
Dynamics

The Geometry of Behavior

Fourth Edition

Part 3

Global Behavior

with 136 illustrations

by

Ralph H. Abraham

&

Christopher D. Shaw

Aerial Press

Aerial Press
P.O. Box 1360
Santa Cruz, California 95061

Design and Typography: Page & Curtis, Santa Cruz, California

Library of Congress Control Number 00-135270

Abraham, Ralph
Dynamics — the geometry of behavior / Ralph H. Abraham and
Christopher D. Shaw — 4th ed.

Part 3, Global Behavior

Four Part eBook
ISBN 0-942344-24-3

©2005 by Aerial Press

All rights reserved.

Part 3: Global Behavior

Mathematical Dynamics Hall of Fame	333
10. Global Phase Portraits	337
10.1 Multiple attractors	339
10.2 Actual and virtual separatrices	344
11. Generic Properties	349
11.1 Property G1 for critical points	351
11.2 Property G2 for closed orbits	354
11.3 Property G3 for saddle connections in 2D	357
11.4 Properties G4 and F	360
12. Structural Stability	363
12.1 Stability concepts	365
12.2 Peixoto's theorem	370
12.3 Peixoto's proof	374
13. Heteroclinic Tangles	377
13.1 Point to point	379
13.2 Outsets of the Lorenz mask	383
13.3 Point to cycle	391
13.4 Cycle to cycle	397
13.5 Birkhoff's signature	400
14. Homoclinic Tangles	407
14.1 Homoclinic cycles	409
14.2 Signature sequence	414
14.3 Horseshoes	420
14.4 Hypercycles	424
15. Nontrivial Recurrence	427
15.1 Nearly periodic orbits	429
15.2 Why Peixoto's theorem failed in 3D	436
15.3 Nonwandering Points	438

Dynamics

The Geometry of Behavior

Fourth Edition

Part 3

Global Behavior

Dedicated to Mauricio Matos Peixoto



PHOTOGRAPH BY CAROLINE BLAKEMORE

Mathematical Dynamics

Hall of Fame

The early days of modern dynamics span half a century, beginning in the 1880's, as described in Part One. At this time in France, Poincaré innovated qualitative methods. More or less simultaneously in Russia, Liapounov pioneered stability methods. These techniques then underwent separate, parallel developments. By the 1930's, important progress had been made in Europe and America, following the lead of Poincaré. Birkhoff, at Harvard, was the outstanding figure. Meanwhile, in Russia, the ideas of Liapounov had grown. Andronov was an important figure in this tradition.

There followed a quiet period. For another quarter century, the tradition of Poincaré dwindled in Europe and America. Developments in Russia were forgotten in the West. During this period, experimental dynamics began in Europe, as described in Part Two.

Eventually, through the efforts of emigré mathematicians familiar with the Russian work, such as Lefshetz and Minorsky, the qualitative theory of dynamical systems was revived in America. Beginning in the 1950's, a vigorous mathematical program picked up steam and continues today. The global behavior of dynamical systems is the main theme of this movement, which we may call mathematical dynamics.

Table 3.1 - The History of Global Theory

Date	Europe/America	Russia
1850		
1900	Poincaré Floquet	Liapounov Mandelshtam
1950	Birkhoff Lefshetz, Minorsky, de Baggis, Peixoto, Markus, Thom, Smale, Pugh	Kolmogorov Arnol'd
2000		

Here are some capsule histories.

Henri Poincaré, 1854-1912. Besides pioneering the new methods of dynamics and topology, Poincaré discovered tangles and bifurcations as we know them today.

Aleksandr Mikhailovich Liapounov, 1875-1918. In his Ph.D. thesis of 1892, Liapounov established the Characteristic Exponents of an equilibrium point as the determinants of its asymptotic stability.

George David Birkhoff, 1884-1944. Birkhoff was fascinated by tangles, and wrote several papers about them. In one, he introduced the *signature* of a tangle, making a first step in the historic struggle to untangle them. In another, he showed that homoclinic tangles are always surrounded by myriad periodic trajectories.

Aleksandr Aleksandrovich Andronov, 1901-1952. With co-workers Leontovich and L.S. Pontrjagin, Andronov pioneered the phase portrait point of view. Andronov and Pontrjagin published a five-page paper in 1937 which revolutionized global dynamics. Its main contribution was the definition of

structural stability. In the same year, Andronov published an influential book, written with C.E. Chaikin, on nonlinear oscillations.

Gaston Floquet. He established the Characteristic Multipliers of a limit cycle as the determinants of stability, parallel to the CE's of Liapounov, in 1879.

Solomon Lefshetz, 1884-1972. In the World War II years, this great innovator of algebraic topology turned his attention to qualitative dynamics. A text on the local theory in 1946 was followed by a global treatment in 1957, in which structural stability was discussed in two-dimensional systems. A native of Russia, he reinjected the tradition of Liapounov into the mainstream of Western mathematics.

Nicolai Minorsky, b. 1883. Like Lefshetz, Minorsky emigrated to the United States in the prewar years. His knowledge of the Russian school of dynamical systems theory, presented in his book of 1952, gave great impetus to the resumption of mathematical dynamics in the United States.

Henry de Baggis, b. 1916. A student of Lefschetz, in 1947 he proved the conjecture of Andronov and Pontrjagin on structural stability in the plane.

Mauricio M. Peixoto. Also a student of Lefschetz, he improved enormously on de Baggis's result in 1959. In doing so, he forged the connection between dynamics and topology which has been so fruitful in recent years.

René Thom, b. 1923. Thom used dynamics in his work in topology, for which he was awarded the Field Medal. In 1960 or so, he began advocating the importance of the concept of structural stability in applications, and his very global view of bifurcations. His program was presented in full in his epochal book, *Structural Stability and Morphogenesis*, in 1966.

Lawrence Markus, b. 1922. Another pioneer in the merger of topology and dynamics, he clarified the meaning of generic property in global dynamics, in 1960. This work is described in Section 11.1.

Stephen Smale, b. 1930. Like Thom, Smale used dynamics in his work in topology, which earned a Fields Medal in 1960. He then went on to study dynamics itself, and produced a series of papers in the 1960's which have been very influential ever since. In one of these, he improved substantially on Birkhoff's results on homoclinic tangles, as we explain in Section 14.4.

10 Global Phase Portraits

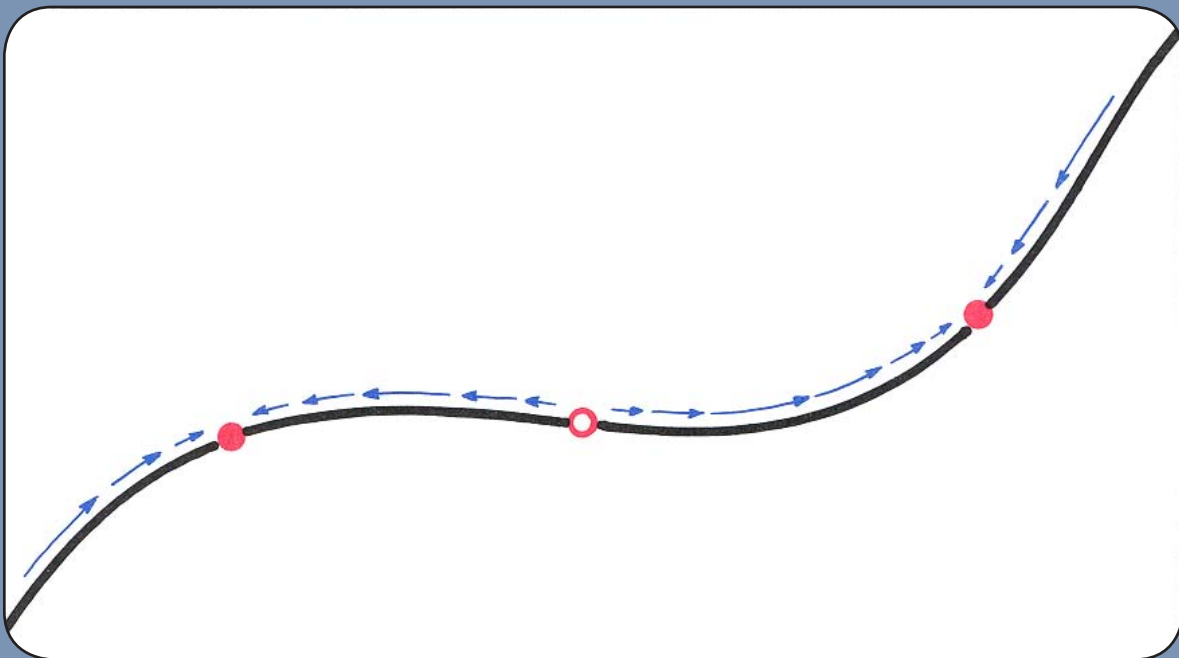
In Part One, we introduced limit points and cycles in dimensions one, two, and three. The decomposition of the state space into basins of attraction, by the separatrices, was emphasized. In Part Two, the inset structure of the separatrices was developed. The geometry of the exceptional limit sets, determined by their Liapounov characteristic exponents, was described. In this chapter, we review all this and assemble it into a global overview of the phase portrait of a typical system.

10.1. Multiple Attractors

For pedagogic reasons, our discussion has often centered on an attractor. However, generic systems commonly have several attractors. So we begin this review chapter with an explicit acknowledgment of this fundamental feature: multiple attractors.

Let's begin with the simplest case, in which the state space is one-dimensional: a curve.

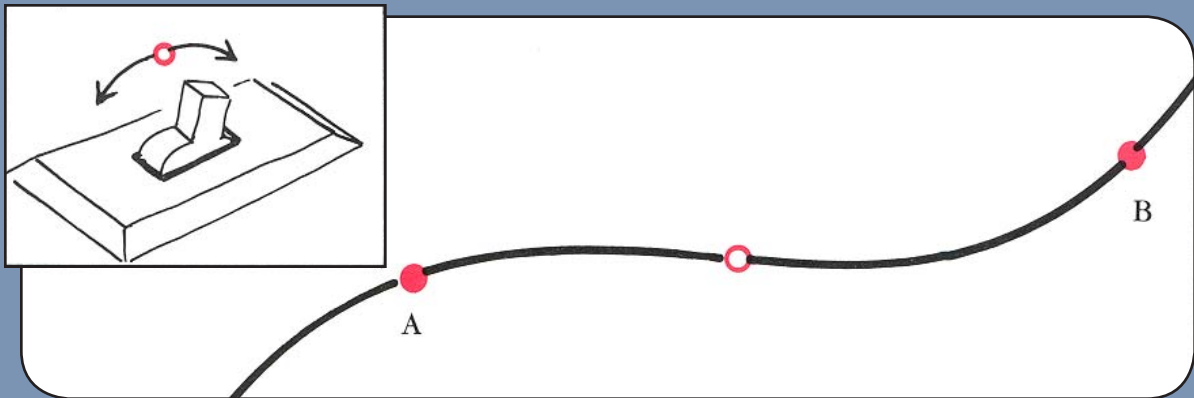
In this context, limit sets are points. Generically, point attractors and point repellers alternate along the curve. The repelling points separate the basins of the attracting points. An initial state, chosen from one of the basins, tends toward the unique attractor in its basin. The different attractors represent the equilibrium states that may be observed in this system.



10.1.1.

In this example, there are *two* attractive points, each in its own basin. The system is *bistable*, in that two distinct stable equilibria are possible.

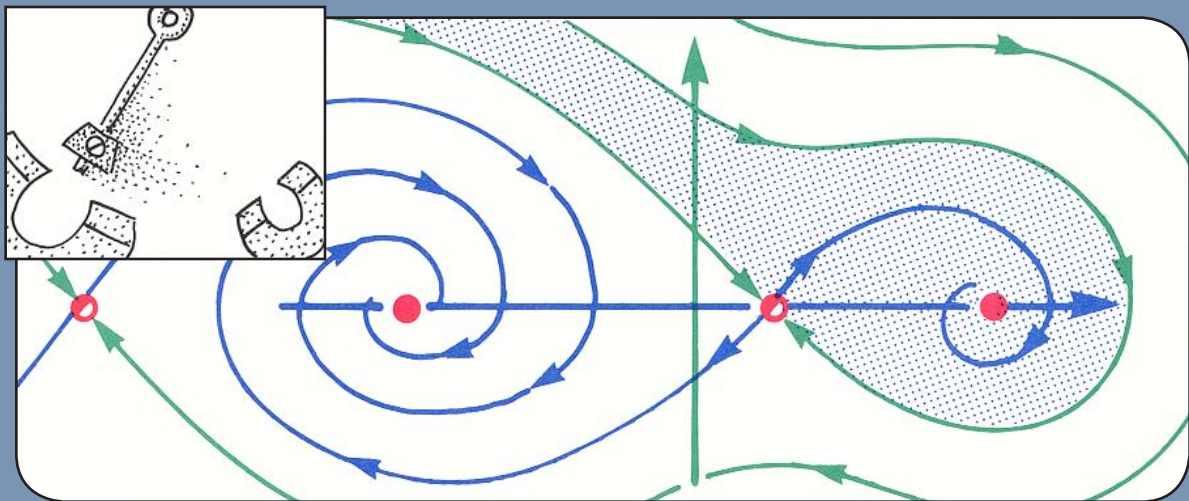
In general, a one-dimensional system is *multi-stable*, in that more than one stable equilibrium point is possible.



10.1.2.

Notice that in this example, the two basins are separated by the point repeller. Initial points slightly to the left of the repeller tend to attractor A, while those slightly to the right tend to attractor B. This behavior is roughly like a mechanical toggle switch.

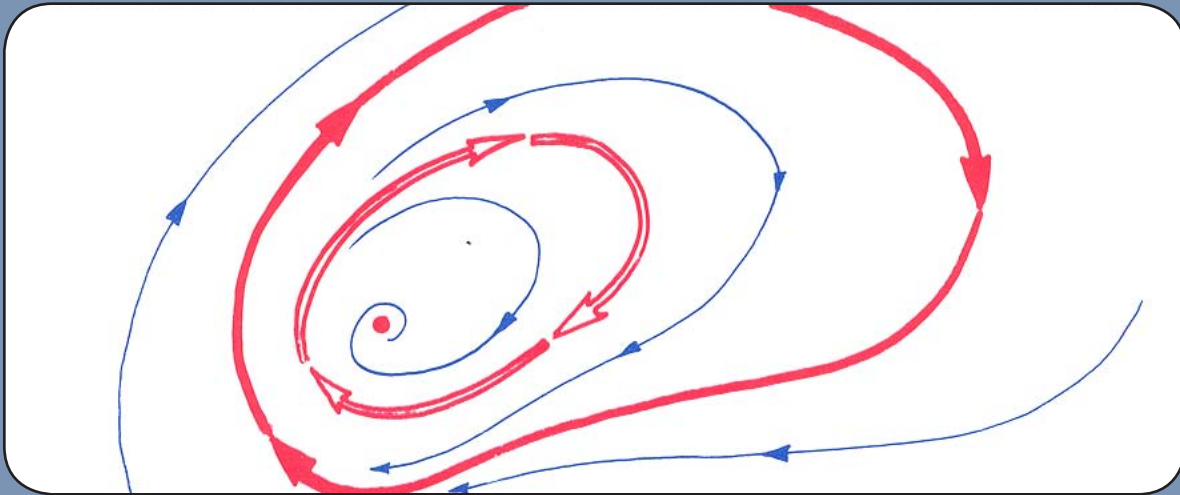
Now let's move on to two dimensions.



10.1.3.

Remember the pendulum? Here is the magnetic bob from Figure 2.1.22. This is also a bistable system. But the two basins are two-dimensional, so the separatrix between them is a curve. This curve is *repelling*, yet not a *repellor*. In fact, it consists of the *inset* of the saddle point between the point attractors. This saddle point represents an *unstable equilibrium* of the bob, balanced between the forces of the two magnets. And its inset represents those improbable initial states which tend to this unstable equilibrium and balance there.

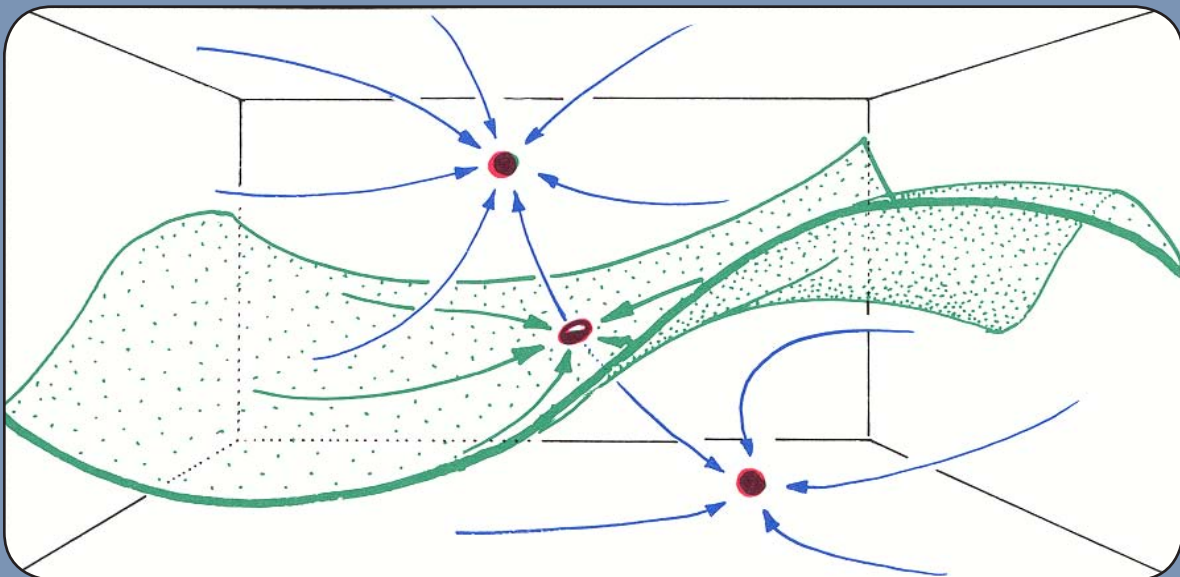
Not every separatrix is the inset of a saddle point.



10.1.4.

Recall this portrait, from Figure 1.5.8. Here, the periodic repeller bounds the two-dimensional basin of an attractive point. It is a separatrix.

Two dimensions are rather special. Let's have a look at the three-dimensional case, which is more typical.

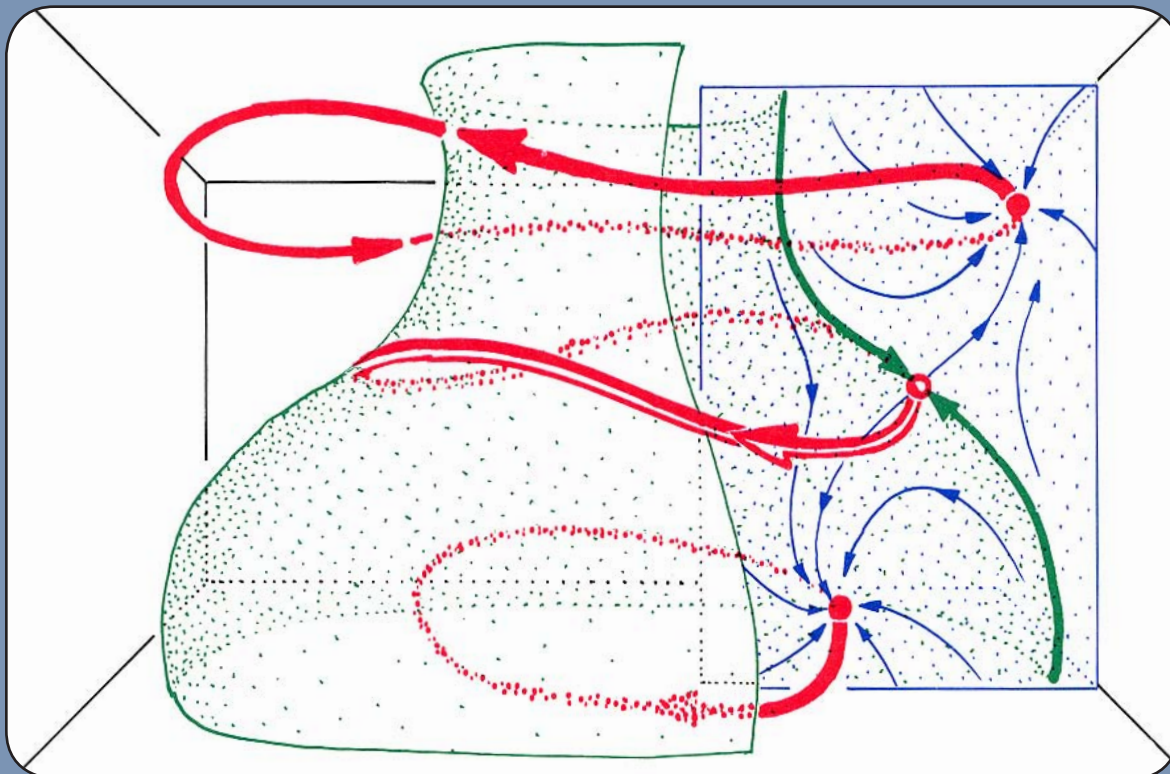


10.1.5.

In this portrait of a simple bistable system in 3D, there are again two attractors. Both are rest points. Their basins are three-dimensional, and are bounded by a surface. This surface, the separatrix in this example, is the inset of a saddle point of index 1.

Recall that the *index* of a saddle point is the dimension of its outset.

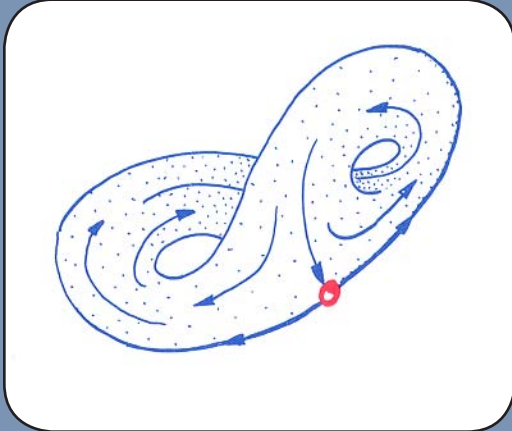
Separatrices need not be insets of a saddle point. They are, usually (but not always), insets of a nonattractive limit set: point, cycle, or chaos.



10.1.6.

Here, for example, is a bistable system with two periodic attractors. Their basins are bounded by a cylindrical surface, the separatrix. It is the inset of a periodic saddle.

Remember that limit sets can be aperiodic, that is, chaotic. Thus, there may be both chaotic attractors and chaotic separatrices in a typical multistable system. Details are given in Part Two.



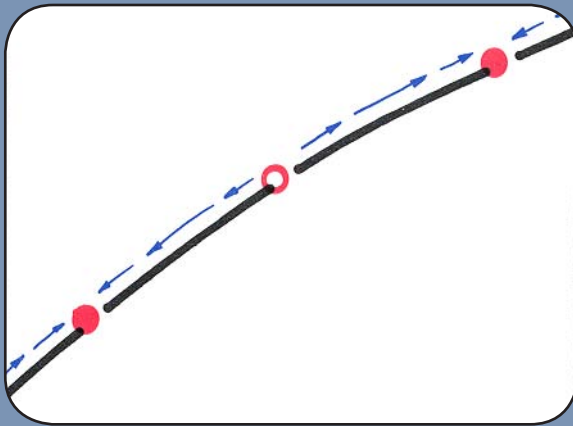
10.1.7.
This is one of
the most
famous chaotic
attractors.

10.2.

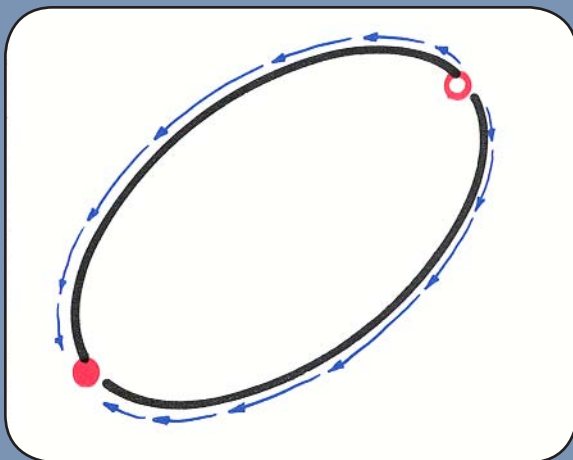
Actual and Virtual Separatrices

In Section 1.5, we defined the separatrix of a dynamical system as the complement of the basins of attraction. That is, an initial state belongs to the separatrix if its future (omega) limit set is not an attractor. According to this agreement, the separatrix consists of the insets of the non-attractive (or exceptional) limit sets. (See Section 1.5.) But do they, in fact, actually separate basins? If so, they are called *actual separatrices*. But, as we shall see, it may happen that they do not separate basins. In this case, they are called *virtual separatrices*.

Here are some examples, beginning with 1D.



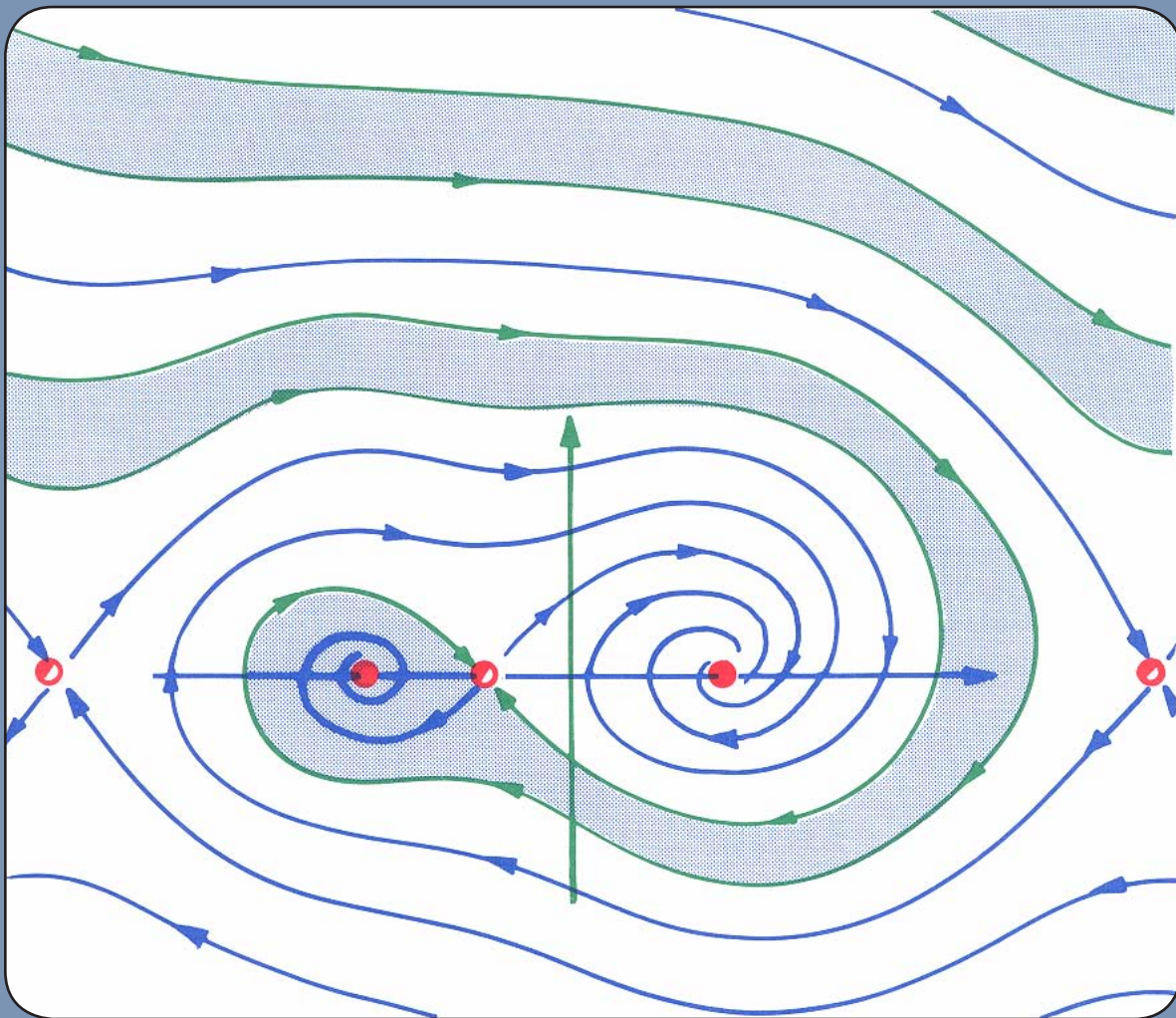
10.2.1.
As we have seen in the preceding section, point repellers may separate basins in one-dimensional state spaces.



