Physica D 59 (1992) 25–38
North-Holland

PHYSICA D

A nonsmooth Coulomb friction oscillator

Brian Feeny[1]
Dept. of Mechanical Engineering, Michigan State University, East Lansing, MI 48824, USA

Received 1 February 1991
Revised manuscript received 30 March 1992
Accepted 30 March 1992
Communicated by J. Guckenheimer

A forced Coulomb friction oscillator, whose frictional force is allowed to vary with displacement, is analyzed geometrically. The equation of motion for the oscillator is piecewise linear. We geometrically observe the nature of the flow in each region of solvability, and then see how these solutions interact at the boundary of the regions. The dynamics of the flow is viewed in terms of a map on the boundary between the regions. For chaotic motion, we geometrically construct the strange attractor, and show that its exact behavior is that of a one-dimensional map. The following dynamical properties arise from the nonsmooth nature of the Coulomb friction law: the flow may not be invertible; the flow may reach its attractor in finite time; the dimension of the attractor may be less than or equal to two; embeddings of an observable may not be diffeomorphic to the full phase flow.

1. Introduction

A friction law that is often used in dynamical applications is the Coulomb friction law, which states that there is a static coefficient of friction μ_s, and a kinetic coefficient of friction μ_k. When there is relative sliding between bodies in contact (with a normal load N), a friction force $\mu_k N$ opposes the motion. When there is no sliding, a friction force of magnitude less than or equal to $\mu_S N$ balances the external forces. This friction law can be written as $F = -Nf(\dot{x})$, where

$$\begin{aligned} f(\dot{x}) &= -\mu_k\,, & \dot{x} &< 0\,,\\ -\mu_s \le f(\dot{x}) &\le \mu_s\,, & \dot{x} &= 0\,,\\ f(\dot{x}) &= \mu_k\,, & \dot{x} &> 0\,. \end{aligned} \tag{1}$$

We apply this friction law to a forced harmonic oscillator. The nondimensional equation of motion is

$$\ddot{x} + 2\zeta\dot{x} + x + n(x)\, f(\dot{x}) = a\cos(\Omega t)\,, \tag{2}$$

where $a\cos(\Omega t)$ is the harmonic forcing function, ζ is the damping coefficient present when the dry friction is removed, and $n(x)$ represents a normal load which may vary with displacement. In particular, we let $n(x) = 1 + kx$ for $x > -1/k$, and $n(x) = 0$ for $x < -1/k$ to prevent friction from occurring when the normal load between the contact surfaces becomes negative (contact is lost). This will be referred to as the *realizability condition*, since negative friction would not be observed in a mechanical system.

An interesting feature of this ordinary differential equation of motion is that, because of the friction function of eq. (1), it is *discontinuous* and *multivalued* at $\dot{x} = 0$. It is multivalued in that $f(0)$ can take on any value between $-\mu_s$ and μ_s. Examples of other systems whose models may have discontinuous functions are those involving

[1]This work was done at the department of Theoretical and Applied Mechanics, Cornell University, Ithaca, New York 14853.

impacts, shock and shock absorbtion, and sticky contacts.

One approach to such a problem is to view it as a piecewise continuous system, examine the continuous pieces, and match them. Alternatively, we might look at the behavior of a sequence of continuous systems which in some sense converges to the discontinuous system. This paper takes the former approach.

A well-known consequence of having this Coulomb function in the ODE is *stick–slip* motion [1, 2]. Sometimes during the motion, when the velocity passes through $\dot{x} = 0$, the static friction may balance the external forces. When this happens, the oscillator remains stuck at zero velocity until the driving function reaches a critical value such that the friction no longer balances the external forces, and motion resumes (slips).

The case where $k = 0$, and hence $n(x) \equiv 1$, has been studied by many researchers. In 1931, Den Hartog [1] solved this system for periodic motion. He did this by breaking the problem into piecewise linear equations and exploiting symmetry in x. More recently, Shaw [2] extended Den Hartog's results and analyzed the stability of these period orbits. Shaw found stable periodic slipping motions, period-one and period-two sticking motions, and beat phenomena for $\zeta < 1$.

Unfortunately, by including $k \neq 0$, thereby causing $n(x)$ to be active and nonconstant in eq. (2), we lose the symmetry in x. Hence, the calculation of periodic orbits and their stabilities in the spirit of [1] and [2] is extremely difficult. We choose another approach for the analysis: we graphically examine the qualitative nature of the system. As a result, we are able to geometrically construct the attractor, and make some further statements about its behavior. This type of analysis has been done in classic nonlinear-vibrations texts [3, 4] for two-dimensional autonomous systems.

Eq. (2) is piecewise integrable, that is, it is solvable in subregions of the state space. We geometrically observe the nature of the flow in each of these regions of solvability, and then see how these solutions interact at the boundary of the regions. The dynamics of the flow are viewed in terms of a map on the boundary between the regions. From a qualitative picture of this map, we can build the attractor, and show that

(1) the dynamical behavior reduces rigorously to a one-dimensional map,

(2) the flow of the Coulomb oscillator may not be invertible (previously reported by Shaw [2]),

(3) the flow may reach its attractor in finite time, and

(4) the attractor has dimension less than or equal to two.

Takens [5] had observed these properties in *constrained systems*, which have been discussed in detail by Oka [6]. These properties are made possible because the Coulomb friction law produces a *discontinuous* and *multivalued* vector field (the same mechanism responsible for stick–slip motion). A consequence is that embeddings of an observable may not be diffeomorphic to the full phase flow. Other properties may arise in other systems with discontinuities. For example, in a problem modeling shock, Antman [7] uncovered nonunique solutions.

2. Piecewise linear equations

To simplify the analysis, we look at the special case of eq. (2) where $\mu_s = \mu_k = 1$. Based on dynamical friction measurements in an experimental oscillator with titanium contacts, this is a reasonable assumption [8]. Further, we neglect the viscous damping term $\zeta\dot{x}$ (in the experimental oscillator, the damping ratio was measured as $\zeta = 0.015$ [9]). Finally, out of interest, we will only look at the case where $k > 1$ (as k passes through 1, the nature of the *sticking region*, described in section 2.1, undergoes a significant transition [9]). Eq. (2) can be written for regions in which they are solvable:

$$\ddot{x} + (1 + k)x = -1 + a\cos(\Omega t)\,,$$
$$\dot{x} > 0,\ x > -1/k\,, \tag{3a}$$

$$\ddot{x} + (1 - k)x = 1 + a\cos(\Omega t)\,,$$
$$\dot{x} < 0,\ x > -1/k\,. \tag{3b}$$

For $x \le -1/k$, the normal load becomes zero due to the realizability condition, and we have

$$\ddot{x} + x = a\cos(\Omega t)\,,\quad x < -1/k\,. \tag{3c}$$

For $-1 < k < 1$, eq. (3b) has harmonic solutions. For $k > 1$, the solution of eq. (3b) consists of a driven saddle.

