this subsequence of elliptic fixed points, and so on. Thus, around each elliptic fixed point there is a simultaneous application of the Poincaré–Birkhoff fixed-point theorem and KAM theorem which leads to a remarkable self-similar structure on all scales, as sketched in Figure 4.16.

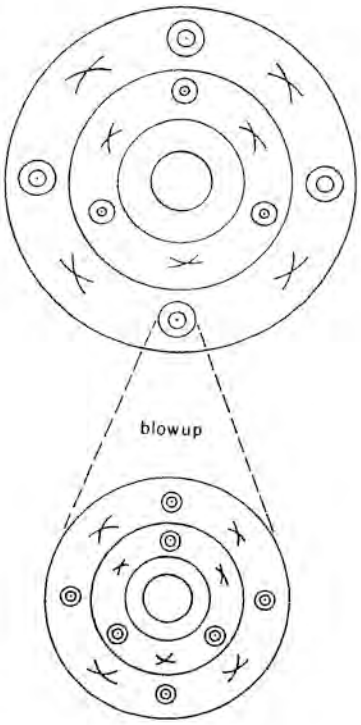


Figure 4.16 Successive applications of KAM and Poincaré–Birkhoff fixed-point theorems leading to self-similar fixed-point structure. (× denotes hyperbolic fixed points.)

4.4 HOMOCLINIC AND HETEROCLINIC POINTS

To complete the picture shown in Figure 4.16, we now have to consider what happens in the neighborhood of the hyperbolic fixed points. As we shall see, the results are very striking. Here we follow the discussion given by Berry (1978).

A hyperbolic fixed point is characterized by four invariant curves or manifolds. These are the two ingoing or *stable manifolds*, H^+, and the two outgoing or *unstable manifolds*, H^-, as shown in Figure 4.17. Points on H^+

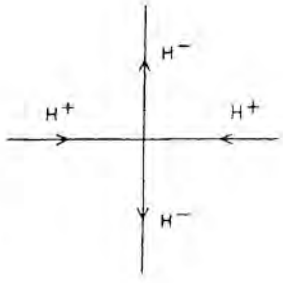


Figure 4.17 Stable manifolds H^+ and unstable manifolds H^- of a hyperbolic fixed point.

approach the fixed point H exponentially slowly, that is,

$$\lim_{s\to\infty} T^s x \to H, \qquad (x \text{ on } H^+) \tag{4.4.1}$$

whereas points on H^- recede from H exponentially slowly, that is,

$$\lim_{s\to\infty} T^{-s} x \to H, \qquad (x \text{ on } H^-) \tag{4.4.2}$$

4.4.a The Intersections of H^+ and H^-

We will now investigate the way in which H^+ and H^- can "interact" with each other. As we have already seen, for integrable systems the H^+ and H^- manifolds emanating from a hyperbolic fixed point (see, e.g., the pendulum) form a separatrix. In Figure 4.18*a* we show the case of an outgoing manifold joining up smoothly with an ingoing manifold to form a single smooth loop. Such a curve is sometimes called a *homoclinic orbit*. In Figure 4.18*b* we show another example in which the H^+ and H^- from a family of three hyperbolic fixed points (i.e., fixed points of T^3) all join up smoothly in the manner shown. This is not dissimilar (only superficially!) to the structure seen in the $E = \frac{1}{12}$ surface of section for the Henon–Heiles system (Figure 4.4*a*).

This smooth joining of manifolds is the exceptional (i.e., nongeneric) situation that can only arise in integrable systems. The generic situation is far more complex. The manifolds H^+ and H^- are not allowed to intersect themselves but, instead, intersect each other as shown in Figure 4.19. If the intersection point(s) involve the H^+ and H^- manifolds from the same fixed point or from the fixed points of the same family (e.g., the three fixed points of T^3 shown in the inner part of Figure 4.16), they are called *homoclinic*

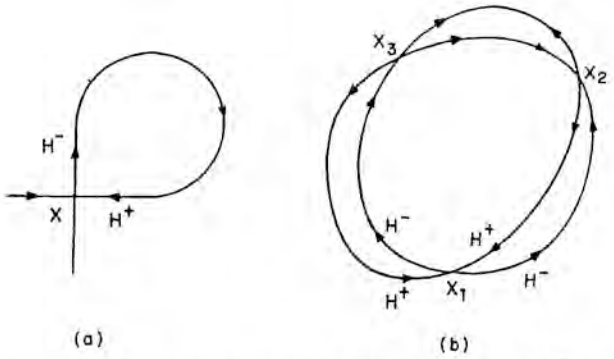


Figure 4.18 (*a*) Smooth joining of H^- to H^+ from same hyperbolic fixed point X leading to homoclinic orbit. (*b*) Family of three smoothly connected hyperbolic points X_1, X_2, X_3.

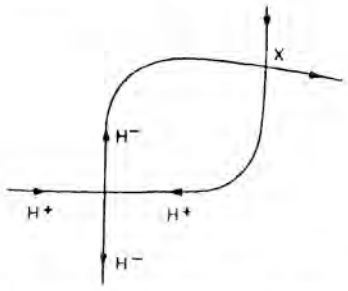


Figure 4.19 Intersection of stable manifold H^+ and unstable manifold H^-, emanating from the same hyperbolic fixed point, to give homoclinic point X. It should be emphasized that the curves drawn here do not correspond to a (single) trajectory but, rather, correspond to the curve(s) drawn through the successive intersections of a trajectory with the plane.

points. If the intersecting H^+ and H^- emanate from the fixed points of different families (e.g., the fixed points of T^3 and T^4 shown in the inner and outer parts of Figure 4.16), they are called *heteroclinic points*. However, these intersections are not at all simple! Consider the homoclinic point X in Figure 4.20*a* and its adjacent points X' and X''. These two points map, as shown, to TX' and TX'', respectively. The problem is that, since X is "ahead" of both X' and X'', its image TX is required, by the continuity of the mapping, to be "ahead" of both TX' and TX''. This is clearly impossible. However, the contradiction can be resolved by making a loop in the manifold as shown in Figure 4.20*b*. Now TX is a new intersection (i.e., homoclinic) point. Continuing the argument, TX must map to a new homoclinic point T^2X via a second loop as shown in Figure 4.20*c*. Furthermore, since T^2X is closer to the hyperbolic point than TX, the distance between T^2X and TX will be less than that between TX and X. By area preservation, the area in the loops between X and TX and T^2X must be the same. This being so, the second loop must be longer and thinner than the first. Further repetition of this argument results in an infinite number of intersections, that is, the whole area becomes dense with homoclinic points with the intervening loops becoming ever longer and thinner. The overall picture (Figure 4.21) is one of incredible complexity. This was appreciated by Poincaré, who wrote about it in his seminal treatise *Les Methodes Nouvelles de la Mechanique Celeste* (Poincaré, 1892) as follows:

> The intersections form a kind of lattice, web or network with infinitely tight loops; neither of the two curves (H^+ and H^-) must ever intersect itself but it must bend in such a complex fashion that it intersects all the loops of the network infinitely many times.

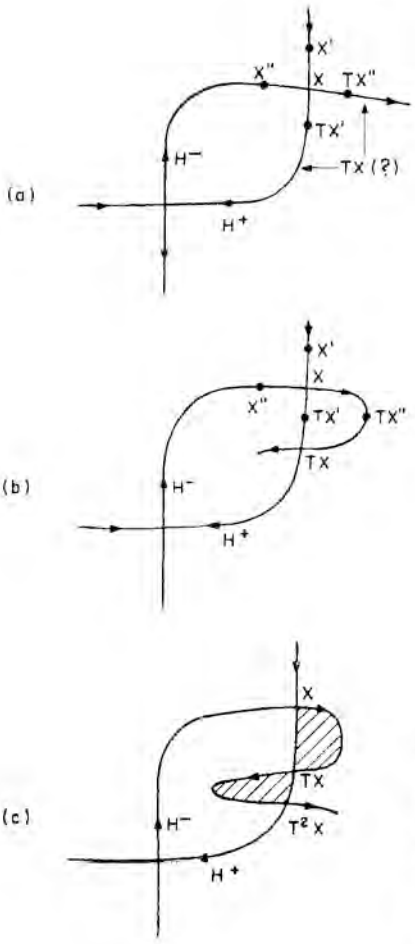


Figure 4.20 (*a*) Neighboring points X' and X'' being mapped to TX' and TX'' and apparently nonunique image, TX, of homoclinic point X. (*b*) Creating a unique image TX by inserting a loop in the manifold. (*c*) Image T^2X of TX creating a longer and thinner loop enclosing the same area.

> One is struck by the complexity of this figure which I am not even attempting to draw. Nothing can give us a better idea of the complexity of the three body problem and of all the problems in dynamics where there is no holomorphic integral and Bohlin's series diverge.

We can now fill in, at least approximately, the details of Figure 4.16 and obtain the following picture (Figure 4.22). It is remarkable to think that this structure repeats on all scales and, furthermore, is the generic situation for nonintegrable systems. Thus, as remarkable as the detailed dynamics (shown in Figure 4.10*b*) in the neighborhood of the hyperbolic fixed point of Henon's map might have seemed at first, it should now not be so surprising.

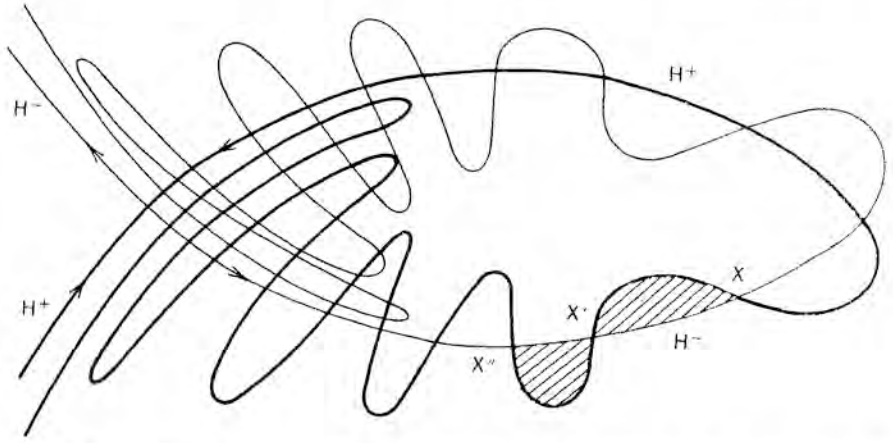
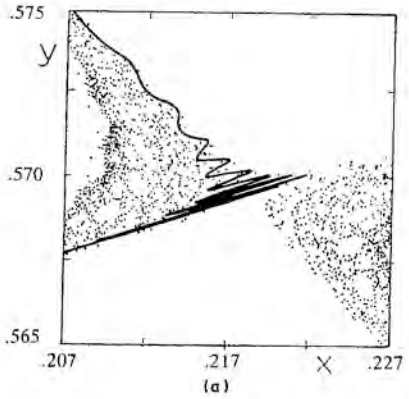


Figure 4.21 Network of intersections of H^+ and H^- leading to dense area of homoclinic points and network of ever longer and thinner loops of the same area (hatched regions).

4.4.b Whorls and Tendrils

In trajectory-by-trajectory studies such as those shown for Henon's map, we see the "seas" of chaos but, of course, we do not see curves corresponding to the wildly intersecting manifolds H^+ and H^-. In order to get a sense of what this structure actually looks like, we have to iterate a whole line element (each point of which corresponds to a different initial condition) under the mapping. Such a computation is shown in Figure 4.23 for a line element in the neighborhood of a hyperbolic fixed point of Henon's map.

Figure 4.22 Typical self-similar network of elliptic and hyperbolic fixed points with associated homoclinic webs. (Reproduced, by permission, from Ford (1975).)



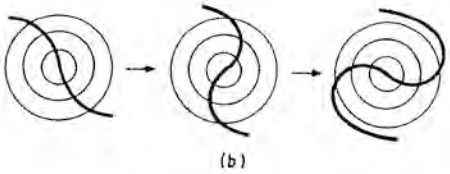


Figure 4.23 (*a*) Iteration of line element in neighborhood of hyperbolic fixed point leading to a "tendril." (Reproduced, by permission, from Berry (1978).) (*b*) Iteration of line element in neighborhood of elliptic fixed point leading to a "whorl."

We call these oscillations of the line element, with their characteristic flailing structure, *homoclinic oscillations*. In fact, in any (strongly) chaotic region (not necessarily the neighborhood of just one hyperbolic fixed point), a line element will evolve in this way—stretching exponentially fast and flailing backwards and forwards. This characteristic feature, found for evolving line elements on the plane, has been termed a *tendril*. This is in contrast to the behavior of a line element evolving in the neighborhood of an elliptic fixed point. On the basis of our discussions of twist maps, it is not difficult to deduce that the line element will form a tightly curling structure that we term a *whorl*. The whorl–tendril features are the "real-world" manifestations of chaos in area-preserving maps. As we will describe in Section 4.8, the lovely patterns that one sees in cream floating on the surface of a cup of coffee, or of gasoline films flowing down the street on a rainy day, can be explained precisely in these terms.

Finally, we make a remark about heteroclinic points. Since these correspond to the intersections of H^+ and H^- from different fixed-point families (e.g., T^3 and T^4 of Figure 4.16), it is likely that some intervening invariant curve (i.e., some torus with (very) irrational winding number) will also have to break down. Clearly, this will require a stronger perturbation. Thus one can expect the appearance of heteroclinic points to herald the onset of fairly widespread chaos. Criteria for the onset of such widespread chaos are important since it is believed that this will, for example, enhance transport processes. A number of techniques to predict this transition have been developed, including the method of overlapping resonances due to

Chirikov and the method of residues due to Greene. These will be discussed in subsequent sections.