10.2.2.
But if we connect the ends of the curve, we have a unstable system! There is only one basin. The separatrix (a single point repeller) bounds it, but does not separate anything. It is a *virtual separatrix*.

Likewise, in 2D, the separatrix consists of curves that are either insets of saddle points or periodic repellers. Examples of both sorts have been shown in the preceding section. But now look at these.

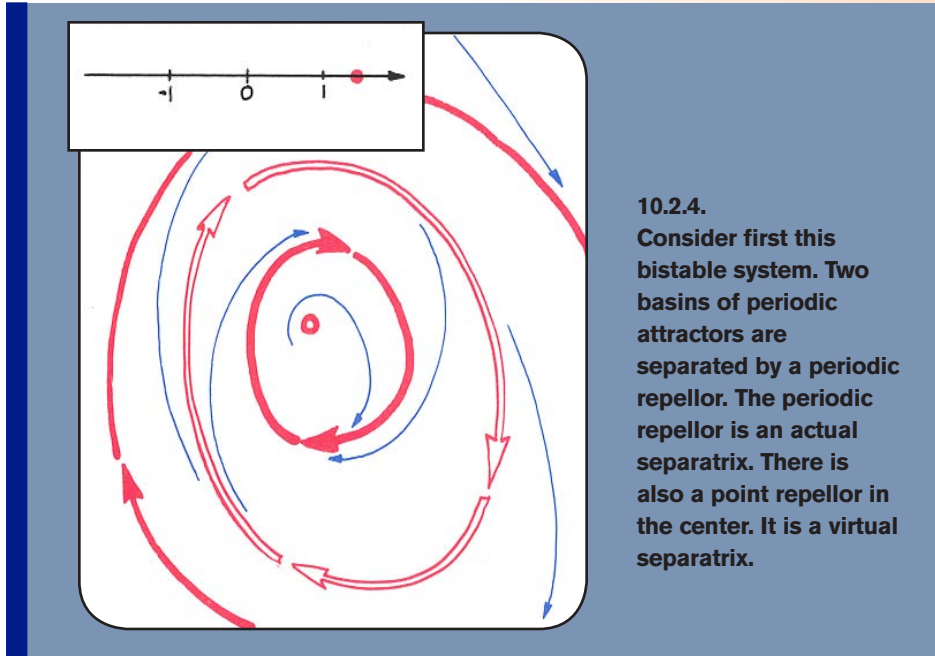
Recall that in the bistable magnetic pendulum, there is a saddle point near the bottom, as shown in Figure 10.1.3. But like the simple pendulum of Section 2.1, there is also a saddle point at the top of the swing.



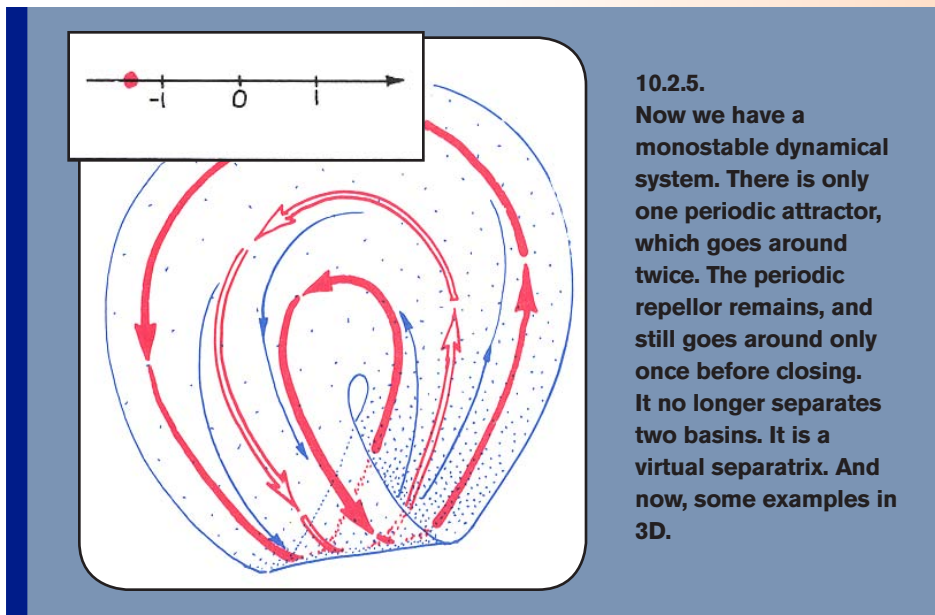
10.2.3.

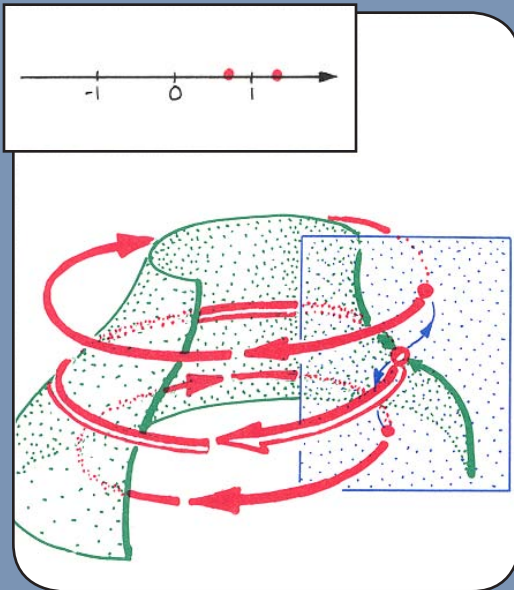
The saddle point at the top of the swing represents the watershed between falling to the right and falling to the left. Its inset consists of those improbable initial states that tend to balance at the top of the swing. As shown here, the initial states close to this inset, to either side, belong to the same (unshaded) basin. Thus, this inset curve is a *virtual separatrix*.

This inset failed to actually separate basins because the state space is a cylinder. Another way an inset may fail to divide basins occurs on the Möbius band.

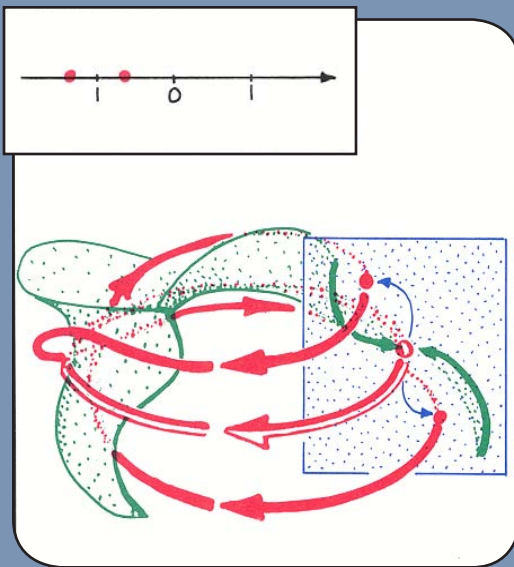


Now remove the point repeller at the center, cut through the remaining strip, give one end a half-twist, and carefully paste the ends together again.



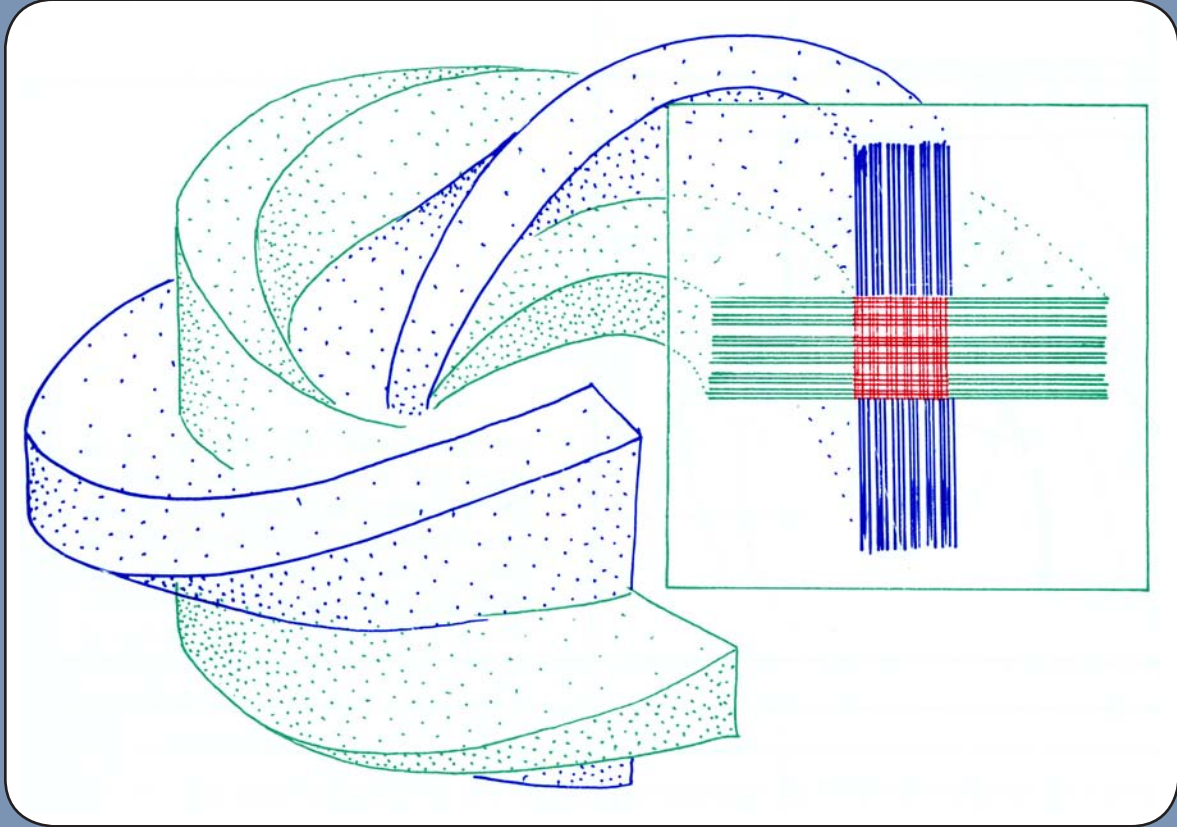


10.2.6.
Recall this portrait from Figure 7.3.9. The inset of a periodic saddle in three-dimensional Euclidean space is twisted an even number of times. It still bounds two basins, and is an actual separatrix.



10.2.7.
On the other hand, as shown in Figure 7.3.10, it may twist an odd number of times. Then it bounds only one basin, and is a virtual separatrix.

Finally, recall that insets may be thick, or chaotic. Our favorite example, Poincaré's solenoid, was constructed step by step in Section 8.1.



10.2.8.

The fractal inset of this periodic saddle of homoclinic type is twisted once, as shown in Figure 8.1.7. It is a virtual separatrix.

11 Generic Properties

We always try to convey the features of typical, garden-variety, dynamical systems. The exceptional cases are more complicated and numerous, and they interrupt the discussion. Moreover, we feel that they should not arise very often in applications, because they are exceptional. This prejudice, shared by all dynamicists, has become a main theme in dynamical systems theory.

The properties characterizing these typical systems are called *generic properties*. Although this name was established early in the program, it turned out that it might have been better to call them *weakly generic properties*. For it has become commonplace to observe exceptional behavior (violating a so-called generic property) very frequently. An explanation for this paradox will be given in Part Four, “Bifurcation Behavior.” Meanwhile, with this warning, we will continue to call these properties generic!

A considerable portion of the history of mathematical dynamics has been dominated by the search for generic properties. These define a class of phase portraits that are far simpler than arbitrary ones. The goal of the search is to narrow down the complexity of the portraits enough to allow a complete classification. This was achieved for dynamical systems in the plane by Peixoto around 1959. This gave the whole program a tremendous boost, but the higher dimensional generic systems are still hopelessly complex. This chapter presents the fundamentals of this program, initiated by Andronov and Leontovich in 1934.

The prototypical results, due to Peixoto, apply to orientable (untwisted) surfaces. An early global result for other state spaces was found by Lawrence Markus around 1960. Definitive results were obtained by Ivan Kupka and Stephen Smale in 1964. Now we will describe the essence of this main theme in the theory.


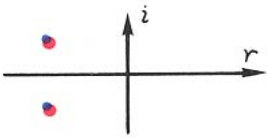
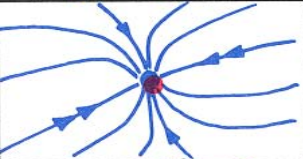
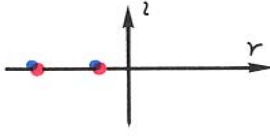

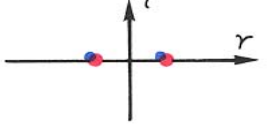
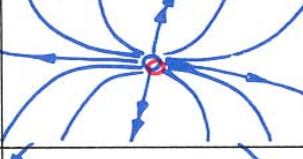
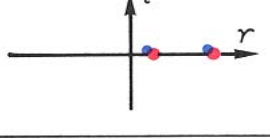


We begin with the definition of the most important global properties of dynamical systems, or vectorfields: G_1 , G_2 , and G_3 . Then, in a final section, we describe the official meaning of *generic property* and state the *Kupka-Smale Theorem: Properties G_1 , G_2 and G_3 are generic.*

11.1.

Property G1 for Critical Points

To begin, let's recall the distinction between hyperbolic and nonhyperbolic critical points.

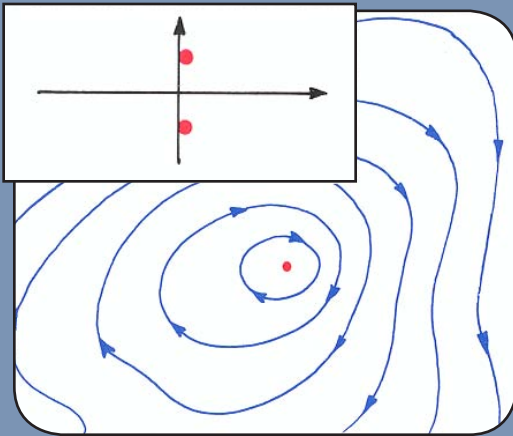
In Chapter 6 we created an atlas of limit points. Using their CE's, we carefully distinguished the hyperbolic and nonhyperbolic cases. We brushed aside the nonhyperbolic cases, claiming they are nondegenerate, exceptional, or *nongeneric*. The global formulation of this assertion is the part of the *Kupka-Smale Theorem* asserting the genericity of property G1, for critical points. In this section, we describe this property of critical points (that is, limit points).

type	index	portrait	C.E.
attractors	0		
	0		
saddle	1		
repellers	2		
	2		

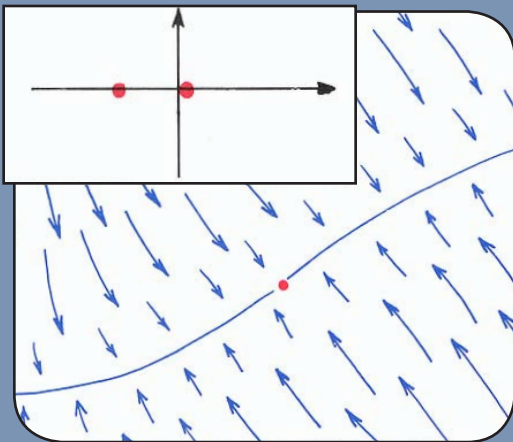
11.1.1. This is Figure 6.4.8, showing the five elementary critical points in 2D. There are seven hyperbolic critical points, namely, these five together with the radial attractor and the radial repeller.

Here, *radial* means that the CE's are real and equal.

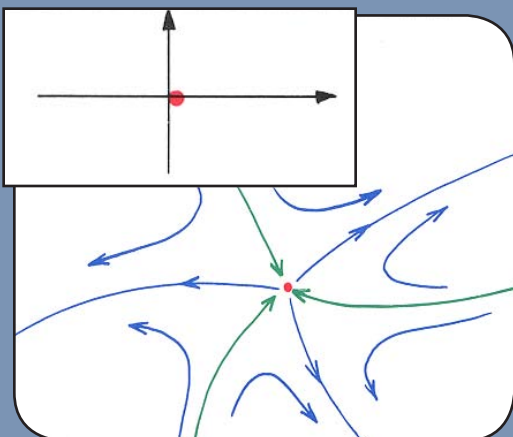
The radial type is intermediate between the spiral and nodal types.



11.1.2.
This is a nonhyperbolic critical point called a *center*. The CE's are shown in the inset window.

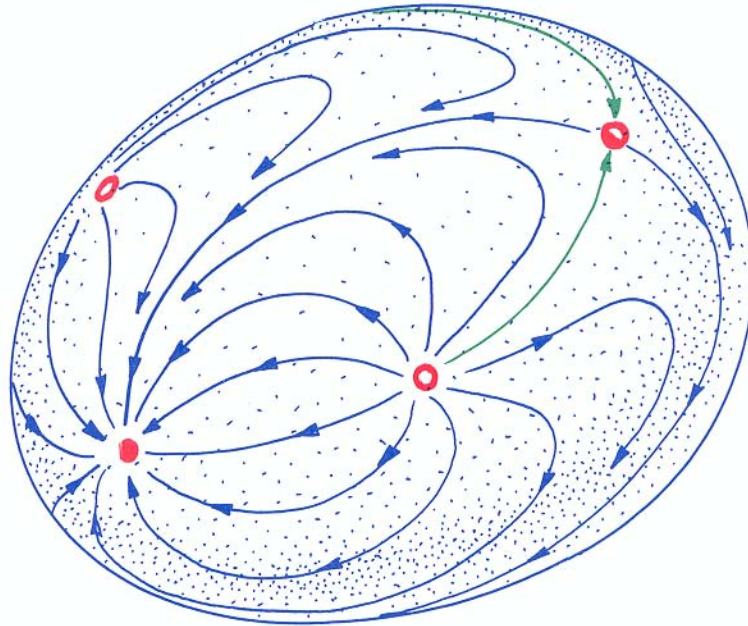


11.1.3.
This is another type of nonhyperbolicity.



11.1.4.
This is the worst case of nonhyperbolicity. Many more different portraits are possible with both CE's zero than in the two cases above.

Now we are ready for property G1.



11.1.5.

A dynamical system has property G1 if all of its critical points are elementary. In this example, each and every critical point is elementary.

In the literature of dynamical systems theory, this definition usually has *hyperbolic* in place of *elementary*. But this version probably results in a more satisfactory theory, from the point of view of the experimentalist, or in the context of applications.

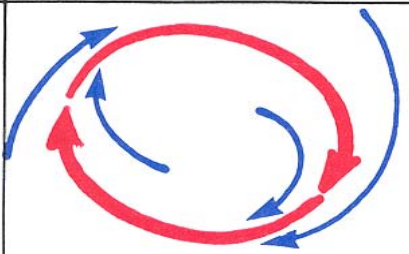

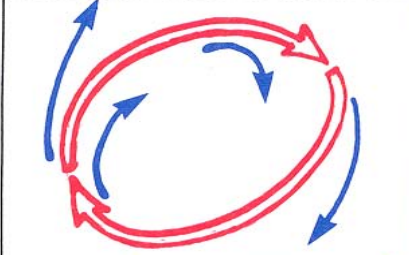
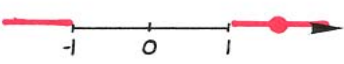
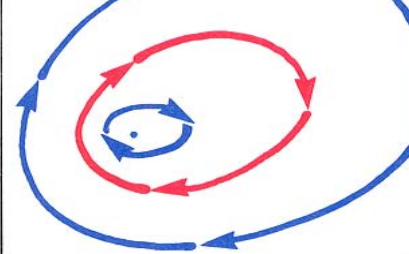
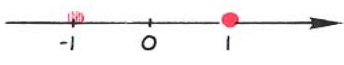
For the eight elementary critical points that occur in 3D, see Figures 6.5.5. and 6.5.6.

11.2.

Property G2 for Closed Orbits

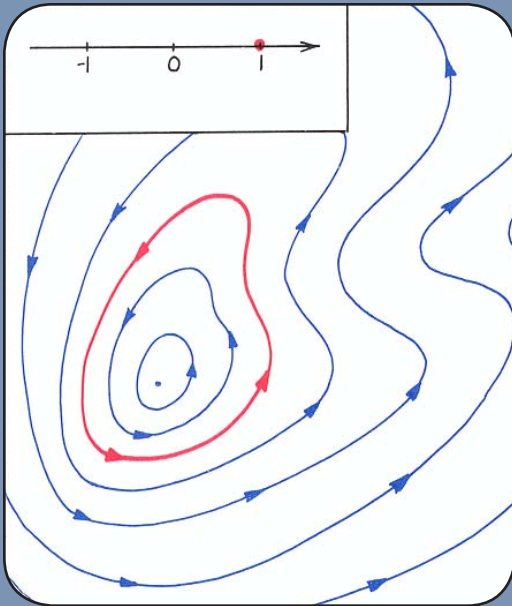
In Chapter 7, we created an atlas of limit cycles. Using their CM's, we carefully distinguished the hyperbolic and nonhyperbolic cases. As in the case of limit points, we neglected the nonhyperbolic cases. The global justification of this neglect is the part of the *Kupka-Smale Theorem* asserting the genericity of property G2, for limit cycles. In this section, we describe this property of limit cycles.

To begin, let's recall the distinction between hyperbolic and nonhyperbolic limit cycles. For 2D, these were shown in Figure 7.2.7.

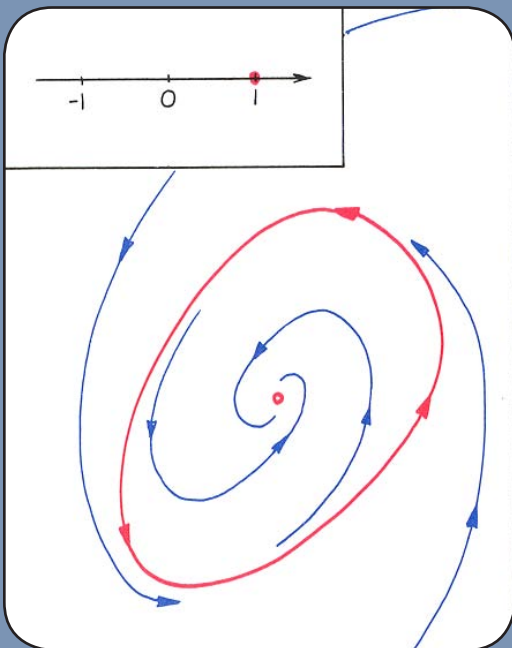
	portrait	C.M.
attractor		 $-1 < C.M. < 1$
repellor		 $C.M. < -1$ or $C.M. > 1$
non-hyperbolic		 $C.M. = -1$ or $C.M. = 1$

11.2.1.
In 2D, a limit cycle has only one characteristic multiplier (CM), which is *real*. These are the only hyperbolic limit cycles in 2D. The absolute value of the CM is smaller than 1 (periodic attractor) or greater than 1 (periodic repellor).

In the nonhyperbolic case, the CM is equal to plus or minus 1, and the limit cycle may be an attractor, a repeller, or neither. Here are two examples, with the CM equal to plus 1.



11.2.2.
This portrait, called a *center*, has more or less concentric limit cycles. Each of them is nonhyperbolic.

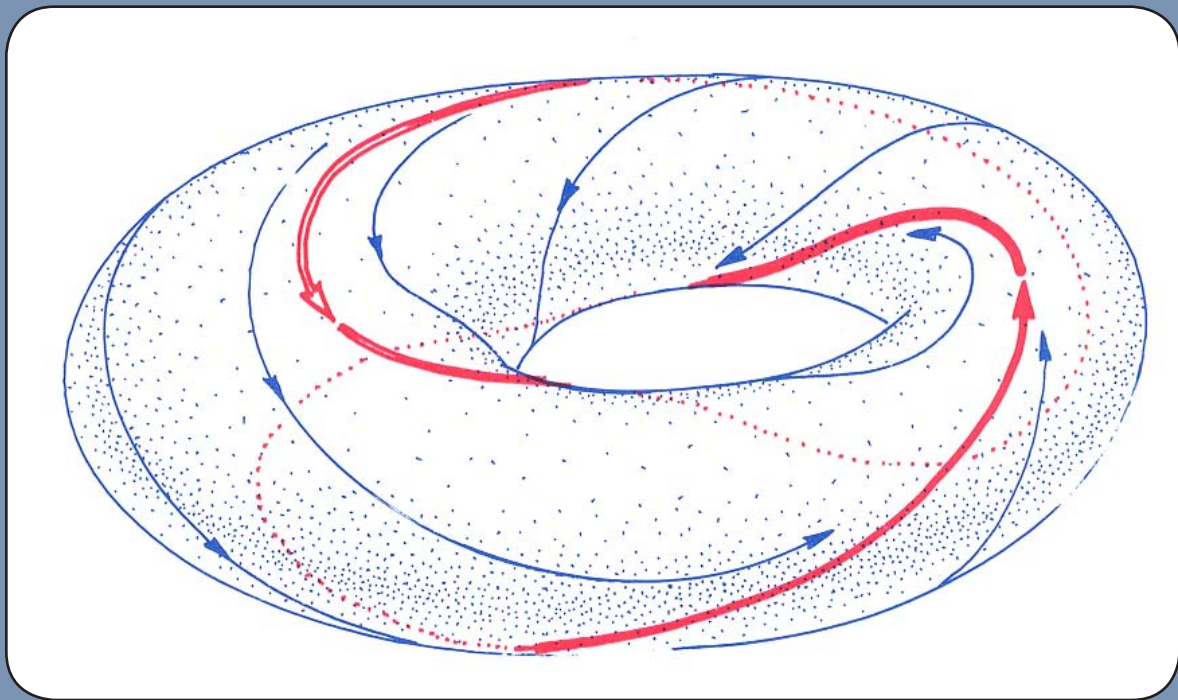


11.2.3.
This portrait has a single limit cycle. It attracts on one side, and repels on the other. Its CM of plus 1 is not enough information to predict its attracting/repelling behavior.

This completes our partial survey of limit cycles in 2D. In 3D, each limit cycle has two CM's. They may be conjugate complex, or both real. If they are both real, they may be distinct or identical. This brings up the distinction between *hyperbolic* and the similar idea, *elementary*. The actual definition of *hyperbolic limit cycle* in any one dimension is: there are no CM's of absolute value 1. *Elementary* is a little stronger. An *elementary limit cycle* is one which is hyperbolic, plus all its CM's are distinct.

All the elementary limit cycles in 3D are shown in Figure 7.5.7.

Here is the definition of G2.



11.2.4.

A dynamical system satisfies property G2 if each and every one of its limit cycles is elementary. In this example on the two-dimensional torus, there are several limit cycles in a braid, and each is elementary.

11.3.

Property G3 for Saddle Connections in 2D

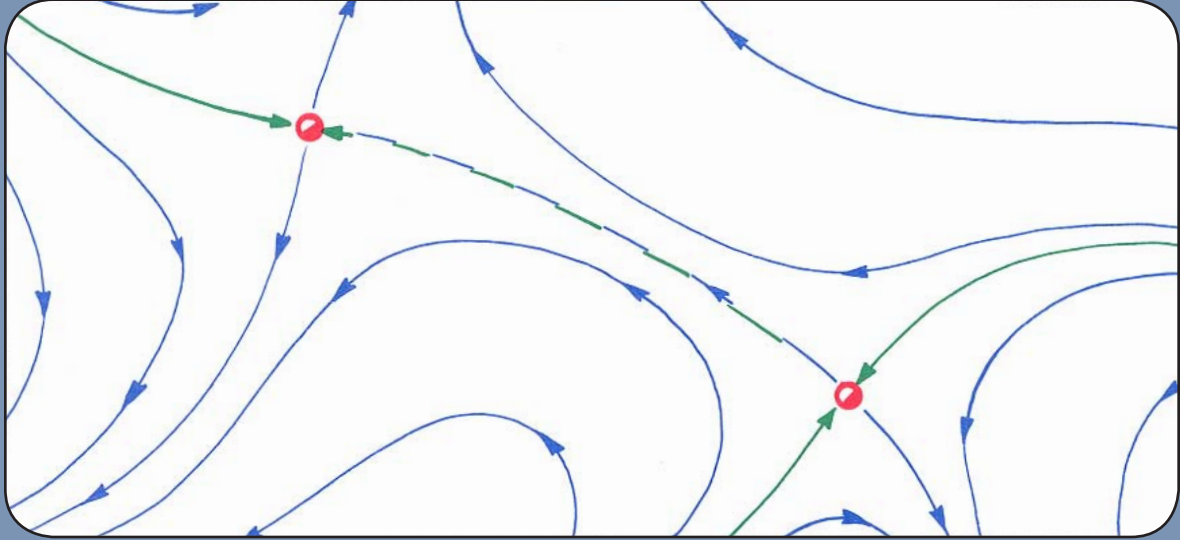
Every trajectory of a dynamical system comes from somewhere and goes somewhere. That is, it has an alpha limit set and an omega limit set. Every trajectory is in the outset of its alpha limit set, and at the same time in the inset of its omega limit set. Thus, outssets and insets normally intersect each other.