Let D denote the (x, t)-plane ($\dot{x} = 0$) and consider the mappings of the flow at D:

$$P^+\colon \oplus \to D\,,\quad \dot{x} > 0\,,$$
$$P^-\colon \ominus \to D\,,\quad \dot{x} < 0\,,$$

where P^+ is the map that arises from the flow of eq. (3a), and P^- is the map that arises from the flow of eq. (3b), and $\oplus \subset D$ and $\ominus \subset D$ denote the domains of P^+ and P^-, respectively. The goal is to describe the maps P^+ and P^- geometrically, and see how they interact with the *sticking region*, described below. The idea is sketched in fig. 1. For simplicity, we will start the discussion by omitting eq. (3c), which arises from the realizability condition. Thus we will discuss the mapping between the flow of eq. (3a) and eq. (3b). This analysis is of interest anyway, since regular and chaotic motions confined to the region $x > -1/k$ are observed in numerical integrations. The results will be completed by including eq. (3c) numerically.

2.1. The sticking region

Analysis of eq. (2) for $\dot{x} = 0$ leads to the concept of *sticking regions*, also discussed in [1] and [2]. First, it is helpful to write eq. (2) as a system of first-order ODEs:

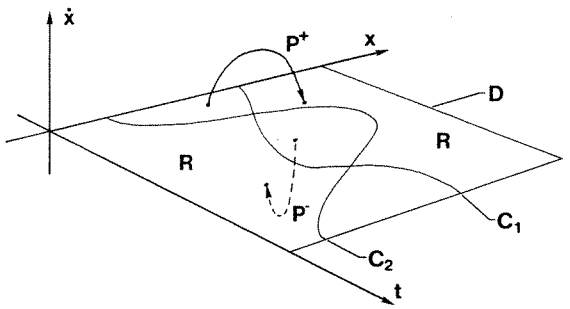

Fig. 1. Trajectories governed by each piecewise linear equation are associated with either the map $P^+\colon \oplus \to D$, for trajectories in $\dot{x} > 0$, or the map $P^-\colon \ominus \to D$, for trajectories in $\dot{x} < 0$. Some orbits get mapped into the sticking region R, where the motion remains constant until such time that the point is no longer in the sticking region. The curves C_1 and C_2 represent the boundaries, in D, between the sticking regions R and the domains $\oplus$ and $\ominus$ of P^+ and P^-, respectively.

$$\dot{x}_1 = x_2\,,$$
$$\dot{x}_2 = -x_1 - n(x_1)\, f(x_2) + a\cos(\Omega t)\,. \tag{4}$$

The sticking region can be found by analyzing eq. (4) for fixed points and using the multivaluedness of $f(x_2)$ at $x_2 = 0$ [2, 9]. However, we will describe the sticking regions using another viewpoint which will help set the mood of this analysis.

If we write eqs. (4) in extended phase space for $\dot{x} > 0$, we have

$$\dot{x}_1 = x_2\,,$$
$$\dot{x}_2 = -(1 + k)x_1 - 1 + a\cos(\Omega t)\,,\quad x_2 > 0\,,$$
$$\dot{t} = 1\,. \tag{5}$$

By looking at the sign of $\dot{x}_2$ in the (x, t)-plane D, where $\dot{x} = 0$, we find regions where the flow of eqs. (5) is upward and regions where the flow is downward. The curve (in the plane D) dividing the regions, called C_1, is given by $\dot{x}_2 = 0$, which yields

$$C_1\colon x_1 = \frac{-1 + a\cos(\Omega t)}{1 + k}\,.$$

We can do the same for the case of $\dot{x} < 0$. The equations in extended phase space are

$$\dot{x}_1 = x_2 ,$$
$$\dot{x}_2 = -(1-k)x_1 + 1 + a\cos(\Omega t) , \quad x_2 < 0 ,$$
$$\dot{t} = 1 . \tag{6}$$

The regions for upward and downward flow on the (x, t)-plane D for eqs. (6) are separated by a curve C_2, which is given by $\dot{x}_2 = 0$, or

$$C_2 : x_1 = \frac{1 + a\cos(\Omega t)}{1-k} .$$

Since the flows of eqs. (5) and (6) meet at D, we try to match the flow above D with the flow below D. There are some regions where the flow of equations (5) is directed from $\dot{x} > 0$ toward D, and simultaneously the flow of eqs. (6) is directed from $\dot{x} < 0$ toward D, producing a conflict in the flow directions. (Similarly, there are regions where the flows from both $\dot{x} > 0$ and $\dot{x} < 0$ are directed away from D.) These regions of conflict are the sticking regions R, shown in fig. 2. Regions where the flow directions agree and are upward (toward $\dot{x} > 0$) are labeled $\oplus$, and regions where the flow directions are downward (toward $\dot{x} < 0$) are labeled $\ominus$. In the region "above" (defined as $\{(x, t) : x > x^{(1)}$ and $x > x^{(2)}$, where $(x^{(1)}, t) \in C_1$, and $(x^{(2)}, t) \in C_2\}$) both C_1 and C_2, the flows of *both* equations (5) and (6) are directed *toward* the (x, t)-plane. Hence, orbits hitting this region are trapped in D, x remaining fixed, as time evolves, until the moment that the direction of flows is in agreement. The multivaluedness of $f(x_2)$ provides a bridge between the flow of eqs. (5) and the flow of eqs. (6), and allows this trapping to take place. In the region "below" both C_1 and C_2, the flow of *both* systems is directed *away* from the (x, t)-plane. Unless we are on the region, orbits will be taken away from it. Thus, it is an unstable sticking region and is not expected to be observed.

