\[ \rho' = \frac{\arg(P^q x) - \arg(P^{q-1} x) + \arg(P^{q-1} x) - \arg(P^{q-2} x) + \cdots + \arg(P x) - \arg(x)}{2\pi q} \]

By the assumption, each parenthesized pair in the numerator is just the angle subtended by the projection of the trajectory during some period of the input. Therefore, the numerator is equal to \(2\pi q\) and our claim is proved.

ACKNOWLEDGMENT

We would like to thank Dr. M. Odyniec for his insightful comments and kind assistance in the preparation of this paper.

REFERENCES


Thomas S. Parker was born in Pasadena, CA, on November 21, 1957. He received the B.S. (with highest honors) and the M.S. degrees in electrical engineering from the University of California, Berkeley, in 1980 and 1982, respectively. He is currently a Hertz Fellow and is pursuing the Ph.D. degree at Berkeley.

His interests include nonlinear dynamical systems, numerical methods, and strange behavior.

Leon O. Chua, for a photograph and biography please see page 510 of this issue.

An Introduction to the Lorenz Equations

COLIN SPARROW

Abstract — Theoretical and numerical results for the three-parameter three-dimensional family of ordinary differential equations known as the Lorenz equations are surveyed. Particular emphasis is laid on the importance of homoclinic bifurcations and their relationship to the appearance of chaotic behaviors.

INTRODUCTION

The Lorenz equations, first presented in 1963 by Lorenz [18] are one of the best-known systems of chaotic ordinary differential equations. In this paper, we begin by examining some basic properties of the equations and discuss the bifurcations which lead to the appearance of a strange attractor in the flow; most of these results have been known for several years. Then we will look at less well-understood parameter ranges and present some more recent results. It should be said, however, that so much is now known about the Lorenz equations that it is not possible to survey all the available results; for a fuller discussion the reader is referred to [33] which also contains a more extensive reference list.

Basic Properties and Derivation

The Lorenz equations

\[ \begin{align*}
    \dot{x} &= \sigma (y - x) \\
    \dot{y} &= rx - y - xz \\
    \dot{z} &= xy - bz 
\end{align*} \]

where \(\sigma, r, \) and \(b\) are three real positive parameters, can easily be shown to be bounded and dissipative (e.g., [18], [33]). In other words, all trajectories eventually enter and thereafter remain within some bounded set in \(\mathbb{R}^3\), and the flow shrinks all volumes. These two remarks, taken together, imply that all trajectories tend towards some bounded set of zero volume. To be precise, the bounded set of zero volume which interest us most is the nonwandering set, \(\Lambda\); this set contains all the recurrent behavior of the flow.\(^1\)

\(\Lambda\) will contain stationary points (one if \(r \leq 1\), three if \(r > 1\)) and may contain periodic orbits (limit cycles) and/or more complicated sets, depending on the parameter values

\(^1\) is the set of all points \(x \in \mathbb{R}^3\) with the following property: for every neighborhood \(U\) of \(x\) and every finite time \(T > 0\) there is a \(y \in U\) with \(\phi(t, y) \in U\) for some \(t > T\). \(\phi\) is the flow determined by the equations.

0998-4094/83/0800-0533$01.00 ©1983 IEEE
chosen for study; all of these objects may be either stable ("sinks") or nonstable ("saddles") but they cannot be unstable ("sources"). This last remark follows from the dissipative (volume contracting) nature of the flow which, it should be noted, also implies there are no invariant tori in Λ. Our major interest in the equations is to discover the topology of Λ, the way this changes as the parameters change (bifurcations), and the behavior of the flow on and near Λ, since these factors determine the long-term properties of all trajectories.

Integration of the Lorenz equations does not present any unusual numerical problems and when approximate solutions to the equations are calculated on a computer we expect these solutions to move rapidly towards some stable component of Λ; if the component in question is either a stationary point or a periodic orbit (of reasonably low period) we expect it to be readily identifiable as such. For some parameter values numerically calculated trajectories do not appear to settle down to either of these possibilities; we can continue the integration for as long as we please and trajectories continue to wander back and forth, never quite closing up, in an apparently "random" way. Fig. 1 shows an example of such a trajectory, and it is such trajectories which lead us to describe the equations as "chaotic". (In this situation we may not be convinced that the numerically calculated trajectory is providing an accurate long-term approximation to a true trajectory of the system; nonetheless, we can be sure that the numerical trajectory is moving close to some component of Λ and, therefore, that Λ has a complicated structure which warrants further investigation.)