However, most of the time, a trajectory comes from a repellor and goes to an attractor. Exceptionally, one comes from a repellor and goes to a saddle, or comes from a saddle and also goes to a saddle. Such a trajectory is called a *saddle connection*, or a *heteroclinic trajectory*. It is even possible for a trajectory to connect a saddle to itself! This is called a *homoclinic trajectory*. Poincaré realized that these trajectories were particularly important in the qualitative behavior of dynamical systems.

Note that a heteroclinic trajectory always belongs to the outset of a saddle (the *donor*), and to the inset of a saddle (the *receptor*) as well. Therefore, the donor outset and the receptor inset must intersect, and their intersection contains the entire heteroclinic trajectory. Generally, the intersection of a saddle outset and a saddle inset contains not just one, but an entire family of heteroclinic trajectories. Property G3 concerns the quality of the intersection of insets and outssets of limit sets of saddle type, especially saddle points and periodic saddles. It requires that these intersections all be *transverse* (that is, cleanly crossing).

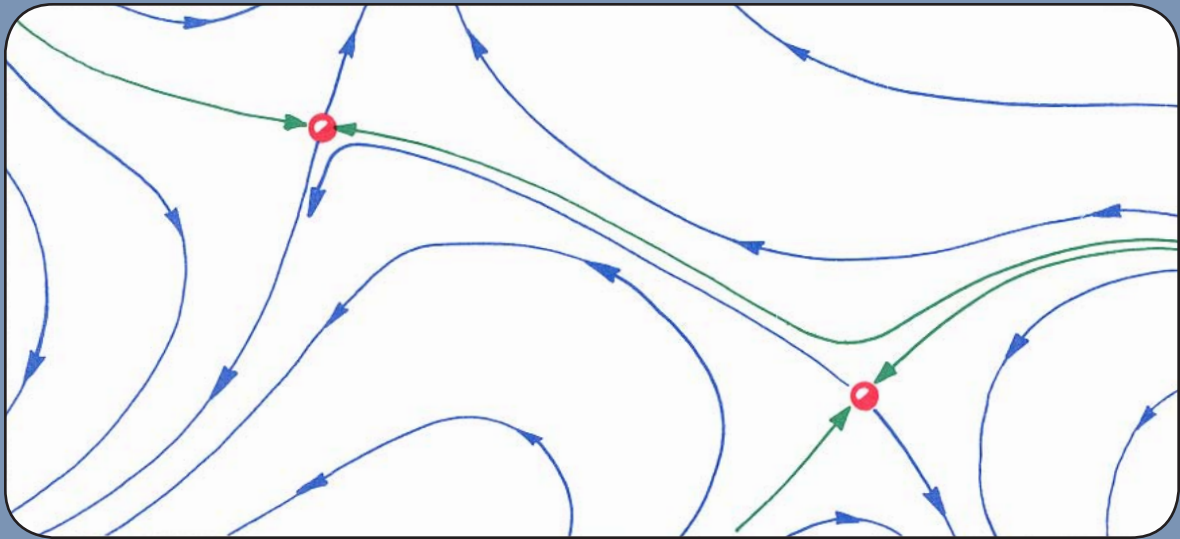
In state spaces of one dimension, there are no saddles. In two dimensions, hyperbolic saddle points have invariant curves as inset and outset. There are no periodic saddles. In this section, we briefly explain property G3 *in dimension two only*. The full story is told in detail in Chapters 13 and 14.

In two dimensions, a dynamical system satisfies property G3 if it has no saddle connections at all.



11.3.1.

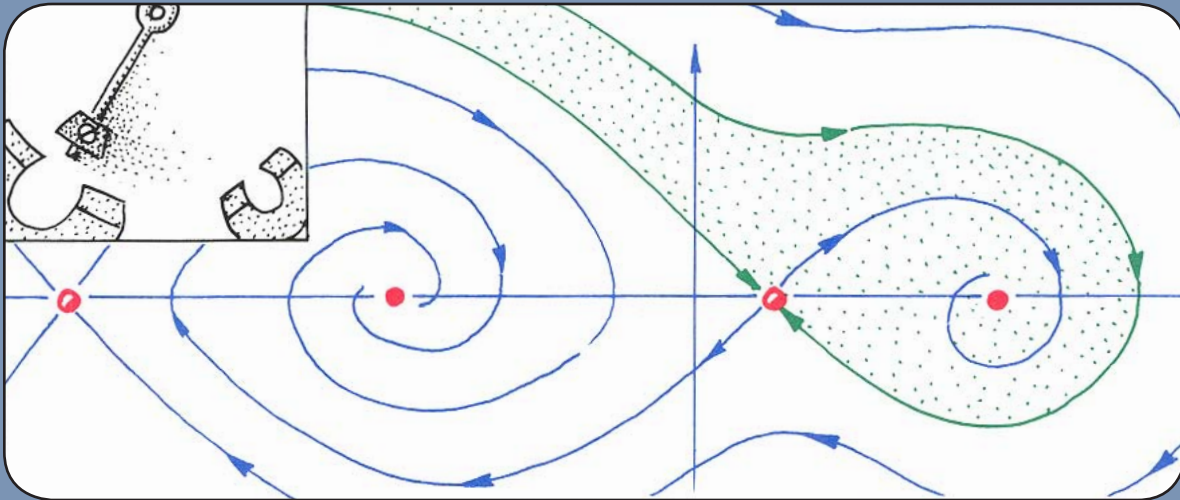
This is a saddle connection in 2D. The dashed trajectory comprises half of the outset of the hyperbolic saddle point on the left, its donor. Simultaneously, it is half of the inset of the hyperbolic saddle point on the right, its receptor. As this system contains a saddle connection, it does not satisfy G3.



11.3.2.

This system has no saddle connection. The outset of the saddle points on the left consists of two trajectories, which go to attractors (not shown). The inset of the saddle point on the right consists of two trajectories, which come from repellers (not shown). One of the trajectories leaving the left saddle narrowly misses one of the trajectories approaching the saddle on the right. This portrait is obtained from the preceding one by a slight perturbation.

Property G3 is a global property. It requires, in two dimensions, that each saddle outset avoid coinciding with any saddle inset.



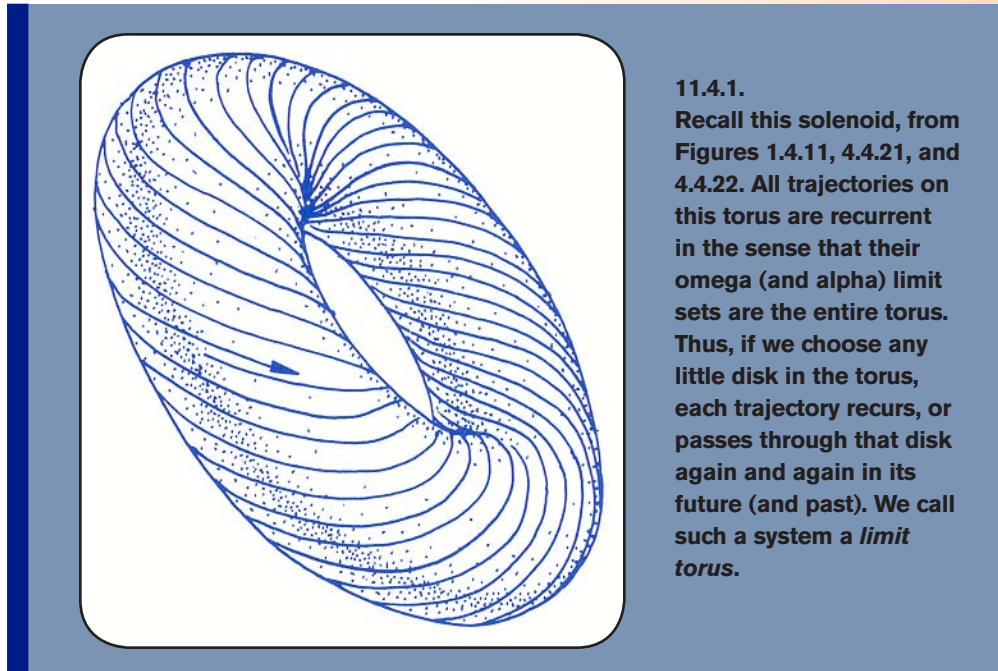
11.3.3.

The magnetic pendulum is a global system satisfying property G3. All four saddle outset trajectories successfully avoid all four saddle inset trajectories. (See Figure 2.1.22.)

11.4.

Properties G4 and F

Another generic property, G4, will be described in Chapter 15. It was originally formulated by Peixoto, in its oriented, two-dimensional version: The system has *no nontrivial recurrence*. Here is the main example of nontrivial recurrence.



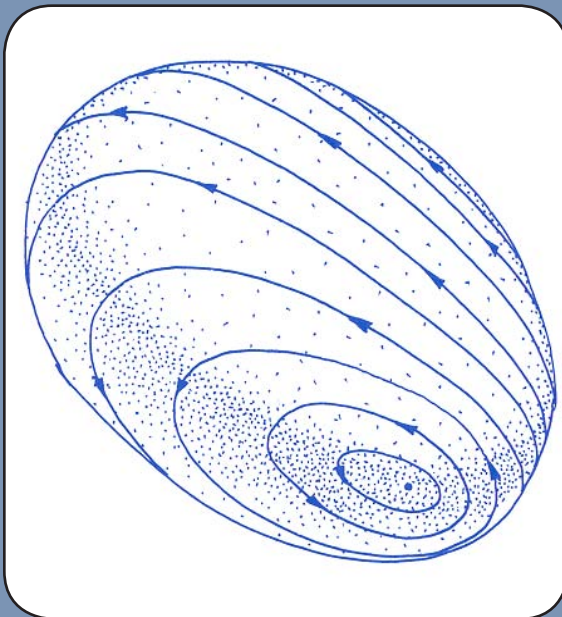
11.4.1.

Recall this solenoid, from Figures 1.4.11, 4.4.21, and 4.4.22. All trajectories on this torus are recurrent in the sense that their omega (and alpha) limit sets are the entire torus. Thus, if we choose any little disk in the torus, each trajectory recurs, or passes through that disk again and again in its future (and past). We call such a system a *limit torus*.

In other words, a limit torus is *topologically transitive*, as described in Figure 9.2.11. It shares this property of all the known chaotic attractors and limit sets. But, it occurs in two-dimensional systems, while chaotic sets do not. So in 2D, the toroidal solenoid is the main example of nontrivial recurrence, while in 3D the situation is much more complicated.

Finally, there is one more generic property we must describe, one which is special to the 2D case.

A dynamical system has *property F* if it has only a finite number of limit sets. In the 2D context, limit sets must be limit points, limit cycles or limit tori. This is a classical result of two-dimensional dynamic systems theory, known as the *Poincaré-Bendixson theorem*. Thus a 2D system satisfying G4 (no limit tori) will also satisfy property F if it has only a finite number of limit points and only a finite number of limit cycles.



11.4.2. Here is a 2D system violating property F. It has a center: an infinite number of limit cycles, arranged as concentric cycles around a limit point. See Figures 2.1.18, 2.2.3, and 2.2.5 for examples.

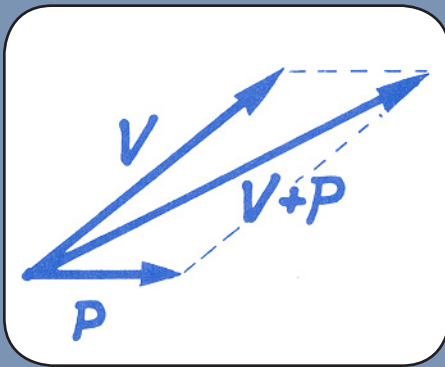
These properties G1, G2, G3, G4 and F, were all introduced by Andronov, de Baggis, and Peixoto in their historical works on structural stability in 2D. We now turn to that subject.

12 Structural Stability

In the applications of dynamics in various fields, the dynamics - that is, the actual vectorfield - can never be specified exactly. In fact, outside of a few cases in theoretical physics, one basically makes a rough guess. The mathematical theory of dynamical systems might be useful anyway, if it can describe features of the phase portrait that persist when the vectorfield is allowed to move around. This idea, now called *structural stability*, emerged early in the history of dynamics.

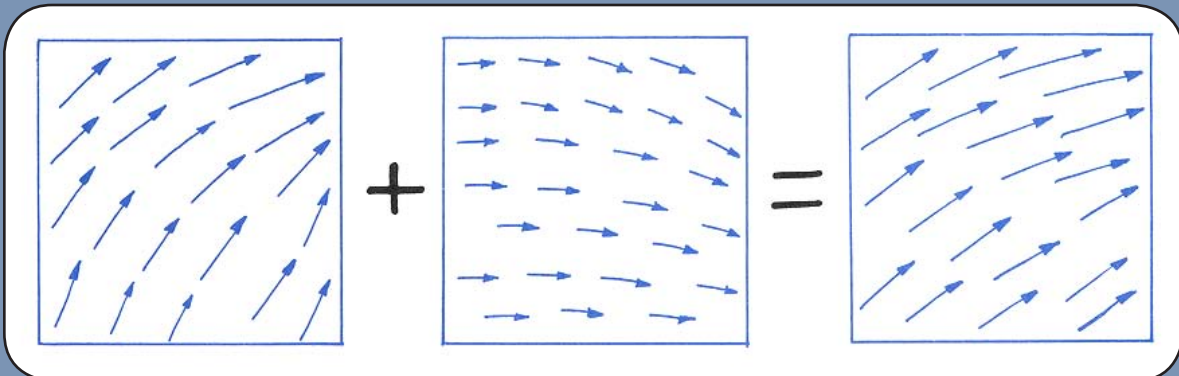
12.1. Stability Concepts

The idea of structural stability seems to have appeared first in the 1930's, in the writings of Andronov and collaborators, in Russia. It was introduced to North America by Lefshetz, the great topologist, and has played a central role in the development of the subject ever since.



12.1.1.

The criteria for structural stability rely upon two supplementary notions: perturbation and topological equivalence. A *perturbation* of a vectorfield means the addition to it of a relatively small vectorfield, frequently unspecified. Here we show the effect of a perturbation, at a single point in the state space.

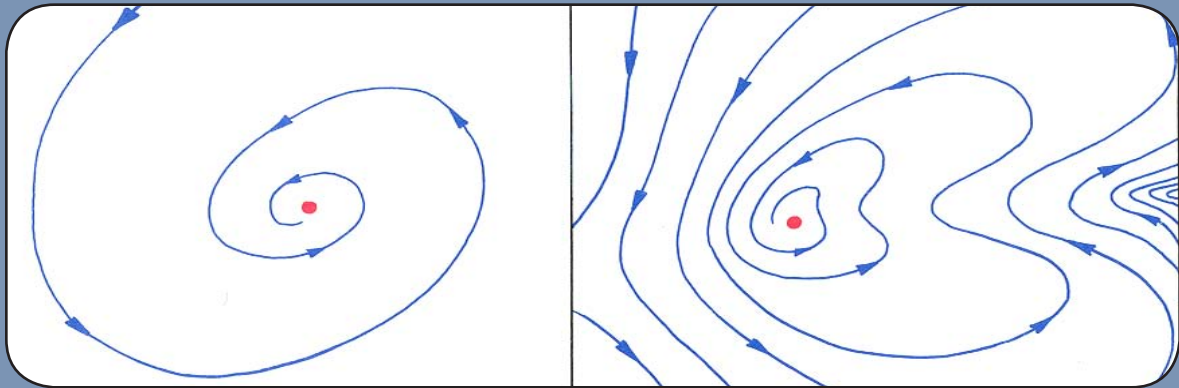


12.1.2.

Here we show the effect of a global perturbation. The perturbation is itself a vectorfield, as shown here. The effect of adding this perturbing vectorfield to the original one (on the left) is to modify it at every point in the state space.

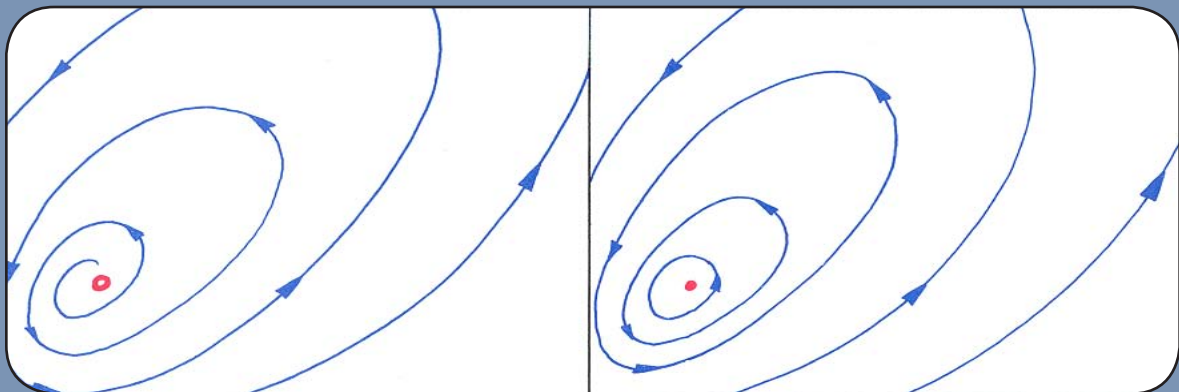
Topological equivalence, or synonymously, *topological conjugacy*, of two phase portraits, means there is a *homeomorphism* of the state space, or continuous “rubber sheet” deformation, which maps one of the portraits to the other, preserving the arrow of time on each trajectory.

Here are some topologically equivalent portraits in two dimensions.



12.1.3.

These two point attractors are *topologically equivalent*. A homeomorphism can deform one into the other, preserving the integral curves.



12.1.4.

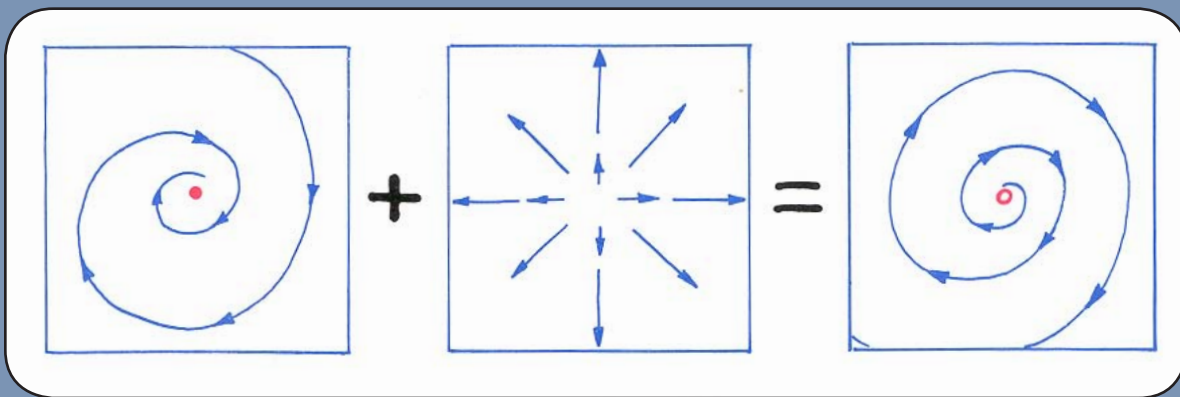
But the point repeller on the left is *not topologically equivalent* to the center on the right. A homeomorphism cannot map a spiral onto a circle.

To be faithful to the theory in higher dimensions, we will need also the concept of *epsilon equivalence*. This is a topological equivalence of dynamical systems, in which the deforming homeomorphism only stretches or slides the state space a small amount (measured by epsilon). Likewise, in the spirit of classical mathematics, we will call a perturbation a *delta perturbation*, if it is small (measured by delta).

Now we use both of these stability concepts, delta perturbation and epsilon equivalence, to introduce the idea of structural stability.

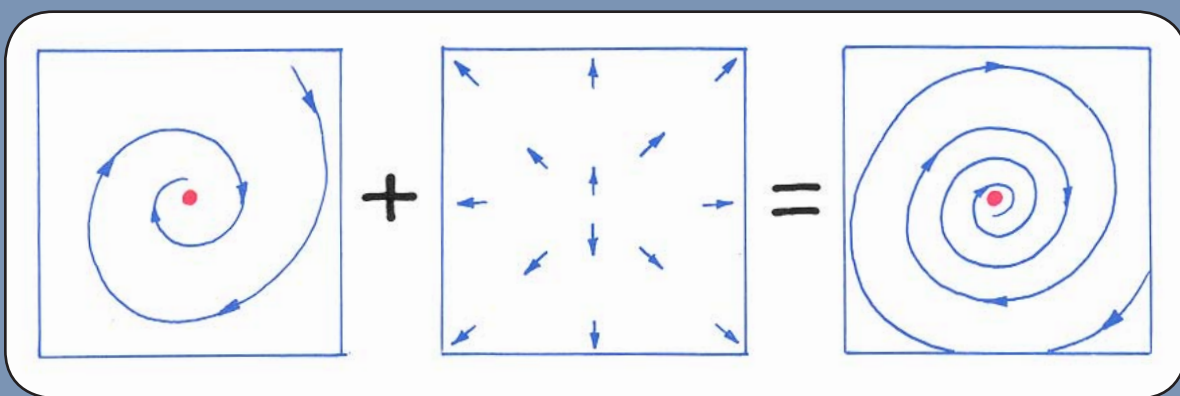
A vectorfield has the property of *structural stability* if (choosing epsilon) all delta perturbations of it (sufficiently small) have epsilon equivalent phase portraits.

Here is a simple example.



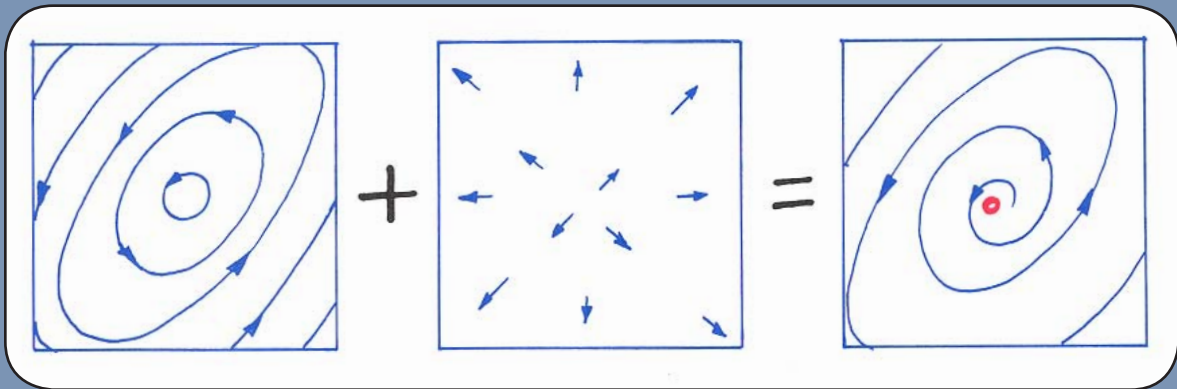
12.1.5.

Imagine a system with a spiral attractor which attracts *very weakly*. By adding a medium-sized perturbation pointing outward, we might be able to change it into a spiral repeller.



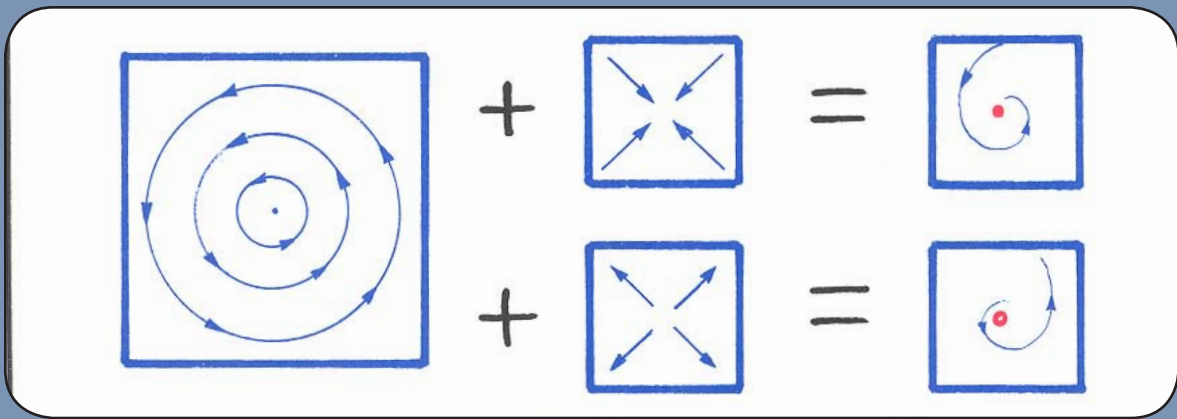
12.1.6.

But adding a delta perturbation pointing outward (sufficiently feeble) may make our attractor weaker, but it still attracts. It is topologically (in fact, epsilon) equivalent to the original system. This is an example of a *structurally stable system*.



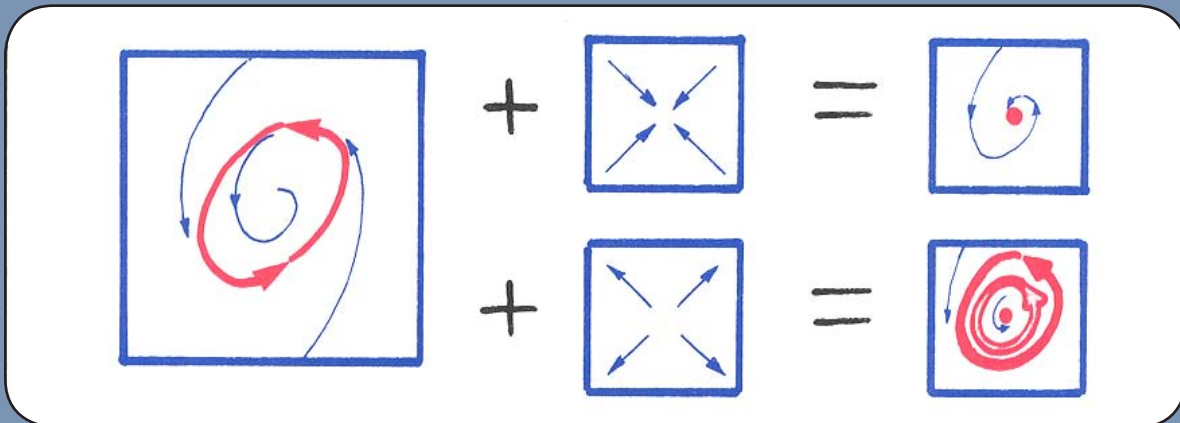
12.1.7.

Now consider this dynamical system, a center. The addition of a delta perturbation, pointing outward (no matter how weak) results in a point repeller, which is not topologically equivalent to the center. This is a primary example of a *structurally unstable system*.



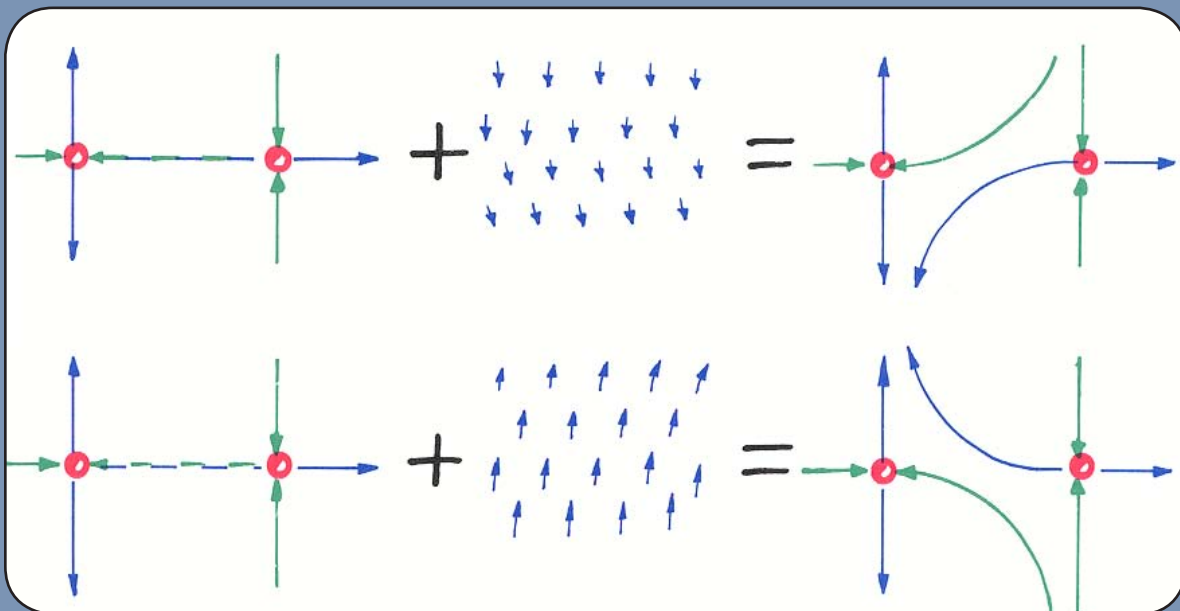
12.1.8.