Conveniently, $x = -1/k$ at the intersection of C_1 and C_2, so the unstable sticking region is not

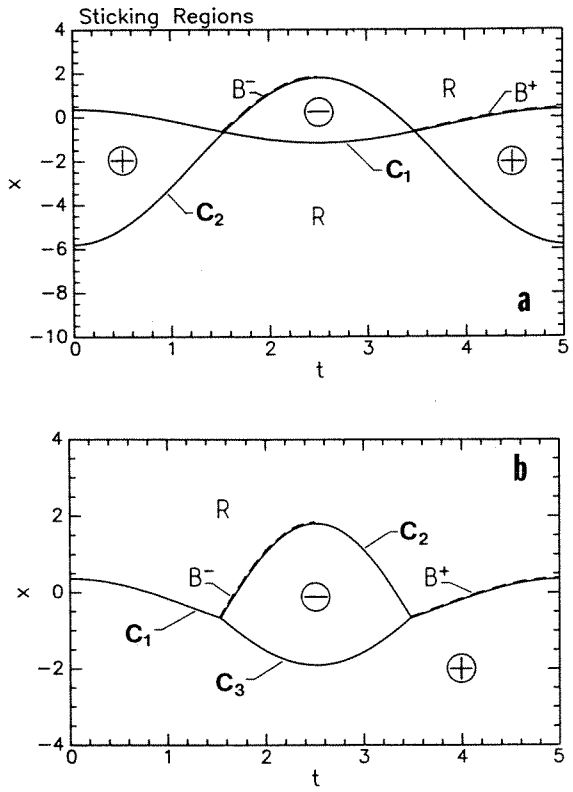

Fig. 2. The sticking regions when (a) the physical-realizability condition is ignored, (b) The physical realizability condition is observed. $\oplus$ and $\ominus$ indicate upward and downward flows. R indicates the sticking region. B^+ labels the portion of C_1 with positive slope, bordering $\oplus$, and B^- labels the portion of C_2 with positive slope, bordering $\ominus$. In (b), C_1 and C_2 are pieces of the same curves in (a), drawn at a different scale. $\Omega = 1.25$, $a = 1.9$, and $k = 1.5$.

physically realizable, and so it will never be observed. (It is, in fact, irrelevant.) If we apply the realizability condition, the direction of flow for $x < -1/k$ is then governed by eq. (3c). The curve C_3 separating regions of upward and downward flow in eq. (3c) is given by

$$C_3 : x_1 = a\cos(\Omega t) , \quad x < 1/k .$$

The curves C_1, C_2, and C_3 all intersect at the same points:

$$x = -1/k , \quad t = \frac{1}{\Omega}\arccos(-1/ak) .$$

The multivaluedness of $f(x_2)$ at $x_2 = 0$ is necessary for the existence of solutions in the sticking region. If f were a single-valued function, and, say, $f(0) = 0$, what would happen? A trajectory in the sticking region in D would be governed by the equation (2) with $f(0) = 0$. This reduces to a case where there is no static friction, and has the form of eq. (3c). An orbit in D would be, with probability one, directed out of D. Such a trajectory, if it were to exist, would have to leave D immediately. But it cannot leave D since the flow outside D is directed toward D. So the trajectory does not exist. The multivaluedness of f provides a bridge for the flow across D through the sticking region.

2.2. *Graphical display of solutions*

The flow of eq. (3a) consists of a homogeneous part and a particular part. The solution is of the form

$$x(t) = A\cos\omega(t - t_0) + B\sin\omega(t - t_0) - \frac{1}{1+k} + \frac{a}{1+k-\Omega^2}\cos(\Omega t),$$

where $\omega = \sqrt{1+k}$, $\Omega^2 \neq 1 + k$, and A and B are determined by the initial conditions. In $(x, \dot{x}, t)$-space, the family of solutions is a set of nested tubes. Exploiting the periodicity of the excitation, the family of solutions is a set of nested tori. Given some initial conditions, the flow of eq. (3a) is confined to some tube (fig. 3). This flow is good for $\dot{x} > 0$. A solution tube interacting with D looks like a worm or eel coming out of the water, and reentering the water. Some worms are so fat that they never get their bellies above the water. The family of worms intersects D, in the set (*not* a Poincaré map) shown in fig. 4. A solution based at a point such as A will flow along a tubular solution surface until it next returns to D at A'. Thus, $P^+(A) = A'$. Similarly, in fig. 4, $P^+(B) = B'$. Fig. 4 also shows that, depending on the values of Ω and k, the orientation of the solution tubes can be flipped, due to the phase angle in the forced response of harmonic oscillator.

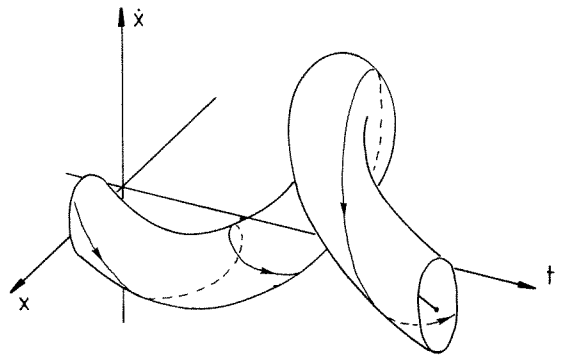

Fig. 3. A solution tube with $\Omega^2 < 1 + k$ for the equation valid for positive x_2. The axis of the tube is the particular solution, about which the homogeneous solution oscillates.

Likewise, the flow of eq. (3b) consists of a "saddle" driven along the particular solution.

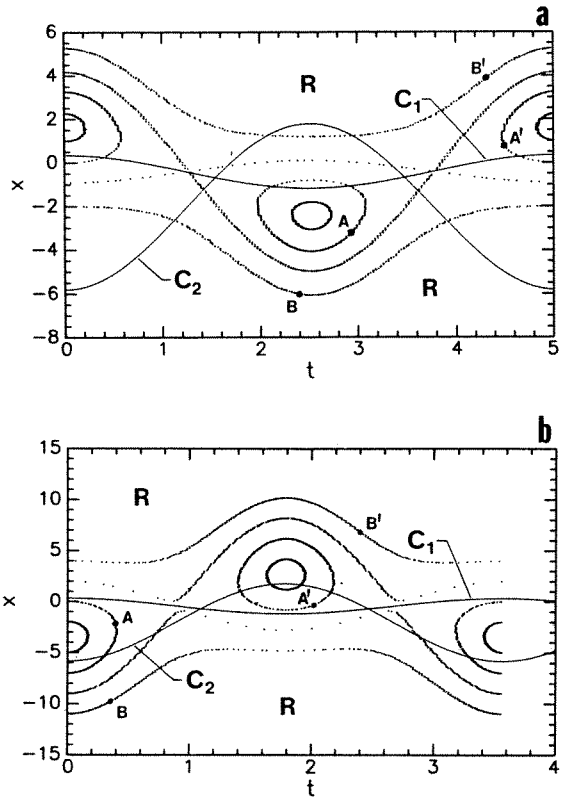

Fig. 4. The pattern of the solution tubes intersecting D with (a) $\Omega^2 < 1 + k$, and (b) $\Omega^2 > 1 + k$.

The solution for eq. (3b) when $k>1$ is of the form

$$x(t) = E\,e^{-r(t-t_0)} + F\,e^{r(t-t_0)} + \frac{1}{1-k} + \frac{a}{1-k-\Omega^2}\cos(\Omega t)\,,$$

where $r=\sqrt{k-1}$, $\Omega^2 \neq 1-k$, and E and F are determined by the initial conditions. Given some initial conditions, the flow is confined to a family of saddle sheets. The intersection of these saddle sheets with D is shown in fig. 5. In the figure, $P^-(A)=A'$ demonstrates how a point may map according to the flow of eq. (3b).