It is this chaotic behavior which has generated so much interest in the Lorenz equations, the more so because the equations can be derived, in several different ways and with more or less plausibility, from various real world problems [2], [8], [9], [12], [18], [20], [23], [24]. In some cases (e.g., Lorenz's original two-dimensional fluid cell convection problem [18]) it is accepted that the interesting parameter ranges—those in which we have chaotic solutions—do not overlap with parameter ranges in which the equations are a reliable model of the original problem. In other cases (e.g., irregular spiking in lasers [9], [12]) it is likely that the equations do successfully model some of the complicated behavior observed in the physical problem. Here we will concentrate on the mathematical properties of the equations, leaving to others (e.g., [27]) discussion of their possible relevance to the physical world.

LOW r BIFURCATIONS AND BEHAVIOR

The Lorenz equations have been studied for many values of the parameters σ, β, and r, but most commonly for values σ = 10, β = 8/3, and r varying between zero and infinity. The choice of r as the varying parameter is conventional—r is related to the Rayleigh number in fluid dynamical derivations—but also makes good mathematical sense; information about the relatively simple behavior of the equations at small and large r-values can be obtained using classical mathematical techniques and we can work towards the more complicated intermediate behavior from both ends. The choice of the two constant parameters, σ and β, is also conventional [18] and turns out to be wise; it is argued in [33] that σ = 10, β = 8/3 lies in the range of a and b values for which the simplest interesting behavior is observed in the Lorenz equations as r varies. Below we summarize the behavior of the equations for a whole range of parameter values, though we concentrate hardest on the types of behavior seen for the parameter values discussed above. In what follows notice that the equations have a natural symmetry which takes points with coordinates (x, y, z) into points with coordinates (−x, −y, z). Though
most of the interesting properties of the equations persist if
this symmetry is slightly perturbed, we will not consider
the more general nonsymmetric case here.

(i) \( r < 1; \) (any \( \sigma, b > 0 \)).

The only stationary point, the origin, is globally stable.
(Consider the Liapounov function \( V = x^2 + \sigma y^2 + az^2 \).)

(ii) \( r = 1; \) (any \( \sigma, b > 0 \)).

There is a simple bifurcation in which the origin be-
comes nonstable by casting off a pair of stable stationary
points, \( C_1 \) and \( C_2 \), as \( r \) increases. The \( C_i \) are given by

\[
C_{1,2} = \left\{ \pm \sqrt{b(r-1)}, \pm \sqrt{b(r-1)}, r-1 \right\}.
\]

(iii) \( r > 1 + \epsilon(\sigma, b); \) (any \( \sigma, b > 0 \)).

There is a small and calculable number, \( \epsilon(\sigma, b) \), such
that for \( r > 1 + \epsilon(\sigma, b) \) the eigenvalues of the linearized
flow near the origin (which are real) satisfy the inequality
\(-\lambda_2 > \lambda_1 > -\lambda_3 > 0 \) where \( \lambda_3 = -b \) is the eigenvalue
associated with the eigenvector along the \( z \)-axis) and there
are a complex pair of eigenvalues of the linearized flow
near the other two stationary points, \( C_1 \) and \( C_2 \). It appears
that there are no interesting behavior changes in the
parameter range \( 1 < r < 1 + \epsilon(\sigma, b) \), and for \( r > 1 + \epsilon(\sigma, b) \)
the basic properties of the Lorenz flow are illustrated in
Fig. 2.

In Fig. 2 we have the origin, nonstable, with a one-
dimensional unstable manifold and a two-dimensional sta-
ble manifold. The relative sizes of the three eigenvalues
at the origin will be important later. Each of the stationary
points \( C_i \) has associated with it a real negative eigenvalue
in addition to the complex-conjugate pair of eigenvalues.
There is a section of a surface \( S, WXYZ \), such that
trajectories started on this surface move downwards (\( z \) decreasing)
and either to the left or right, depending on
whether they start to the left or right of \( AB \). Thereafter
they swing around \( XW \) or \( YZ \) before striking \( S \) again. \( AB \)
is the intersection of part of the stable manifold of the
origin with \( S \), and trajectories started on \( AB \) tend towards
the origin and never return to \( S \). Within the framework
illustrated in Fig. 2 and described above, a great number of
different behaviors are possible (see, for example, Figs. 1
and 6). These may be studied either in terms of the
complete three-dimensional flow or by considering the first
return map (Poincaré return map) of the surface \( S \) to itself;
we will use both approaches below. For all interesting
values of the parameters it seems that the nonwandering
set \( \Lambda \) lies entirely within \( z > 0 \).