In fact, the center may be perturbed into either a point repeller or a point attractor, depending on the inclination of the perturbation.



12.1.9.

On the other hand, this portrait is structurally stable. The inclination of the perturbation may make the periodic attractor smaller or larger, but the perturbed portraits are all topologically equivalent.



12.1.10.

Here is another important example. Consider a system with a saddle connection, as in Figure 11.3.1. Adding a delta perturbation pointing downward (or upward), we destroy the saddle connection. The resulting phase portrait is not topologically equivalent. These two examples illustrate all basic types of structural instability in 2D.

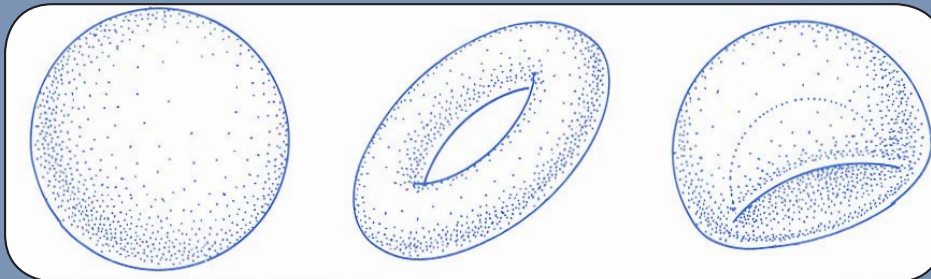
12.2.

Peixoto's Theorem

Now we go on to Peixoto's historic theorem, relating the generic properties of the preceding chapter to structural stability in 2D.

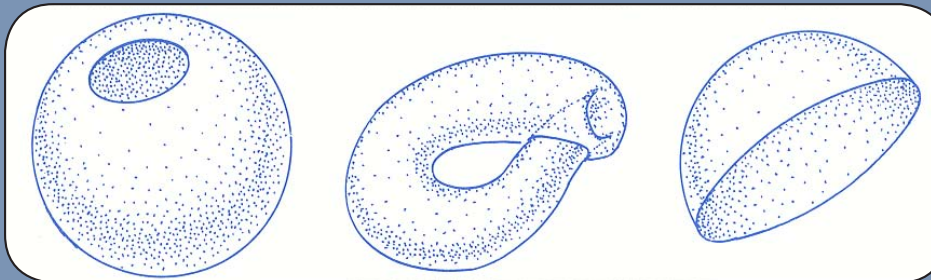
A watershed in the history of dynamics, Peixoto's work brought together different topology and classical dynamics, ushering in a new age of mathematical dynamics. The attempts to extend his 2D results to 3D and beyond characterized the early days of this new approach, in the 1960's.

Peixoto's result applies to a very restricted class of state spaces, called *compact, orientable surfaces*. We start with these.



12.2.1.

A state space is called *compact* if it can be described as a surface (of whatever dimension) in a finite region of Euclidean space (of a higher dimension) which is a closed set. Here, *closed* means no holes or loose ends. A surface is *orientable* if it has two sides inside and outside. The surfaces shown here are all compact and orientable. *All state spaces in this section will be assumed to be compact, orientable 2D surfaces.*



12.2.2.

This excludes a sphere with a hole, the Klein bottle, the upper hemisphere, and so on. Nevertheless, the theory described here has been extended to many of these spaces as well.¹

Now we are ready to state Peixoto's theorem. We will use *property S* as a synonym for structural stability.

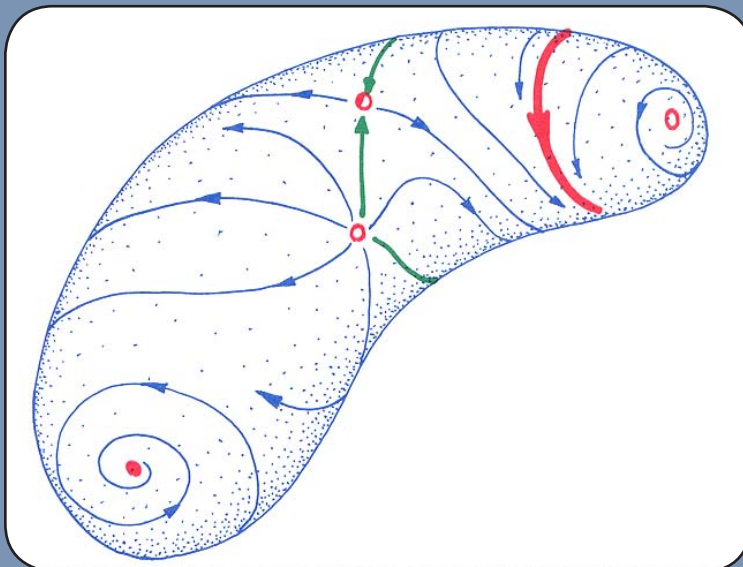
Peixoto's theorem: among all smooth dynamical systems on a compact, orientable surface,

- A. properties G1, G2, G3, G4 and F are generic,**
- B. property S is equivalent to these properties (A), and**
- C. property S is generic.**

Clearly C follows from A and B, but this is the most exciting aspect of the theorem. For it says that in applications, this strong kind of stability is to be expected as the typical case, while structural instability is pathological.

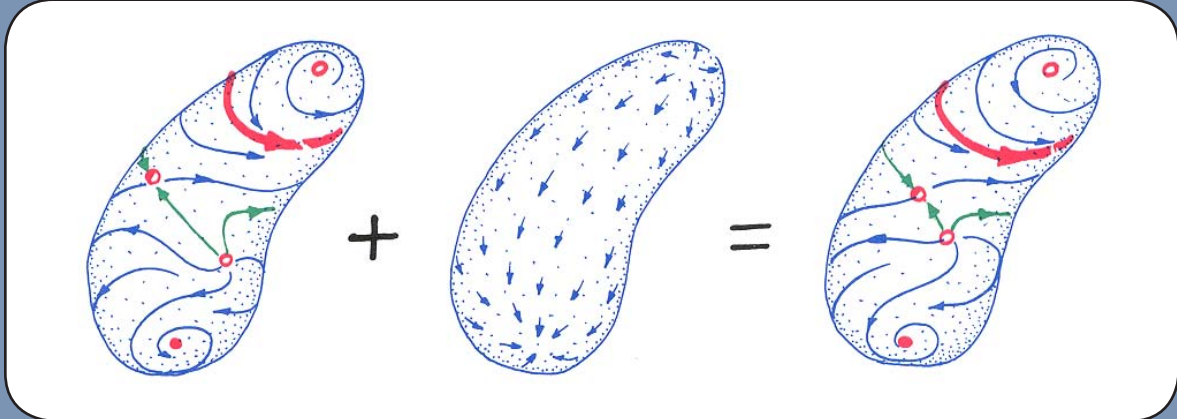
Here, *generic* is a technical term, which we translate as *typical* sometimes. However, the atypical cases (especially those in which property G4 is violated) are so frequently observed in experiments that we should use *weakly generic* as the technical term, and understand *typical* as meaning *slightly more probable* than the exceptional cases. The reason for this paradox is that the Kronecker (solenoidal) flows on the torus (See Part One) occur for a *fat fractal* or *thick Cantor* set of leaves in Thom's big picture.² This will be explained in more detail in Part Four.

Part A was generalized promptly to higher dimensions, except for the genericity of E, which failed, along with C. Part B also was generalized, by Smale and Palis. More about this in later chapters.



12.2.3.
Here is a system exhibiting G1-G4 plus F, and thus S.

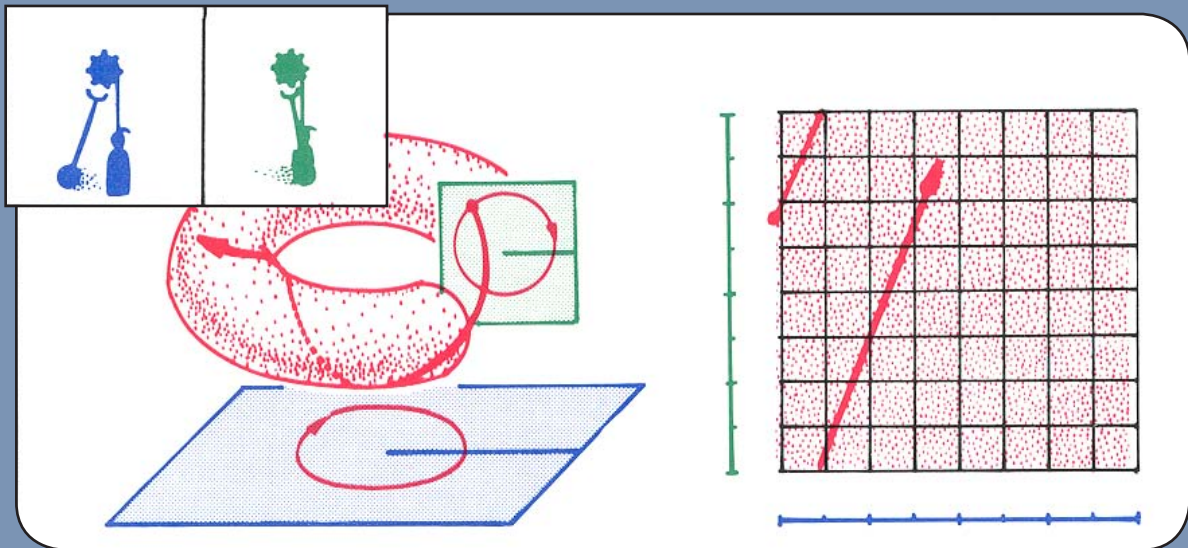
Peixoto's proof is outlined in the next section. Here, we give some examples.



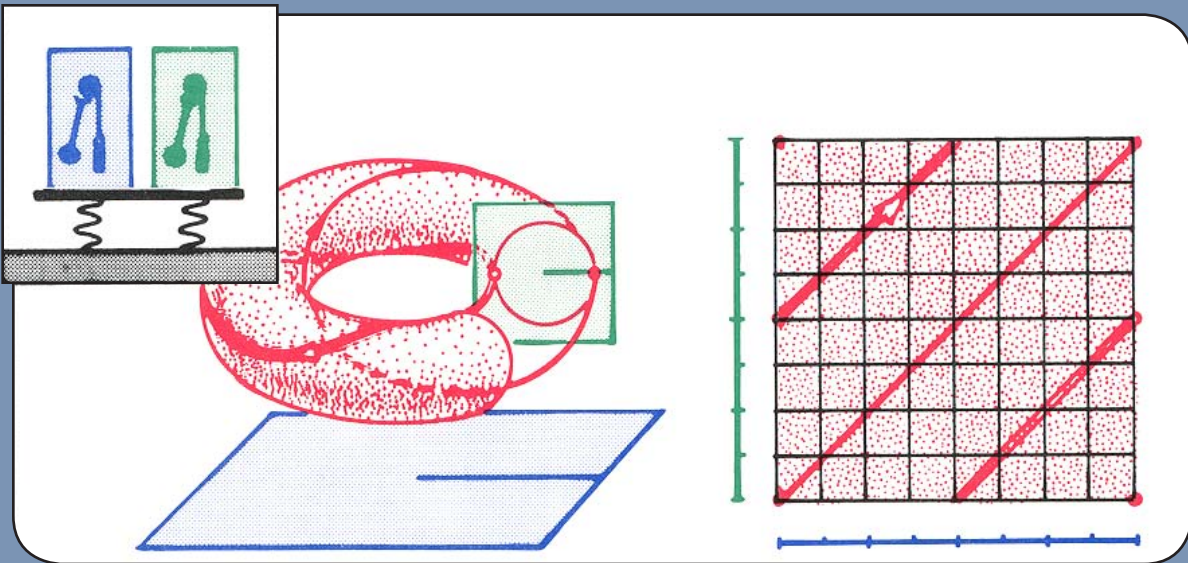
12.2.4.
A delta perturbation yields an epsilon equivalent portrait.

But Peixoto's theorem says more: saddle connections are structurally unstable, as we saw in Section 11.3.

Peixoto's theorem says still more: nontrivial recurrence (solenoidal flow on a torus) can be perturbed (in Thom's big picture) into a structurally stable system.



12.2.5.
Here is a torus with a solenoidal flow. It violates property G4, so by part B of Peixoto's theorem, it is not structurally stable. By part C, it can be changed to an S system by a delta perturbation. *Warning:* This delta perturbation may be rare, or hard to find, since it belongs to the complement of a thick Cantor set, as explained in Part Four.



12.2.6.

This S system will not have any limit points or limit tori, but it must have limit sets. So, there are some limit cycles, braided around the torus. They occur in pairs, alternately attracting and repelling. The implications for frequency entrainment of coupled oscillators are discussed in detail in Chapter 5.

12.3.

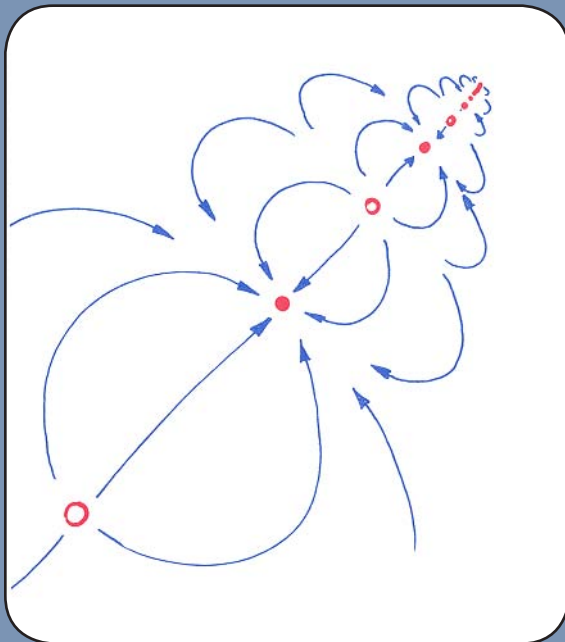
Peixoto's Proof

We break the proof into five steps:

1. G1 implies FP (finite number of limit points).
2. G2 implies FC (finite number of limit cycles).
3. G4 implies no limit tori.

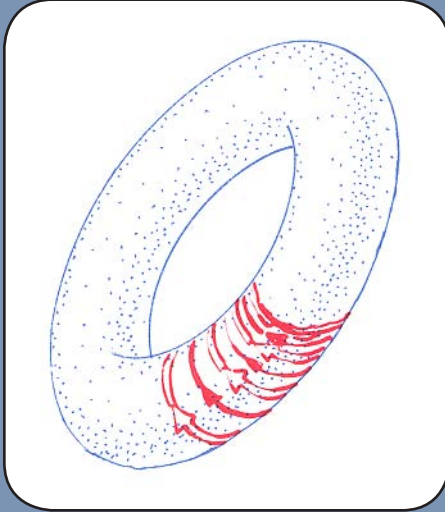
Therefore, G1, G2, and G4 imply E.

4. G1, G2, G3, and G4 (and hence F) imply S.
5. S implies G1, G2, G3, and G4.



12.3.1.

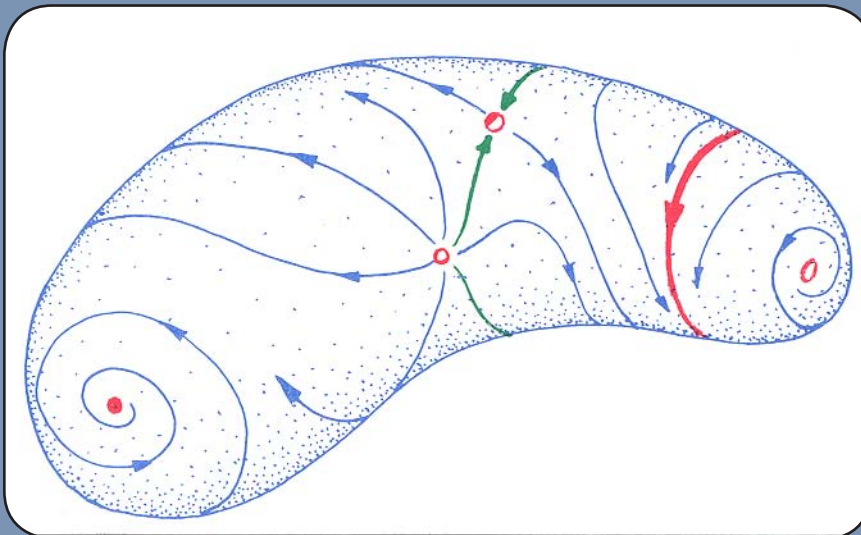
Step 1: Generic property G1 implies there are only a finite number of limit points. For in the compact state space, an infinite number of critical points would have to contain a convergent sequence as shown here. And the critical point at the end of the sequence will have to violate G1.



12.3.2.

Similarly, generic property G2 implies there are only a finite number of limit cycles. This is special to two dimensions, where an infinite number of limit cycles would be forced to “pile up.” That is, either they must converge to a limit cycle, as shown here (violating the generic condition G2 - hyperbolic limit cycles), or they must accumulate at a limit point (violating G1 - hyperbolic limit points).

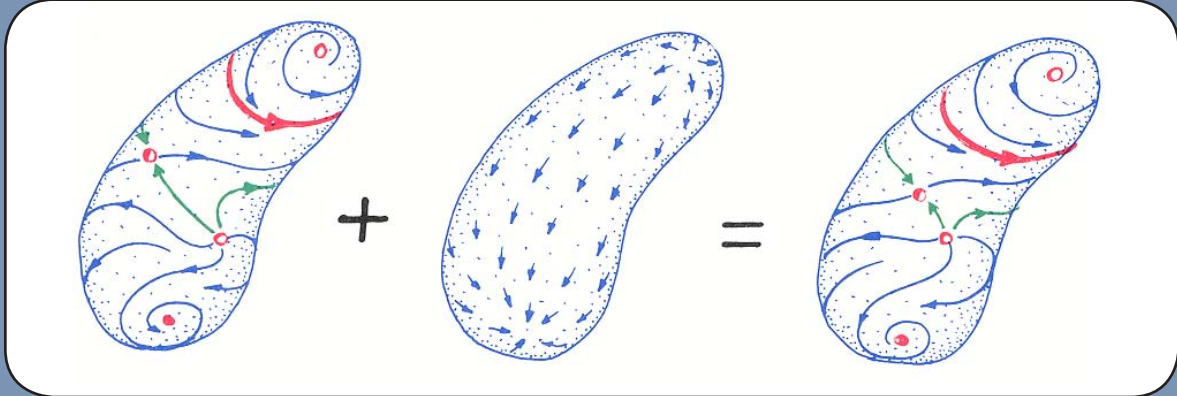
The proof of this step used topology and calculus, and is not terribly difficult.



12.3.3.

Step 2: If the system is generic (G1, G2, G3, and G4), then it has only a finite number of limit points, a finite number of limit cycles, and no other limit sets. This is called *property F*. Further, they are all hyperbolic, and there are no saddle connections. Here is a typical portrait of this type.

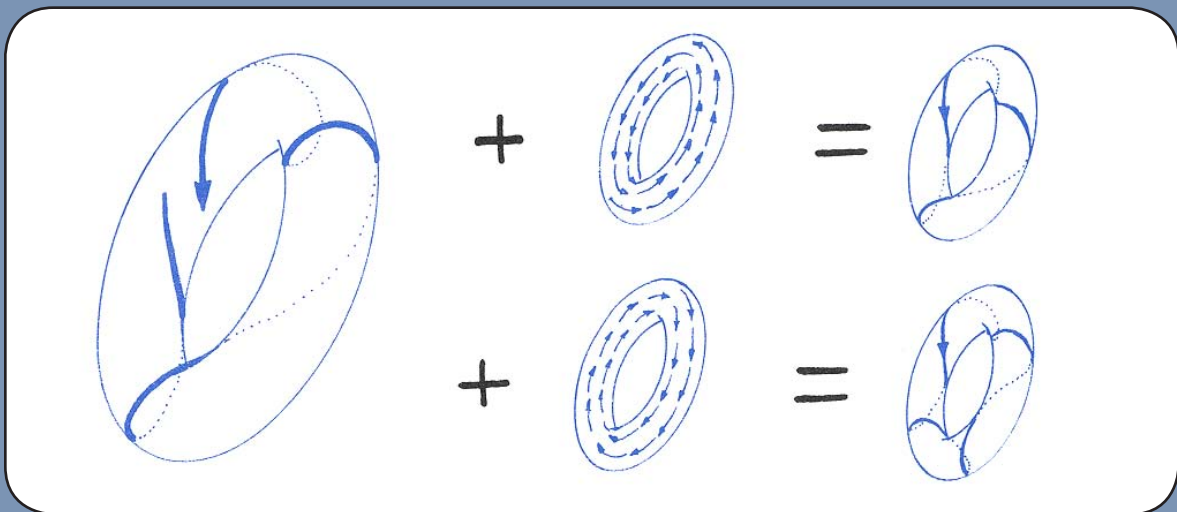
The proof of Step 2 requires the infamous *Closing Lemma*. This is used to eliminate the possibility of a toroidal limit set. First proved in the present context by Peixoto, it has been wonderfully generalized by Pugh and Robinson.³



12.3.4.

Step 3: These generic properties (G_1 , G_2 , G_3 , G_4 , and F) ensure structural stability. An arbitrary small perturbation of the portrait shown in the preceding panel produces an equivalent portrait.

The proof of this step requires the actual construction of a topological deformation from the original portrait to the perturbed one, but is not too difficult.



12.3.5.

Step 4: Structural stability ensures the generic properties (G_1 , G_2 , G_3 , G_4 , and necessarily, F). The preceding section gives examples showing how structural stability ensures the first three properties. Here is an example showing how G_4 is ensured. The center portrait has a toroidal limit set with no limit cycles or limit points. The only limit set is the entire state space, a torus. Small perturbation can produce the two portraits shown below, which are not topologically equivalent. The difficult Closing Lemma is used in this step also. *Warning:* Again, the perturbations producing these structurally stable (braided) flows from the solenoidal flow can be rare, or hard to find, because of belonging to the complement of a thick Cantor set.

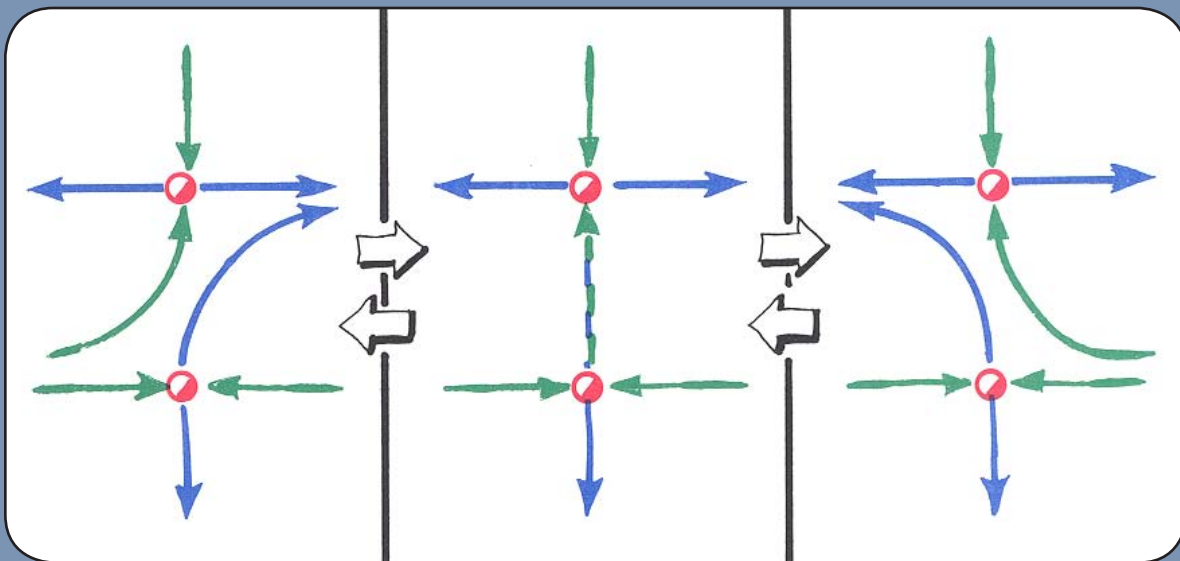
13 Heteroclinic Tangles

Limit points and cycles of saddle type may be distributed throughout the state space. Each has insets and outlets, which wander around near each other. Intersections are not unlikely. These, called *saddle connections*, consist of trajectories of the dynamical system that lead from one saddle (called the *donor*) to another (the *receptor*). This connecting curve is called a *heteroclinic trajectory* if the donor and receptor saddles are different, or a *homoclinic trajectory* if they are the same. This chapter is devoted to saddle connections by heteroclinic trajectories which satisfy the generic property G_3 , or *transversality*. The *homoclinic case* (a trajectory connects a saddle to itself) is described in the next chapter.

In state spaces of one dimension, there are no saddles. In two dimensions, there are generic saddle points with one-dimensional insets and outlets. In the three-dimensional cases, there are generic saddle points and cycles, of which the insets and outlets may be surfaces. In this chapter, we will describe all of the transverse heteroclinic saddle connections in two and three dimensions: limit point to limit point, limit point to limit cycle, and cycle to cycle.

13.1. Point to Point

First, consider phase portraits in the plane, with two hyperbolic limit points of saddle type. The insets of each are curves, likewise their outlets. These curves are trajectories of the dynamical system.

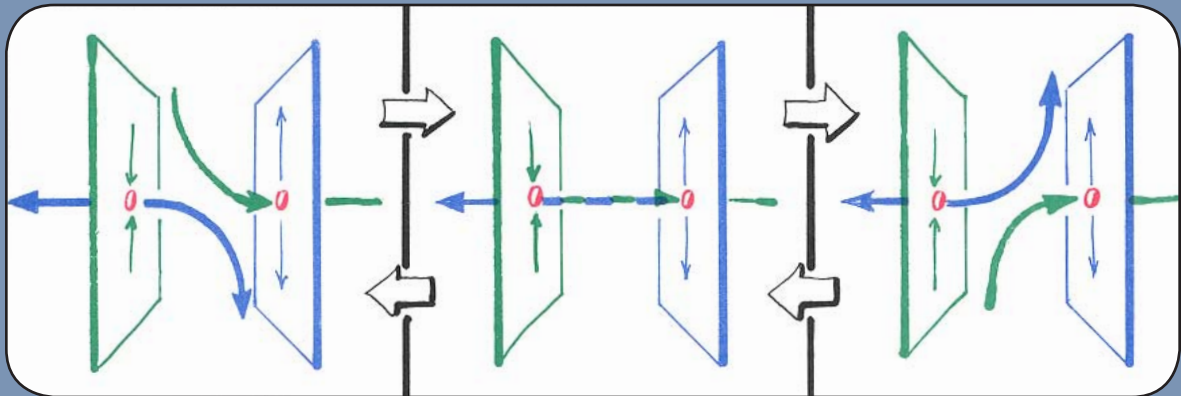


13.1.1.

These three phase portraits each have two hyperbolic limit points of saddle type. The end ones have no saddle connection, while the one in the center has a single heteroclinic trajectory. The sequence has occurred previously in Part One, under the name *saddle switching*. It represents the actual coincidence of the outset from the left saddle and the inset to the one on the right. The transverse intersection of two curves in the plane must be in isolated points. Therefore, this intersection is not transverse. It is a *nongeneric* saddle connection. There are not transverse saddle connections in the two-dimensional case.