4.5 CRITERIA FOR LOCAL CHAOS

4.5.a Lyapunov Exponents

An important characteristic of chaotic motion is the great sensitivity of the motion to small changes in initial conditions. Closely neighboring trajectories are found to diverge *exponentially*, whereas regular trajectories are found to separate only *linearly* in time. (Note, of course, that this divergence cannot go on forever in a bounded phase space.) The rate of divergence can be precisely quantified in terms of *Lyapunov exponents*, which measure the mean rate of exponential separation of neighboring trajectories. In fact, Lyapunov exponents are an extremely useful way of characterizing dynamical systems, and their use is by no means restricted to the Hamiltonian systems considered in this chapter. So, for generality, we will consider some (autonomous) system governed by the differential equations

$$\frac{dx_i}{dt} = F_i(x_1, \ldots, x_n), \qquad i = 1, \ldots, n \tag{4.5.1}$$

In our examinations of the stability of a given fixed point, we linearized the equations of motion about that point. Now we linearize the equations about any reference *orbit* $\bar{\mathbf{x}} = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)$ to yield the tangent map

$$\frac{d\delta x_i}{dt} = \sum_{j=1}^{n} \delta x_j \left(\frac{\partial F_i}{\partial x_j}\right)_{\mathbf{x}=\bar{\mathbf{x}}(t)} \tag{4.5.2}$$

The norm

$$d(t) = \sqrt{\sum_{i=1}^{n} \delta x_i^2(t)} \tag{4.5.3}$$

provides a measure of the divergence of two neighboring trajectories, that is, the reference trajectory $\bar{\mathbf{x}}$ and its neighbor with initial conditions $\bar{\mathbf{x}}(0) + \delta\mathbf{x}(0)$. The mean rate of exponential divergence is defined as

$$\sigma = \lim_{\substack{t\to\infty \\ d(0)\to 0}} \left(\frac{1}{t}\right) \ln\left(\frac{d(t)}{d(0)}\right) \tag{4.5.4}$$

where $d(0) = \sqrt{\sum_{i=1}^{n} \delta x_i^2(0)}$. In addition, as will be explained below, it can be shown that there exists a set of n such quantities σ_i, $i = 1, \ldots, n$. These σ_i are called the *Lyapunov characteristic exponents* and they can be ordered by size, that is

$$\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \cdots \geq \sigma_n \tag{4.5.5}$$

For regular motion, however, the exponents are zero since $d(t)$ grows only linearly (or possibly algebraically) with time

In order to understand these ideas more fully, it is useful to first of all consider the Lyapunov exponents of mappings. Indeed, the simplest possible case is a *one*-dimensional map of the form

$$x_{i+1} = f(x_i) \tag{4.5.6}$$

where $f(x)$ is some (simple) nonlinear function of x, for example, $f(x) = 4\lambda x(1-x)$. The remarkable dynamical properties of such a map will be described at greater length in Chapter 5. The tangent map is simply

$$x_{i+1} = f'(x_i)\,\delta x_i = \prod_{j=0}^{i} f'(x_j)\,\delta x_0 \tag{4.5.7}$$

where $f'(x_j)$ is the derivative of $f(x)$ evaluated at each point x_j along the given trajectory. The associated Lyapunov exponent is easily deduced by analogy with (4.5.3) to be

$$\sigma = \lim_{N\to\infty} \frac{1}{N} \ln\left[\prod_{j=1}^{N} f'(x_j)\,\delta x_0\right]$$

$$= \lim_{N\to\infty} \frac{1}{N} \sum_{j=0}^{N} \ln|f'(x_j)| \tag{4.5.8}$$

The exponent σ is independent of the initial point x_0 (apart from a set of measure zero initial conditions). In the case of $f(x) = x^2$ the reader should have little difficulty in verifying that $\sigma = \ln 2$.

In the case of multidimensional mappings

$$\mathbf{x}_{i+1} = \mathbf{F}(\mathbf{x}_i)$$

where $\mathbf{x}$ and $\mathbf{F}$ are n-dimensional vectors, there will be a set of n characteristic exponents corresponding to the n eigenvalues of the associated tangent map. Introducing the eigenvalues $\lambda_i(N)$, $i = 1, \ldots, n$, of the matrix

$$(TM)_N = (M(\mathbf{x}_N)M(\mathbf{x}_{N-1}) \cdots M(x_1))^{1/N} \tag{4.5.9}$$

where $M(\mathbf{x}_i)$ is the linearization of $\mathbf{F}$ at the point $\mathbf{x}_i$, the exponents are defined as

$$\sigma_i = \lim_{N\to\infty} \ln|\lambda_i(N)|, \qquad i = 1, \ldots, n \tag{4.5.10}$$

It should be clear that for area-preserving maps and Hamiltonian flows, the sum of the exponents must be zero in order to ensure that phase volume is preserved.

We now turn to the case of flows, as governed by Eqs. (4.5.1), and write

(4.5.2) in the vector form

$$\frac{d}{dt}\boldsymbol{\delta}\mathbf{z} = M\boldsymbol{\delta}\mathbf{z} \tag{4.5.11}$$

where $\boldsymbol{\delta}\mathbf{z} = (\delta x_1, \ldots, \delta x_n)$ and M is the linearized matrix with elements $(M)_{ij} = (\partial F_i/\partial x_j)_{\mathbf{x}=\mathbf{x}(t)}$. There will exist a set of basis vectors $\hat{e}_i$ $(i = 1, \ldots, n)$ such that $\boldsymbol{\delta}\mathbf{z} = \sum_{i=1}^{n} a_i\hat{e}_i$. The stretching (or contracting) rates in each of the directions $\hat{e}_i$ provide us with the set of exponents σ_i $(i = 1, \ldots, n)$, which can be ordered as in (4.5.5). Clearly, as time evolves, a small volume element will be stretched most in the direction $\hat{e}_i$ with the largest exponent. Thus, in practice, (4.5.4) will yield just this exponent (σ_1, according to (4.5.5)). For Hamiltonian systems with n degrees of freedom, the vector $\boldsymbol{\delta}\mathbf{z}$ becomes $2n$-dimensional (i.e., $\boldsymbol{\delta}\mathbf{z} = (\delta q_1, \ldots, \delta q_n, \delta p_1, \ldots, \delta p_n)$), and there will be $2n$ exponents. However, now there is a special symmetry between the σ_i, namely,

$$\sigma_i = -\sigma_{2n-i+1} \tag{4.5.12}$$

Thus any stretching in one "direction" is canceled by contraction in another, thereby ensuring Liouville's theorem. If the exponents are calculated on a given energy shell, the space is $2n-1$-dimensional. Thus it follows from (4.5.12) that two (or more, depending on the dynamics) of the σ_i must be zero.

The actual computation of exponents for n-dimensional flows (rather than mappings) is nontrivial. Consider, for example, working with the definition (4.5.4). If the norm $d(t)$ increases exponentially, there will be the risk of computer overflow and related errors. Instead, one uses a scheme suggested by Benettin et al. (1976). Here one starts with the initial norm $d(0)$ normalized to unity and computes the divergence over some interval τ, which is then renormalized back to a norm of unity. In this way, one computes (see Figure 4.24) the sequence of quantities

$$d_j = \|\boldsymbol{\delta}\mathbf{x}^{(j-1)}(\tau)\| \tag{4.5.13}$$

where $\| \ \|$ denotes Euclidean norm, and

$$\boldsymbol{\delta}\mathbf{x}^{(j)}(0) = \frac{\boldsymbol{\delta}\mathbf{x}^{(j-1)}(\tau)}{d_j} \tag{4.5.14}$$

where $\boldsymbol{\delta}\mathbf{x}^{(j)}(\tau)$ is computed from (4.5.2), with the initial values $\boldsymbol{\delta}\mathbf{x}^{(j)}(0)$, along the reference trajectory $\bar{\mathbf{x}}$ from $\bar{\mathbf{x}}(j\tau)$ to $\bar{\mathbf{x}}((j+1)\tau)$. By analogy with (4.5.4), one defines

$$\sigma_N = \frac{1}{N\tau}\sum_{j=1}^{N} \ln d_j \tag{4.5.15}$$

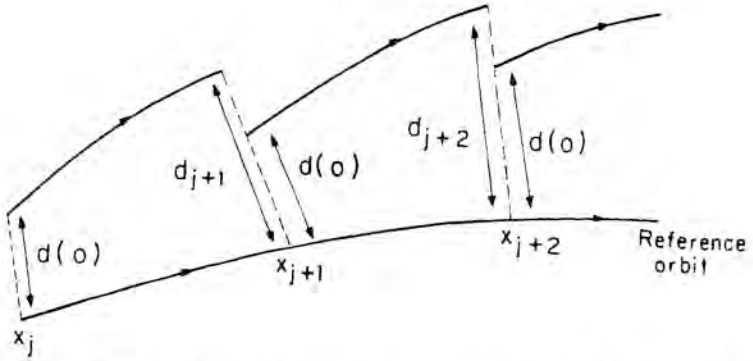


Figure 4.24 Computation of the largest Lyapunov exponent. After each period τ, the distance from the reference orbit is rescaled back to $d(0)$.

Furthermore, for τ not too large, it can be shown that the limit $N \to \infty$ exists and is independent of τ. Indeed, one can show that

$$\lim_{N\to\infty} \sigma_N = \sigma_1 \tag{4.5.16}$$

where σ_1 is the largest of the set of exponents (4.5.5). The computation of the complete spectrum of Lyapunov exponents $\sigma_1, \ldots, \sigma_n$ requires more sophisticated techniques; a discussion of this important problem is better left to the experts (see Benettin et al. (1980)).

A quantity related to Lyapunov exponents, but more difficult to compute in practice, is the so-called *Kolmogorov entropy* (KS entropy). Formally, it is defined somewhat like entropy in statistical mechanics (i.e., it involves partitions of phase space, etc.) and gives a measure of the amount of information lost or gained by a system as it evolves. A remarkable result of Pesin (1977) shows that it can be computed from the Lyapunov exponents by the relation

$$h_k = \int_P \sum_{\sigma_i > 0} \sigma_i \, d\mu$$

which represents the sum of all positive Lyapunov exponents averaged over some (connected) region of phase space P with measure $d\mu$. A very readable introduction to this idea, which includes computation of h_k for the Henon–Heiles system, has been given by Benettin et al. (1976).

4.5.b Power Spectra

Another valuable characterization of orbits is in terms of their Fourier transform or power spectrum. Indeed, in many experimental systems (e.g., fluid dynamics), the data are often recorded as the Fourier transform rather than as the real time (or space) signal.

Firstly, consider the case of regular motion. Since the trajectories are confined to tori, they can be represented in the standard form

$$\mathbf{q}(t) = \sum_{\mathbf{m}} \mathbf{q}_{\mathbf{m}}\, e^{i\mathbf{m}\cdot(\boldsymbol{\omega}t+\boldsymbol{\delta})} \tag{4.5.17}$$

where $\mathbf{q}_m$ are the vector of Fourier coefficients associated with the variables $\mathbf{q} = (q_1, \ldots, q_n)$, and $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_n)$ are the frequencies of the associated torus. Clearly, the Fourier transform of (4.5.17)—or rather its squared modulus—will just be a set of δ-functions at the fundamental frequencies $\boldsymbol{\omega}$ and various overtones. The term *power spectrum* is, strictly speaking, defined to be the Fourier transform of the correlation function of a particular variable, say q_i, namely

$$C(t) = \langle q_i(0) q_i(t) \rangle \tag{4.5.18}$$

where $\langle\ \rangle$ denotes some ensemble average. The power spectrum is thus

$$I(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} C(t)\, e^{i\omega t}\, dt \tag{4.5.19}$$

a result which is an example of the Weiner–Khinchine theorem.

For regular motion, a natural choice of ensemble is just the associated torus. Since it is covered by a one-parameter (the initial phase $\boldsymbol{\delta}$) family of trajectories, the ensemble average is just the (phase) average over all $\boldsymbol{\delta}$. In this case it is easy to show that

$$C(t) = \sum_{\mathbf{m}} |q_{\mathbf{m}}^{(i)}|^2\, e^{i\mathbf{m}\cdot\boldsymbol{\omega}t} \tag{4.5.20}$$

and hence

$$I(\omega) = \sum_{\mathbf{m}} |q_{\mathbf{m}}^{(i)}|^2 \boldsymbol{\delta}(\mathbf{m}\cdot\boldsymbol{\omega} - \omega) \tag{4.5.21}$$

where the $q_{\mathbf{m}}^{(i)}$ are the Fourier coefficients associated with the variable q_i. Furthermore, for tori with incommensurable frequencies, the flow is ergodic on the torus. Thus, phase average equals time average and one can compute $I(\omega)$ just using a single trajectory, that is,

$$I(\omega) = \frac{1}{2\pi} \lim_{T\to\infty} \frac{1}{T} \left| \int_{-T}^{T} q_i(t)\, e^{i\omega t}\, dt \right|^2 \tag{4.5.22}$$

(thus defined, $I(\omega)$ is essentially the spectral line-shape function much used in spectroscopy).

In practice, a time series is only available for a finite time, and the spectrum is really a convolution between the actual motion and the "box-function" $h(t) = 1$, $-T \leq t \leq T$. For example, consider the simple case of the periodic motion

$$q(t) = Ae^{-i\Omega t} \tag{4.5.23}$$

Using (4.5.22), it is easy to show that

$$I(\omega) = \lim_{T\to\infty} \frac{2|A|^2}{\pi} \frac{\sin^2(\omega - \Omega)T}{(\omega - \Omega)^2 T} \tag{4.5.24}$$

which, for finite T, exhibits a maximum at $\omega = \Omega$ and a series of sym-

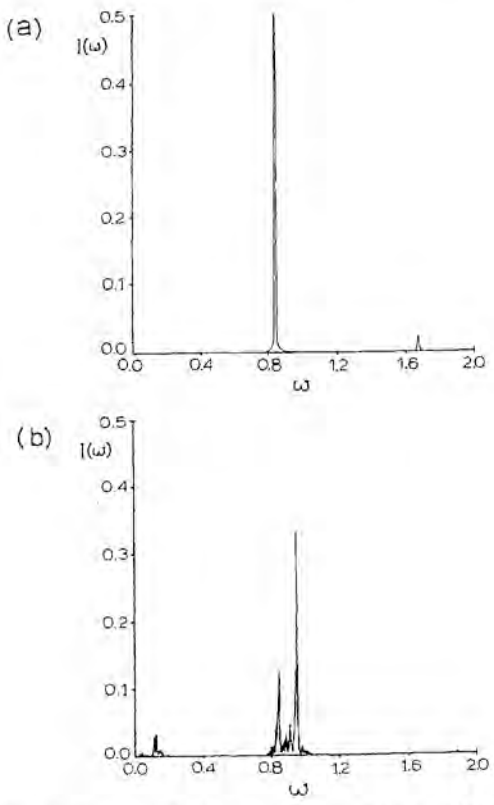


Figure 4.25 Power spectra (of the quantity $(x(t) + y(t))$ for individual trajectories of the Henon–Heiles system at $E = \frac{1}{8}$: (*a*) regular trajectory; (*b*) irregular trajectory.

metrical, decaying side-bands. In the limit $T \to \infty$, (4.5.24) reverts to the expected δ-function at $\omega = \Omega$.