For a range of \( r \)-values larger than one it seems that all
trajectories started to the left of \( AB \) spiral in to \( C_1 \) (which
is stable) and all trajectories started to the right of \( AB \)
spiral into \( C_2 \) (also stable); the stable manifold of the
origin divides \( \mathbb{R}^3 \) into two basins of attraction (one each
for each \( C_i \)) in a fairly simple way and \( \Lambda \) contains only the
three stationary points. When \( 3\sigma < 2b + 1 \) this also appears
to be the behavior when \( r \) is large (see below) and we
henceforth ignore this parameter range.

**HOMOClinic Explosions**

The most important possibility inherent in Fig. 2 is that
there may be, for certain values of the parameters, homo-
clinic orbits associated with the origin. Homoclinic orbits
are trajectories which tend towards a stationary point in
both forwards and backwards time, and they occur in the
Lorenz equations when the unstable manifold of the origin
is part of the stable manifold of the origin. Two possible
homoclinic situations are illustrated in Fig. 3 and both
actually occur in the Lorenz system.

Each homoclinic situation can be analyzed quite rigor-
ously for parameter values close to the critical ones and in
a region of phase space close to the homoclinic orbit [31],
[33]. The analysis involves asking whether there are any
trajectories which remain forever (in both forwards and
backwards time) within a region \( D \), which is the union of a
small box around the origin and two tubes surrounding, at
the critical parameter values, the symmetric pair of homo-
clinic orbits. The results of the analysis, if we assume that \( r \)
is the varying parameter and \( r^* \) is the homoclinic parameter
value, are as follows.

On one side of \( r^* \) there are no trajectories which remain
forever within \( D \). On the other side of \( r^* \) there is a
complicated “strange invariant set” (which is part of the nonwandering set) which contains an uncountable number of trajectories which do remain within \( D \) forever. This set actually consists of exactly one trajectory corresponding to each possible sequence of passages through the two tubes; thus there are a countable infinity of periodic orbits corresponding to the different repeating sequences of passages through the two tubes, an uncountable infinity of trajectories which eventually terminate in the origin, and an uncountable infinity of trajectories which wander aperiodically through the tubes forever. The different trajectories in the set are each individually nonstable, each comes arbitrarily close to other trajectories in the set (and hence, in some sense, we can think of the set as a single object), and the set as a whole is also nonstable (almost all trajectories started within \( D \) eventually leave \( D \)). This bifurcation, associated with the homoclinic connection at \( r = r^* \), we call a homoclinic explosion since it adds a great deal of material to the nonwandering set \( \Lambda \).

The details of determining which side of the bifurcation \((r < r^* \text{ or } r > r^*)\) the strange invariant set exists, and of identifying the behavior of the unstable manifold of the origin on both sides of the bifurcation (which will determine which other homoclinic explosions may occur for nearby \( r \)-values) are described in [33]; we can proceed without them.

The First Homoclinic Explosion and Preturbulence

It appears that the first significant change in the behavior of the equations as \( r \) increases is that there is a homoclinic explosion like that illustrated in Fig. 3(a), at an \( r \)-value \( r_1 \) \(( \approx 13.926 \text{ with } \sigma = 10 \text{ and } b = 8/3)\), which produces a strange invariant set as \( r \) increases. Though the rigorous homoclinic analysis only applies in a neighborhood of \( r_1 \), the strange invariant set appears to continue in existence, topologically unchanged though it moves in phase space, in an interval of \( r \)-values \( r_1 < r < r_4 \). In this parameter range we have a return map on the surface \( S \) (which can be taken as the plane \( z = r - 1 \) at these \( r \)-values), as illustrated in Fig. 4(a). Here trajectories started on \( S \) in \( ABNM \) and in \( JKBAX \) next return to the surface \( S \) somewhere within the two long thin sectors \( RM'N' \) and \( LK'J' \), respectively, where \( R \) is the point where the right-hand branch of the unstable manifold of the origin first strikes \( S \). \( L \) is the equivalent point for the left-hand branch (and so, for instance, trajectories started near to \( AB \) but to the left of it next strike \( S \) near to \( L \)), and \( MN \) and \( JK \) are chosen so that they map into themselves. Notice that the map stretches directions “perpendicular” to \( AB \) (so trajectories started close together on \( S \) arrive back at \( S \) further apart in the direction perpendicular to \( AB \)) and strongly contracts directions “parallel” to \( AB \) (computer simulations of the equations produce maps in which no thickness can be discerned in the sectors \( RM'N' \) and \( LK'J' \)). Analysis of the map, providing we make certain assumptions about its properties (such as that the stretching occurs everywhere within \( JKNM \)), allows us to deduce that the strange invariant set still exists (e.g., [17]), intersecting \( S \) within \( JKNM \) in a collection of points given by the intersections of two transverse Cantor sets of lines. Trajectories started on \( S \) in the regions \( XJKNW \) and \( MNZY \) spiral into the stable stationary points \( C_1 \) and \( C_2 \), respectively.