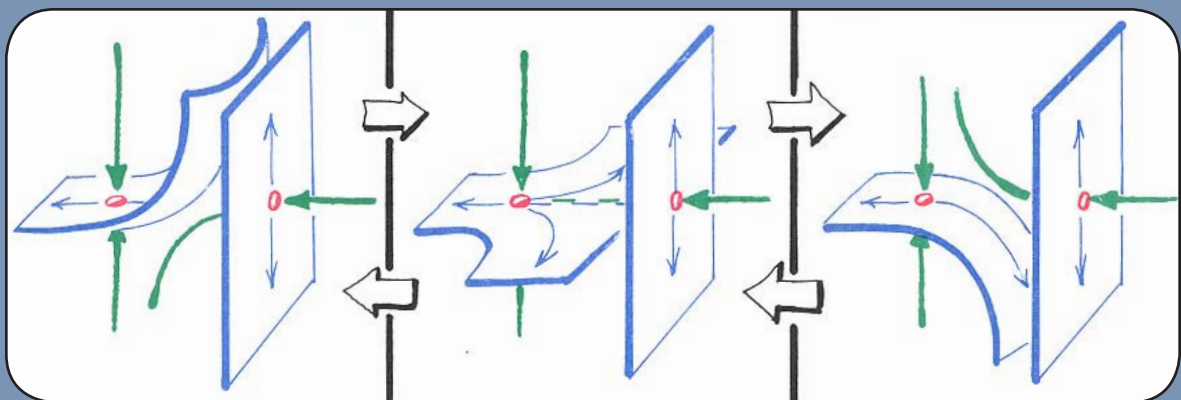
And now, on to two hyperbolic saddle points in 3D.

In the three-dimensional case, there are several possibilities. There are two types of topologically distinct hyperbolic saddle points: index 1 (inset two-dimensional) and index 2 (inset one-dimensional, outset two-dimensional). Each can be a donor or receptor of a saddle connection. But *transverse* saddle connections, in 3D, only occur between two-dimensional outlets and two-dimensional insets. Such an intersection consists of a single curve, a trajectory.



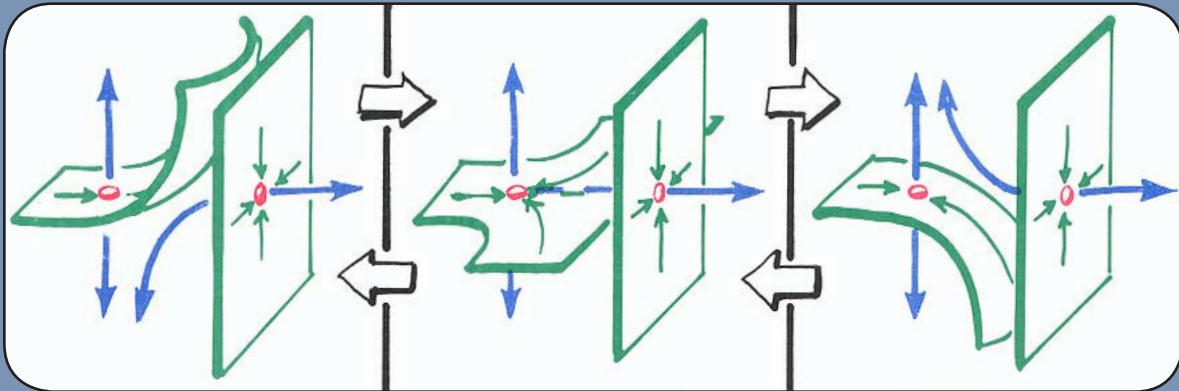
13.1.2.

A saddle point of index 1 cannot have a transverse connection to a saddle point of index 2, in three dimensions. Three closely related portraits are shown here, in analogy to saddle switching in the two-dimensional case. The one in the center has a nontransverse heteroclinic trajectory connecting the two saddle points.

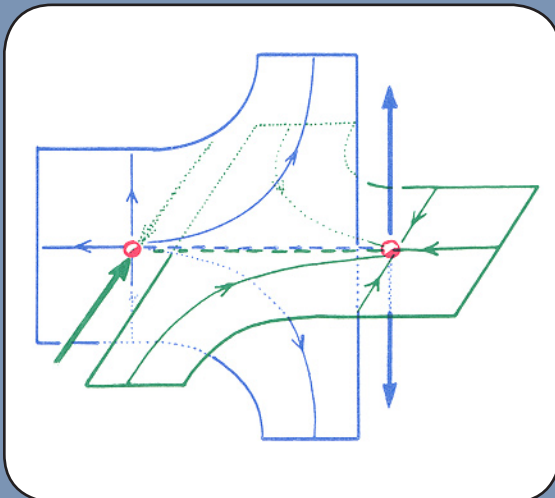


13.1.3.

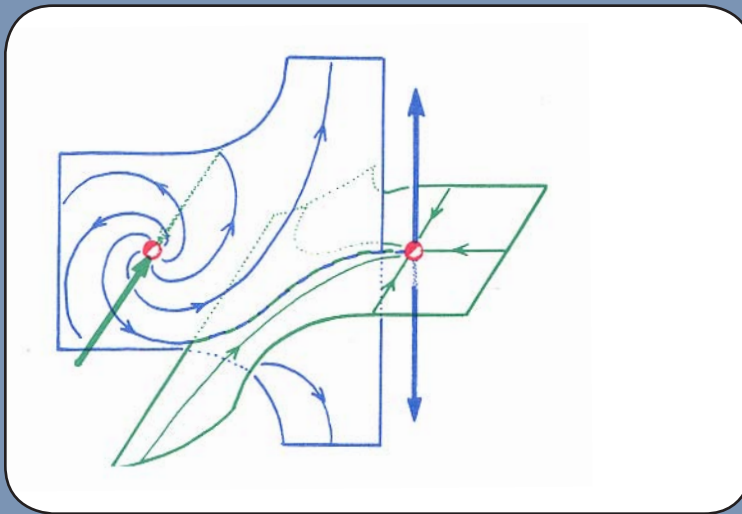
The next donor, a saddle point of index 2, cannot have a transverse connection to a saddle point of index 2 (same receptor as above), in three dimensions. Here again, three similar portraits are shown. The one in the center is an example of a nontransverse heteroclinic trajectory.



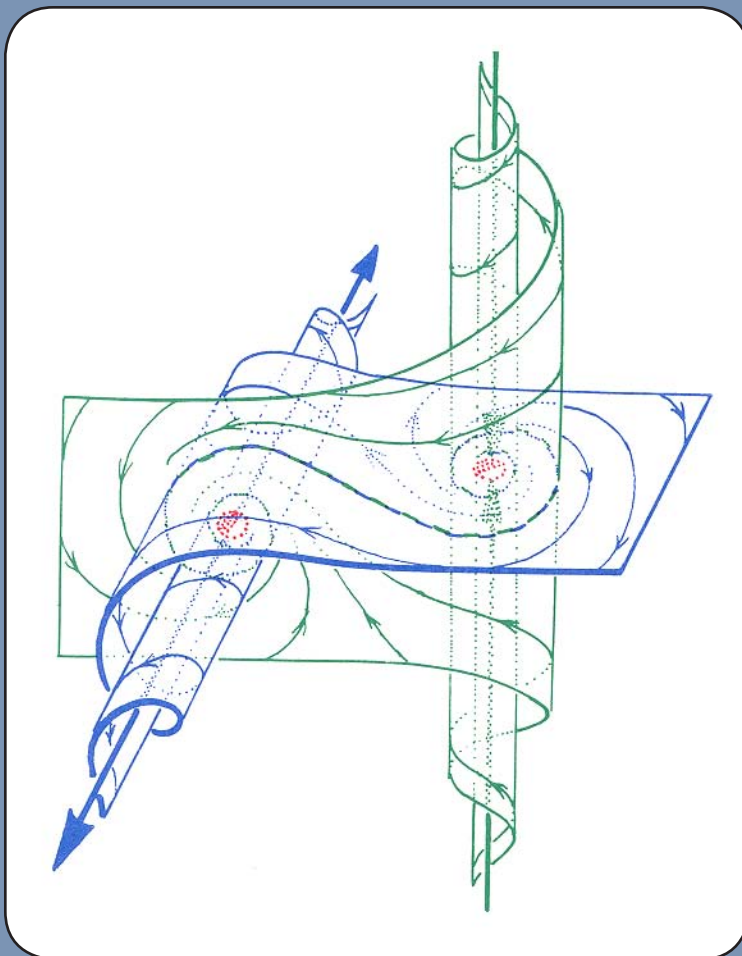
13.1.4. Transverse connection from a saddle point of index 1 to a saddle point of index 1 (like the case of index 1 to index 2, and index 2 to index 2, described above) cannot occur in three dimensions.



13.1.5. In this fourth case, a heteroclinic trajectory leads from a saddle point of index 2 to one of index 1. The outset of the donor and the inset of the receptor are both two-dimensional. Thus, a transverse intersection of them in a one-dimensional curve (necessarily a trajectory of the dynamical system) is possible. A nontransverse intersection along a heteroclinic trajectory is also possible - for example, the two surfaces could be tangent to each other, along their intersection. Here, the transverse case is illustrated. This is the only generic (transverse) connection between saddle points in three dimensions.



13.1.6.
The preceding illustration shows the transversely connected saddle points, assuming both are the radial (nonspiral) type. Here, the donor has been replaced by a spiral type. This is topologically equivalent to the preceding portrait.

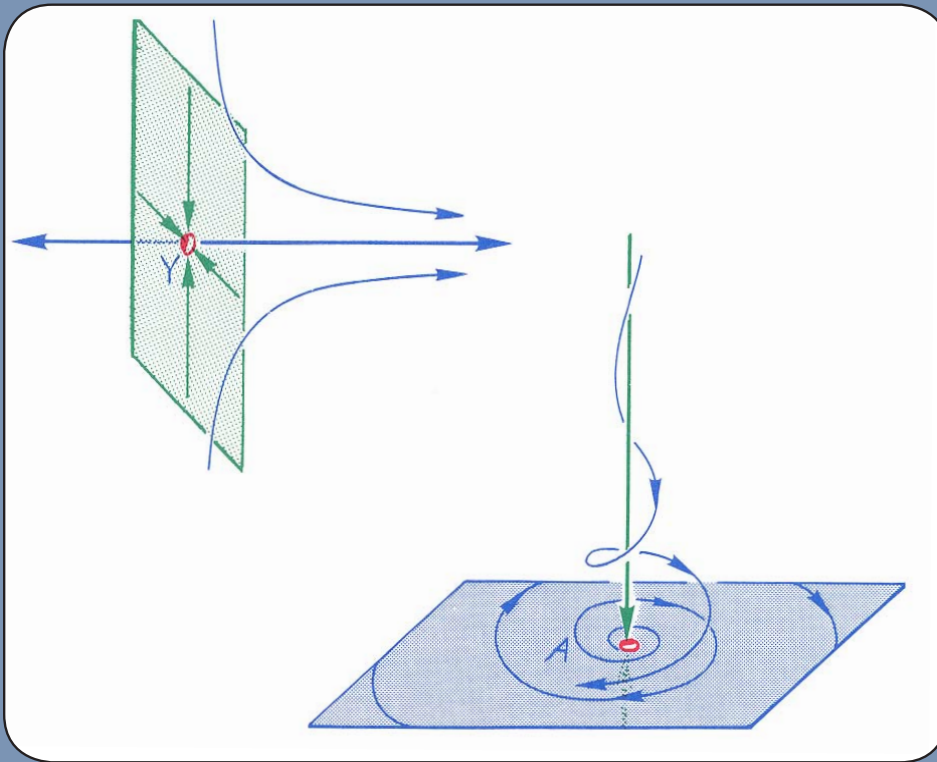


13.1.7.
In this example, both the donor and the receptor are of the spiral type. Again, this is topologically equivalent to the preceding portraits.

13.2.

Outsets of the Lorenz Mask

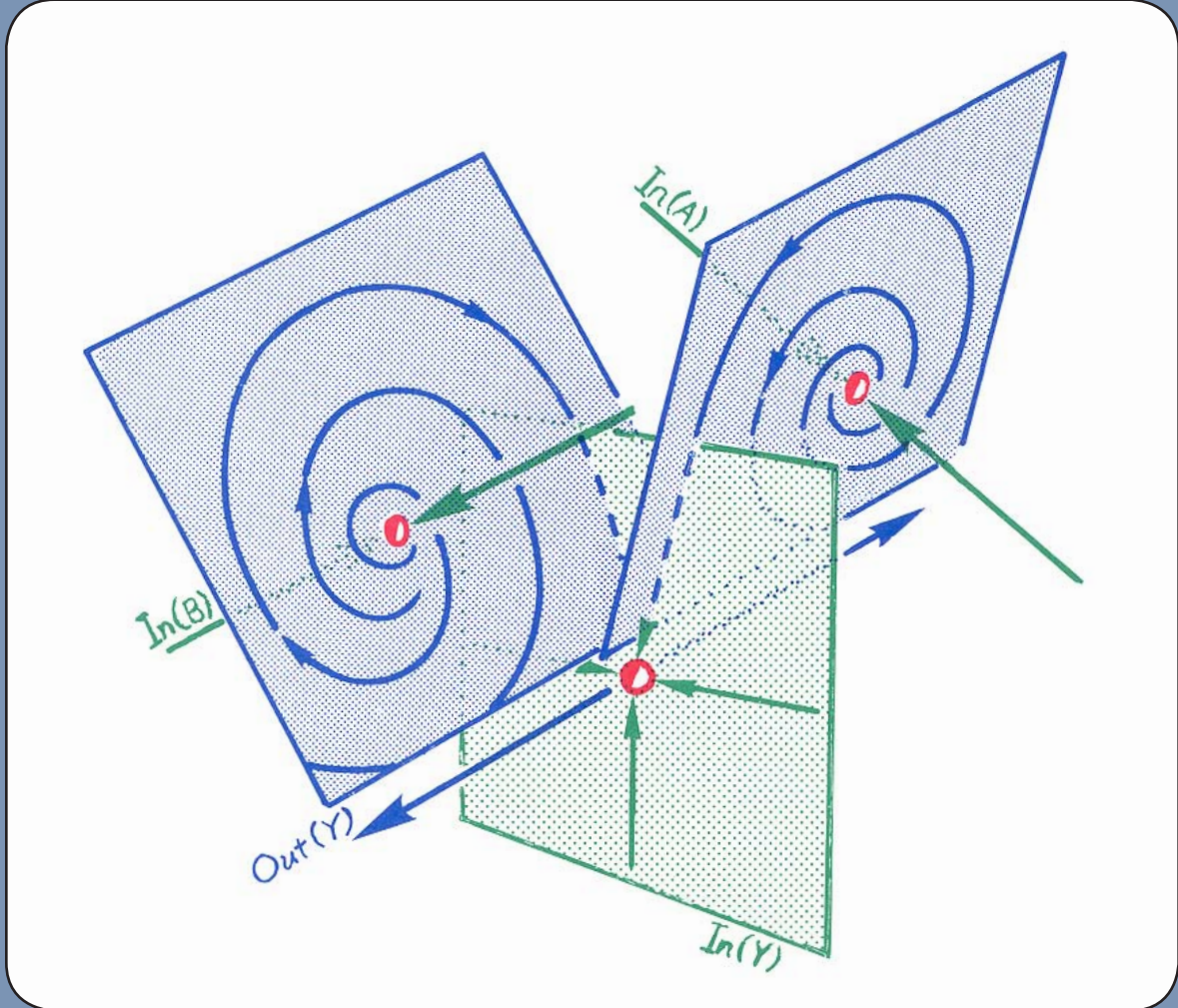
Recall the Lorenz mask, from Part Two. This was the first chaotic attractor to be firmly established in experimental dynamics. It is actually made of tangled outsets. Here, developed in stages, is the complex of point-to-point tangles found in the Lorenz system.¹ There is a radial saddle point of index 1 (the *receptor*) situated between two spiral saddle points of index 2 (the *donors*). The outset surfaces of the two donors are heteroclinically incident to the inset surface of the receptor.



13.2.1. Here are two saddle points, A and Y. They are hyperbolic, in three dimensions. One, A, has index 2, with spiral dynamics on its planar outset (shaded), $\text{Out}(A)$. The other, Y, has index 1, with nodal dynamics on its planar inset

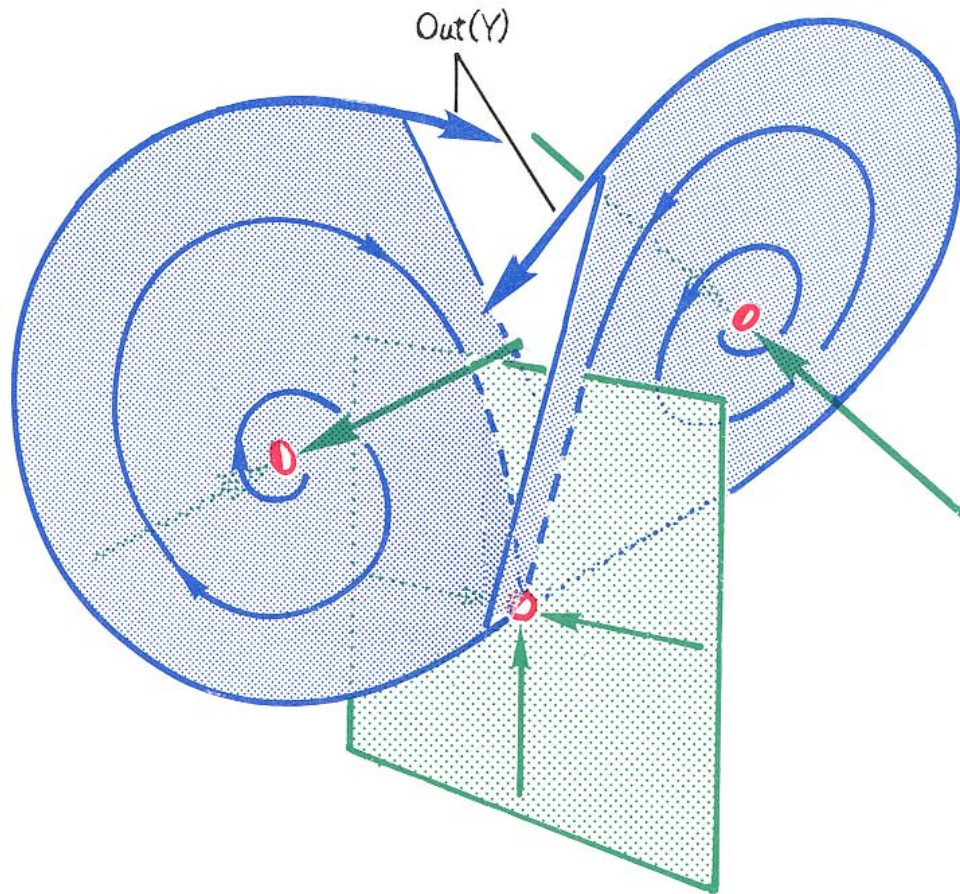
(dotted), $\text{In}(Y)$. The two outsets are attractive, as shown by the neighboring trajectories. As $\text{Out}(A)$ and $\text{In}(Y)$ are both two-dimensional, they could intersect transversely in three space. If they did, the transversal intersection would have to be a trajectory, called a *heteroclinic trajectory*.

Next, we will build up this complex, step by step.

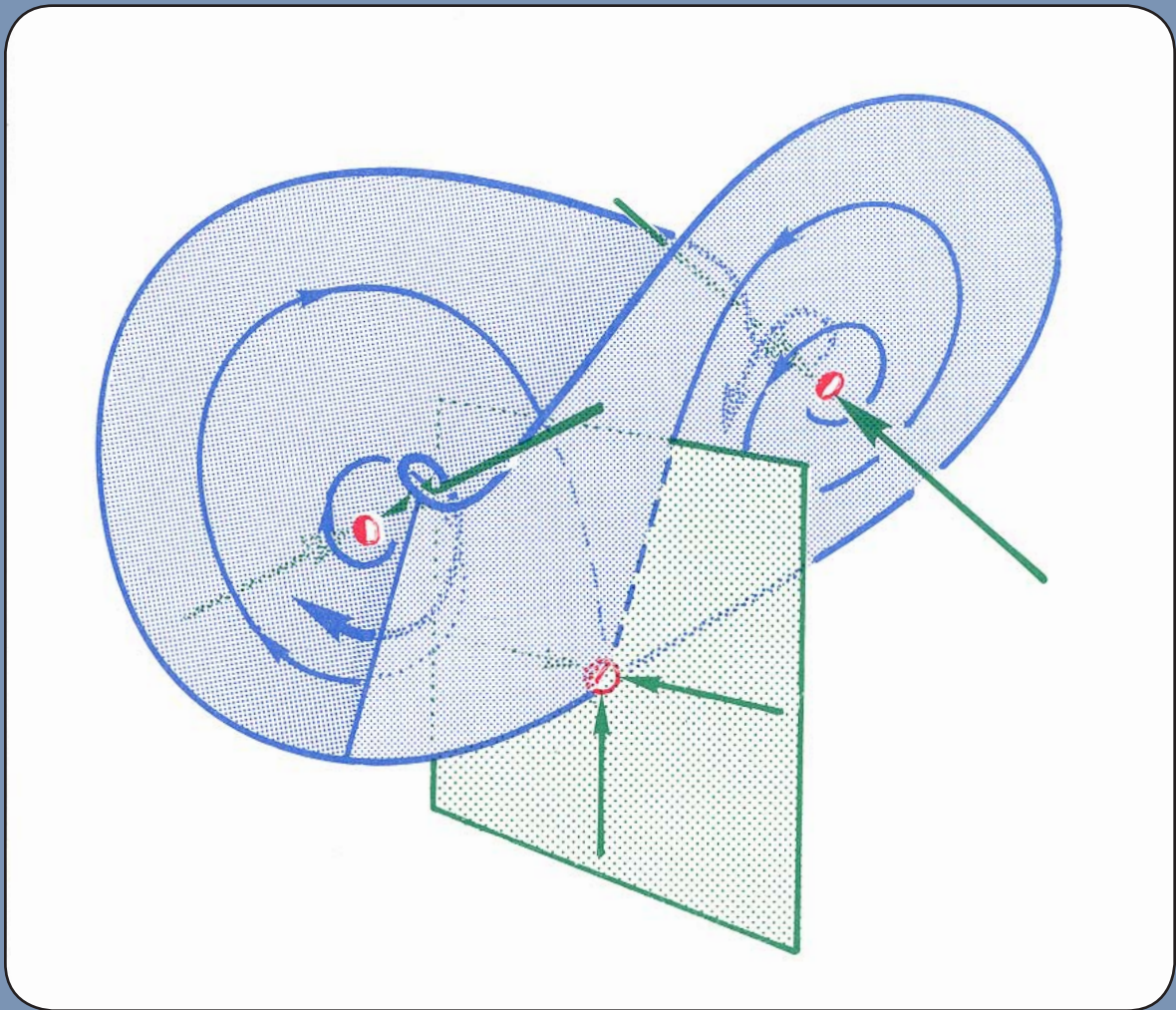


13.2.2.

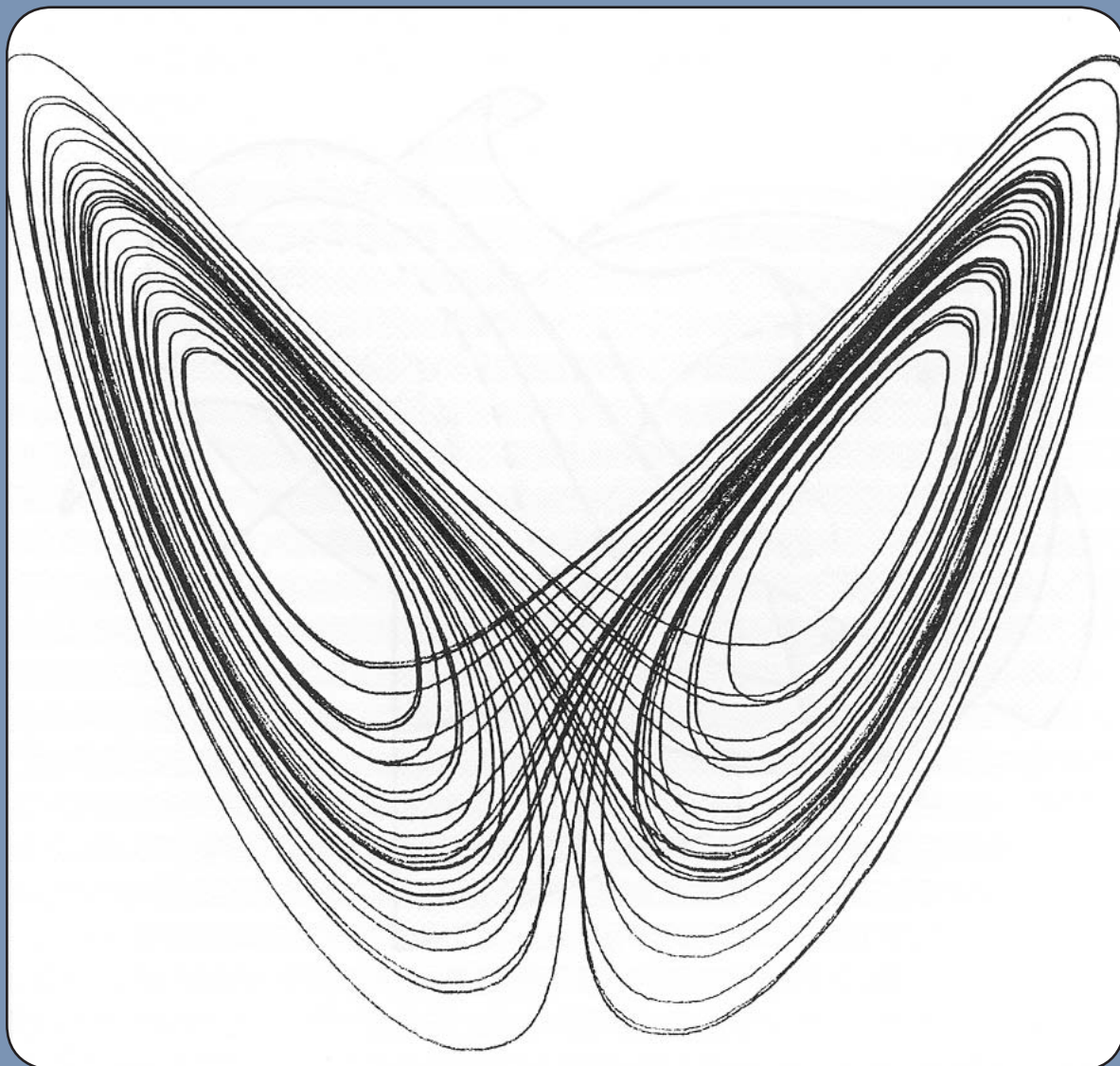
Adding another saddle point, B, essentially identical to A, we make a yoke like this. Both A and B are heteroclinic to Y. They are *transversely heteroclinic*, as the two planar insets (shaded) intersect the planar inset (dotted) transversely. There are *two heteroclinic trajectories* in this yoke. Note that the arriving insets are incident upon the departing inset, at Y. We call this a *neat yoke*. Next, we will see where these insets end up.