Any orbit lies on some solution tube for $\dot{x}>0$. When the trajectory passes through D, it may interact with the sticking region. If the flow goes below D, where $\dot{x}<0$, it will be confined to a saddle sheet until it passes through D again. (It is possible that a trajectory on a saddle sheet will approach ∞, never returning to D. By enforcing the realizability condition, such an orbit will eventually enter a region of frictionless oscillation, and return to the region of friction.) For these flows, the mappings $P^+: \oplus \rightarrow D$ and $P^-: \ominus \rightarrow D$ can, in principle, be constructed. We would like to see how the active regions (non-sticking regions) map under P^+ and P^-, and how they interact with the sticking regions.

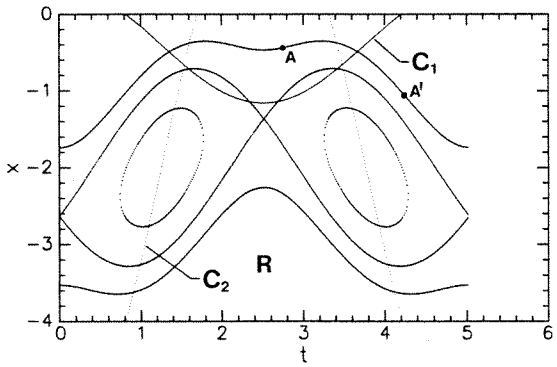

Fig. 5. The pattern (*not* a Poincaré map) created by the intersection of saddle sheets with D for the flow of the equation valid for negative velocity.

3. Qualitative mapping of regions

In this section, results are presented for parameter values of $a=1.9$, $\Omega=1.25$, and $k=1.5$, unless otherwise stated. They were chosen as parameter values which qualitatively modeled the experiment, although they do not correspond to measured parameter values.

Referring to fig. 2a and neglecting the realizability condition, we consider the mapping of the region $\oplus$ via P^+, and the mapping of the region $\ominus$ via P^-. If R is the sticking region, then $\oplus \cup \ominus \cup R = D$. Motions in R either stay in R forever, or, through the evolution of time, exit R into $\oplus$ or $\ominus$ via the map S: $R \rightarrow B^+ \cup B^-$, where B^+ and B^- are part of the boundaries of $\oplus$ and $\ominus$ as shown in fig. 2. (The sets B^+ and B^- are the components of C_1 and C_2, respectively, through which sticking orbits originating in R must pass as they exit R.) Therefore, the system can be understood through the mappings of $\oplus$ and $\ominus$. Certainly $P^+(\oplus)\cap\oplus=\emptyset$, and $P^-(\ominus)\cap\ominus=\emptyset$. It is also likely that $P^+(\oplus)\cap R\neq\emptyset$, and $P^-(\ominus)\cap R\neq\emptyset$.

We want to qualitatively describe the images $P^+(\oplus)$ and $P^-(\ominus)$. Some information lies in the flow on a solution surface, which is locally equivalent to flow on a plane. For example, consider the solution surface M shown in fig. 6a, and its intersection with the curve C_1 at point C. Flow in the neighbourhood of C on M, as seen from outside of the worm, is locally equivalent to flow on the plane, either in fig. 6b or 6c, depending on the direction of flow. We can determine the direction of flow on M near C by its geometry and the forward nature of evolving time. In the case pictured in fig. 4a, the forward nature of time near C indicates that the flow on M (as seen from outside the tube) is like that of fig. 6b. The flow folds the (x, t)-plane about C_1 like a hinge, in that points in the neighborhood of C get mapped into a neighborhood of C.

If the mapping of $\oplus$ has a hinge on part of C_1, consider the mapping P^+ of a neighborhood $U \subset \ominus$ of the hinge. $P^+(U)$ is also a neighbor-

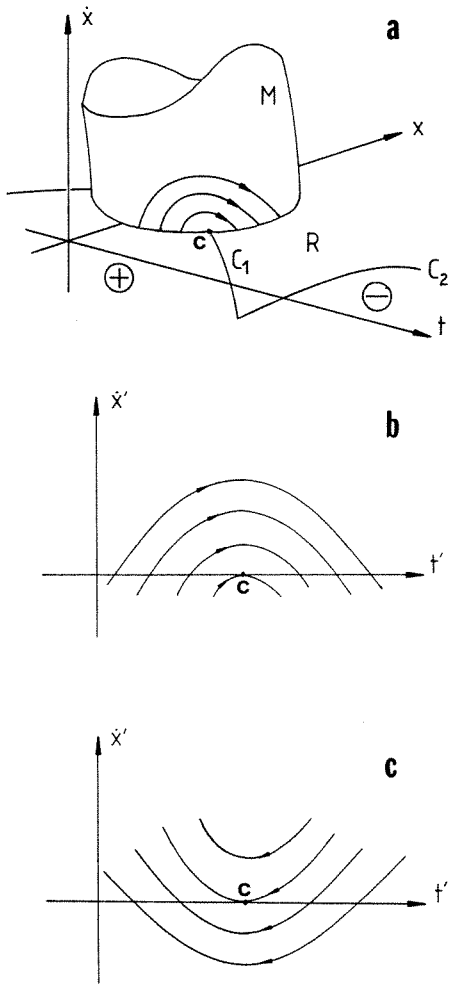

Fig. 6. (a) The intersection of a solution tube M with curve C_1 at point C is shown in a left-handed coordinate system. (b), (c) The possible topologically equivalent flows in $\mathbb{R}^2$. The axes $\dot{x}'$ and t' correspond locally to $\dot{x}$ and t.

hood of the hinge in $R \cup \oplus$. Therefore, since the hinge is adjacent to R, $P^+(U) \cap R = W \neq \emptyset$. Points in the set W remain in R until time evolves such that they are in B^-. The map S: $R \rightarrow B^- \cup B^+$ is singular – regions in R are condensed to B^+ and B^-. Hence, the map is not invertible, and the flow is not invertible. This was also discovered by Shaw [2]. This noninvertibility is the key to the other properties we will observe.

Similarly, flow on a solution surface such as

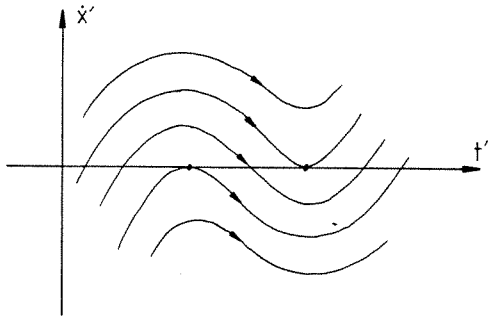

Fig. 7. The trajectories on a saddle sheet near its intersection with curve C_2 are represented by this equivalent flow in a plane. The axes $\dot{x}'$ and t' correspond locally to $\dot{x}$ and t.

that which intersects D as in fig. 5 is locally similar to the two dimensional flow in fig. 7. This picture varies depending upon the distance of the saddle surface to the saddle's stable and unstable manifolds.