For \( r \) near \( r_3 \), \( MN \) and \( JK \) lie very close to \( AB \), the invariant set all passes very close to the origin, and no immediate numerical evidence of its existence is easily available. As \( r \) increases, \( MN \) and \( JK \) move away from \( AB \), the invariant set “spreads out”, the average stretching perpendicular to \( AB \) becomes less severe, and the size of the regions \( RM'N' \cap XJKNW \) and \( LK'J' \cap MNZY \) (the “escape regions”) decreases in size; these developments give rise to a phenomenon known as “preturbulence” [17] in which some trajectories, started within \( JKNM \), wander near the strange invariant set for long periods of time before eventually returning to \( S \) within one or other of the escape regions from whence they spiral into \( C_1 \) or \( C_2 \). The flow near the nonwandering set is chaotic since the stretching in directions perpendicular to \( AB \) causes nearby trajectories to diverge; this ensures that almost all trajectories do eventually end up in one or other escape region.

Strange Attractor

At a parameter value \( r_4 \) \(( \approx 24.06 \text{ with } \sigma = 10, b = 8/3)\) the average stretching decreases sufficiently that the “escape regions” shrink away completely. The area \( JKNM \) is now mapped into itself and trajectories started on \( S \) within \( JKNM \) always strike \( S \) again within \( JKNM \). We now have an attractor intersecting \( S \) within \( JKNM \) and we call it a strange attractor since the stretching still ensures that we have no stable periodic orbits. For \( r > r_4 \) we have a return map like that shown in Fig. 4(b) and behavior like that shown in Fig. 1. Notice that there is a fixed point of the return map somewhere on each of \( JK \) and \( MN \) (these lines are mapped onto themselves with a contraction); these are points where nonstable periodic orbits which go once around \( C_1 \) or \( C_2 \) intersect \( S \). For some parameter values \((\sigma > b + 1)\) these periodic orbits shrink down towards \( C_1 \)
and $C_2$ as $r$ increases, eventually coalescing with them in a subcritical Hopf bifurcation which removes the orbits and leaves the $C_i$ nonstable [13]. This occurs at an $r$-value $r_y = (\sigma(a + b + 3)/(\sigma - b - 1)) \approx 24.74$ with $\sigma = 10$, $b = 8/3$, but is irrelevant for our consideration of the development of the strange attractor; there are $r$-values $(\sigma > b + 1, 3\sigma > 2b + 1)$ for which we expect the whole of the $\sigma = 10, b = 8/3$ type development to occur despite the $C_i$ remaining stable for all $r$ however large, and others (see below) for which many interesting changes occur while the $C_i$ are stable even though they do lose stability at finite $r$-values. The proximity of the $r_y$ and $r_z$ values when $\sigma = 10$ and $b = 8/3$ can be considered accidental.

The strange attractor that results from return maps like that shown in Fig. 4(b) (whether or not the areas $XWKJ$ and $MNZY$ have shrunk to zero) has been the subject of several papers [10], [11], [25], [34]. To be precise, what is usually studied is a mathematical model flow which has, by definition, those properties which we hope the Lorenz flow possesses. The most important of these are (a) a suitable condition on the stretching in the direction perpendicular to $AB$, and (b) the existence of a contracting foliation. This latter condition means that $S$ can be filled with a continuum of arcs ("parallel" to $AB$), each of which is contracted and taken into another of the arcs by the return map. If this is the case, we can identify each contacting arc with a single point and study a one-dimensional noninter- vable map of an interval to itself, and this study is usually a part of the papers mentioned above. The major theorems on these model "Lorenz attractors" tell us that the flow, in an interval of $r$-values, has a strange attractor, but that the topology of the attractor, which is completely determined by the behavior of the unstable manifold of the origin, changes in every neighborhood of every $r$-value. These theorems can be understood intuitively by noting that in $r > r_4$ the origin and its unstable manifold are part of the strange attractor and that for many $r$-values we may expect this to imply that there will be homoclinic connections with the origin. These connections will be more complicated than those in Fig. 3; the unstable manifold will wander back and forth many times before striking $S$ on $AB$ and returning to the origin. However, the remarks we made earlier about homoclinic explosions will apply at each of these $r$-values; each homoclinic explosion either adds or subtracts a strange invariant set from the nonwandering set, thus changing the topology of the strange attractor.