In the chaotic regime, one can still go ahead and compute $I(\omega)$, using a single trajectory, as defined in (4.5.22). (Strictly speaking, this is no longer a "line-shape" function since it is not clear what the ensemble is in (4.5.18).) One finds that the spectrum of an irregular trajectory is much more complicated than for a regular one. Typically, one sees some dominant peaks surrounded by a lot of "grass" (see Figure 4.25). On the basis of numerical evidence alone, it is not clear whether this grassy portion of the spectrum is truly continuous for the irregular trajectories of generic Hamiltonian systems. Nonetheless, the difference in spectrum between regular and irregular motion is usually fairly striking and again provides a valuable characterization of dynamical systems. Indeed, there are a number of important rigorous results available which tell us that a system will only have a discrete spectrum if it is "ergodic," whereas to have a continuous spectrum it must be "mixing." (See Section 4.7 for a brief discussion of these concepts.)

4.6 CRITERIA FOR THE ONSET OF WIDESPREAD CHAOS

So far our discussions of chaos have been of a local nature in that we have concentrated on the chaotic behaviour of individual trajectories as well as the means of identifying them and quantifying their behavior. As useful as these techniques are, it would clearly be valuable if there were also methods for estimating when—as a function of energy or some nonlinear coupling parameter—the bulk of the trajectories become chaotic. This is what we term "widespread" chaos—although we hasten to add that in all of the techniques to be described, the notion of "widespread" is somewhat subjective, and a precise measure of this does not seem to be quantifiable. Nonetheless, estimates of when widespread chaos first appears are most useful. For example, in theories of unimolecular decomposition, it can indicate the validity of statistical, rather than dynamical, theories. In other situations, it can similarly indicate when it becomes valid to think in terms of transport processes. In this section we will briefly describe two such techniques. The first is the method of *overlapping resonances* due to Chirikov (1979), which is capable of giving crude analytical estimates of the onset of widespread chaos. The other is due to Greene (1979) and is able to predict when individual tori will breakdown. More detailed reviews of these and other methods have been given elsewhere (Tabor, 1981; Lichtenberg and Lieberman, 1983).

4.6.a Method of Overlapping Resonances

In order to understand Chirikov's method, we must first understand what is meant by a resonance. Consider some integrable, n-degree-of-freedom

Hamiltonian H_0 perturbed by some H_1, that is,

$$H(\mathbf{I}, \boldsymbol{\theta}) = H_0(\mathbf{I}) + \epsilon H_1(\mathbf{I}, \boldsymbol{\theta}) \tag{4.6.1}$$

where $\mathbf{I}$, $\boldsymbol{\theta}$ are the n-component action-angle vectors. In the usual way, we write H_1 as a Fourier series in the angle variables, namely,

$$H(\mathbf{I}, \boldsymbol{\theta}) = H_0(\mathbf{I}) + \epsilon \sum_{\mathbf{m}} H_{\mathbf{m}}(\mathbf{I}) e^{i\mathbf{m}\cdot\boldsymbol{\theta}} \tag{4.6.2}$$

where $H_{\mathbf{m}}(\mathbf{I})$ are the Fourier coefficients and $\mathbf{m} = (m_1, \ldots, m_n)$. The equations for the unperturbed variables are, of course,

$$\begin{aligned} I_i &= I_i(0) \\ \theta_i &= \omega_i(\mathbf{I})t + \theta_i(0) \end{aligned} \tag{4.6.3}$$

where $\omega_i = \partial H_0/\partial I_i$ and $I_i(0)$ and $\theta_i(0)$ are the initial values. Now consider the situation in which H_0 is perturbed by only one term in the Fourier sum, that is,

$$H(\mathbf{I}, \boldsymbol{\theta}) = H_0(\mathbf{I}) + \epsilon H_{\mathbf{m}}(\mathbf{I})\, e^{i\mathbf{m}\cdot\boldsymbol{\theta}} \tag{4.6.4}$$

In this case the perturbed variables satisfy the equations of motion

$$\dot{I}_i = -i\epsilon m_i H_{\mathbf{m}}(\mathbf{I})\, e^{i\mathbf{m}\cdot\boldsymbol{\theta}} \tag{4.6.5a}$$

$$\dot{\theta}_i = \omega_i(\mathbf{I}) + \epsilon H'_{\mathbf{m}}(\mathbf{I})\, e^{i\mathbf{m}\cdot\boldsymbol{\theta}} \tag{4.6.5b}$$

where the prime denotes differentiation with respect to I_i. To first order in ϵ, Eqs. (4.6.5) can be integrated by simply substituting, where appropriate, the unperturbed solutions (4.6.3), in which case (4.6.5a) yields

$$I_i \simeq I_i(0) - \frac{\epsilon m_i H_{\mathbf{m}}(\mathbf{I}(0)) e^{i(\mathbf{m}\cdot\boldsymbol{\omega})t + i\delta}}{\mathbf{m}\cdot\boldsymbol{\omega}} \tag{4.6.6}$$

where δ is a phase factor. Clearly, such an approximation will break down for any $\mathbf{m}\cdot\boldsymbol{\omega}(\mathbf{I}) \leq \epsilon$. This is known as a *resonance*, and the reader will see that this is essentially the same as the small divisor effect discussed in Chapter 3. Indeed, we have already seen an example of a resonance in Chapter 1, in the case of the driven linear oscillator (1.6.4) in which the solution has a denominator proportional to $\omega^2 - \Omega^2$ (1.6.5), where ω is the intrinsic frequency and Ω is the external driving frequency. For a single perturbing term (i.e., Eq. (4.6.4)), such a resonance is not a disaster and can be removed by a suitable canonical transformation as described below. Indeed, the system (4.6.4) is still completely integrable since it is possible to

construct new integral of motion of the form

$$F = \hat{\mathbf{m}} \cdot \mathbf{I} \tag{4.6.7}$$

where $\hat{\mathbf{m}}$ is a vector orthogonal to $\mathbf{m}$ (i.e., $\mathbf{m} \cdot \hat{\mathbf{m}} = 0$). This is easily verified by taking the Poisson bracket of F with H, namely,

$$[F, H] = \sum_{i=1}^{n} \left(\frac{\partial H}{\partial I_i} \frac{\partial F}{\partial \theta_i} - \frac{\partial H}{\partial \theta_i} \frac{\partial F}{\partial I_i} \right)$$

$$= -i\epsilon \sum_{j=1}^{n} \hat{m}_j m_j H_{\mathbf{m}} e^{i\mathbf{m} \cdot \boldsymbol{\theta}} = 0 \tag{4.6.8}$$

Thus, isolated resonances, although they can cause a considerable distortion of tori in their neighborhood, do not introduce any chaos into a system. However, when two or more resonances are simultaneously present, they will render a system nonintegrable. Furthermore, when they are sufficiently "close" to each other, they will result, as is now described, in the appearance of widespread chaos.

In a very readable illustration of the resonance problem, Walker and Ford (1969) took the integrable Hamiltonian

$$H_0(I_1, I_2) = I_1 + I_2 - I_1^2 - 3I_1 I_2 + I_2^2 \tag{4.6.9}$$

and investigated the effect of adding a 2:2 resonance and a 3:2 resonance, that is,

$$H(\mathbf{I}, \boldsymbol{\theta}) = H_0(\mathbf{I}) + \alpha I_1 I_2 \cos(2\theta_1 - 2\theta_2) + \beta I_1^{3/2} I_2 \cos(2\theta_1 - 3\theta_2) \tag{4.6.10}$$

The effect of each resonance in isolation and together is shown in Figure 4.26. At low energies the two resonant zones are well separated. As the energy of the system is increased, the two zones overlap and a "macroscopic zones of instability" appears. By this term, Walker and Ford simply meant a clearly visible splatter of points on the surface of section. The size of this zone increases with increasing energy. The structure is further complicated by the appearance of "secondary" resonant zones as the two principal zones approach each other. By means of a numerical investigation, the authors were able to predict the energy at which the overlap of the resonances first occurred; that is, they were able to predict (successfully) the onset of widespread chaotic motion. The overlapping of resonances would appear to play a key role in such an onset, as well as giving a great deal of physical insight. When principal zones start overlapping, many higher-order resonances are also involved and thus one may be moderately confident that fairly large areas of phase space have had (most of) their tori destroyed and that the ensuing chaos will indeed be "wide-

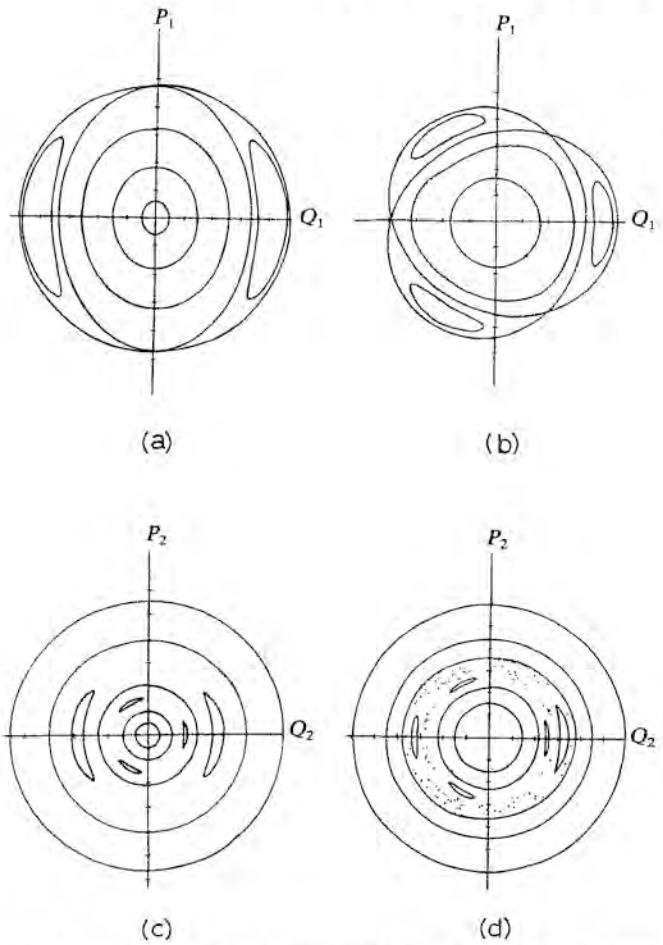


Figure 4.26 Surfaces of section of resonant Hamiltonian (4.6.10): (*a*) only 2:2 resonance acting ($\beta = 0$); (*b*) only 2:3 resonance acting ($\alpha = 0$); (*c*) both resonances present ($\alpha = \beta = 0.02$) but widely separated at $E = 0.18$; and (*d*) the two resonant zones overlapping at $E = 0.2905$; the erratic splatter of points is generated by a trajectory started in the region of resonance overlap. (Reproduced, by permission, from Walker and Ford (1969).)

spread," since trajectories are now free to wander between regions that were previously separated by nonresonant tori.

A means of predicting approximately when resonant overlap will occur has been proposed by Chirikov (1979). This method works best for forced one-dimensional oscillators, a model very useful in the design of accelerators or in the study of molecular bonds or atoms subjected to radiation fields. We consider a one-dimensional nonlinear oscillator (e.g.,

$H_0 = \frac{1}{2}(p^2 + \frac{1}{2}q^4))$ perturbed by an external periodic force, for example $V = q\cos(\phi)$, where $\phi = \Omega t + \phi_0$ is the external phase. The unperturbed system, being one-dimensional, can always be solved in action-angle variables (I, θ), and we can then express any (reasonable) external field as a Fourier series in these variables, that is,

$$H = H_0(I) + \epsilon \sum_{m,n} V_{mn}(I)\, e^{i(m\theta + n\phi)} \tag{4.6.11}$$

In what follows, the results are only significant for nonlinear oscillators. In the linear case, we know that when the external frequency Ω equals the oscillator frequency ω, the motion "blows up." In the nonlinear case, there is also a resonance in the vicinity of $\Omega = \omega$. However, as discussed in Chapter 1, when the amplitude of the oscillator increases, the frequency, which is energy dependent, changes and the system comes out of resonance. In the above case, the oscillator frequency is given by the usual equation, $\omega(I) = \partial H_0/\partial I$, and there is a resonance at those values of $I = I_r$ such that

$$\frac{\omega(I_r)}{\Omega} = \frac{k}{l} \tag{4.6.12}$$

In this case the corresponding phase (and its harmonics) $l\theta - k\phi$ are slowly varying compared to other terms in the series (4.6.11). By virtue of the nonlinearity of H_0, there will, of course, be other values of I_r giving rise to other resonances, and, generally speaking, the set of resonances is everywhere dense. However, to simplify matters, we start by considering the resonance (4.6.12) in isolation and examine the behavior of the Hamiltonian in its vicinity. In the next few paragraphs we show how to reduce (4.6.11) to a simple form from which a "resonance width" can be estimated.