Computer-generated mappings of $\oplus$ for parameter values which produce a hinge in the flow about the curve C_1 discussed above are shown in fig. 8.

The dynamics of the Coulomb oscillator can be described by successive mappings of $\oplus$ or $\ominus$, under P^+, P^-, and S, when appropriate. Fig. 9 shows the computer-generated sequence of mappings of $\ominus$ for a particular set of parameter values. The process accounts for the realizability condition. Within one period of time, the entire region of initial conditions has been crushed into a set of curves. This smashing is *not asymptotic*! It occurs suddenly in the sticking regions. The attractor lies in the images of B^+ and B^- (shown in fig. 2). Many initial points will be condensed onto the attractor in finite time.

The result in fig. 9 suggests that we only need to understand mappings of B^+ and B^- to understand the long-term behavior of the entire system. When we study the system in this way, we are implicitly assuming that all orbits eventually pass through the sticking region, exiting onto B^+ and B^-. We must therefore ask: under what conditions do all orbits pass through the sticking region?

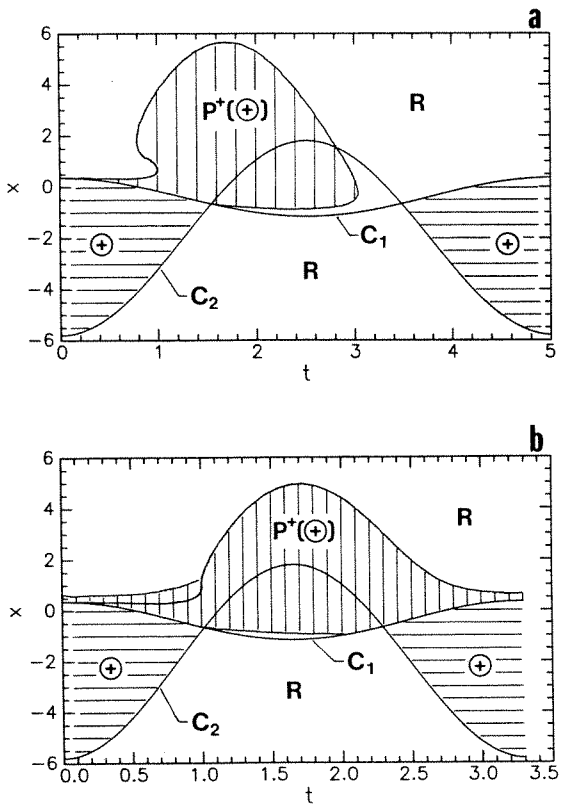

Fig. 8. Computer-generated mappings of $\oplus$ for (a) $\Omega^2 < 1 + k$, and (b) $\Omega^2 > 1 + k$. The physical-realizability condition is not used.

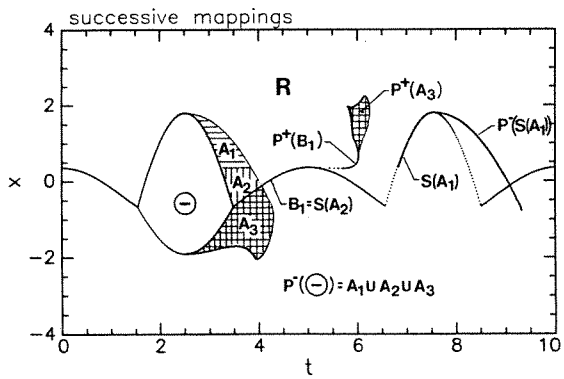

Fig. 9. Successive mappings of $\ominus$. Within one period of excitation, the entire set of points has condensed to a line. The physical-realizability condition is enforced. The parameter values are $\Omega = 1.25$, $a = 1.9$, and $k = 1.5$.

Proposition. If the inverse image and the forward image of $\ominus$ (or $\oplus$) do not intersect in D, then all orbits will pass through the sticking region.

Proof. Suppose $P^-(\ominus) \cap P^{+-1}(\oplus) = \emptyset$. If a point $r \notin P^{+-1}(\ominus)$, and $r \in \oplus$, then $r_1 = P^+(r) \notin \ominus$. Since $P^-(\oplus) \cap \oplus = \emptyset$, $r_1 \in R$. On the other hand, if $r \in P^{+-1}(\ominus)$, and $r \in \oplus$, then $r_1 = P^+(r) \in \ominus$, and $r_2 = P^-(r_1)$. By the hypotheses that $P^-(\ominus) \cap P^{+-1}(\oplus) = \emptyset$, $r_2 \notin P^{+-1}(\ominus)$. Since $r_2 \notin P^{+-1}(\ominus)$, we have seen in the above argument that $P^+(r_2) \in R$.

In other words, if we follow any point s starting in $\ominus$, its mapping $r = P^-(s)$ will be in the image of $\ominus$, $P^-(\ominus)$. If $P^-(\ominus)$ does not intersect the preimage of $\ominus$, then r is not in the preimage of $\ominus$. If r is not already in R, then its mapping $P^+(r)$ will certainly be in R, and the motion will stick. In the case when the initial $r \in R$, the motion begins trivially in the sticking region.

The proposition provides a sufficient condition for all trajectories to eventually pass through the sticking regions. We need not prove the proposition of the images and preimages of $\oplus$ because all orbits in $\oplus$ either map to R or to $\ominus$.

Fig. 10 numerically shows that $P^-(\ominus)$ does not

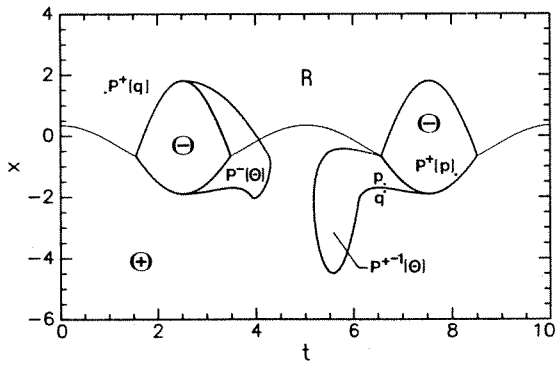

Fig. 10. Due to the periodicity of time, the two regions $\ominus$ are identified. The preimage of $\ominus$ and the forward image of $\ominus$ do not intersect, therefore all orbits pass through the sticking region. In reference to section 5, a discontinuity in the mapping on D is illustrated by the mappings $P^+(p)$ and $P^+(q)$, which start arbitrarily close, but get mapped far apart.

intersect $P^{+-1}(\ominus)$ for the chosen set of parameters, thus all orbits pass through the sticking region.