It should be noted that the intuitive remarks above apply whether or not the model assumptions are appropriate for the true Lorenz flow, and if it turns out that we do not actually have a well-understood model-type attractor in the flow it is nonetheless certain that homoclinic explosions are changing the topology of the attracting set at an infinite number of $r$-values (though these may not be densely spread throughout all $r$-intervals as in the model). It seems unlikely that we will know, in the near future, if the true Lorenz flow does satisfy the model conditions. So far there has been no evidence to the contrary and, in any case, the model will remain of interest.

Fig. 5. Schematic return maps on the surface $S$ in the period doubling regime. $R$ and $L$ are the first intersections of the unstable manifold of the origin with $S$. $R'$ and $L'$ are the first returns of $R$ and $L$. The shaded regions contain the first returns of points in $DEFG$. (a) $30 < r < 54.6, (\sigma = 10, b = 8/3)$. (b) $54.6 < r < 500, (\sigma = 10, b = 8/3)$.

Period Doubling and Chaos

As $r$ increases beyond the well-understood parameter ranges discussed in the preceding sections of this paper, return maps on the surface $S$ change their form from that shown in Fig. 4(b) to those shown in Fig. 5(a) and (b). (Note that the return surface $z = r - 1$ may no longer be appropriate at these larger $r$-values [33]; other appropriate surfaces can be found which, while still including the $C_i$, seem to avoid the problem of important trajectories becoming tangential to $S$ as $r$ changes.) The maps still involve a good deal of stretching and so it is not surprising that chaotic behavior is observed for most $r$-values. However, the existence of the roles in the return maps, which introduces the possibility of contraction in both directions and hence stable periodic orbits, presents us with the same problems that students of most systems of chaotic differential equations have to face (see, for example, Rossler's paper).

Numerical experiments in the appropriate parameter ranges show alternate parameter intervals of stable periodic and chaotic behavior. (With $\sigma = 10$ and $b = 8/3$ the appropriate ranges are approximately $30 < r < 54.6$ for Fig. 5(a) and $54.6 < r < 500$ for Fig. 5(b); there is not, in fact, any important dynamical difference between the behaviors expected in these two intervals. Fig. 5(b) is not strictly accurate for the approximate interval $200 < r < 500$ since here the curves LE' and RF' no longer intersect AB; however, the curves RG' and LD' continue to intersect AB until, approximately, $r = 500$.) The stable periodic behavior appears in the form of the now well-known period-doubling cascades (with the higher period orbits in each cascade existing at lower $r$-values) and many authors have written on aspects of this behavior as it applies to the Lorenz equations [7], [16], [21], [22], [28], [29]. Rigorous analysis of the flow, even under best case model assumptions, is no longer possible. The bends in the return map remove any chance that there is a contracting foliation of the kind which was essential for the analyses of the Lorenz strange attractors discussed earlier. This means we can no longer
answer theoretically questions of the sort, "Are there parameter values for which there are no stable orbits?", "Are there parameter values for which we have many stable orbits?", etc. In addition, numerical experiments cannot be expected to answer these questions where stable orbits of relatively high period are concerned. Though many stable orbits have now been located [15], [33]—some are shown in Fig. 6—and though it is known that there are actually an infinite number of period-doubling cascades involving stable periodic orbits (some of very high period), the majority of the behavior observed numerically in the range $30 < r < 215$ ($\sigma = 10$, $b = 8/3$) is chaotic, and no stable orbits have been observed in the range covered by Fig. 5(a). In $215 < r < 313$ we see what appears to be a final period doubling cascade and for $r > 313$ we see only one stable symmetric orbit which persists for all larger $r$-values.

Most attempts to understand the period-doubling/chaotic regime do not use any of the special properties of the Lorenz equations and many of the results obtained by various authors apply equally well to other quite different chaotic systems. For example, intermittent chaos (where trajectories spend a long time on almost periodic behavior before wandering chaotically for a time and then returning to almost periodic behavior [21], [22]) and noisy periodicity (where the behavior looks like a fairly simple period orbit plus noise [14], [19]) are two commonly observed numerical phenomena, occurring in chaotic parameter intervals at either end of an observed period-doubling cascade. Some properties of these behaviors (which, as we have seen, are defined very loosely by noting certain gross features of the chaos) can be understood in terms of return maps, etc., though others, for instance the fact that the chaotic behavior known as noisy periodicity appears phase coherent [3], [19], cannot. Certainly, both phenomena occur in a wide range of systems. Other approaches, such as the numerical computation of Liapounov exponents [4], fractal dimension [4], spectra [3], and various types of probabilistic dimension [4], all of which are designed, in some sense, to work towards an understanding of chaotic systems in general,
signally fail to add to our comprehension of the Lorenz equations in particular.