Drawing on our study of canonical transformations in Chapter 2, we introduce the generating function

$$F = F(J, \theta, \phi) = (l\theta - k\phi)J + \theta I_r \tag{4.6.13}$$

where J is the new momentum and the term θI_r provides, as we see below, a convenient shift in the origin of the new action variable J. From the generating function, we obtain the relations

$$I = \frac{\partial F}{\partial \theta} = lJ + I_r \tag{4.6.14}$$

and

$$\psi = \frac{\partial F}{\partial J} = l\theta - k\phi \tag{4.6.15}$$

where ψ is the new "resonant" phase conjugate to the new momentum J. The time derivative of F is also required, that is,

$$\frac{\partial F}{\partial t} = -k\Omega J \tag{4.6.16}$$

Performing the canonical transformation the Hamiltonian becomes, in terms of the new variables,

$$H = H_0(J) + \epsilon \sum_{m,n} V_{mn}(J) \exp\left[i \frac{1}{l}\{m\psi + (km + nl)\phi\}\right] - k\Omega J \tag{4.6.17}$$

The transformed Hamiltonian has almost the same form as the original one (4.6.11). However, by performing this transformation, one is effectively putting the observer in a rotating frame in which the rate of change of the new phase measures the slow deviation from resonance. Although it is not necessarily true that $\dot{\psi} \ll \dot{\phi}$, we assume that it is near the resonance and, hence, that during one complete cycle of ψ, ϕ will have passed through many cycles. The average contribution of these rapidly oscillating terms is zero, and this leads us to the next stage, namely, averaging the Hamiltonian over the fast-phase variables, that is,

$$\bar{H}(J, \psi) = \frac{1}{2\pi} \int_0^{2\pi} H(J, \psi, \phi)\, d\phi \tag{4.6.18}$$

The resulting "averaged" Hamiltonian, H, takes the form

$$\bar{H} = H_0(J) + \epsilon \sum_p V_{pl,-pk} \cos(p\psi) - k\Omega J \tag{4.6.19}$$

where we have gone over to real arithmetic (assuming $V_{-l,k} = V_{l,-k}$ and $V_{00} = 0$ as well as absorbing a factor of 2 in the Fourier coefficients $V_{pl,-pk}$). There are still all the harmonics to deal with, but at this stage we make the assumption

$$V_{pl,-pk} \ll V_{l,-k} \qquad \text{for } p = 2, 3, \ldots \tag{4.6.20}$$

and (4.6.19) reduces to

$$\bar{H} = H_0(J) + \epsilon V_{l,-k} \cos\psi - k\Omega J \tag{4.6.21}$$

The last stage is as follows. Having worked on the assumption that $\psi \ll \phi$ in the region of the resonance (4.6.12), we expand (4.6.21) about $I = I_r$ to second order, although we assume that the coefficient $V_{l,-k}(J)$ is only a

slowly varying function of I. Thus we obtain

$$\bar{H} = H_0(I_r) + lJ\left(\frac{\partial H_0}{\partial I}\right)_{I=I_r} + l^2\frac{J^2}{2}\left(\frac{\partial^2 H_0}{\partial I^2}\right)_{I=I_r} + \epsilon V_{l,-k}(I_r)\cos\psi - k\Omega J \tag{4.6.22}$$

Since $(\partial H_0/\partial I)_{I=I_r} = \omega(I_r)$, the terms that are linear in J conveniently cancel by the resonance condition (4.6.12). Dropping the constant term $H_0(I_r)$, we are left with the "resonant" Hamiltonian

$$H_r = \frac{J^2}{2M} + \epsilon V_{l,-k}\cos\psi \tag{4.6.23}$$

where the "mass" M is given by

$$M^{-1} = l^2\left(\frac{\partial^2 H_0}{\partial I^2}\right)_{I=I_r} \tag{4.6.24}$$

The "resonant" Hamiltonian (4.6.23) has exactly the form of a pendulum Hamiltonian whose separatrix (see Figure 4.27) is determined by the equation

$$J_{sx} = \pm(4M\epsilon V_{l,-k})^{1/2}\cos\left(\frac{\psi}{2}\right) \tag{4.6.25}$$

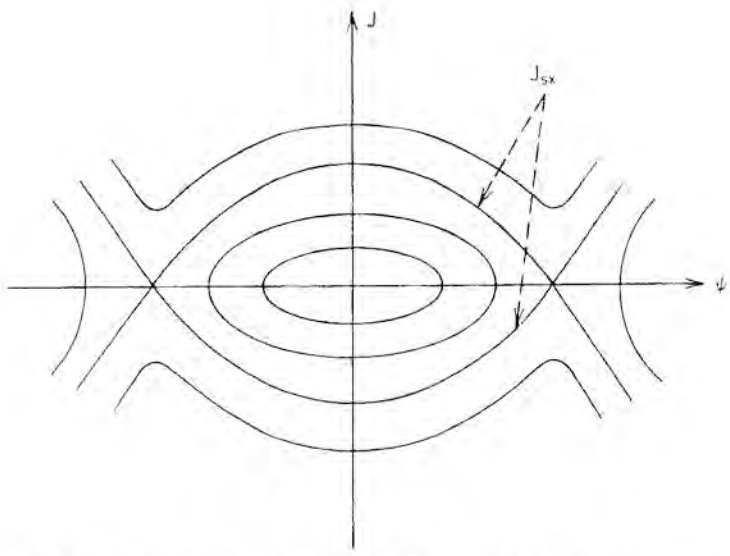


Figure 4.27 Phase plane for pendulum Hamiltonian (4.6.23) showing separatrix.

In the old (I, θ) variables, this is simply

$$I_{sx} = I_r \pm (\Delta I_r) \cos\left(\frac{l\theta - k\phi}{2}\right) \tag{4.6.26}$$

where

$$(\Delta I_r) = 2l(\epsilon M V_{l,-k})^{1/2} \tag{4.6.27}$$

The quantity (ΔI_r) is the resonance "half-width," which can also be expressed in terms of frequency, that is,

$$(\Delta\omega_r) = \frac{\partial\omega}{\partial I}(\Delta I_r) = \frac{1}{l^2 M} 2l(\epsilon M V_{l,-k})^{1/2} = \frac{2}{l}\left(\frac{\epsilon V_{l,-k}}{M}\right)^{1/2} \tag{4.6.28}$$

Notice that since $M^{-1} = l^2(\partial\omega/\partial I)$ the resonance half-width's dependence on the order of the resonance is only in the Fourier coefficients $V_{l,-k}$. From (4.6.27) and (4.6.28), we can see that the effect of a resonant perturbation is of $O(\epsilon^{1/2})$.

It is important to note that several assumptions have gone into the derivation of the resonant Hamiltonian (4.6.23). Nonresonant terms in H are neglected by assuming they are rapidly oscillating, and hence their average value over a cycle of motion is zero. The resonance harmonics have been neglected by assuming $V_{l,-k} \ll V_{pl,-pk}$; furthermore, the higher-order terms in the expansion about I_r have been dropped. Chirikov (1979) has suggested that these assumptions can be summarized in the "moderate nonlinearity condition"

$$\epsilon \ll \alpha \ll \frac{1}{\epsilon} \tag{4.6.29}$$

where $\alpha = (I/\omega)(\partial\omega/\partial I)$. This condition is discussed further by Chirikov.

So far, though, the "resonant" Hamiltonian (4.6.23) is still integrable, since it consists of only one resonance in isolation. Chirikov's "criterion of overlapping resonances" is obtained by evaluating the width of another (principal) resonance and then finding the coupling strength ϵ at which the two resonances touch; that is, we find the ϵ for which

$$(\Delta\omega_r)_1 + (\Delta\omega_r)_2 = \Delta\Omega \tag{4.6.30}$$

where $(\Delta\omega_r)_1$ and $(\Delta\omega_r)_2$ are the widths of the two resonances and $\Delta\Omega$ is their separation. The width of each resonance zone is calculated independently of all the others; clearly, this is a major approximation and one simply hopes that the "moderate nonlinearity condition" (4.6.29) will ensure that the error is not too great.

Chirikov (1979) has tested his method out on a number of simple systems. Of particular interest is the one with model Hamiltonian

$$H(I, \theta, t) = \frac{I^2}{2} + K \sum_{n=-\infty}^{\infty} \cos(\theta - nt) \tag{4.6.31}$$

Physically, it represents a rotor (of unit mass) being acted on by an infinite series of resonances or, alternatively, when written in the equivalent form

$$H(I, \theta, t) = \frac{I^2}{2} + 2\pi K \cos\theta \sum_{m=-\infty}^{\infty} \delta(2\pi m - t) \tag{4.6.32}$$

as a pendulum subject to a series of "kicks" at times $t = 2\pi m$. The reader may confirm that integration of Hamilton's equations over one period yields precisely the "standard map" (4.2.26) with $p \equiv I$ and $q \equiv \theta$ and $k = K/(2\pi)^2$. Each term in the series of resonances provides us directly with a resonant Hamiltonian (cf. (4.6.23))

$$\bar{H}^{(n)} = \frac{I^2}{2} + K \cos\psi_n \tag{4.6.33}$$

where ψ_n is the slowly varying phase $(\theta - nt)$. One can immediately write the resonance half-width from (4.6.28), that is,

$$(\Delta\omega_r)_n = 2K^{1/2} \tag{4.6.34}$$

Resonances occur at every integer value of $\omega = I = I_r = n$ (this is represented schematically in Figure 4.28). The spacing between resonances is unity $(\Delta\Omega = I_r^{(n+1)} - I_r^{(n)} = (n+1) - n = 1)$, and they touch when

$$(\Delta\omega_r) = \frac{\Delta\Omega}{2} = \frac{1}{2} \tag{4.6.35}$$

This enables one to predict the critical value of the perturbation parameter K at which resonance overlap occurs, that is, since

$$(\Delta\omega_r) = 2K^{1/2} = \tfrac{1}{2} \tag{4.6.36}$$

this gives

$$K_{\text{crit}} = \tfrac{1}{16} \tag{4.6.37}$$

Numerical studies of this system suggest that widespread chaos sets in at about $K = \frac{1}{40}$; that is, the overlap criterion is out by about a factor $2\frac{1}{2}$. Further refinements, such as including higher harmonics of the resonant phases, give an improved estimate of $K \simeq \frac{1}{30}$.

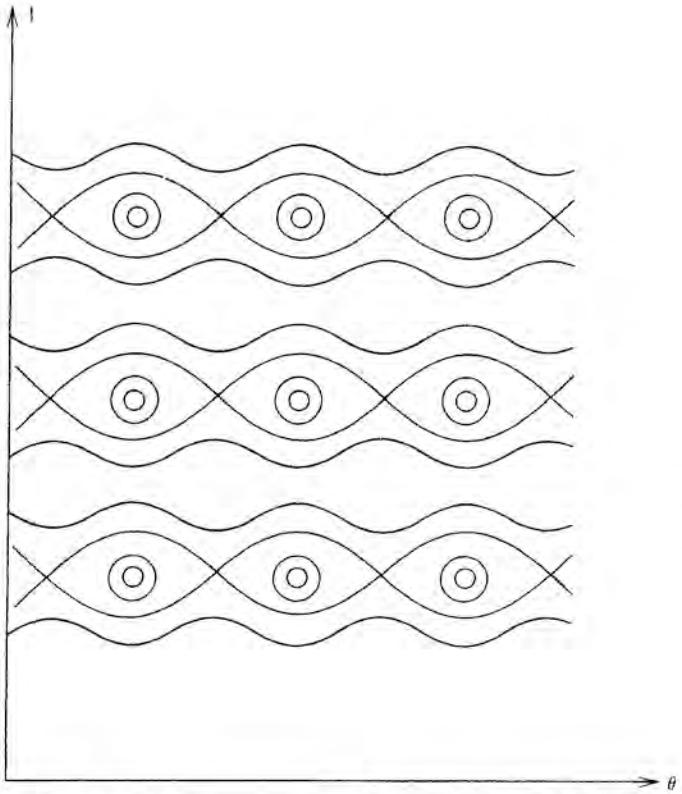


Figure 4.28 Phase plane of resonances for "kicked" pendulum Hamiltonian (4.6.32).

Chirikov's method has been the object of much research activity, and sophisticated versions of it (to deal, for example, with the effect of "secondary" resonances) have been developed which are able to give far more accurate predictions for the onset of widespread chaos. The technique can also be applied to autonomous systems with many degrees of freedom (the theory is "cleanest" for driven one-degree-of-freedom systems). The underlying ideas are the same, but the analysis becomes more complicated; the interested reader is referred elsewhere to follow up this topic.

4.6.b Greene's Method

We now turn (albeit too briefly) to discuss an important method, developed by Greene (1979), for predicting the onset of chaotic motion based on the stability properties of closed orbits. It is based on the hypothesis that the dissolution of an invariant curve (torus) can be associated with the sudden change from stability to instability of nearby closed orbits. To see this more

precisely, imagine a weakly perturbed integrable system. According to the KAM theorem, those invariant curves with "sufficiently" irrational winding number are preserved. The neighboring rational (and close-to-rational) curves break up in the manner described previously (the Poincaré–Birkhoff fixed-point theorem), that is, into equal numbers of elliptic (stable) and hyperbolic (unstable) fixed points. Greene's method is based on the observation that when the perturbation is made sufficiently strong (or the energy high enough), the set of stable fixed points also becomes unstable (they become "hyperbolic-with-reflection" fixed points). The contention is that this then signals the dissolution of an invariant curve "close" to that set of fixed points. Recalling the discussion of the KAM theorem in Chapter 3, a rather nice way of estimating the closeness of a closed orbit to a given invariant curve is by expressing that curve's winding number in the form of a continued fraction, that is,

$$\alpha = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cdots}}} \tag{4.6.38}$$

where $a_0, a_1, a_2, \ldots$ are positive integers. Thus the successive trunctions of this representation of an irrational winding number yield the winding numbers of the closed orbits that become ever "closer" to the chosen invariant curve. By following the stability properties of these sequences of closed orbits, as they "close in" on an invariant curve, Greene (1979) was able to predict the breakup of that curve.