The sequence of mappings of the boundary of $\ominus$ can be viewed as a one-dimensional tent-like map. Depending upon the degree of stretching and folding, we may have periodic or chaotic dynamics. The dynamics of this oscillator has been shown to be associated with a one-dimensional single-humped map [8, 9]. Unlike the usual case where the one-dimensional map is an approximation which exploits a strong stable foliation, this one-dimensional map arises exactly.

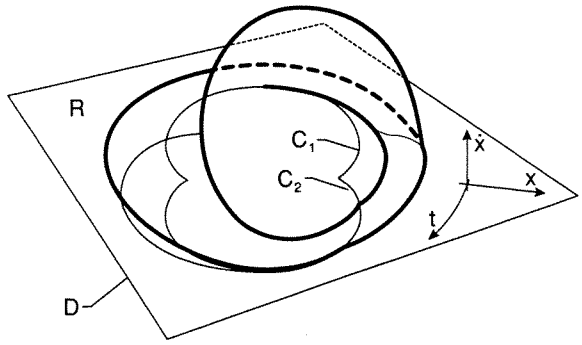

Fig. 11. The attractor is constructed by identifying periodicity in the time variable, and extrapolating the previous map sequence into a flow.

4. Construction of the attractor

The sequence of mappings of B^- can be viewed by wrapping $t \bmod(2\pi/\Omega)$ around and back to itself (in S^1). These mappings can then be extrapolated into a flow in $(x, \dot{x}, t \bmod(2\pi/\Omega))$-space. The resulting template (fig. 11) resembles the numerical solution displayed in the same space (fig. 12). The motion is in a branched manifold. Whether the branched manifold produces a strange attractor or a periodic attractor depends upon the degree of stretching, which depends on parameter values. Whether all the motions go through the sticking region also depends upon parameter values.

The Lorenz attractor [10] and the Rössler attractor [11] lead to prime examples of branched manifolds. (Guckenheimer and Holmes [10]

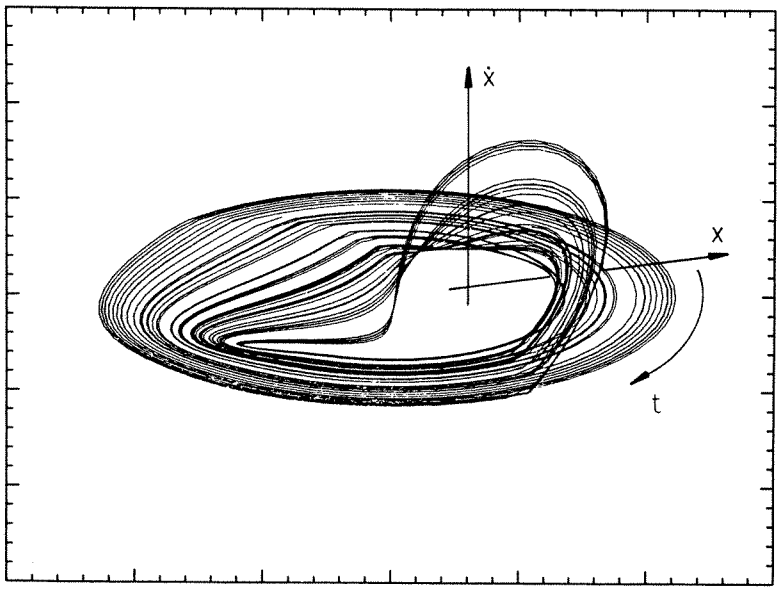

Fig. 12. A numerical solution of eq. (2) with $\Omega = 1.25$, $a = 1.9$, and $k = 1.5$, displayed in three dimensions. The radial coordinate axis is x, the circumferential coordinate axis is $t \bmod(2\pi/\Omega)$, and the cylindrical axis is $\dot{x}$.

demonstrate the construction of the Lorenz branched manifold, and further geometric analysis.) Their motions are strongly asymptotically attracted to the branched manifolds. The semi-flow on the branched manifold is an approximation, and does not fully represent the detailed behavior of the system.

In the Coulomb oscillator, however, the motion is condensed suddenly (not asymptotically) onto the branched manifold. This noninvertible condensation, caused by the discontinuity and multivaluedness of the Coulomb friction law, gives the Coulomb oscillator the opportunity to have a strange attractor in a two-dimensional manifold. The dynamics in the branched manifold fully represents the long-term behavior of the system.

When analyzed using constrained equations, chaos in Lorenz- or Rössler-like attractors is also confined to a two-dimensional manifold [12]. Another oscillator which has some similarities in its structure is the self-excited friction oscillator which can move through a one-way trap door, modeled by Troger [13].

The entire attractor for this system consists of the attracting branched-manifold, and of the permanent-sticking region which surrounds this toroidal structure. the permanent-sticking region exists on D for all x_1 greater than the maximum of $(-1+a)/(1+k)$ and $(1-a)/(1-k)$. These values represent the maximum values of curves C_1 and C_2. All points landing in this part of the sticking region are stuck forever.

The dynamics on branched manifolds may be viewed via one-dimensional maps, i.e. maps of the form $s_{n+1} = g(s_n)$. A map that would account for both the dynamics on the branched manifold, and the permanently sticking motions, would consist of a component ($s_n \geq 0$) resembling a single-humped map, and a component ($s_n < 0$) coinciding with the identity line, respectively (fig. 13). The component coinciding with the identity line produces an infinite locus of fixed points for all $s_n < 0$. Motions on $s_n > 0$ are dynamic, and may have periodic or chaotic attractors. Motions that get mapped to $s_n < 0$ are trapped there forever. The map defined for $s_n > 0$ was obtained from a return map on a coordinate defined in the Poincaré section of the numerical solution (fig. 12).

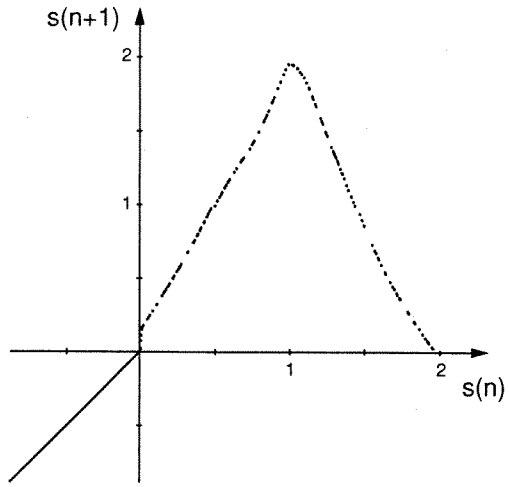

Fig. 13. A one-dimensional map schematically represents motion on the branched manifold, and permanently sticking motion. The logistic-like component, defined for $s_n > 0$, produces dynamics like that on the branched manifold. The component confined to the identity produces a locus of fixed points for all $s_n < 0$.