Fortunately, the Lorenz equations offer another approach to the period-doubling/chaotic regime. We can think of the period-doubling cascades as being a mechanism for removing (as \( r \) increases) periodic orbits from the nonwandering set \( \Lambda \). We can then ask by what mechanism these orbits were added to \( \Lambda \) in the first place. It appears that they were all added (at least for \( a = 10 \) and \( b = 8/3 \)) in homoclinic explosions. This observation provides us a powerful tool with which to investigate the less easily observed behavior. It is relatively easy to identify the simpler homoclinic connections which occur as \( r \) increases; if the way in which the unstable manifold of the origin takes its successive loops around the two points \( C_1 \) and \( C_2 \) changes between two \( r \)-values this implies the existence of a homoclinic connection at an intermediate \( r \)-value. In addition, the analysis of homoclinic connections in general provides certain information about the pattern of behavior of the unstable manifold of the origin on each side of each homoclinic explosion and, by fitting this information together for successive connections, it is possible to deduce the existence of infinite successions of connections between those that are easily observed [33]. That homoclinic connections do actually continue to occur in the parameter ranges covered by Fig. 5 is easy to see. For example, the change over from Fig. 5(a) to Fig. 5(b) occurs when the points \( R' \) and \( L' \) lie on \( AB \) and this gives us the homoclinic connection shown in Fig. 3(b); it is, incidentally, in the explosion associated with this connection that the orbit shown in Fig. 6(a) is produced.

This approach immediately leads to observations that seem to have escaped many researchers despite the large amount of time spent studying the Lorenz equations. For instance, we may conjecture that the occurrence of homoclinic connections for parameter values up to \( r = 500 \) implies that there are period doubling cascades (and presumably attracting chaotic behavior) coexisting with the stable symmetric periodic orbit which is the only attracting orbit observed numerically in \( r > 313 \). Slightly different arguments along the same lines provide fairly detailed information about the period doubling cascades expected in the complete chaotic/period-doubling parameter range. See [33] for more details. It is true, however, that much of this information can also be conjectured from a careful examination of numerically computed versions of Fig. 5. The more important implications of the new approach are described below.

**Large \( r \) and the Lorenz Equations for General \( \sigma \) and \( b \) Values**

It is possible to analyze the large \( r \) behavior (\( r \to \infty \)) for all positive values of \( \sigma \) and \( b \). This analysis depends on the fact that the equations become conservative in the limit \( r = \infty \). Here we restrict ourselves to outlining the simpler results and some of their implications [26], [30], [33].

(i) \( 3\sigma < 2b + 1 \)

For all large enough \( r \) the nonwandering set contains only the three stationary points and it is possible that there are no bifurcations at all in the parameter range \( 1 < r < \infty \).

(ii) \( 3\sigma > 2b + 1 \)

For all large enough \( r \) there is a stable symmetric orbit which winds once around the \( z \)-axis. If \( \sigma > b + 1 \), \( C_1 \) and \( C_2 \) are nonstable for large enough \( r \) and the symmetric orbit is the only orbit which persists for all large \( r \). If \( \sigma < b + 1 \), \( C_1 \) and \( C_2 \) remain stable but there are also two nonstable periodic orbits (which do not wind around the \( z \)-axis) at all large \( r \)-values. We assume that the only difference between \( \sigma > b + 1 \) and \( \sigma < b + 1 \) is the occurrence or nonoccurrence of the Hopf bifurcation at a finite \( r \)-value; the remainder of the development is likely to be similar.

In a large part of \( 3\sigma > 2b + 1 \) (including \( \sigma = 10 \), \( b = 8/3 \)), the behavior just described is the only large \( r \) behavior. This being the case, it is possible to construct a conjectural argument that in this parameter range we must always see, for \( r \) varying between zero and infinity, a series of bifurcations and behavior changes very similar to those observed with \( \sigma = 10 \) and \( b = 8/3 \) [33]. This argument depends on our knowledge of the behavior of the unstable manifold of the origin at low \( r \)-values, the fact that we can deduce something of its behavior at large \( r \)-values by noting that it must be attracted to the stable symmetric orbit which is the only attracting set at large \( r \)-values, and the observation that our understanding of the series of homoclinic explosions which must occur in order to get from the low \( r \) behavior to the large \( r \) behavior suggests that material must be added to and removed from the nonwandering set in the manner that is actually observed (with some well-defined variations in detail allowed) when \( \sigma = 10 \) and \( b = 8/3 \). Some of the conjectural elements in the argument will almost surely remain unproven for some time; others, particularly those which can be studied in the context of a generalized Lorenz model flow, might be relatively easy to prove. It is also possible that further study of the large \( r \) behavior will provide additional clues.