The two essential ingredients of this method are (1) finding the closed orbits and (2) determining their stability characteristics. A detailed discussion of the former problem is outside the scope of this chapter. Suffice it to say that there are now a variety of well-developed and efficient methods for finding closed orbits of any desired topology (winding number). Of these we mention the approach described by Greene (1979) and a method developed by Helleman and Bountis (1979). We describe the stability analysis by using the example studied in detail by Greene, namely, the "standard map" (on the unit torus), that is,

$$I_{n+1} = I_n + \frac{k}{2\pi} \sin 2\pi\theta_n, \qquad \text{mod } I = 1 \tag{4.6.39a}$$

$$\theta_{n+1} = \theta_n + I_{n+1}, \qquad \text{mod } \theta = 1 \tag{4.6.39b}$$

The parameter k can be regarded as a perturbation parameter; for $k = 0$, the mapping takes the trivial form

$$I_{n+1} = I_n \tag{4.6.40a}$$

$$\theta_{n+1} = \theta_n + I_{n+1} \tag{4.6.40b}$$

In this case, the mapping is clearly "integrable," since all the orbits lie on straight lines. These are just the invariant curves of the unperturbed mapping. Returning to Figure 4.11, which shows the standard map computed at $k = 0.97$, we see some strongly irregular orbits, filling up substantial portions of the phase plane, as well as the typical alternating hyperbolic and elliptic fixed-point structure. Notice also that there are still invariant curves remaining that divide the phase space. These curves prevent a trajectory from wandering over the whole phase plane. Clearly, it will not be until these curves are destroyed that the "chaos" will be truly widespread.

The stability of a given closed orbit is determined by evaluating the tangent-space mapping. This corresponds to linearizing the mapping at each iteration. Thus if we denote the "tangent-space" variables as $(\delta I, \delta\theta)$, we have the tangent map

$$\begin{bmatrix} \delta I_{n+1} \\ \delta\theta_{n+1} \end{bmatrix} = M \begin{bmatrix} \delta I_n \\ \delta\theta_n \end{bmatrix} \tag{4.6.41a}$$

where

$$M = \begin{bmatrix} 1 & -k\cos 2\pi\theta_n \\ 1 & 1 - k\cos 2\pi\theta_n \end{bmatrix} \tag{4.6.41b}$$

The tangent mapping is, of course, area preserving since

$$\det|M| = 1 \tag{4.6.42}$$

For an orbit that closes after Q iterations of the mapping, the eigenvalues, $\lambda_\pm$, of the 2×2 matrix

$$M^{(Q)} = \prod_{n=1}^{Q} \begin{bmatrix} 1 & -k\cos 2\pi\theta_n \\ 1 & 1 - k\cos 2\pi\theta_n \end{bmatrix} \tag{4.6.43}$$

give the stability indices, or Floquet multipliers, of the orbit. Denoting the matrix elements of $M^{(Q)}$ by $M_{ij}^{(Q)}$, we have explicitly

$$\lambda_\pm = \tfrac{1}{2}(M_{11}^{(Q)} + M_{22}^{(Q)}) \pm \tfrac{1}{2}((M_{11}^{(Q)} + M_{22}^{(Q)})^2 - 4)^{1/2} \tag{4.6.44}$$

where we have made use of condition (4.6.42), that is,

$$M_{11}^{(Q)} M_{22}^{(Q)} - M_{12}^{(Q)} M_{21}^{(Q)} = 1 \tag{4.6.45}$$

From our previous discussions we know that if the eigenvalues are complex,

the orbits are stable, whereas if the eigenvalues are real, the orbits are unstable. Greene (1979) introduces a quantity called the *residue*, which is defined as

$$R = \tfrac{1}{4}(2 - \mathrm{Tr}(M^{(Q)})) \tag{4.6.46}$$

where Tr denotes trace.

From (4.6.44), it is easy to see that if $0 < R < 1$, the eigenvalues are imaginary and hence the orbit is stable, that is, the fixed points are elliptic. If $R < 0$ or $R > 1$ the eigenvalues are real and hence the orbit is unstable. More precisely, for $R < 0$ the fixed points are hyperbolic and for $R > 1$ they are "hyperbolic-with-reflection." (For parabolic fixed points, $R = 0$.) It can be shown that for an orbit of "length" Q, the residue is proportional to k^Q for both large and small k. (Recall that k is the perturbation parameter for the system studied here.) A quantity, called the *mean residue* f, is then introduced that scales away this exponential dependence on Q, that is,

$$f = \left(\frac{R}{\beta}\right)^{1/Q} \tag{4.6.47}$$

where β is some arbitrary constant introduced for practical convenience. We can now proceed to characterize the stability properties of the sequence of closed orbits converging on a chosen invariant curve. Each successive closed orbit (determined through the successive truncations of the continued fraction representation of the winding number of the chosen curve) has a larger Q, corresponding to increasing topological complexity of that orbit. The remarkable thing is that the corresponding sequence of mean residues is found to converge to some finite value. The rate of convergence seems to be determined by the value of β; for this problem the optimum value was found to be $\beta = \frac{1}{4}$. For further discussion of this matter, the reader is referred to Greene's original papers (Greene, 1979). It is then demonstrated (empirically) that when the converged mean residue becomes greater than unity (now assuming $\beta = 1$), the invariant curve associated with that sequence of closed orbits is destroyed.

This criterion then enables one to find the value of the perturbation parameter k at which any chosen invariant curve breaks up. For the above system the method has been found to give very accurate results. Furthermore, Green (1979) has made an ingenious extension of his method to predict the onset of widespread chaos. It is based on the conjecture that the more closely an irrational curve can be approximated by a sequence of rationals, the smaller the perturbation (k) required to destroy it. Thus one might reasonably assume that the last invariant curve to be destroyed will be the one whose winding number is least closely approximated by a sequence of rationals. This is the invariant curve whose winding number has the continued fraction representation

$$\alpha = 1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \cdots}}}$$
$$= \frac{(\sqrt{5}-1)}{2} \tag{4.6.48}$$

which is the famous "golden mean." Thus by the time k is sufficiently large so that this invariant curve breaks up, one may fairly confidently assume that all the other curves have also been destroyed. There will then be no impediment to an irregular trajectory wandering over the whole of the phase plane, and widespread chaos will ensue. The critical value of k corresponding to the breakup of the golden mean curve was found to be about unity, in close agreement with the observed onset of widespread chaos.

4.7 STATISTICAL CONCEPTS IN STRONGLY CHAOTIC SYSTEMS

We have already mentioned the concept of ergodicity several times, and in this section we will enlarge the discussion to include other, related concepts that are useful in understanding the properties of strongly chaotic systems. Excellent, introductory discussions of these ideas have been given by Lebowitz and Penrose (1973) and Zaslavskii and Chirikov (1972).

4.7.a Ergodicity

A simple illustration of ergodicity is provided by recalling the case, given in Section 2.5, of flow on a torus with irrationally related frequencies. Con-

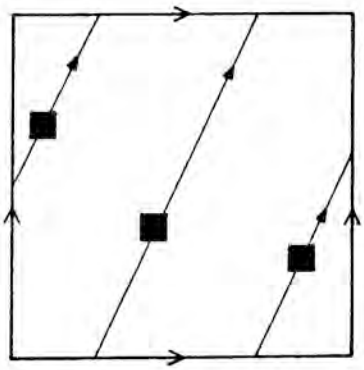


Figure 4.29 Ergodicity on the torus: Small-area element exploring the torus by uniform translation.

sider a two-dimensional torus, for which we write the flow as

$$\phi_1 = 2\pi\omega_1 t + \phi_1(0), \qquad \text{mod } \phi_1 = 1 \tag{4.7.1a}$$

$$\phi_2 = 2\pi\omega_2 t + \phi_2(0), \qquad \text{mod } \phi_2 = 1 \tag{4.7.1b}$$

where we have introduced the variables $\phi_i = \theta_i/2\pi$ which have period 1. This 2-D torus is topologically equivalent to the unit square with identified edges, and, as already described, it is easy to see that the flow is ergodic—and hence time average equals phase average—if ω_1/ω_2 is irrational. Notice (of course) that a small-area element only explores the torus by uniform translation without itself ever undergoing any distortion (see Figure 4.29). It should be clear from this example that ergodicity does not imply chaotic behavior.

4.7.b Mixing

Chaotic behavior is associated with the exponential divergence of nearby trajectories and hence positive Lyapunov exponents. In this case a small-area element will clearly undergo considerable distortion as it evolves—this leads to the concept of *mixing*. A simple system which exhibits mixing is the famous "cat map" of Arnold (linear automorphism of the unit torus), which is nothing more than the linear, area-preserving transformation T

$$T: \begin{bmatrix} x_{n+1} \\ y_{n+1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_n \\ y_n \end{bmatrix} \qquad \begin{matrix} \text{mod } x = 1 \\ \text{mod } y = 1 \end{matrix} \tag{4.7.2}$$

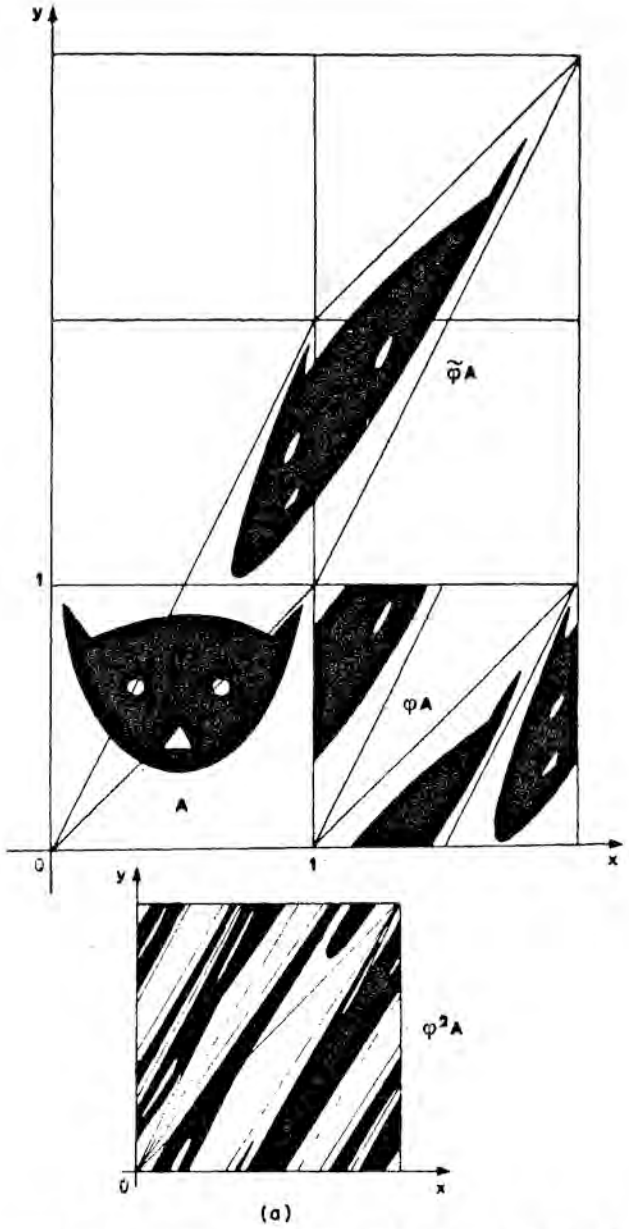
As illustrated in Figure 4.30 after only two iterations of the map, an area element undergoes considerable distortion. The difference between (4.7.1) and (4.7.2) is that the latter has a shearing component and it is this plus the periodicity of the torus that lead to the observed effect of both stretching and *mixing*. By contrast with (4.7.1), a small-area element not only translates around the torus but rapidly becomes a long thin filament. Clearly mixing implies ergodicity, but ergodicity does not imply mixing.

The eigenvalues of T are easily computed to be

$$\lambda_\pm = \frac{3 \pm \sqrt{5}}{2} \tag{4.7.3}$$

with $\lambda_+\lambda_- = 1$ since the transformation is area preserving. The two real eigenvalues λ_+ and λ_- lead to exponential stretching and contracting, respectively—the former in the direction of the eigenvector

$$\xi_+ = \begin{bmatrix} 1 \\ \dfrac{1+\sqrt{5}}{2} \end{bmatrix} \tag{4.7.4a}$$



(a)

Figure 4.30 (a) Two iterations of the Arnold cat map. (Reproduced, by permission, from Arnold and Avez (1968).) (*b*) Mixing on the torus: Small-area element exploring the torus by translation and stretching.

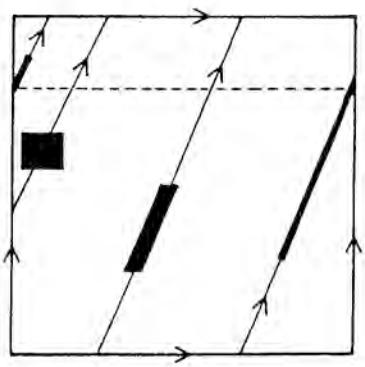


Figure 4.30 *(Continued)*

and the latter in the direction

$$\xi_- = \begin{bmatrix} 1 \\ \dfrac{1-\sqrt{5}}{2} \end{bmatrix} \tag{4.7.4b}$$

For this simple linear mapping, it is easy to see that the (positive) Lyapunov exponent is just $\sigma = \ln[(3+\sqrt{5})/2]$.