**13.2.3.**

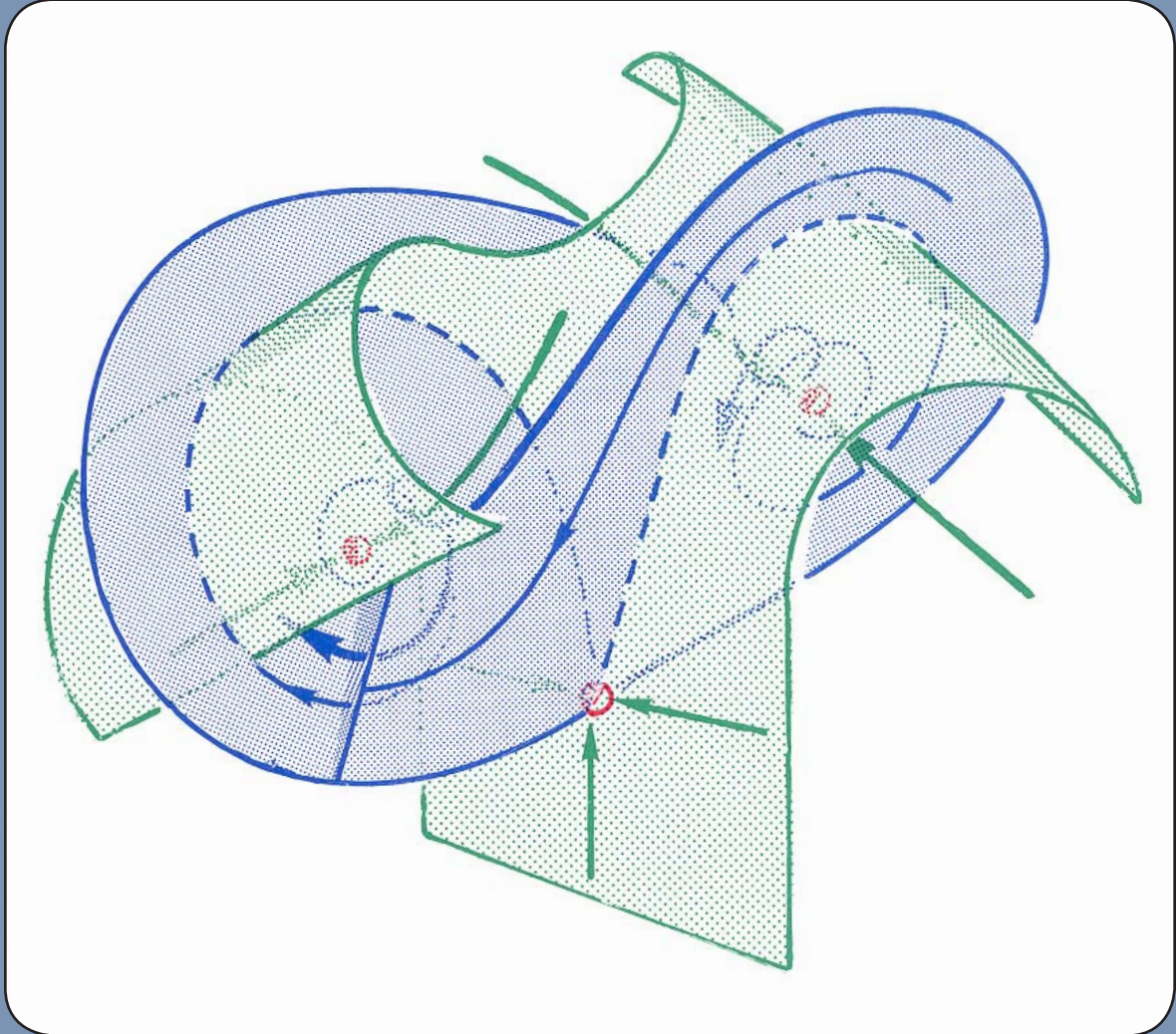
As the arriving outsets, Out (A) and Out (B), both have spiral dynamics, the departing outset that bounds them, Out (Y), swirls around and reinserts, as shown here. It cannot go off to infinity, as the Lorenz system has a repeller at infinity.

**13.2.4.**

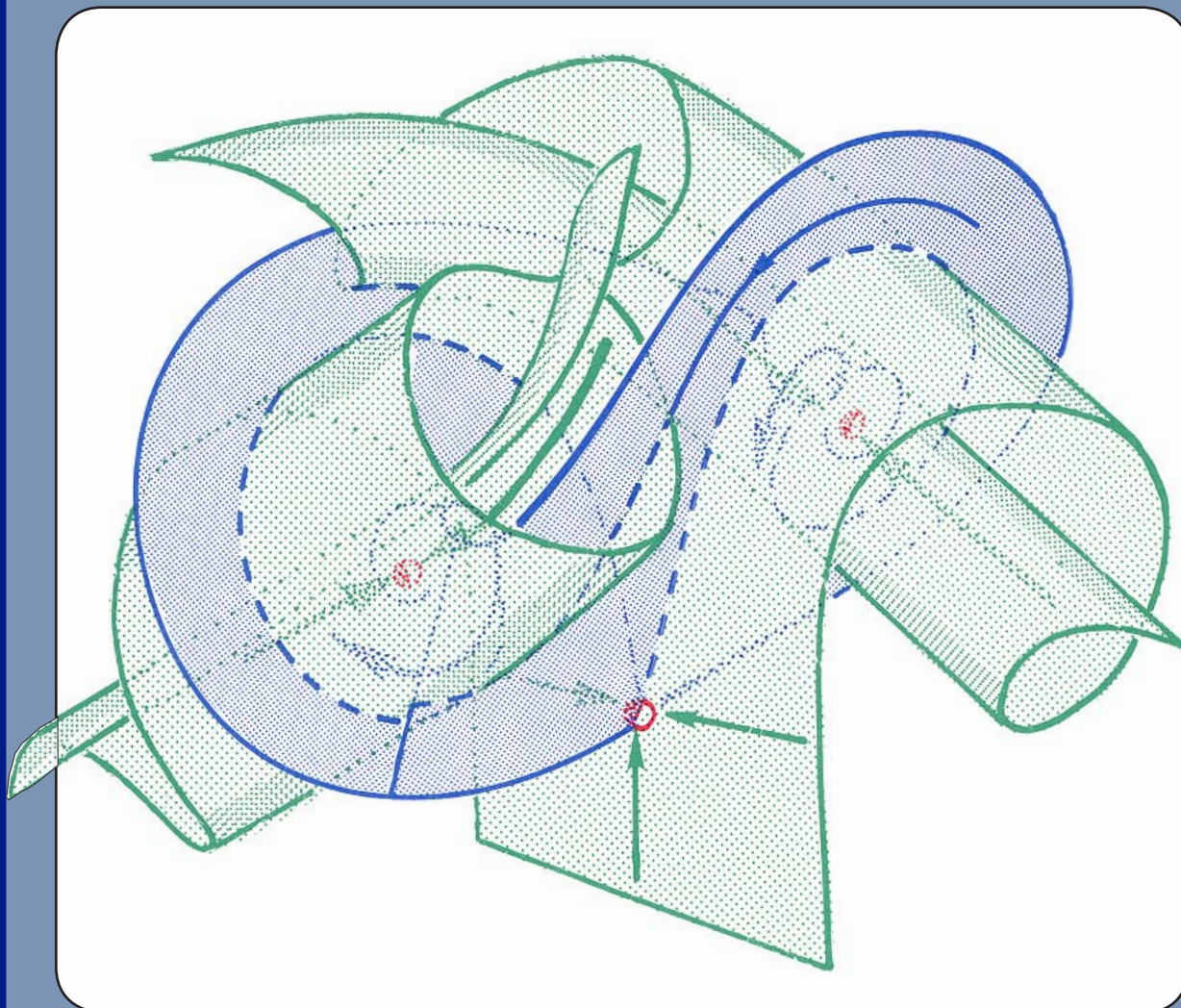
The result of reinserting is this: as each branch of $\text{Out}(Y)$ swirls around one of the shaded outsets, it approaches near the other shaded outset. It gets attracted, as outsets are attractive. Thus, the omega limit set of $\text{Out}(Y)$ is within the closure of the union of the three yoked outsets.

**13.2.5.**

And here, for comparison, is a computer drawing by Robert Shaw of the Lorenz attractor. Inspection of the equations reveals the three distinguished saddle points, right where we want them. But the planar inset of the saddle point in the lower center is qualitatively invisible. It is a kind of a separatrix. Now we will add it to the picture, with its full extension.

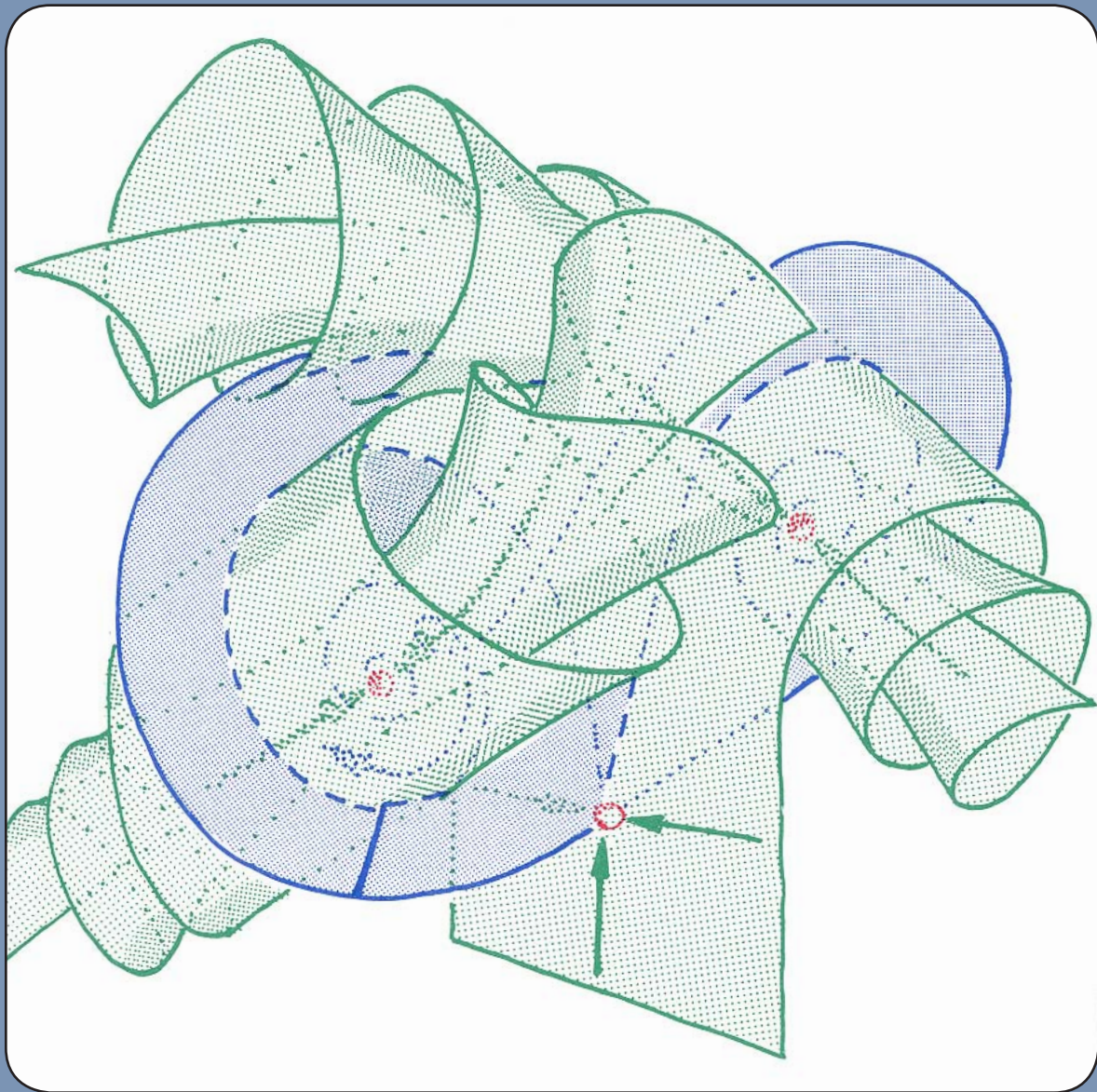
**13.2.6.**

Referring to Figure 13.2.4, we run the flow backwards in time, to extend the planar (dotted) inset outward from Y. It follows the heteroclinic trajectories (dashed) back to the yoked saddles, A and B, scrolling as it goes.



13.2.7.

Extending the dotted inset farther backwards in time, it scrolls up tightly around the *one-dimensional insets* of A and B, $\text{In}(A)$ and $\text{In}(B)$.



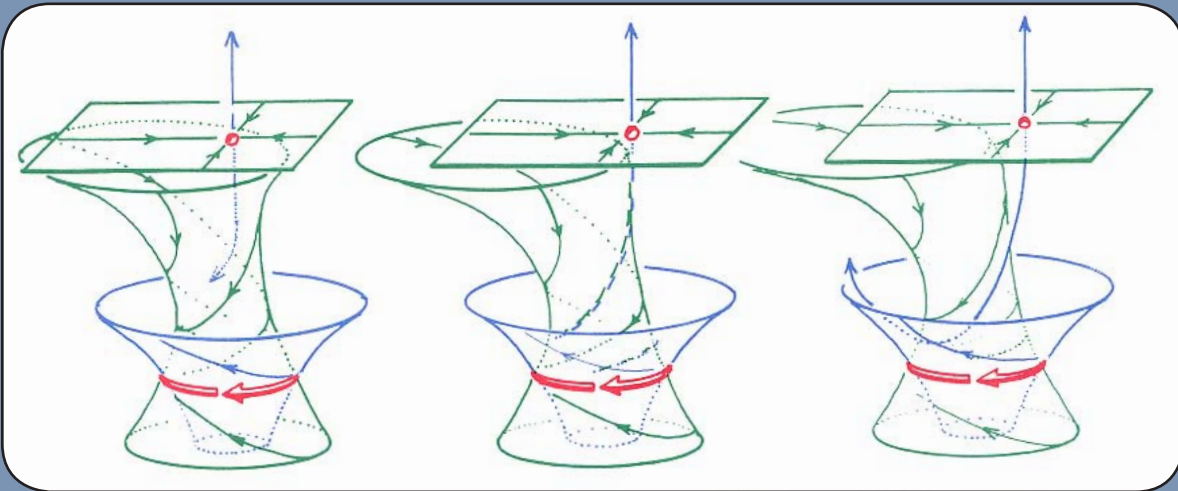
13.2.8. Extending the dotted inset farther backwards still, the four ends of the scrolls are pulled out along the curves, in (A) and in (B), toward their source at infinity.

The chaotic Lorenz attractor is composed of a yoke of tangles, folded into itself. Perhaps all of the familiar chaotic attractors have such an *outset structure*. But even in nonchaotic systems, the tangles are very important features.

We resume now our excursion into tangles.

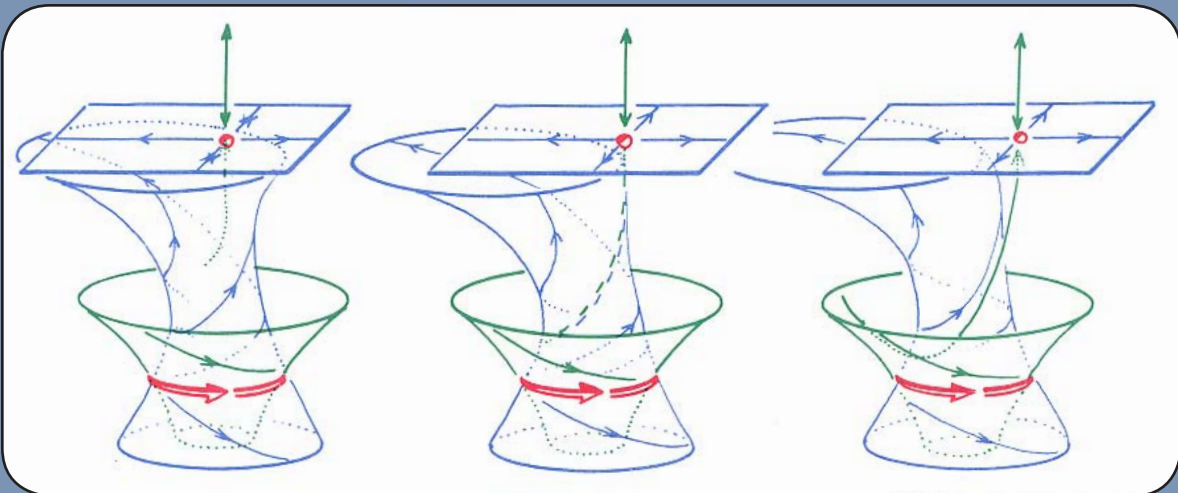
13.3. Point to Cycle

There is only one kind of hyperbolic saddle cycle in 3D: index 1 (two-dimensional inset and outset). The two-dimensional outset of a hyperbolic limit point of index 2 can have a transverse intersection with the two-dimensional inset of such a limit cycle.



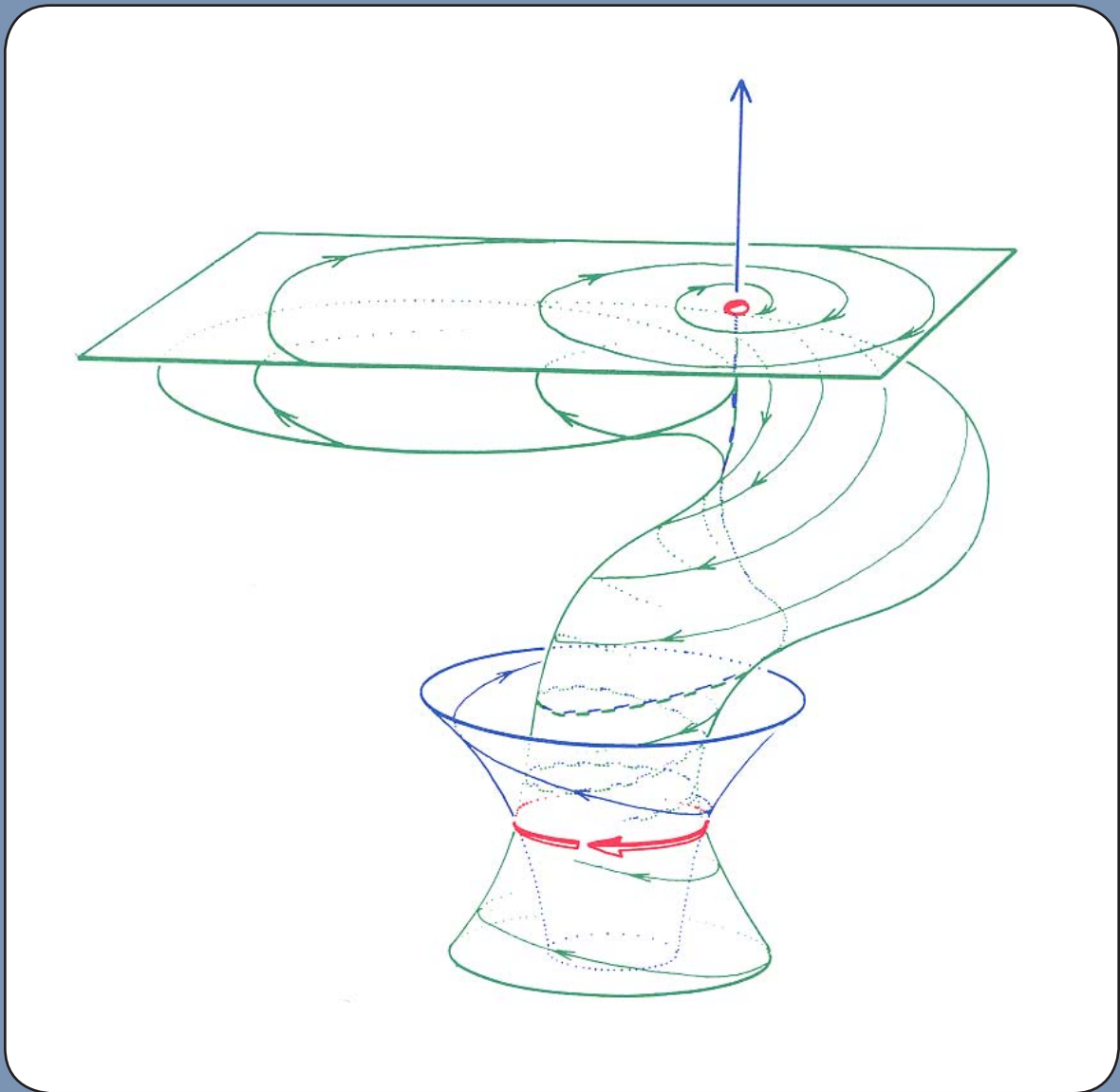
13.3.1.

A heteroclinic trajectory from a saddle point of index 1 to a saddle cycle can never be transverse in three dimensions. Here is a nongeneric portrait, in the center, flanked by two nearby generic ones.

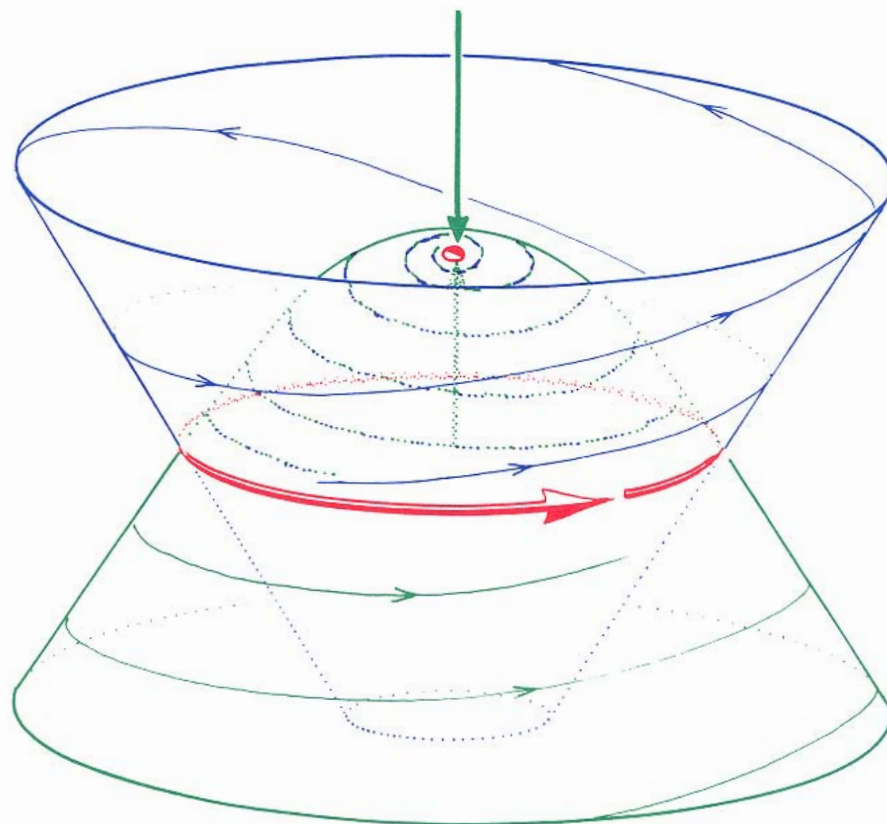


13.3.2.

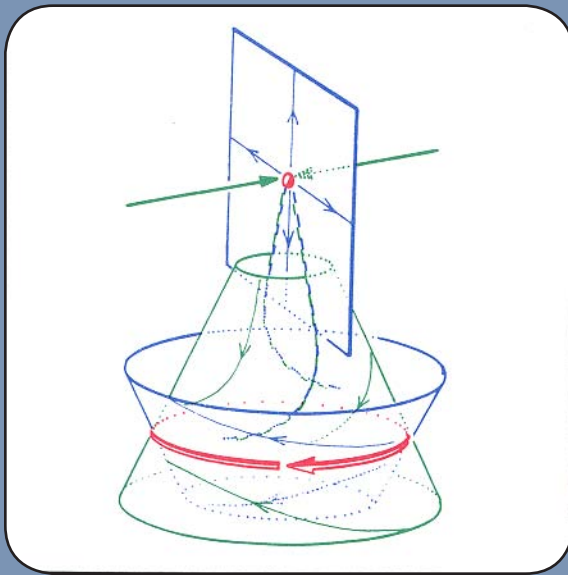
Similarly, a heteroclinic trajectory from a saddle cycle to a saddle point of index 2 is nongeneric.



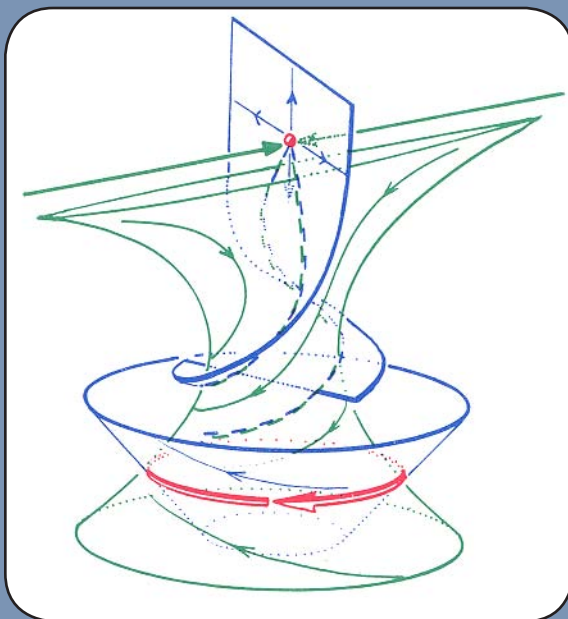
13.3.3.
The two preceding panels illustrate nongeneric connections between a saddle cycle and a saddle point of the radial type. Here is an analog, with the radical point replaced by a spiral.

**13.3.4.**

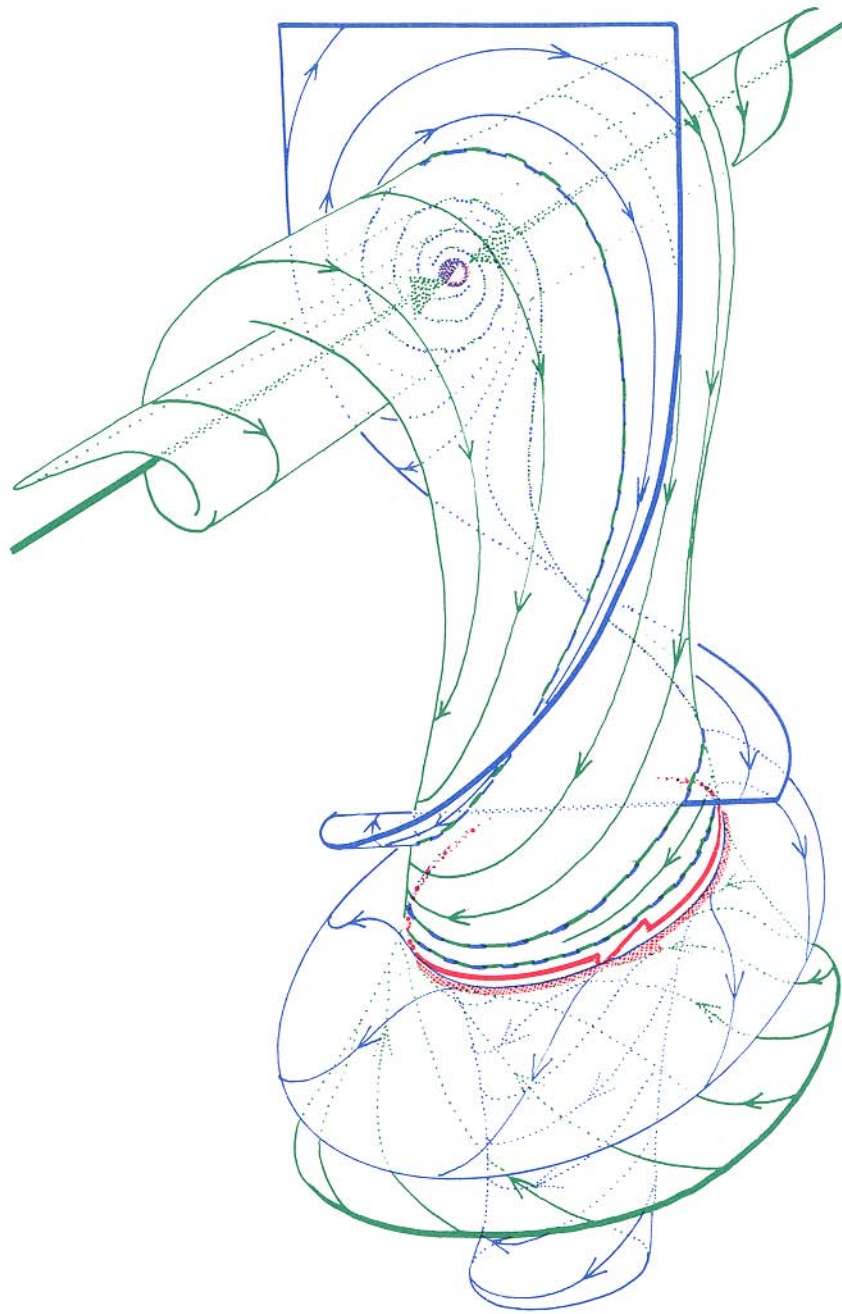
In this example, the outset of a saddle point of index 2 actually coincides with the inset of a saddle cycle. These nongeneric examples illustrate a *degeneracy of order 1*: only one condition of genericity has been violated.



13.3.5. Nevertheless, heteroclinic connection from a saddle point of index 2 to a saddle cycle can occur generically in three dimensions. Here is the first step in the visualization of this configuration.