The power of the argument outlined above lies in its wide area of application. Any Fig. 2 type system with similar large \( r \) behavior (whether symmetric or not) will be a candidate for similar arguments. In addition, as we see below, it may be useful in more complicated situations. Its major weakness is that it predicts only a minimum "necessary" behavior. It is entirely possible that other bifurcations, not associated with homoclinic or large \( r \) behavior, produce and then destroy elements of the nonwandering set in such a way that much of the behavior described above is rendered numerically unobservable; it is perhaps surprising that this does not seem to be the case in the Lorenz equations themselves.

**Small \( b \) or Large \( \sigma \)**

When the parameters \( a \) and \( b \) are chosen so that \( \sigma \) is much larger than \( b \) (the exact relation has not been estab-
lished) the behavior of the Lorenz equations is qualitatively more complicated than that described above. The additional complexity arises because the unstable manifold of the origin spirals around the z-axis before passing downwards through the suitably chosen return surface S. This spiraling is illustrated in Fig. 7 and does not occur when, for example, $\sigma = 10$ and $b = 8/3$. In fact, the number of these spirals increases without limit as $r$ increases (if it occurs at all). It is fairly easy to see that as additional turns are added the point $R$ where the manifold intersects $S$ must cross and recross $AB$ (the stable manifold of the origin) giving rise to an unending series of homoclinic explosions. Each of these produces a strange invariant set containing orbits and trajectories which also spiral many times around the z-axis between passages through $S$. This behavior can be predicted from the large $r$ analysis [33] and Fig. 8 shows an example of chaotic behavior involving spiraling around the z-axis.

For low $r$-values, before the spiraling of the manifold commences, the development of the behavior is as usual. More precisely, the arguments of the previous section allow us to conjecture [33] that there will be a development very close to that seen in the entire range $\sigma = 10$, $b = 8/3$, and $1 < r < \infty$, all occurring at low $r$-values. Numerical experiments confirm this conjecture and for $b = 0.25$ with $\sigma = 10$ one can see this whole development in a parameter range (roughly $1 < r < 12$) in which the stationary points $C_1$ are still stable. This observation reinforces the remarks made earlier to the effect that the Hopf bifurcation is irrelevant for the development of the rest of the behavior.

At larger $r$-values, once the manifold does commence Fig. 7 type spiraling, the exact sequence of behavior changes has hardly been investigated at all and is certainly extremely complicated. However, three main processes have been identified. These are:

1) The never ending sequence of homoclinic explosions, described above, which continually adds ever more topologically complicated orbits and trajectories to the nonwandering set. This process is given by the large $r$ analysis.

2) A never ending sequence of period doubling windows which removes material from the nonwandering set as $r$ increases. This process is predicted by the large $r$ analysis [33] and some of its more detailed properties can be deduced in the particular case that $\sigma$ and $r$ are both large from the work of Fowler and McGuinness [5], [6].

3) A heteroclinic connection between the points $C_1$ (see Fig. 9) must, it is conjectured in [33], occur at some $r$-value whenever the above two processes occur. For $b = 0.25$ and $\sigma = 10$ the relevant $r$-value lies between 480 and 500. This conjecture was produced as the only way of resolving a difference between the large $r$ analysis (which tells you what you have left at large $r$) and the conjectural arguments about sequences of homoclinic explosions (which tells you what is produced by the sequence of explosions which occurs) and, if it is true as appears to be the case for the parameter values just mentioned, it has complicated implications for the behavior near the relevant parameter values (see, for example, [1], [32] which treat a nonsymmetric case).

Further results will doubtless be obtained by direct numerical means but difficulties can be expected. Much of the interesting behavior will occur close to the z-axis and to the origin (which presents numerical problems) and, for many parameter values, experiments will be complicated.
by the existence of theoretically uninteresting attractors; at low $r$-values the $C_r$ will be stable and at large $r$-values the symmetric orbit which winds once around the $z$-axis and which persists for all large enough $r$ will be stable, and these attractors will capture most numerically calculated trajectories.