5. Basins of attraction

The flow through some initial regions of phase space will eventually be attracted to the branched manifold, while flow through other regions will be imprisoned in the permanent-sticking region. Which points will go where? What is the nature of the basin of attraction of the branched manifold?

In fig. 14a we have plotted points in the $(x, \dot{x})$-plane, at initial time $t = 0$, which flow onto the branched manifold. Trajectories in the grey area flow to the branched manifold. Therefore, initial conditions in the grey area lead to solutions which are controlled by the dynamics of the branched manifold. Orbits in the white area flow

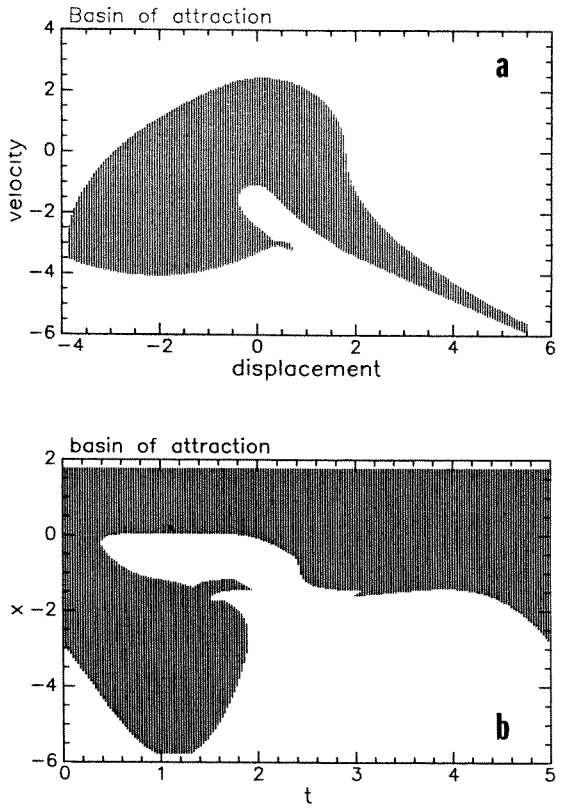

Fig. 14. The basin of attraction of the branched manifold, for $\Omega = 1.25$, $a = 1.9$, and $k = 1.5$, is shown in grey. (a) The initial conditions in x and $\dot{x}$ are chosen at time $t = 0$. (b) The initial conditions in x and t chosen on D.

to the permanent-sticking region, where they will stay forever. Fig. 14b shows the points in D which map to each part of the attractor. This system is an example of a chaotic system which has smooth basin boundaries. Another example of such a system is the Hénon attractor [14].

In dynamical systems, basin boundaries [14, 15] are defined by stable manifolds (of unstable objects) which separate flows to one basin from another. However, in this case, it is helpful to visualize the basin boundaries through mappings in D. A map of a region in D arises from a linear flow which takes the region from D and returns it to D (or enters another subzone, as in the enforcement of physical realizability). Some chunks of this original region in D simply return to D in the permanent-sticking region, and some do not. Those that do not land in the permanent-sticking region get mapped again, according to a linear flow. Again, of these continuing images, some land in the permanent-sticking region, and some do not. There is a simple boundary separating points which land in different regions. There is no asymptotic repulsion from this simple boundary. Trajectories on either side of this boundary merely glide along next to each other, unaware of any cultural differences, until they land on opposite sides of a fence, one being static, the other dynamic.

What happens to orbits which land directly on the fence? The permanent-sticking region is defined by maximum x_m of both curves C_1 and C_2, for all t. Whether x_m is defined to be in the permanent-sticking region, or in the dymanic region, is arbitrary. If this maximum is defined as in the sticking region, then points in the basin boundary map to the permanent-sticking region.

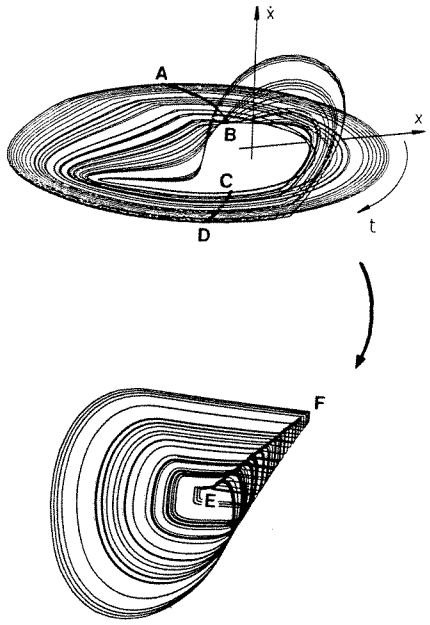

Fig. 15. Embeddings of a sticking observable may not be homeomorphic to the flow in phase space. During the embedding, points in the sticking region between AB and CD in phase space get crushed noninvertibly to EF, which lies on the identity line in a rotated embedding space.

If this maximum is defined to be free to map through P^+ or P^-, then it is considered to be a point in the basin of the branched manifold. If, for example, the permanent-sticking region is defined by the maximum of curve C_2, denoted as r, then $P^-(r) = r$, making r a fixed point of P^- and thus a periodic point in the flow. Points on the basin boundary would collapse onto this periodic orbit. Regardless of the definition of the status of x_m, its trajectories will coincide with the boundary between the permanent-sticking region and the active region.

Is it possible to have basins of attraction that are not simply connected in D? (If one basin of attraction were not simply connected, then the other would be disconnected.) Basically, the basin boundaries are preimages of the boundary of the permanent-sticking region. This boundary is connected. Thus, if the combination of mappings P^+, P^-, and S, of the plane D, which may occur during a motion, compose a continuous mapping, then the basin boundary would be connected, and each basin of attraction would be simply connected. However, these mappings may indeed be discontinuous. This is illustrated in fig. 10. The points p and q may be chosen arbitrarily close to each other, yet on either side of the boundary of $P^{+-1}(\ominus)$. An image $P^+(p)$ lands near the border of $\ominus$. But the mapping $P^+(q)$ lands in the sticking region, far from the region $\ominus$. Thus, it would not be startling if we observed disconnected basin boundaries in D. However, the discontinuity in the map $P^+: \oplus \to D$ does not necessarily imply that there be disconnected basin boundaries in D.

Furthermore, having a discontinuous mapping on D does not mean that we have a discontinuous *flow*. The mapping P^+ is, in fact, defined by a linear flow, which is continuous. Thus, a linear, continuous flow can produce a discontinuous map on D. What about piecewise continuous flows involving sticking regions? As part of the flow strikes the sticking region, and the neighboring flow passes through D, could we see voids, or separation, in the flow? For the oscillator studied here, this will not happen. Since the flows about the borders of the sticking regions are hinge-like, suspected voids are continually being filled.