The mapping T^n has many fixed points. They are easily determined by

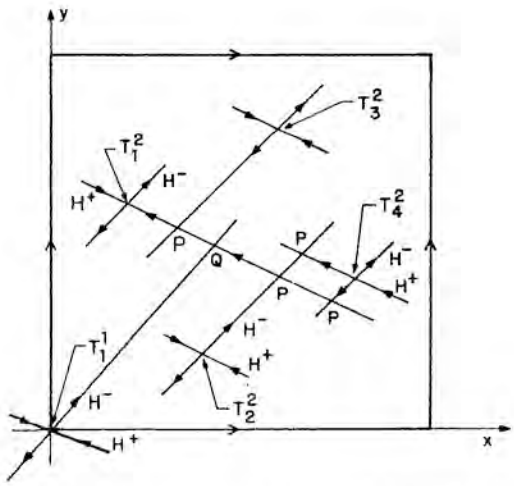


Figure 4.31 Schematic representation of homoclinic (P) and heteroclinic (Q) points of the cat map. T_i^2 ($i = 1, \ldots, 4$) are the four fixed points of T^2, and T_1^1 is the fixed point of T. The homoclinic points P are due to intersections of the H^+ and H^- manifolds of the T_i^2, and the heteroclinic point Q is due to intersection of H^- manifold of T_1^1 and H^+ manifold of T_1^2.

solving the equation

$$\begin{bmatrix} x \\ y \end{bmatrix} = T^n \begin{bmatrix} x \\ y \end{bmatrix} - \begin{bmatrix} k \\ l \end{bmatrix} \tag{4.7.5}$$

where k and l are the integers required to mod the iterates of T^n back onto the unit square. Obviously, the only fixed point of T itelf is just $(x, y) = (0, 0)$. For T^2 the fixed points are (1/5, 3/5), (2/5, 1/5), (3/5, 4/5), and (4/5, 2/5), which are all hyperbolic with eigenvalues $(\lambda_\pm)^2 = (7 \pm 3\sqrt{5})$. It is also rather easy to identify the homoclinic and heteroclinic points of the cat map. For example, the fixed point (0, 0) of T has stable (H^+) and unstable (H^-) manifolds which wrap around the torus in irrational directions, given by ξ_- and ξ_+, respectively, and therefore intersect each other (but never themselves) infinitely often. The same argument also applies to the stable and unstable manifolds of the fixed points of T^2, but these will also intersect the H^+ and H^- of the fixed point of T, thereby leading to an infinity of heteroclinic points as well (see Figure 4.31).

4.7.c The Baker's Transformation and Bernoulli Systems

Another simple transformation with striking properties is the so-called *Baker's transformation*, which can be written as the following mapping on the unit square:

$$\begin{bmatrix} x_{n+1} \\ y_{n+1} \end{bmatrix} = \begin{bmatrix} 2x_n \\ y_n/2 \end{bmatrix}, \qquad 0 \le x_n < \tfrac{1}{2}$$

$$= \begin{bmatrix} 2x_n - 1 \\ y_n/2 + \tfrac{1}{2} \end{bmatrix}, \qquad \tfrac{1}{2} \le x_n < 1 \tag{4.7.6}$$

which corresponds to repeated doublings in the x direction and halvings in the y direction. The mapping is completely reversible; and if it is run backwards, the doubling occurs in the y direction and the halving occurs in the x direction. As Figure 4.32 shows, this area-preserving transformation is reminiscent of a baker rolling out dough. Again it is clear that just a few iterations of the mapping will lead to rapid mixing.

The strongly random nature of this simple transformation can be seen by representing the iterates (x_n, y_n) in binary notation, that is, as strings of zeros and ones. Simple examples of binary numbers are $\frac{1}{16} = 0.0001000\ldots$, $\frac{1}{8} = 0.001000\ldots$, $\frac{1}{4} = 0.01000\ldots$, and so on. Less trivial numbers, such as irrationals, consist of infinite, nonrepeating strings of zeros and ones. Note, however, the basic property that the doubling of a number corresponds to moving the decimal point one place to the right and halving a number corresponds to moving it one place to the left. This is ideal for the baker's transformation. An initial condition $X_0 = (x_0, y_0)$ is represented by the strings

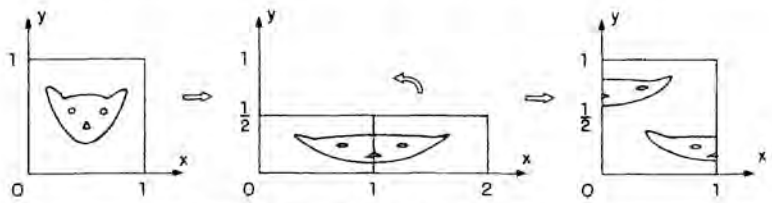


Figure 4.32 Baker's transformation.

$$x_0 = .\, a_1 a_2 a_3 \ldots a_i \ldots \tag{4.7.7a}$$

$$y_0 = .\, b_1 b_2 b_3 \ldots b_i \ldots \tag{4.7.7b}$$

where the a_i and b_i are all either 0 or 1. The position of this point in the unit square can be conveniently represented by putting these two strings back to back, that is,

$$X_0 = \ldots b_i \ldots b_3 b_2 b_1 \,.\, a_1 a_2 a_3 \ldots a_i \ldots \tag{4.7.8}$$

Since (forward) iteration of the map corresponds to doubling of x and halving of y, X_1 is obtained by simply moving the decimal point (4.7.8) one place to the right, that is,

$$X_1 = \ldots b_i \ldots b_3 b_2 b_1 a_1 \,.\, a_2 a_3 \ldots a_i \ldots \tag{4.7.9}$$

and so on, for the successive X_i. This process is known as a *Bernoulli shift*. Now consider a more "coarse-grained" description of the motion in which an orbit (or some function of the orbit) is identified by 0 if $0 \leq x_n < \frac{1}{2}$ and by 1 if $\frac{1}{2} \leq x_n < 1$. This identification simply corresponds to taking just the first digit in the (full) binary representation of the x_n. In this way, the coarse-grained history of the motion (i.e., the sequence of iterates $\bar{X}_0, \bar{X}_1, \bar{X}_2, \ldots, \bar{X}_i, \ldots$) will just be the sequence $a_1, a_2, a_3, \ldots, a_i, \ldots$, and so on. If the map is run from $n = -\infty$ to $n = +\infty$, the history is then the doubly infinite sequence

$$\ldots b_i \ldots b_3 b_2 b_1 a_1 a_2 a_3 \ldots a_i \ldots \tag{4.7.10}$$

(Recall that in the reverse direction (i.e., $-\infty \leq n \leq 0$), x and y deformations are interchanged.) The crucial point is that for typical, irrational, initial coordinates (x_0, y_0), the associated binary representations (4.7.7) are infinite nonrepeating sequences of zeros and ones and hence the doubly infinite sequence (4.7.10) is as random as that which would be obtained by a fair coin toss (1 for heads, 0 for tails). Thus a completely deterministic dynamical system (4.7.6) generates motion that appears to be completely random! Such a system is known as a *Bernoulli system* and represents the

ultimate in randomness. A most significant result in dynamical systems theory is the demonstration that near any homoclinic point of a mapping, the motion can be locally represented by a mapping with the Bernoulli property. We have already seen that homoclinic points are dense in the neighborhood of hyperbolic fixed points in nonintegrable Hamiltonian systems; so the above result reinforces the idea that the chaotic, yet deterministic, trajectories observed in these systems have a truly random nature. Some nice illustrations of this concept have been given by Berry (1978).

4.7.d Hierarchies of Randomness

From the above examples it is clear that different types of dynamical system can have greater or lesser degrees of randomness. An approximate classification of this hierarchy is as follows

(i) Ergodic System. This is the "weakest" type of behavior in which phase average equals time average, that is,

$$\lim_{T\to\infty} \int_{-T}^{T} f(x, t)\, dt = \langle f(x, t) \rangle \tag{4.7.11}$$

where $\langle\ \rangle$ denotes ensemble average over the manifold being considered. Simple examples are irrational flow on a torus or, for one-dimensional systems, flow on the energy shell (see Section 2.5). It is important to recall that generic n-dimensional Hamiltonians ($n > 1$) are not ergodic over the entire energy shell, since this is typically divided by (surviving) tori.

(ii) Mixing Systems. As illustrated, this is a much stronger property than ergodicity. By contrast with (4.7.11), mixing implies that

$$\lim_{t\to\infty} f(x, t) = \langle f(x, t) \rangle \tag{4.7.12}$$

that is, no time averaging is required to achieve "equilibrium." It may be shown that the spectrum of mixing systems is continuous, whereas that of ergodic systems is discrete. (There is, in fact, an intermediate state termed *weakly mixing* which is sufficient for a continuous spectrum.)

(iii) K-Systems. These are systems with positive Kolmogorov entropy. This means that a connected neighborhood of trajectories must exhibit a positive average rate of exponential divergence.

(iv) C-Systems (also called *Anosov Systems*). These are systems which are *globally* unstable, that is, every trajectory has a positive, Lyapunov exponent. The cat map is an example of a C-system.

(v) *Bernoulli Systems*. These are systems whose motion is as random as a fair coin toss (e.g., the baker's transformation).

Any member of the hierarchy also exhibits the properties of "lower" members. So, for example, the cat map, which is a C-system, also has the properties of a K-system (in this case the KS entropy is $\ln[(3+\sqrt{5}/2)]$ and exhibits mixing, which, in turn, implies ergodicity).

4.8 HAMILTONIAN CHAOS IN FLUIDS

So far, all our discussions of chaos have been illustrated by computer studies of simple model systems. Furthermore, we have done little more than pay lip-service to the physical contexts in which these models might arise (e.g., nonlinear oscillations, accelerators, wave guides, molecules in radiation fields, etc.). Clearly, the reader would be much happier if there was some "real-life" physical situation in which he could actually *see* chaos—possibly even a surface of section—with the naked eye. As it turns out, this wish can be fulfilled for certain classes of fluid dynamical problems, which we shall now describe.

4.8.a Fluid Mechanical Background

The first step is to understand the difference between the "Eulerian" and "Lagrangian" descriptions of fluid dynamics. The former specifies the velocity field, $\mathbf{u} = (u, v, w)$, of a fluid with respect to a fixed coordinate frame, that is,

$$u = u(x, y, z, t) \tag{4.8.1a}$$

$$v = v(x, y, z, t) \tag{4.8.1b}$$

$$w = w(x, y, z, t) \tag{4.8.1c}$$

where (u, v, w) are obtained (in principle) by solving the fluid dynamical equations (see Section 5.1) subject to the specified boundary conditions. If the velocity field is explicitly dependent on time, it is termed *unsteady*, in contrast to the time-independent *steady* fields. On the other hand, the Lagrangian description involves the following individual fluid "particle" trajectories. Thus for a given velocity field, one follows a particle by solving the set of ordinary differential equations

$$\dot{x} = u(x, y, z, t) \tag{4.8.2a}$$

$$\dot{y} = v(x, y, z, t) \tag{4.8.2b}$$

$$\dot{z} = w(x, y, z, t) \tag{4.8.2c}$$

subject to the initial conditions $(x(0), y(0), z(0))$.

For an incompressible two-dimensional fluid, one has

$$u_x + v_y = 0 \tag{4.8.3}$$

which tells us that there must be an exact differential $d\psi$ such that

$$u = \frac{\partial \psi}{\partial y} \tag{4.8.4a}$$

$$v = -\frac{\partial \psi}{\partial x} \tag{4.8.4b}$$

The function $\psi = \psi(x, y, t)$ is termed the *stream function*. If one is using the Lagrangian description, one can then write the equations of motion (4.8.2) in the form

$$\dot{x} = \frac{\partial \psi}{\partial y}(x, y, t) \tag{4.8.5a}$$

$$\dot{y} = -\frac{\partial \psi}{\partial x}(x, y, t) \tag{4.8.5b}$$

which has a Hamiltonian structure, with ψ playing the role of the Hamiltonian and x and y being the canonical variables. It is worth emphasizing that this Hamiltonian structure stems from the incompressibility condition (4.8.3) and is valid whether the fluid is viscous or not. Thus, in two dimensions, one is able to determine the fluid particle paths by following the phase-space dynamics of Eqs. (4.8.5).

The precise dynamics will, of course, depend on the nature of ψ. For steady flows, ψ is time independent and Eqs. (4.8.5) reduce to the autonomous system

$$\dot{x} = \psi_y(x, y) \tag{4.8.6a}$$

$$\dot{y} = -\psi_x(x, y) \tag{4.8.6b}$$

which we know, from chapter 1, to be completely integrable with the trajectories lying on smooth curves—termed *stream lines* in fluid dynamics—in the (x, y) phase plane. However, for unsteady flows (i.e., time-dependent ψ), we know that there is the possibility of chaotic behavior. In order to actually follow the fluid-particle trajectories, they must be tagged in some "passive" way, that is, in such a way that they still

follow the dynamics of (4.8.5) without affecting the velocity field. This can usually be achieved with dye, but of course in practice this leads to a distribution ("passive scalar distribution") of tagged fluid particles (e.g., a streak or blob being tagged rather than an individual particle). Thus one is then observing the phase-space evolution of a whole family of trajectories. The deformations of such a family (line element or curve) were discussed in Section 4.4—namely, that line elements on the plane can develop two principal structures, "whorls" and "tendrils," corresponding to the presence of elliptic and hyperbolic fixed points, respectively. The appearance of tendrils is thus a manifestation of fluid-particle chaos in the fluid—a phenomenon that is also referred to as *chaotic advection* or *Lagrangian turbulence*, the latter term being invoked to imply the idea of chaos within the Lagrangian picture of fluid dynamics. (Notions of *Eulerian turbulence* will be briefly discussed in Chapter 5.) Finally we mention that whereas in two dimensions an unsteady flow is needed for the appearance of chaos, in three dimensions chaos can occur for steady flows (i.e., Eqs. (4.8.2) with time independent right-hand sides).