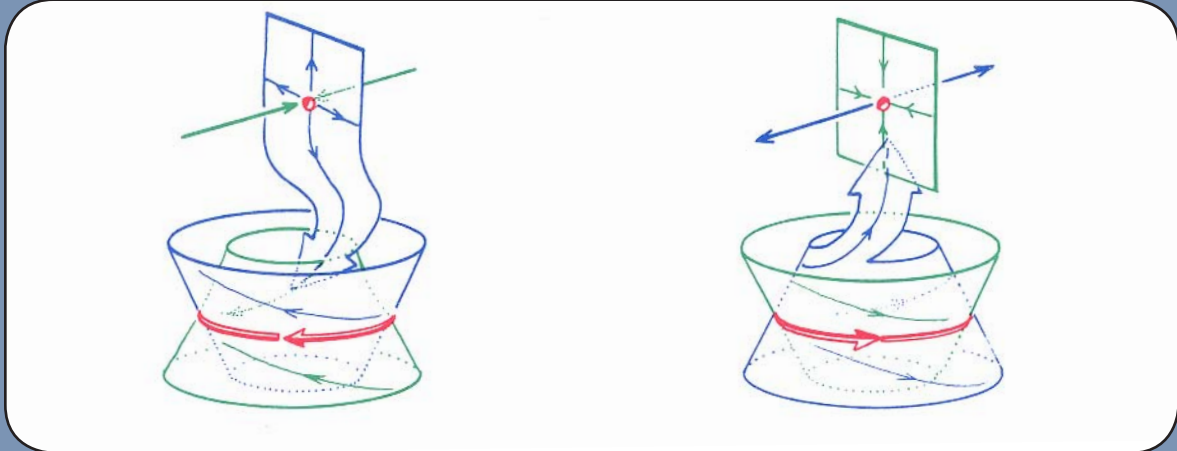


13.3.6. To generate more of the picture, the inset of the limit cycle (upper cone above) must be extended further into the past, to see how the trajectories spiraling into the limit cycle must have come from near the inset trajectories of the limit point.

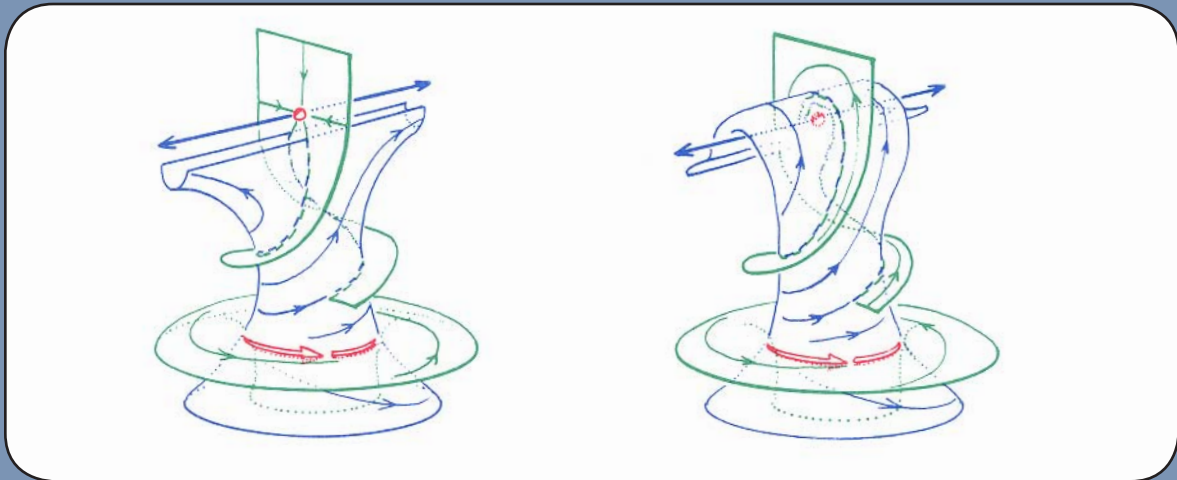


13.3.7.

Before, the saddle point of radial type was shown. Here, it has been replaced by a spiraling one. These two distinctive types of heteroclinic behavior are topologically equivalent, however.



13.3.8.
 The heteroclinic portraits just described can be transformed into two other generic portraits by reversing the direction of time. Thus, the prior connection, on the left, suggests a new sort, on the right, in which the heteroclinic trajectory goes from a saddle cycle to a saddle point of index 2.



13.3.9.
 These two forms, radial and spiral, of the generic saddle connection result. As above, they are topologically equivalent.

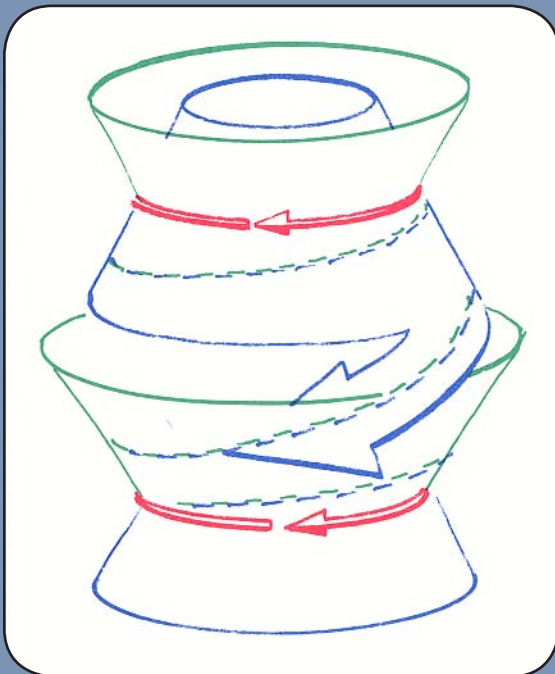
All of the forms of this section could be reversed, by changing the direction of time, to provide examples of heteroclinic tangles from a limit cycle to a limit point: *cycle to point*.

13.4. Cycle to Cycle

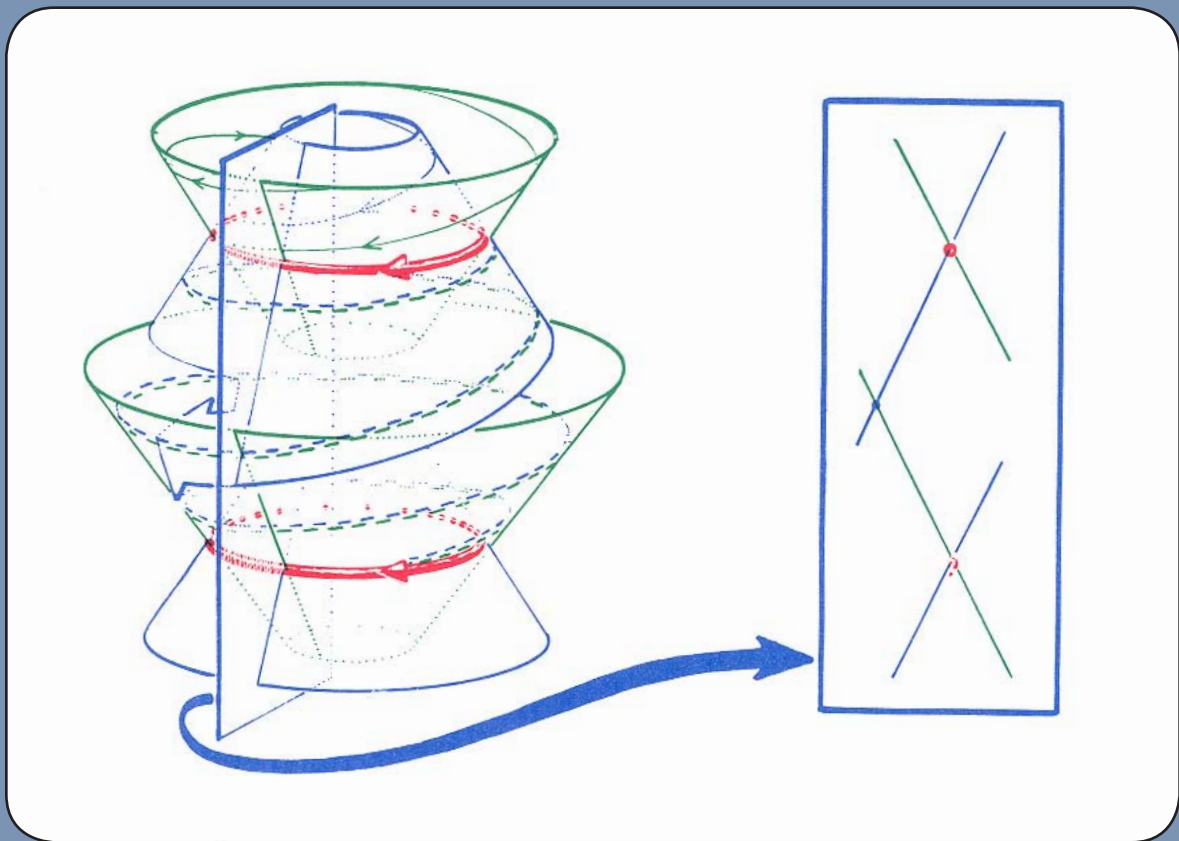
Thus far, three generic and topologically distinct saddle connections have been described:

- saddle point index 2 to saddle point index 1,
- saddle point index 2 to saddle cycle,
- saddle cycle to saddle point index 1.

In three dimensions, there is just one more.

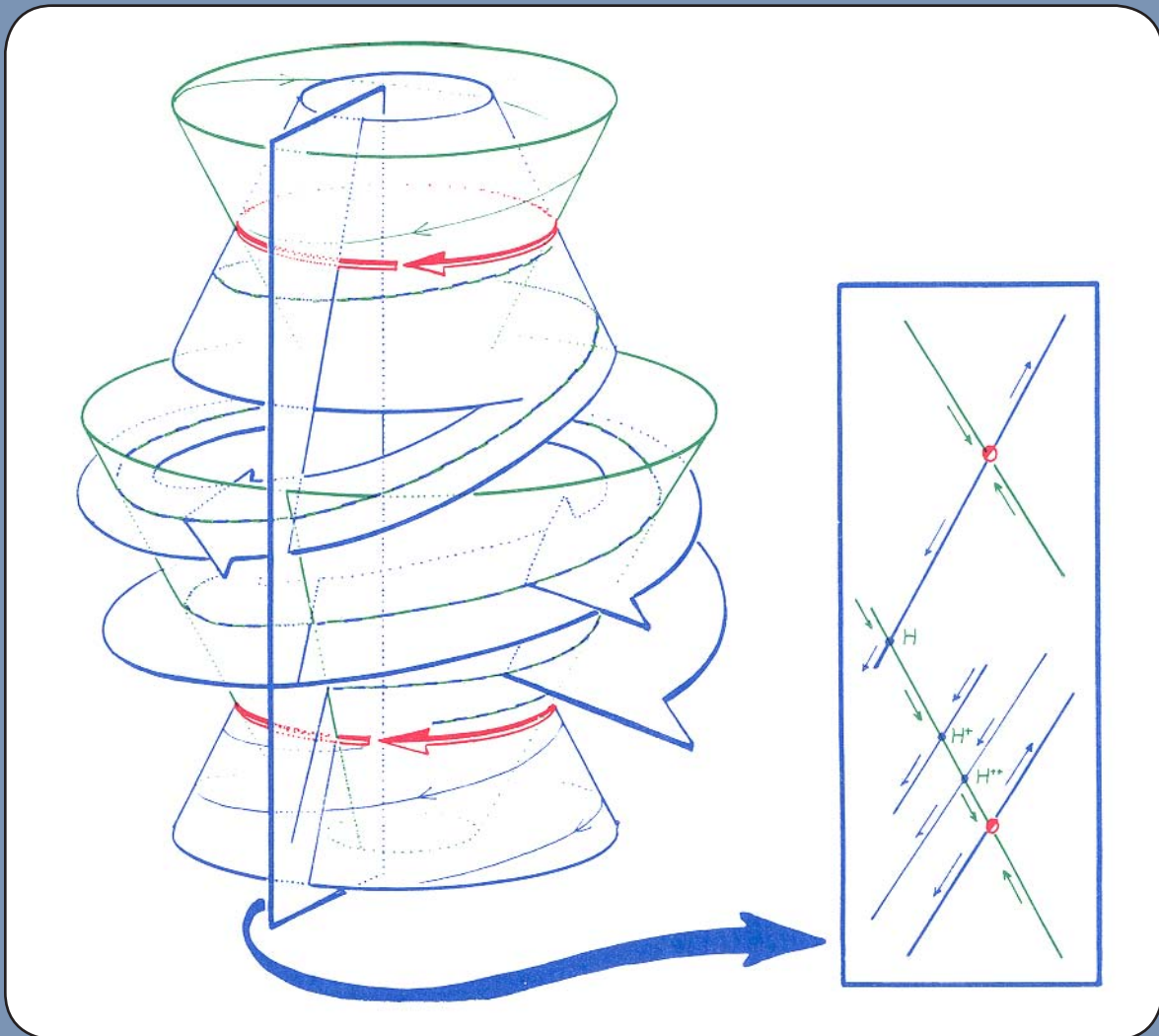


13.4.1. The outset of a saddle cycle (two-dimensional) can intersect the inset of another saddle cycle (also two-dimensional) transversely, in a (one-dimensional) curve of intersection, necessarily a spiraling trajectory. This fourth type of generic heteroclinic behavior is decidedly complicated.



13.4.2.

To dissect the complicated structure of such a connection between limit cycles, Poincaré introduced the *transverse section*, and the *first return map*. Within the cross-section (the *Poincaré section*) the two limit cycles are represented by points, and their insets and outlets by curves. The intersection of the outset of the donor cycle (above) and the inset of the receptor cycle (below) is a heteroclinic trajectory, represented in the Poincaré section by the point designated *H*.



13.4.3.

This picture, understood by Poincaré and fully analyzed by Birkhoff and Smith,³ involves a doubly infinite sequence of intersections of the curves representing the inset and outset. For the marked point, H , representing the heteroclinic trajectory, is mapped by the Poincaré first return map into another point, $H+$, which is also in both curves. This point, $H+$, is actually on the same heteroclinic trajectory as H , at a later time. Further, the image of $H+$ is another point, $H++$, through which both curves must cross.

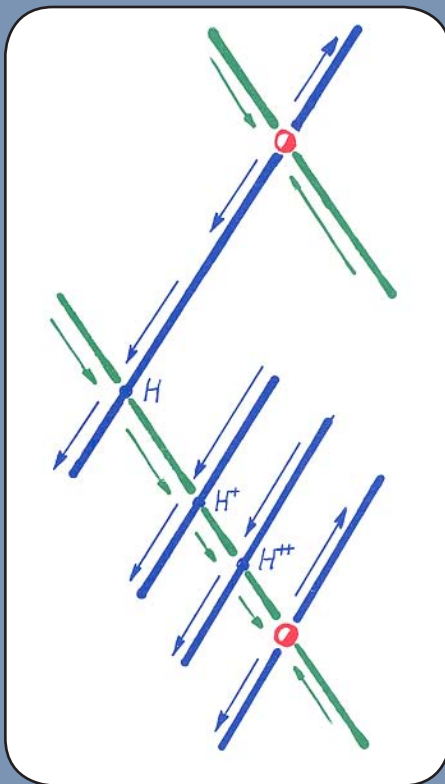
The completion of this drawing, showing the full tangle of curves within the Poincaré section, was carried out brilliantly by Birkhoff. His topological analysis of this picture reveals that between the points of intersection, H and $H+$, there must be, assuming $G3$, an odd number of others.

This construction of Birkhoff is carried out in the next section.

13.5.

Birkhoff's Signature

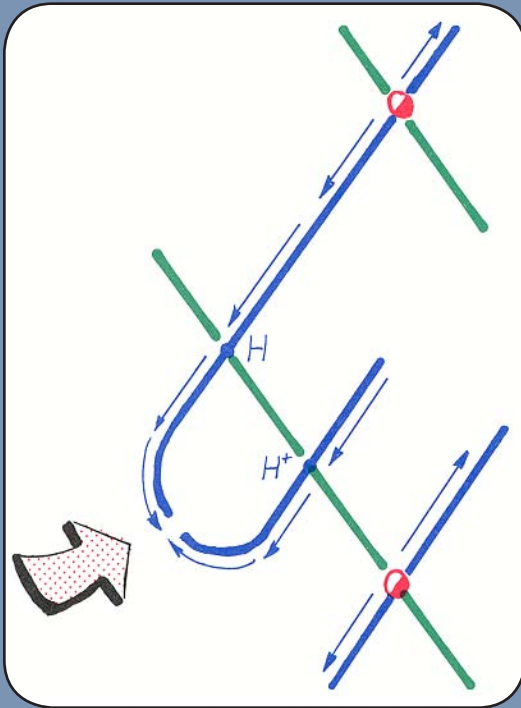
The successive intersections of the inset and outset, curves within the Poincaré section, shown above, are all points belonging to a single heteroclinic trajectory. However, there may be (in fact, must be) other intersections, belonging to other heteroclinic trajectories. Our task now is to chart all of these, and the course of the inset and the outset curves between intersection points.



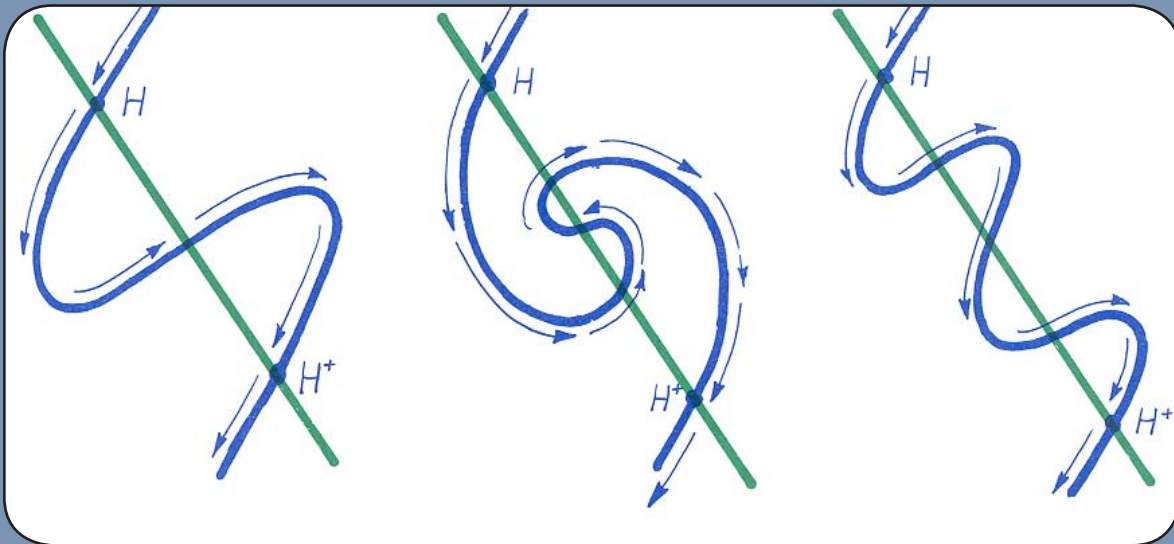
13.5.1.

Here is a close-up view of two successive intersections, H and H^+ , belonging to a single heteroclinic trajectory. They are shown here on a piece of the inset curve of the saddle point on the right, representing the receptor saddle cycle. Through H^+ passes a short piece of the outset curve of the saddle point on the left, representing the donor saddle cycle. How can we fill in the entire donor outset curve, connecting these short segments?

Notice the arrows on the outset segments, indicating the out-directions on the outset curve, away from the donor.

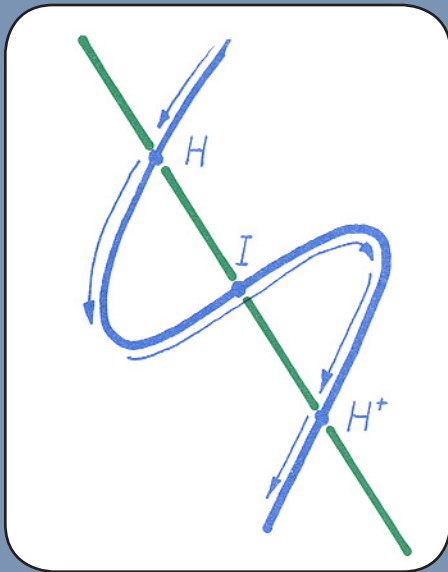


13.5.2. The simplest solution might be just to connect up the loose ends, as shown here. Unfortunately, this does not work. The out-directions must connect properly, without conflict.

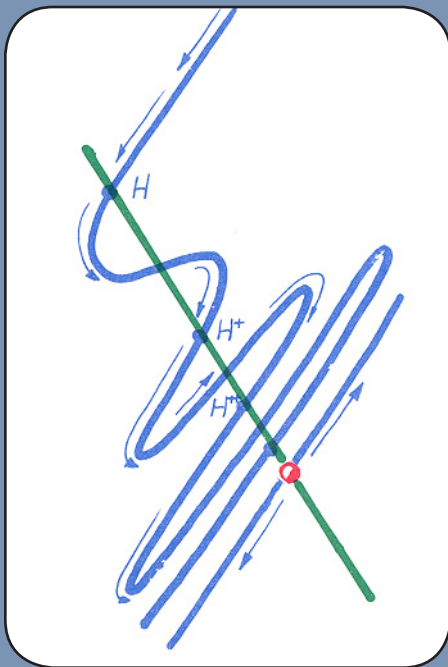


13.5.3. This drawing shows three possible connections for the outset curves, joining the short segments without conflict of the out-directions. The complete outset segment, joining two successive points corresponding to the same heteroclinic trajectory, H and H^+ , cuts through the inset segment joining the same two points in an odd number of points, all heteroclinic, but belonging to different heteroclinic trajectories. The two complete segments, joining H and H^+ , comprise the figure Birkhoff called the *signature* of the saddle connection.

Some more complicated examples are given in the next chapter.

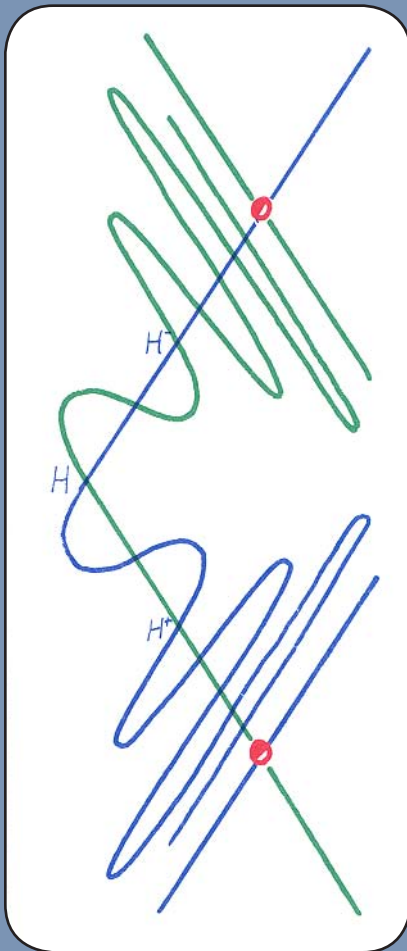


13.5.4.
This shows the simplest possible Birkhoff signature. The odd number of interpolated heteroclinic points is only 1. This point, 1, represents another heteroclinic trajectory, sharing the same donor and receptor, and possessing its own signature (not shown).



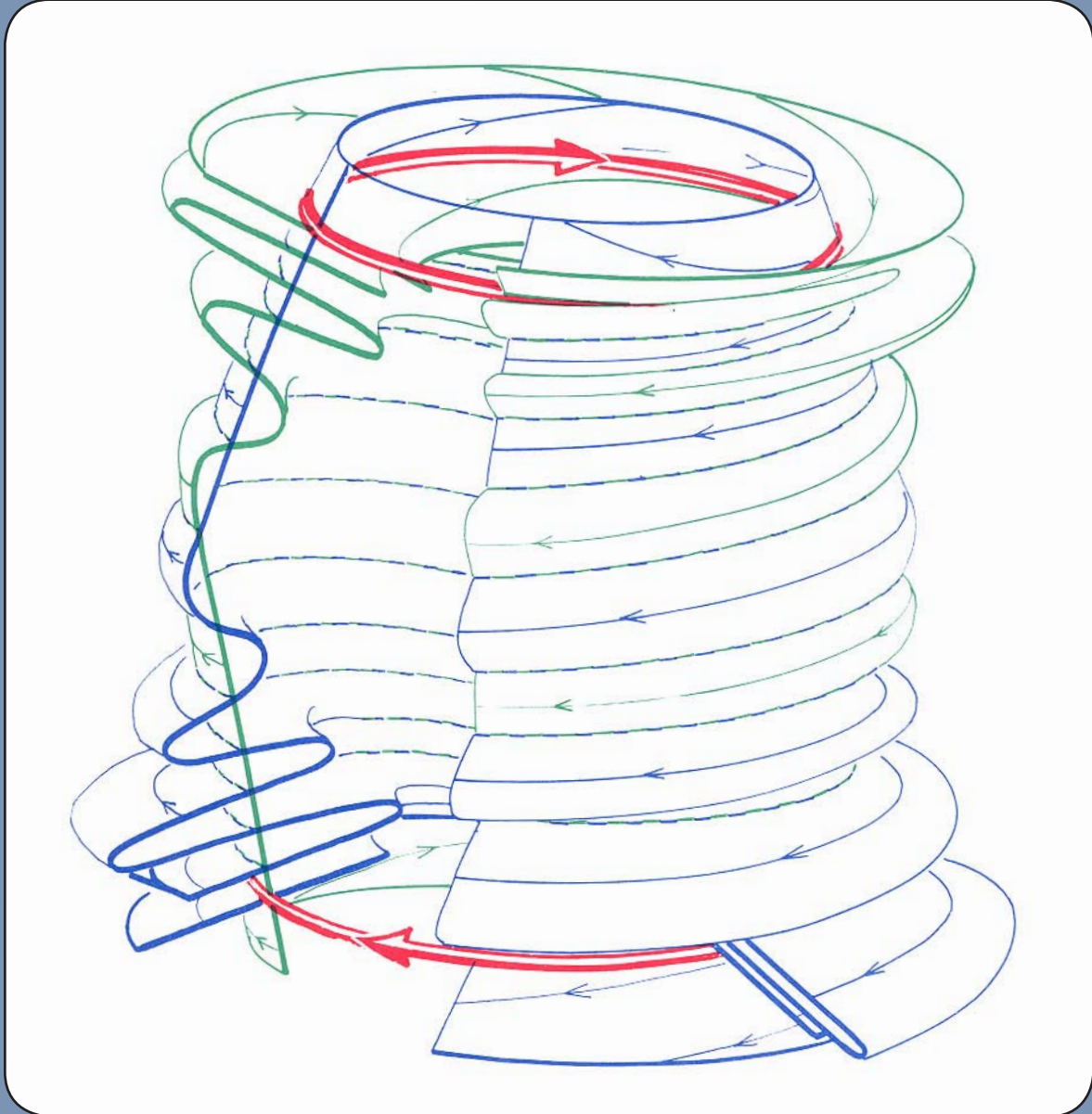
13.5.5.
Reinserting this Birkhoff signature into the starting picture of this section, together with two of its forward images under the first return map, we have a roughly complete idea of the donor outset. There are many possibilities for the future of the outset, but here we have used only the simplest signature, as shown in the preceding panel. In this case, there is an infinite sequence of points of intersection, H, H^+, H^{++}, \dots , all belonging to a single heteroclinic trajectory.

Meanwhile, the inset curve of the receptor is still only half-drawn. Where is its past?



13.5.6.

Extending the receptor's inset backwards in time, we obtain the predecessor of H , H^- , its predecessor, H^+ , and so on. This completes a doubly infinite sequence, corresponding to one full heteroclinic trajectory. Likewise, the interspersed heteroclinic trajectory contributes a complementary doubly infinite sequence as shown here, in the Poincaré section.