6. A cautionary note on embeddings

Takens' embedding theorem [16, 17] states that, if basic hypotheses are satisfied, the method of delays of an observable produces an embedding. However, our problem does not satisfy the hypotheses for Takens' embedding theorem. One of the basic hypotheses is that the observable $y \in C^r$, where $r \geq 2$. In our case, we choose $y = x$. The vector field has a component which corresponds to acceleration. Hence the discontinuous vector field yields discontinuous accelerations, hence $y \notin C^2$. If the observable is $y = \dot{x}$, then $y \notin C^1$.

Problems occur when orbits go through the sticking region. The observable x is constant while in the sticking region. An arbitrary sampling rate will produce samples of x_n, many of which are captured during a time interval in which x is momentarily constant, giving rise to many consecutive samples of the same value. To perform a three-dimensional reconstruction, we construct pseudovectors $\boldsymbol{x}_n = (x_n, x_{n+k}, x_{n+2k})$. For arbitrarily many values of k, there will be many vectors coming from the sticking region $\boldsymbol{x}^s_m$, $m = i, \ldots, j$, such that $x^s_{m+2k} = x^s_{m+k} = x^s_m$. The result is that the reconstructed sticking motions will occasionally pile up on the identity line in psuedo phase space. This is shown in fig. 15. In such case, the map which takes the real manifold M_r into the "embedded" manifold M_e is not invertible, and thus not a diffeomorphism. If this happens, calculations for the characterization of chaos, such as fractal dimension and Lyapunov exponents, which typically rely on embeddings, will be meaningless. (In fact, for this system, both numerically and experimentally, correlation-dimension calculations from embeddings do not converge!)

As a remedy for the dimension calculation in this situation, we can limit the calculation to points which are some distance r from the identity line (in pseudo phase space). For the Coulomb oscillator with $\Omega = 1.25$, $a = 1.9$, and $k = 1.5$, the correlation dimension was computed using the Grassberger–Procaccia method [18, 19] for the attractor described by its three-dimensional phase variables, and for the reconstructed attractor. The attractor described by its phase variables had a correlation dimension of 1.93 ± 0.04. The reconstructed attractor, with $r = 0.25$ in the box norm (the observable was in the range of -1.0 to 1.5), had a correlation dimension of 1.87 ± 0.02.

Such a calculation assumes that the topology of the attractor is uniform, and that no critical information about dimension is lost by omitting the portion of the attractor that got crushed to a line. However, if the attractor were to have nonuniform dimension, dimensional information may be lost when portions of the attractor are omitted.

The potential failure of embeddings may also expose a need for caution in control theory. For a control theorist, the analogy of an embedding is an *observer construction*, in which the states of a dynamical system are estimated from information contained in a small number of observables. Loss of information could conceivably introduce uncertainty in applying control forces.

8. Conclusions

A qualitative technique has been used to describe the dynamics of a discontinuous, multi-valued Coulomb oscillator. The technique is a three-dimensional extension of that previously employed in classical texts on two-dimensional automonous systems.

The analysis reconstructs an attractor similar to that seen in numerical integrations. During the reconstruction, it was shown that infinitely strong contraction takes place during sticking motion. Because of this condensation, a one-dimensional map can describe the long-term dynamics exactly. This is unlike the usual case in which a one-dimensional map is an approximation which exploits a strong stable foliation.

Some other properties of the behavior which result from the discontinuous and multi-valued nature the vector field (the same mechanism which gives rise to stick–slip motion) are:

- the flow may not be invertible,
- the flow may reach the attractor in finite time,
- strange attractors may have dimension less than or equal to two,
- "embeddings" of an observable may not be topologically equivalent to the phase flow.

The embedding property is practical since embedding techniques are often used on experimental data. A good model of an experiment is sometimes unknown, thus the presence of discontinuities may be unknown. If an experimentalist were to apply embedding techniques when hypotheses of the embedding theorem are unknowningly violated, the results may be meaningless.

While there may be some debate over whether discontinuities really exist in physical systems, certainly near discontinuities exist, and to the resolution of measurements, they may be indistinguishable from actual discontinuities.

Acknowledgements

I am grateful to Professor Francis Moon for support, and inspiring ideas through related works; Professor Philip Holmes for discussions and a meticulous critique; Professor Joe Cusumano for his dimension-calculation program, discussions and interest; Professor John Guckenheimer for discussions. Tim Kiemel offered stimulating discussions. Bill Feeny refined an illustration.

Grants from IBM, the ARO, and the AFOSR, provided financial support.

References

[1] J.P. Den Hartog, Trans. ASME 53 (1931) 107.
[2] S.W. Shaw, J. Sound Vib. 108 (1986) 305.
[3] A.A. Andronov, A.A. Vitt and S.E. Khaikin, Theory of Oscillators (Dover, New York, 1966).
[4] J.J. Stoker, Nonlinear Vibration (Interscience, New York, 1950).
[5] F. Takens, Lecture Notes in Mathematics, 535 (Springer-Verlag, Berlin, 1976) p. 237.
[6] H. Oka, Jpn. J. Appl. Math. 4 (1987) 393.
[7] S.S. Antman, Quart. App. Math. XLVI(3) (1988) 569.
[8] Brian Feeny, Chaos and Friction, Ph.D. Thesis, Cornell University, August 1990.
[9] B. Feeny and F.C. Moon, Chaos in a dry-friction oscillator: experiment and numerical modeling, J. Sound Vib., submitted.
[10] J. Guckenheimer and P. Holmes, Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields (Springer, New York, 1983).
[11] A.V. Holden and M.A. Muhamad, in: Chaos, ed. A.V. Holden (Princeton Univ. Press, Princeton, 1986) p. 15.
[12] H. Oka and H. Kokubu, Japan J. Appl. Math. 2 (1985) 495.
[13] H. Troger, Z. Angew. Math. Mech. 62 (1982) T18.
[14] S.W. McDonald, C. Grebogi, E. Ott and J.A. Yorke, Physica D 17 (1985) 125.
[15] C. Grebogi, E. Ott and J.A. Yorke, Science 238 (October 30, 1987) 632.
[16] F. Takens, Lecture Notes in Mathematics, 898 (Springer, Berlin, 1981) p. 366.
[17] N. Gershenfeld, in: Directions in Chaos, II, ed. B.-L. Hao (World Scientific, Singapore, 1988) p. 310.
[18] P. Grassberger and I. Procaccia, Physica D 9 (1984) 189.
[19] J.P. Cusumano, Low-Dimensional, Chaotic, Nonplanar Motions of the Elastica: Experiment and Theory, Ph.D. thesis, Cornell University, January 1990.