4.8.b The Model System

In practice, we would like to find a fluid dynamical system that is two-dimensional, unsteady, *and* one for which the associated stream function is known explicitly—thereby enabling us to perform computer studies to compare with laboratory experiments. Although this might sound somewhat like a tall order, such a system can be constructed and was the subject of a recent, detailed study by Chaiken et al. (1986). The system in question is closely related to what is known as a *journal bearing*. This corresponds to two cylinders, one placed within the other with the gap between them filled with a viscous fluid (traditionally a heavy lubricant). Both cylinders can rotate independently about their axes, which are typically set not to coincide with each other (i.e., the cylinders are eccentric). If the fluid depth is sufficiently large compared to the (outer) cylinder radius, the system can be considered as approximately two-dimensional. If the fluid viscosity is sufficiently high and the cylinder rotation rates are sufficiently low (leading to the so-called *Reynolds number* being very small), the fluid equations can be solved in the Stokes' approximation, that is,

$$\nu \nabla^2 \mathbf{u} = \nabla p \tag{4.8.7}$$

where $\mathbf{u} = (u, v)$ is the two-dimensional velocity field, ν is the kinematic viscosity, and p is the pressure. The boundary conditions are that the component of fluid velocity normal to inner and outer cylinder boundaries must equal the corresponding component of the associated cylinder velocities. Provided that the cylinders (one or both) are rotating at a *steady*

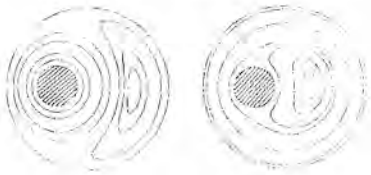


Figure 4.33 Typical streamlines for (a) inner-cylinder-only rotation and (*b*) outer-cylinder-only rotation. Curves plotted are exact solutions to Eq. (4.8.8) with appropriate boundary conditions. Hatched regions correspond to inner cylinder.

(angular) velocity, (4.8.7) is a steady-state problem. Furthermore, by noting that the curl operation kills gradients (i.e., $\nabla \times (\nabla p) = 0$) and that $\nabla \times \mathbf{u} = -\nabla^2 \psi$, (4.8.7) can be reduced to the biharmonic equation

$$\nabla^4 \psi = 0 \tag{4.8.8}$$

For the particular geometry of the system discussed here, it has been known for some time that this equation admits an exact, closed form solution in bipolar coordinates. Some typical results are shown in Figure 4.33 for the cases of inner-cylinder-only and outer-cylinder-only rotation. What is plotted here are curves of constant ψ in the (x, y)-plane—streamlines in fluid dynamical terms, invariant tori in dynamical terms. For these steady-state solutions, the fluid particles behave in an entirely regular way following their respective streamlines.

We now need to introduce some (controlled) unsteadiness into the system. This can be achieved by modulating the cylinder rotations. The simplest case of this is to rotate the cylinders alternately in time. The solution to (4.8.8) is then a piecewise linear combination of the solutions corresponding to the separate cylinder motions. Thus as one cylinder

Figure 4.34 Typical surface of section computed from Eqs. (4.8.5). (Reproduced, with permission, from Chaiken et al. (1986).)

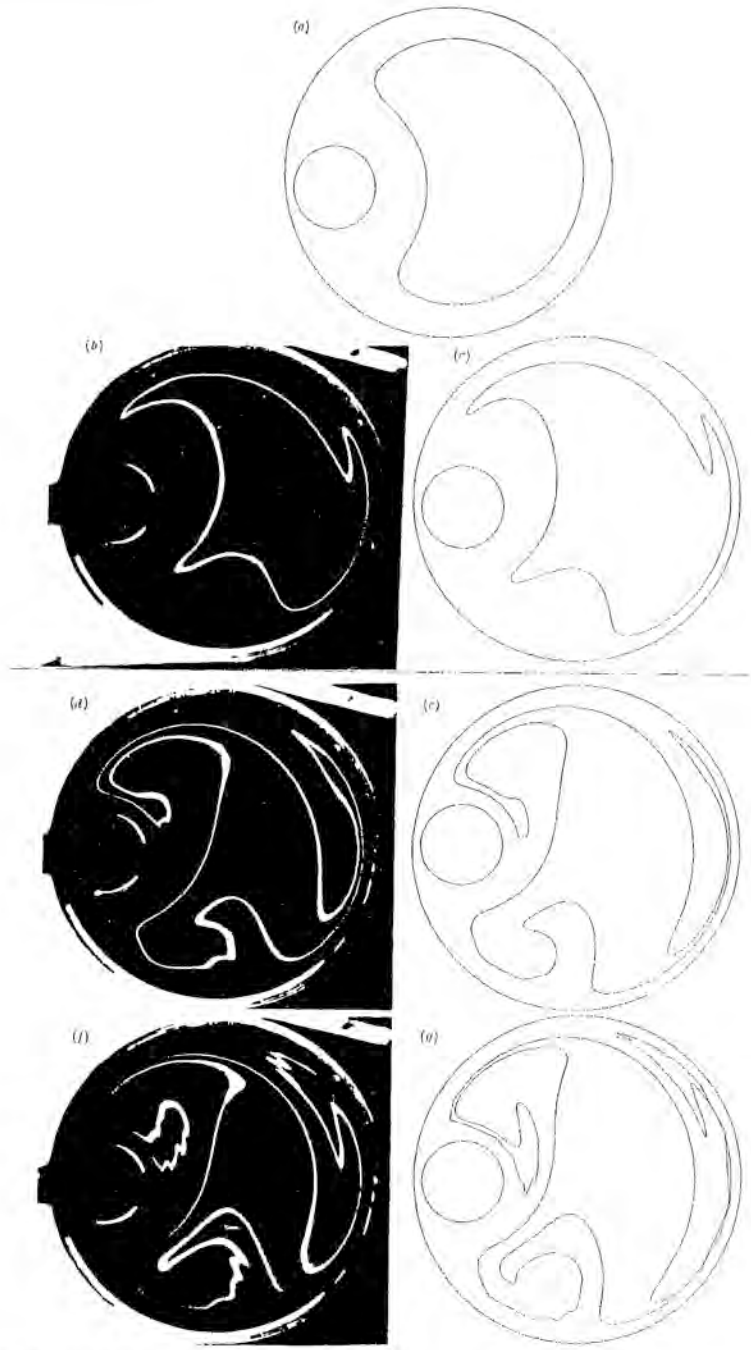


Figure 4.35 (a) Initial contour generated by rotation of outer cylinder. Comparison of experiment and computation after: 7 periods (*b*) and (*c*); 15 periods (*d*) and (e); 21 periods (*f*) and (*g*). (Reproduced, with permission, from Chaiken et al. (1986).)

rotates, a fluid particle follows its streamline and then jumps (virtually instantaneously in a Stokes' flow) to a different streamline as the cylinder motions are switched. It is this mechanism that can lead to chaos as the particles evolve in the three-dimensional phase space (x, y, t). The motion is conveniently followed by constructing a surface of section by taking *strobosopic* "snapshots" of the (x, y) phase plane—the time interval being the sum of the two cylinder rotation times (see Figure 4.6). A typical such surface of secton, generated by solving Eqs. (4.8.5) numerically, is shown in Figure 4.34 exhibiting the generic mixture of regular and irregular motions.

4.8.c Experimental Results

The formation of whorls and tendrils can be used to identify regions of regular and irregular motion, respectively. For example, in the above case an evolving line element (i.e., a streak of dyed fluid) will form whorls in the neighborhood of the main three-island chain and tendrils in the intervening chaotic regions. In Figure 4.35 we show the evolution of an initial curve embracing these regions. As the sequence clearly shows, three large whorls are formed on which are superimposed small homoclinic oscillations (tendrils). The computer simulations are compared directly with the laboratory experiment in which a matching curve of dyed fluid (glycerine in this experiment) placed on the fluid surface evolves under exactly the same cylinder motions (alternate piecewise rotations) as used in the computer studies. (The apparatus is shown in Figure 4.36.)

As the cylinder parameters (i.e., rotation rates and eccentricity) are varied, phase planes exhibiting different degrees of chaos are found. In Figure 4.37 we show the case of a strongly chaotic phase plane with only a few tiny island regions left. In this situation we would expect to see the formation of giant tendrils. This is demonstrated in the sequence in Figure 4.38. Computer simulations become difficult in this case since the exponential separation of neighboring points leads to a rapid breakdown of numerical resolution. (To overcome this the initial curve would have to be packed with an enormous number of initial points.) By contrast, the laboratory experiment has no such problems and in Figure 4.39 we show the striking result of running the experiment for several periods longer than can be followed numerically. As it turns out, the laboratory experiment can also be used to construct surfaces of section. (The reader is referred to the original paper (Chaiken et al., 1986) for the technical details.) In Figure 4.40 we show a laboratory surface of section obtained at the parameters corresponding to the numerical simulation shown in Figure 4.34. Quite good agreement is obtained, with the laboratory result clearly showing the correct elliptic and hyperbolic regions. This is one of the few examples of a surface of section of a chaotic Hamiltonian flow being directly visualized in the laboratory.

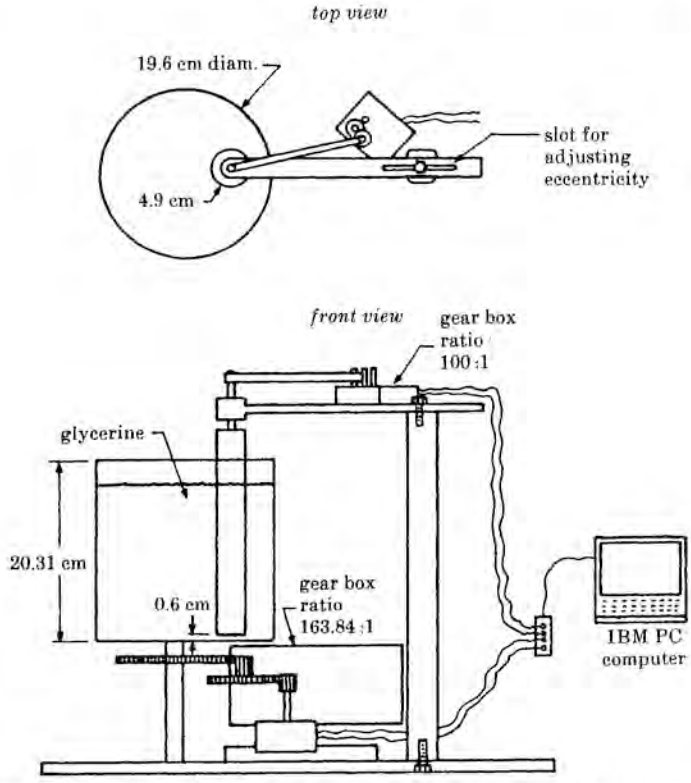


Figure 4.36 Schematic diagram of experimental apparatus. Contours of dyed fluid are placed on surface of glycerine, and their evolution is followed under controlled cylinder motions. (Reproduced, with permission, from Chaiken et al. (1986).)

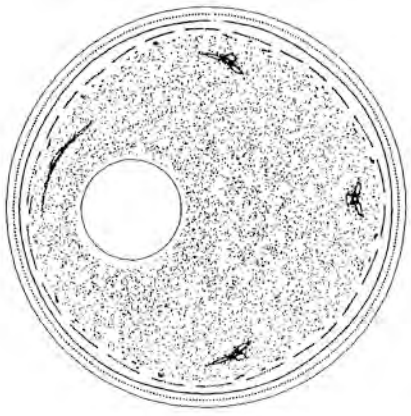


Figure 4.37 Strongly chaotic surface of section computed from Eqs. (4.8.5). (Reproduced, with permission, from Chaiken et al. (1986).)

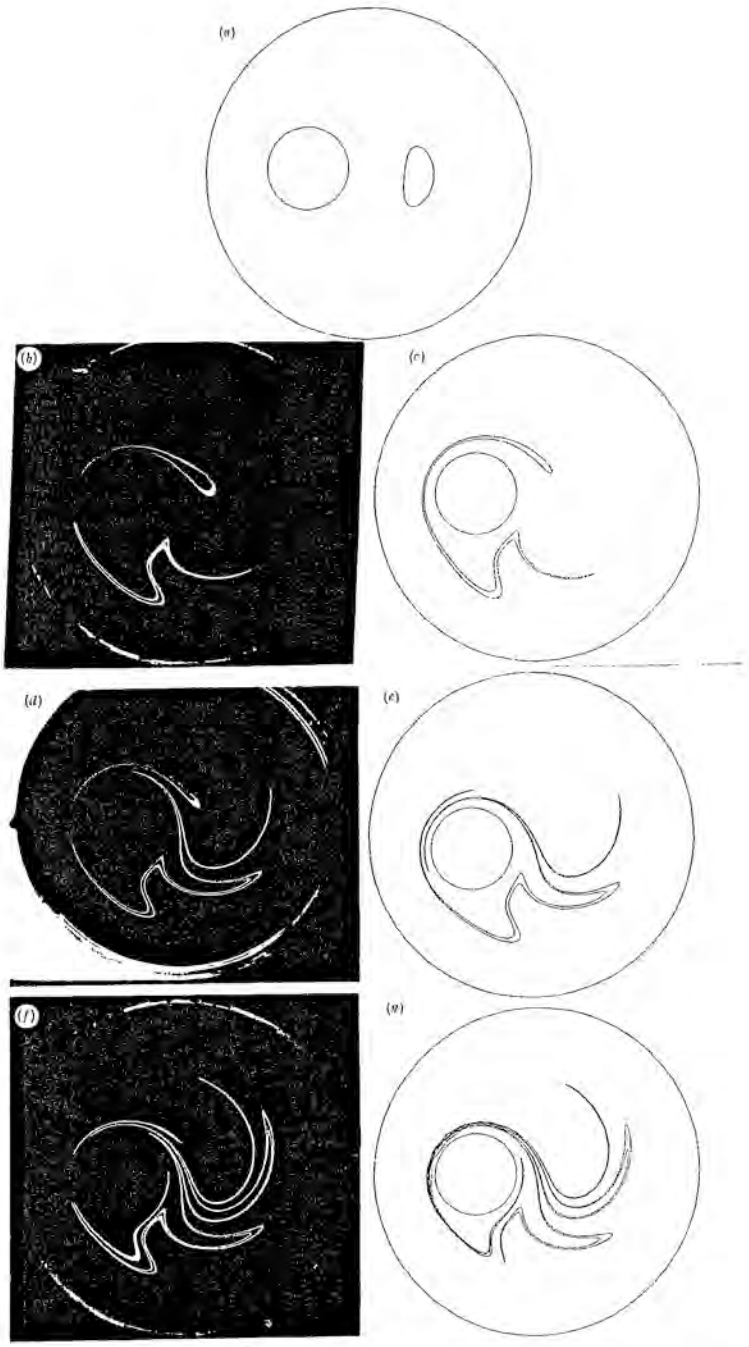


Figure 4.38 Initial contour (*a*) generated by rotation of outer cylinder. Comparison of experiment and computation after: 4 periods (*b*) and (*c*); 5 periods (*d*) and (*e*); 6 periods (*f*) and (*g*). (Reproduced, with permission, from Chaiken et al. (1986).)