CONCLUSION

The Lorenz equations show, for different values of the parameters, an extremely wide variety of behaviors. These range from a well-understood strange attractor, occurring for reasonable values of the parameters, to more exotic, less easily observed and less well-understood behaviors occurring for more extreme values of the parameters. What is perhaps most interesting about the equations, given that the question of their physical relevance is unlikely to be resolved in the near future, is that they provide one of the first examples of a system where some measure of understanding, with a predictive capability, can be arrived at by combining the simplest available picture of the flow (Fig. 2) with analyses (rigorous and conjectural, local and global) that deal with well-separated parameter ranges. Thus, for example, almost all the results of the section on large $r$ and small $b$ were first obtained from arguments developed in studying easier parameter ranges and then confirmed, where this was possible, by numerical experiments.

It is reasonable to ask whether similar approaches may add to the understanding of other chaotic systems. The features of the Lorenz flow of most interest were:

(i) the possibility of analysing the behavior for parameter values (low $r$ and large $r$) lying on either side of the chaotic behavior;

(ii) the observation that most of the behavior was associated in some way or other with homoclinic connections at the origin;

(iii) the fact that the eigenvalues of the linearized flow at the origin were such as to make analysis of the bifurcation associated with the homoclinic connections relatively easy; and

(iv) the observation that the different behaviors of Lorenz equations appear to actually occur in one of the simplest possible patterns compatible with the results of (i)–(iii) above.

Most other systems have only been studied in small parameter ranges and the author believes that (i) and some equivalent of (ii) (involving homoclinic or heteroclinic connections between stationary points) are more often true than is currently supposed. Unfortunately, the stationary points in many systems are such that (iii) is clearly false, though progress in the analysis of homoclinic connections of a saddle-focus type (where a pair of the eigenvalues are complex) may improve matters. As for (iv), it is perhaps a matter of faith that things are not actually much worse than they appear; there are enough unanswered questions about most chaotic systems as it is and it seems reasonable to assume that most “naturally occurring” (i.e., plausibly derived and easy to write on the blackboard) systems behave with at least relative simplicity until it is shown otherwise.

ACKNOWLEDGMENT

The author would like to thank P. Swinney-Dyer, A. Mees, and P. Glendinning for helpful comments and suggestions.

REFERENCES

Nonlinear Lattice and Soliton Theory

MORIKAZU TODA

Abstract — Since the notion of stable pulses, known as solitons, plays a central role in the phenomena of wave propagation in nonlinear systems, an exposition of this topic is developed in some detail. It is known that the equations of motion of the one-dimensional lattice of particles with exponential interaction are integrable, namely, they admit exact solutions, and this system is equivalent to an LC circuit with certain nonlinear capacitance. In addition, a closely related partial differential equation called the Korteweg-de Vries (KdV) equation is also integrable. Special emphasis is placed on these integrable systems.

I. INTRODUCTION

If the waves in a medium are subject to the superposition principle, the medium is said to be linear, and otherwise nonlinear. A vacuum is perfectly linear for the waves of light, and air is nearly linear for the waves of sound. In many cases, materials are linear for electro-magnetic waves, but the speed of propagation or the refractive index depends on the frequency, that is, materials are dispersive. Moreover, when the intensity of light is very strong, as in the case of strong laser beams, the speed of propagation may depend on the intensity; then the superposition principle does not hold and the medium is nonlinear. The medium is thus characterized by dispersivity and nonlinearity. Of course the dissipation affects the speed of propagation and the wave-profile; but since it usually has no drastic effect on the mode of propagation, we shall disregard it for a while until the latter part of this article.

Because of the dispersion effect, a superposed wave, such as a pulse, will collapse losing its identity as it propagates down the medium. It has been thought, in general, that nonlinearity will yield a similar collapse of the initial wave form by deforming and scattering the wave.

However, in the course of a numerical study of the ergodicity of a weakly nonlinear lattice, Fermi, Pasta, and Ulam (FPU) [1] found a recurrence phenomenon: for a variety of sinusoidal initial conditions, they observed that the wave recovered its initial state after some lapse of time. Zabusky and Kruskal [2] simplified the problem by adopting an equation which is a continuum approximation to the weakly nonlinear lattice. The equation is what is called the Korteweg-de Vries (KdV) equation, originally proposed for shallow water waves [3]. They discovered, by a computer simulation, the appearance of stable nonlinear dispersive wave entities, which they called solitions.

Some sort of balancing between the flattening effect of dispersion and steepening effect stabilizes each soliton. A soliton of high amplitude travels faster than one of low amplitude, and solitons pass through one another without losing their identity. During the overlap time interval their

Manuscript received December 1, 1982. The author’s home address is 5-29-8-108 Yoyogi, Shibuya-Ku, Tokyo 151, Japan.

0018-9383/83/0800-0542$01.00 ©1983 IEEE