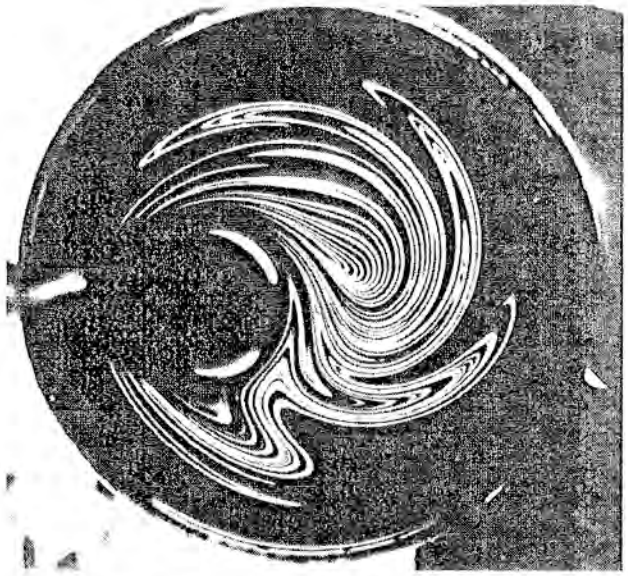


Figure 4.39 Same experiment as in Figure 4.38 after 12 periods. (Reproduced, with permission, from Chaiken et al. (1986).)

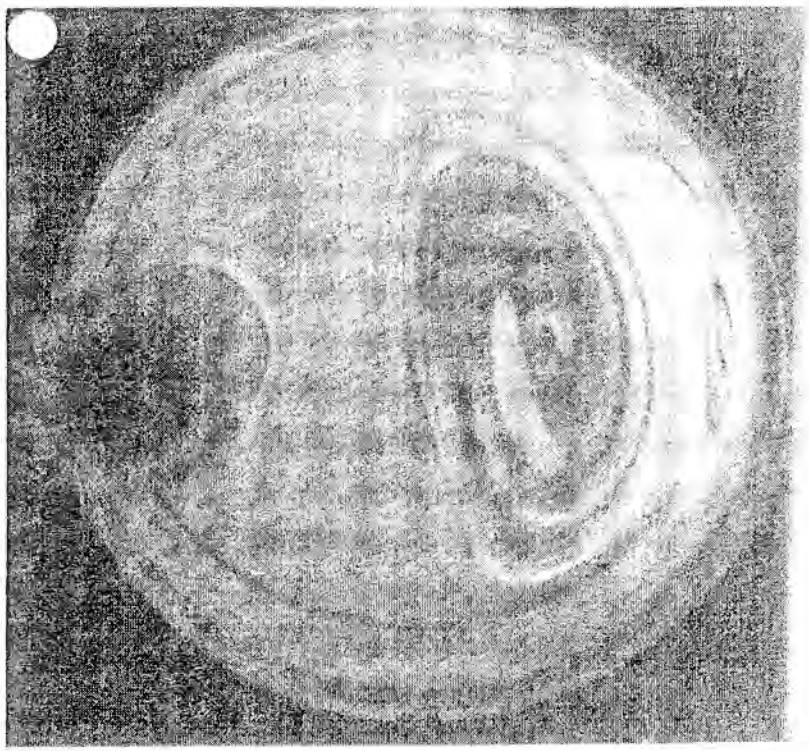


Figure 4.40 Experimental surface of section to be compared with numerically obtained result shown in Figure 4.34. (Reproduced, with permission, from Chaiken et al. (1986).)

APPENDIX 4.1 SURFACE OF SECTION AS A SYMPLECTIC MAPPING

The special "area-preserving" property of the surface of section for two-degree-of-freedom, conservative systems can be seen most elegantly in the geometric language introduced in Appendix 2.2.

From the invariance properties of the Poincaré–Cartan 1-form $\sum_{i=1}^{n} p_i\, dq_i - H\, dt$ we have (cf. Eq. (2.A.29))

$$\oint_{\mathscr{C}} \sum_{i=1}^{n} p_i\, dq_i - H\, dt = \oint_{\mathscr{C}'} \sum_{i=1}^{n} p_i\, dq_i - H\, dt \tag{4.A.1}$$

where $\mathscr{C}$ and $\mathscr{C}'$ are any two closed curves enclosing the same tube of phase-space trajectories. If $\mathscr{C}$ and $\mathscr{C}'$ are curves taken at constant time slices, $t = 0$ for curve $\mathscr{C}$ and $t = T$ for curve $\mathscr{C}'$ (now denoted as $\mathscr{C}_T$), there is no contribution from the $H\, dt$ terms and we obtain the result (4.1.9) (or Eq. (2.A.31)), namely,

$$\oint_{\mathscr{C}} \sum_{i=1}^{n} p_i\, dq_i = \oint_{\mathscr{C}_T} \sum_{i=1}^{n} p_i\, dq_i \tag{4.A.2}$$

Now consider the two-degree-of-freedom case with the 1-form $p_x\, dx + p_y\, dy$ and define an "initial" curve $\mathscr{C}$ as a set of initial conditions on the (p, x) plane at $y = 0$ at a fixed energy E. That is, $\mathscr{C}$ is a curve of points on the surface of section. Under the Hamiltonian flow these points will evolve, forming a "tube" of trajectories in the "extended" phase space $(p_x, p_y, x, y; t)$. For bounded motion this tube will eventually pass back through the surface of section at $y = 0$. However, there is no reason to assume that each point on the tube will pass through $y = 0$ at the same time. Thus, although the reintersection of the tube with the surface of section will be some closed curve, say $\mathscr{C}'$, it will not be of the "fixed time slice" type $\mathscr{C}_T$ used in (4.A.2). Thus we must consider the Poincaré–Cartan invariant in (4.A.1), namely,

$$\oint_{\mathscr{C}} p_x\, dx + p_y\, dy - H\, dt = \oint_{\mathscr{C}'} p_x\, dx + p_y\, dy - H\, dt \tag{4.A.3}$$

However, for a family of trajectories on a given energy shell $E = H =$ constant we note that

$$\oint_{\mathscr{C}} H\, dt = \oint_{\mathscr{C}'} H\, dt = 0 \tag{4.A.4}$$

Furthermore, since the curves $\mathscr{C}$ and $\mathscr{C}'$ are defined on planes with fixed y

$(y = 0)$,

$$\oint_{\mathscr{C}} p_y \, dy = \oint_{\mathscr{C}'} p_y \, dy = 0 \tag{4.A.5}$$

Thus we are left with the result

$$\oint_{\mathscr{C}} p_x \, dx = \oint_{\mathscr{C}'} p_x \, dx \tag{4.A.6}$$

namely, that the area on a surface of section is conserved under the Hamiltonian flow. That is why we can speak of the surface of section, for a two-degree-of-freedom system, as an *area*-preserving map.

For more than 2 degrees of freedom the same idea holds but now the "surface of section" is a $(2n-2)$-dimensional surface embedded in the $(2n-1)$-dimensional energy shell (of a conservative n-degree-of-freedom Hamiltonian). If the "surface" is defined at $q_1 = \text{constant}$, then (4.A.3) generalizes to

$$\oint_{\mathscr{C}} \sum_{i=2}^{n} p_i \, dq_i + p_1 \, dq_1 - H \, dt = \oint_{\mathscr{C}'} \sum_{i=2}^{n} p_i \, dq_i + p_1 \, dq_1 - H \, dt \tag{4.A.7}$$

By analogy with (4.A.4) and (4.A.5) we have vanishing contributions from the terms $\oint p_1 \, dq_1$ and $\oint H \, dt$ about both $\mathscr{C}$ and $\mathscr{C}'$ and are hence left with the result

$$\oint_{\mathscr{C}} \sum_{i=2}^{n} p_i \, dq_i = \oint_{\mathscr{C}'} \sum_{i=2}^{n} p_i \, dq_i \tag{4.A.8}$$

These integrals no longer correspond to simple areas but rather the projected areas on the various (p_i, q_i) planes (see (2.A.50)), namely,

$$\oint_{\mathscr{C}} \sum_{i=2}^{n} p_i \, dq_i = \sum_{i=2}^{n} \int\int_{A_i} dp_i \, dq_i$$

$$= \oint_{\mathscr{C}'} \sum_{i=2}^{n} p_i \, dq_i = \sum_{i=2}^{n} \int\int_{A_i'} dp_i \, dq_i \tag{4.A.9}$$

where the A_i are the various projections of $\mathscr{C}$ and A_i' the projections of $\mathscr{C}'$. Thus we see that the surface of section is a symplectic mapping.

SOURCES AND REFERENCES

Texts and General Review Articles

Arnold, V. I., and A. Avez, *Ergodic Problems of Classical Mechanics*, Benjamin, New York, 1968.

Berry, M. V., Regular and irregular motion, AIP Conference Proceedings, No. 46, *Topics in nonlinear dynamics*, AIP, New York, 1978.

Birkhoff, G. D., *Dynamical Systems*, American Mathematical Society, Providence, RI, 1927.

Ford, J., The statistical mechanics of analytical dynamics, in E. D. G. Cohen, Ed., *Fundamental Problems in Statistical Mechanics*, Vol. 3, North-Holland, Amsterdam, 1975.

Helleman, R. H. G., Self generated chaotic behaviour in nonlinear mechanics, in E. D. G. Cohen, Ed., *Fundamental Problems in Statistical Mechanics*, Vol. 5, North-Holland, Amsterdam, 1980.

Lichtenberg, A. J., and M. A. Lieberman, *Regular and Stochastic Motion*, Springer-Verlag, New York, 1983.

MacKay, R. S., and J. D. Meiss, *Hamiltonian Dynamical Systems*, Adam Hilger, Bristol, 1987. This volume contains a valuable compilation and extensive bibliography of many fundamental research papers in Hamiltonian dynamics.

Moser, J., *Stable and Random Motions in Dynamical Systems*, Annals of Mathematical Studies, Vol. 77, Princeton University Press, Princeton, NJ, 1973.

Poincaré, H., *Les Methods Nouvelles de la Mechanique Celeste*, Gauthier-Villars, Paris, 1892.

Section 4.1

Casati, G., and J. Ford, Stochastic transition in the unequal-mass Toda lattice, *Phys. Rev.*, **A12**, 1702 (1975).

Contopoulos, G., and C. Polymilis, Approximations of the 3-particle Toda lattice, *Physica*, **24D**, 328 (1987).

Ford, J., S. D. Stoddard, and J. S. Turner, On the integrability of the Toda lattice, *Prog. Theor. Phys.*, **50**, 1574 (1973).

Henon, M., and C. Heiles, The applicability of the third integral of motion: some numerical experiments, *Astron. J.*, **69**, 73 (1964).

Henon, M., Integrals of the Toda lattice, *Phys. Rev.*, **B9**, 1921 (1974). See also accompanying paper by H. Flaschka.

Section 4.2

Henon, M., Numerical study of quadratic area preserving mappings, *Quart. Appl. Math.*, **27**, 291 (1969).

MacKay, M. S., J. D. Meiss, and I. C. Percival, Transport in Hamiltonian systems, *Physica*, **13D**, 55 (1984). Cantori and their applications are discussed.

MacKay, R. S., Introduction to the dynamics of area preserving maps, *Proceedings*

of the Spring College on Plasma Physics, Trieste, 1985, World Scientific, Singapore, 1985. A good geometrical introduction to area preserving maps is given.

Percival, I. C., Variational principles for invariant tori and cantori, AIP Conference Proceedings, No. 57, *Nonlinear Dynamics and the Beam–Beam Interaction*, AIP, New York, 1980. Discrete Lagrangians are introduced.

Section 4.5

Benettin, G., L. Galgani, and J. M. Strelcyn, Kolmogorov entropy and numerical experiments, *Phys. Rev.*, **A14**, 2338 (1976).

Benettin, G., L. Galgani, A. Giorgilli, and J. M. Strelcyn, Lyapunov characteristic exponents for smooth dynamical systems and for Hamiltonian systems; a method for computing all of them, *Meccanica*, **15**, 9 (1980).

Pesin, Ya. B., Characteristic Lyapunov exponents and smooth ergodic theory, *Russ. Math. Surveys*, **32**(4), 55 (1977).

Section 4.6

Chirikov, B., A universal instability of many dimensional oscillator systems, *Phys. Reports*, **52**, 263 (1979).

Escande, D. F., Stochasticity in classical Hamiltonian systems: universal aspects, *Phys. Reports*, **121**, 165 (1985). Further developments of the Chirikov method are described.

Greene, J. M., A method for determining a stochastic transition, *J. Math. Phys.*, **20**, 1183 (1979).

Helleman, R. H. G., and T. Bountis, Periodic solutions of arbitrary period, variational methods, Lecture Notes in Physics, Vol. 93, *Stochastic Behavior in Classical and Quantal Hamiltonian Systems*, Springer-Verlag, New York, 1979.

Tabor, M., The onset of chaos in dynamical systems, *Adv. Chem. Phys.*, **46**, 73 (1981).

Walker, G. H., and J. Ford, Amplitude instability and ergodic behavior for conservative nonlinear oscillator systems, *Phys. Rev.*, **188**, 416 (1969).

Section 4.7

Lebowitz, J., and O. Penrose, Modern ergodic theory, *Phys. Today*, **26**, 23 (1973).

Zaslavskii, G. M., and B. Chirikov, Stochastic instability of nonlinear oscillations, *Sov. Phys. Usp.*, **14**, 549 (1972).

Section 4.8

Chaiken, J., R. Chevray, M. Tabor, and Q. M. Tan, Experimental study of Lagrangian turbulence in a Stokes' flow, *Proc. R. Soc. London A*, **408**, 165 (1986).