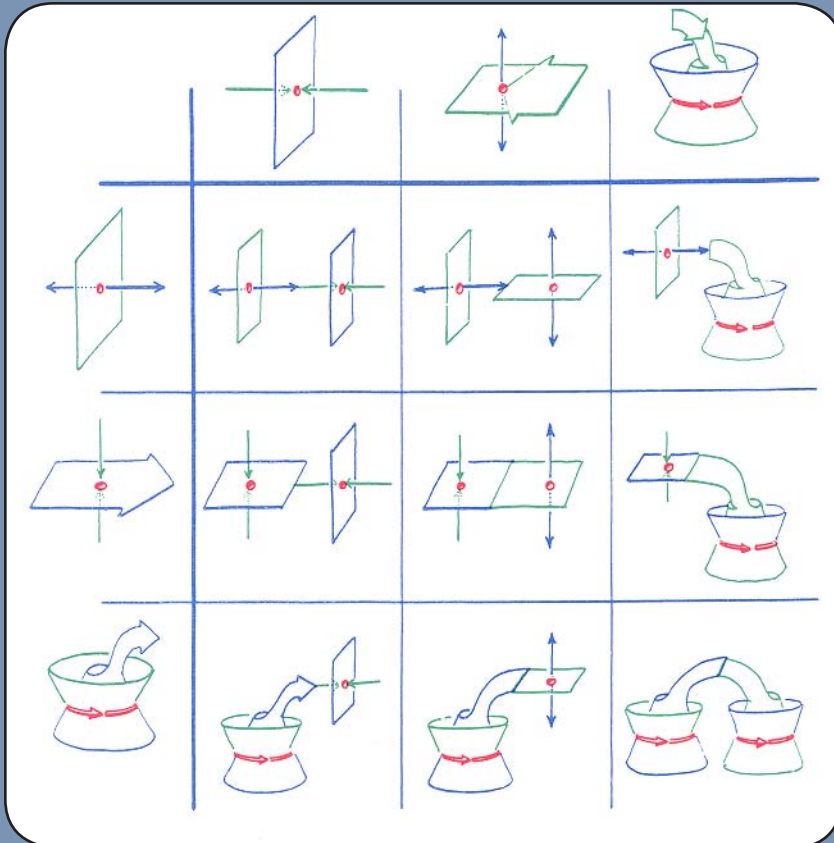


13.5.7.

The doubly infinite sequences each correspond to a heteroclinic trajectory of intersection of the donor's outset and the receptor's inset, in the original three-dimensional context. Here, the generic connection of saddle cycles in three dimensions is shown, with all its complex structure. A section has been removed here, for improved visibility.

If this object were set down upon a rotating phonograph turntable, it would look rather like a bolt being screwed down.

The behavior of the trajectory passing by a cycle-to-cycle heteroclinic tangle is a spiraling asymptotic approach along the inset of the donor, followed by a period of entrapment, spiraling along the screw thread of the heteroclinic tangle, and finally an asymptotic escape, along the outset of the receptor. Thus, the heteroclinic tangle provides a model for *transient oscillation*.



13.5.8.

In the three-dimensional case, there are several possibilities, summarized in this table. The two types of topologically distinct hyperbolic saddle points (of index 1 and 2) and the unique hyperbolic saddle cycle are each possibly donors, or receptors, of a saddle connection. The nine possibilities are pictured here, with the donors down the left, and the receptors along the top. Note the order and orientation of the donors is not the same as those of the receptors.

In summary, there are no generic saddle connections in two-dimensional dynamical systems. In three dimensions, there are four topologically distinct types. In higher dimensions, the situation is even more complicated.

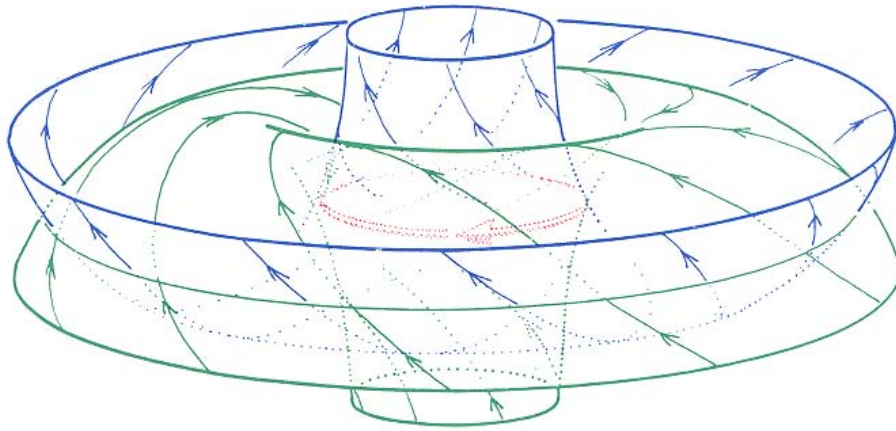
The generic property G3 for dynamical systems is this: all inset and outset intersections are transverse. The genericity of this property, like the properties G1 and G2, is established by the theorem of Kupka and Smale.

14 Homoclinic Tangles

In addition to the four kinds of transverse saddle connections described in the preceding chapter, there is one more that can occur in three dimensions. This is the connection from a saddle cycle to itself, called a *homoclinic connection*. Homoclinic connections are much more important than heteroclinic ones, as *they occur as exceptional limit sets within separatrices*. Further, as shown by Birkhoff and Smith,¹ they are full of limit cycles. The study of this complicated case, initiated by Poincaré, is still in progress. An advance was made by Smale² in 1963. Many topologically different forms are possible. This chapter describes the main ideas of the three-dimensional context, including some constructions not previously published.

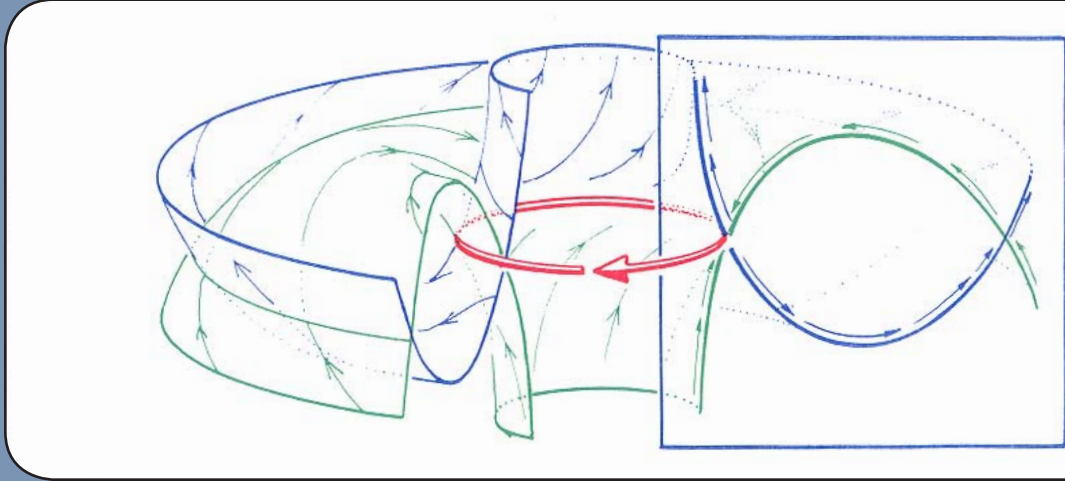
14.1. Homoclinic Cycles

By definition, a homoclinic trajectory must belong to the inset and outset of the same limit set. In the generic context of *properties G1, G2, and G3*, this limit set may not be a point. The simplest generic case is a limit cycle of saddle type, in three dimensions. In this section, we dissect this case.



14.1.1.

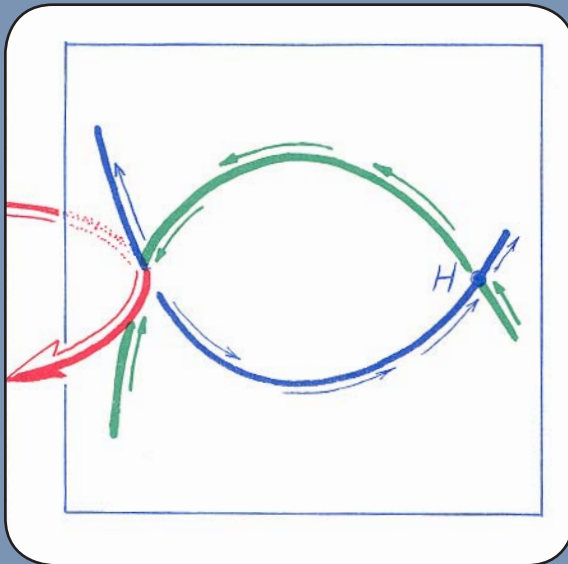
Here the outset of the limit cycle, at the top, is pulled down like a sleeve turned inside out. The inset, below, is likewise pulled up. Then, they are pushed through each other, to produce the beginning of an extensive intersection.



14.1.2.

To visualize the intersection, we cut through it with a Poincaré section. The procedure is the same as the heteroclinic case, described in the preceding chapter (see 13.4.2.).

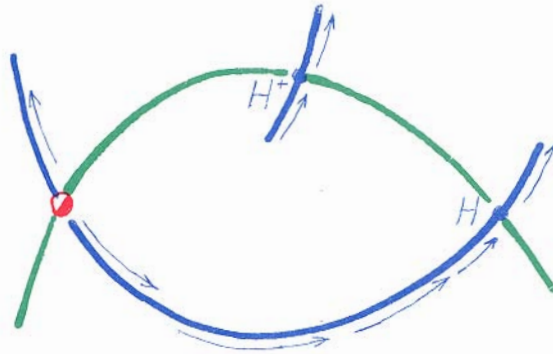
The key to the analysis is the first return map, which maps the Poincaré section into itself, corresponding to one revolution around the limit cycle.



14.1.3.

As in the preceding chapter (see 13.5.1.), the outset surface of the receptor limit cycle (in this case, they are the same cycle) intersect the Poincaré section in two curves, the outset and inset curves. These curves intersect once at the point cut by the limit cycle (shown as a curved arrow here), and again at a point cut by the homoclinic trajectory, such as the homoclinic point H , shown here.

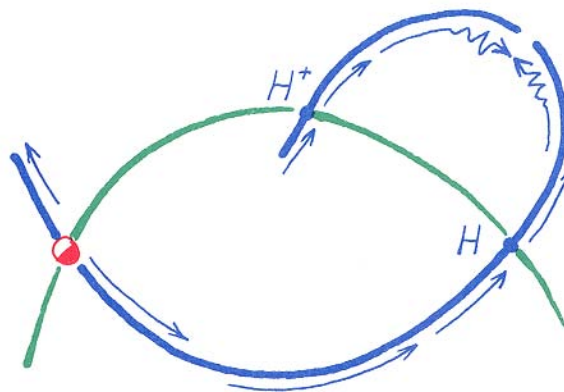
What happens to the homoclinic point after another revolution around the limit cycle?



14.1.4.

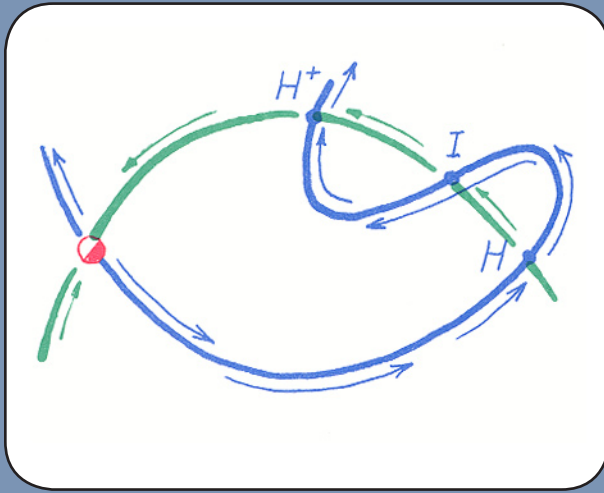
As in the heteroclinic case (again, see 13.5.1.), this point is mapped to another, H^+ , closer to the limit point. This image point is on the inset curve, as this curve is mapped into itself by the first return map. Further, this curve consists of all the incoming points. However, the image point must also be on the outset curve, which is also mapped into itself by the first return map, and which consists of all outgoing points. The homoclinic points, H and H^+ , are both outgoing and incoming, by assumption. Thus through the image point, H^+ , there must also pass a piece of the outset curve, shown here with its out-direction indicated by an arrow.

How may these outset segments be connected, so as to obtain the entire outset?

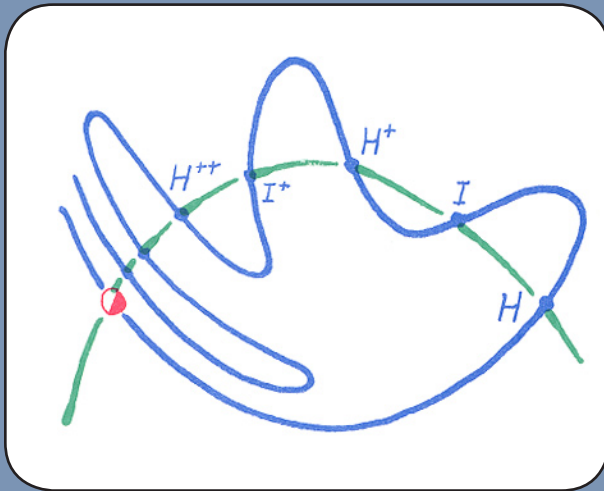


14.1.5.

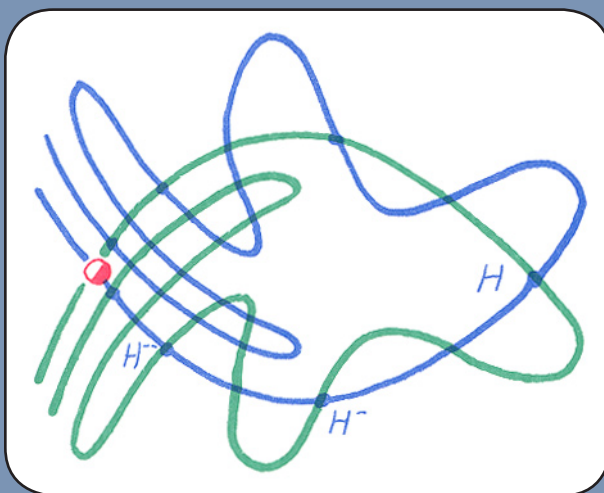
As in the heteroclinic case (see 13.5.2.), direct connection leads to a conflict of out-directions. Thus ...



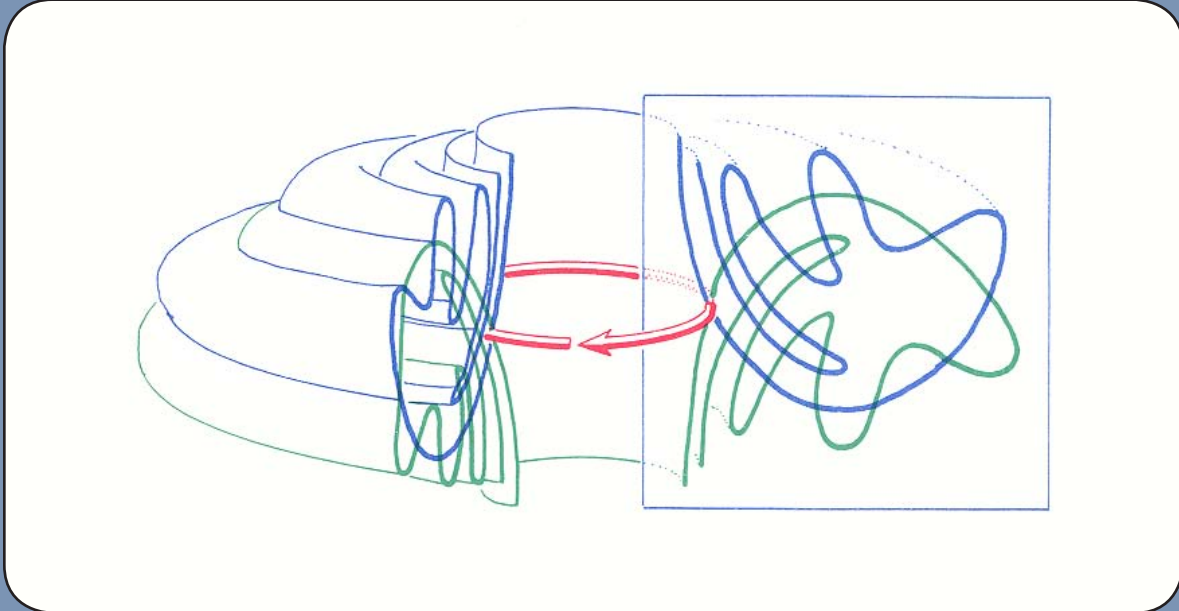
14.1.6. ... as in the heteroclinic case (see 13.5.3.), the outset segment from H to H^+ must cross the inset segment (between the same two points) an odd number of times. This is the simplest legal construction, illustrating the *Birkhoff signature* in the homoclinic case.



14.1.7. Reiterating the first return map again and again, the outset segments push up against the inset curve, near the limit point.



14.1.8. Repeating the construction for negative times (iterating the *prior return map*), the inset segments pile up against the outset curve, again near the limit point. Thus, we obtain a full picture of the entire *homoclinic tangle*, as shown in this drawing of a tangle studied by Hayashi,³ the greatest master of experimental tangle art.



14.1.9.

Here is the tangle within the Poincaré section, replaced within the original 3D context (compare with 13.5.7.). The behavior of a nearby trajectory is a spiraling asymptotic approach, along the non-tangled half of the inset surface, followed by a period of chaotic motion, entrapped within the tangle, and finally a spiraling asymptotic escape, along the non-tangled half of the outset surface. Thus, the homoclinic tangle provides a model for *transient chaos*.

This tangle, based on the simplest Birkhoff signature, reveals additional intersections of inset and outset loops. This deeper structure is not determined by the Birkhoff signature. Thus, to fully describe the structure of the tangle, additional signatures must be specified.

In the next section, we introduce a sequence of signatures, published here for the first time, for the full description of a homoclinic tangle in 3D.

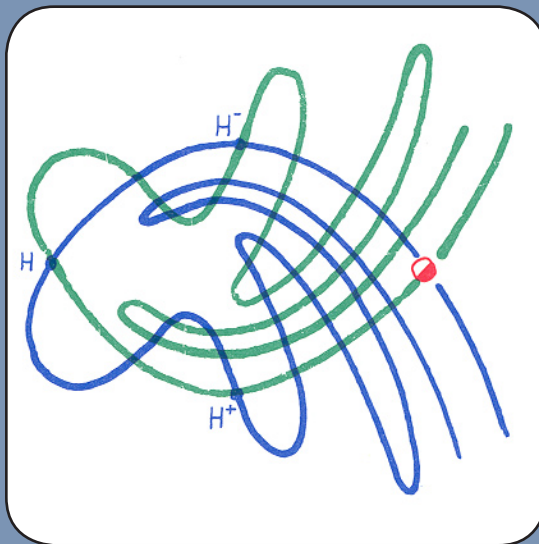
14.2.

Signature Sequence

During the preparation of a preliminary edition of this work in 1980, we tried to deform the Hayashi tangle (shown in the preceding panel, 14.1.8.) into the Smale horseshoe (described in the next section; see also Figure 8.1.10.). Although the two homoclinic tangles have the same Birkhoff signature, we were unable to deform the Hayashi tangle into the horseshoe.

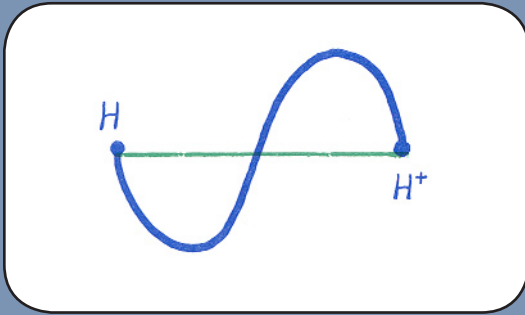
In trying to understand the difference between these two exemplary tangles, we developed an infinite sequence of signatures. The first of these is the Birkhoff signature, which is the same for the two examples. The second, however, is different. Thus, they could not be deformed, one into the other. This led to our signature conjecture: if two tangles have the same signature sequence they are topologically equivalent.

In this section, we construct the signature sequence, step by step, for the Hayashi tangle. In the next section, we will apply it to the Smale horseshoe tangle.



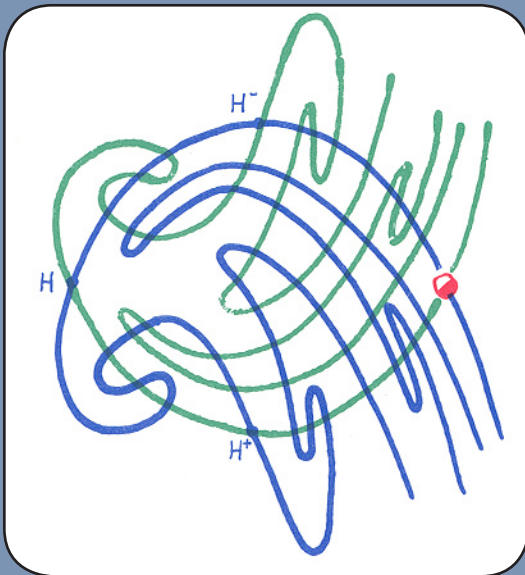
14.2.1.

Here, again, is the Hayashi tangle. It is not a mathematician's pipe dream, but was laboriously drawn by Hayashi, from extensive simulations of the Duffing system (for the forced pendulum, see Part One) with an electronic analog computer. How can we give a full characterization of this tangle? Let's single out a homoclinic point, such as H , and its image H^+ .



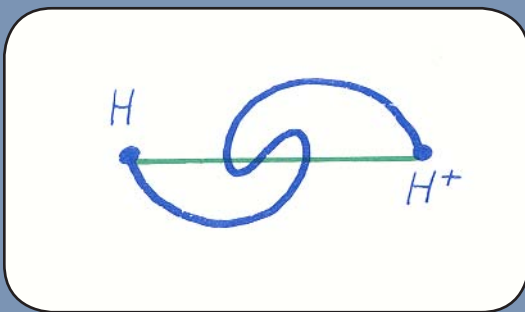
14.2.2.

Extracting the inset and outset curve segments bounded by these two points, we obtain the Birkhoff signature. Again it is the simplest possible one (see Figure 13.5.3 for three alternatives). To some extent, it characterizes the chief feature of the tangle.



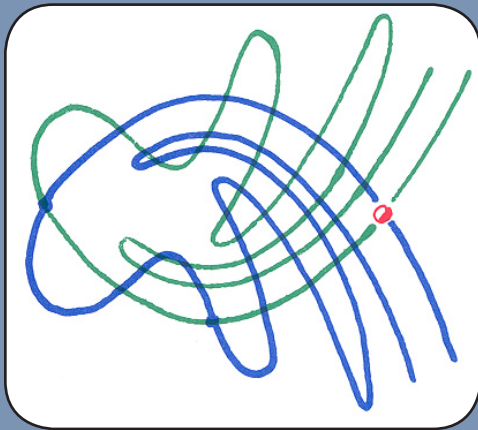
14.2.3.

For example, here is another tangle. At first glance, it appears significantly different from the preceding one.

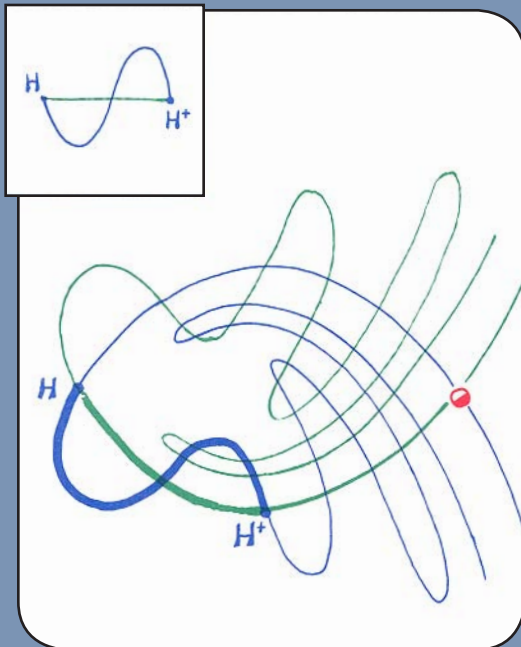


14.2.4.

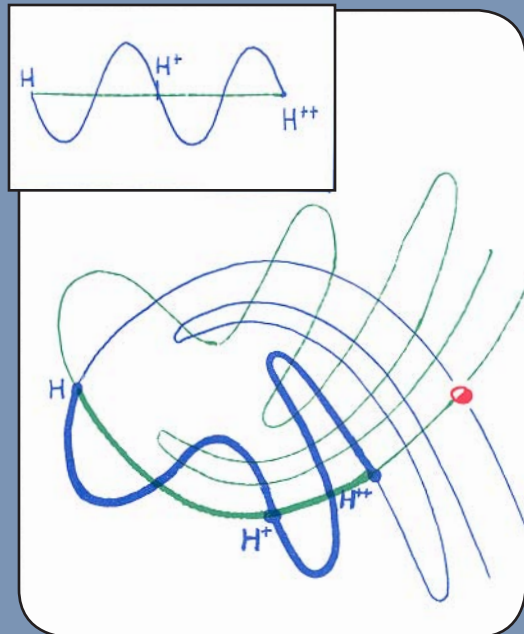
Extracting a Birkhoff signature, we see that it is indeed different. And this does seem to capture the chief feature of this new tangle.



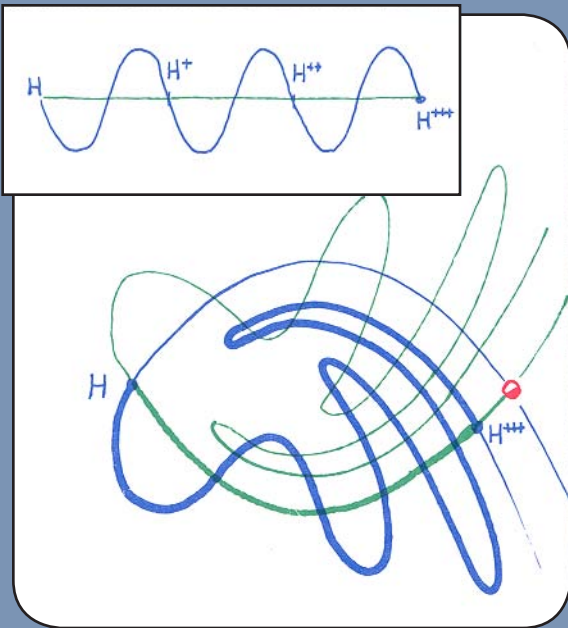
14.2.5.
 Now let's return to the old tangle. Notice how inset loops may cross through several outset loops. We want to capture a signature of this larger-scale behavior of the figure, corresponding to its minor features. We will proceed in steps.



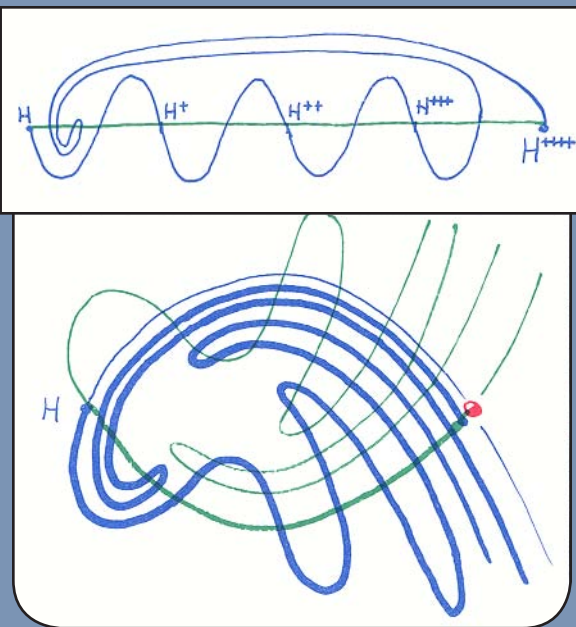
14.2.6.
Step 1. Single out a homoclinic point and its image, and draw the Birkhoff signature they determine. Draw it again, straightened out, as shown in the inset.



14.2.7.
Step 2. Extend the signature another iteration, by applying the first return map. Here we see both the Birkhoff signature and its entire image. Straighten out this figure also, as shown in the inset.

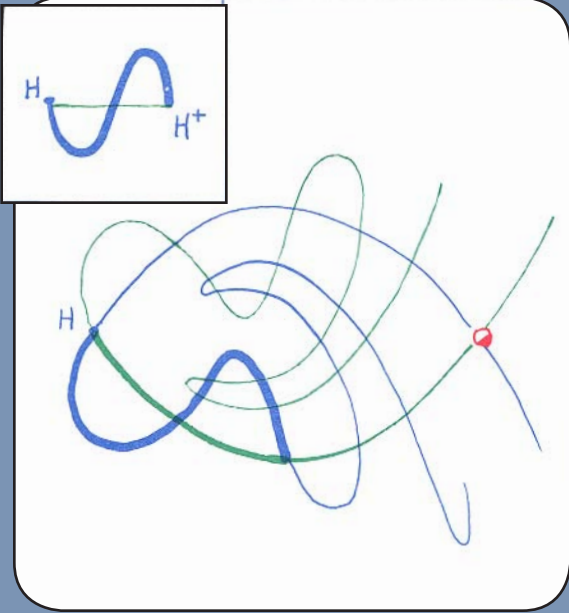


14.2.8.
Step 3. Repeat the preceding step again, and in general, as many times as you can, as long as the elongating outset loops never come back to cross a part of the original signature, or its forward iterations.

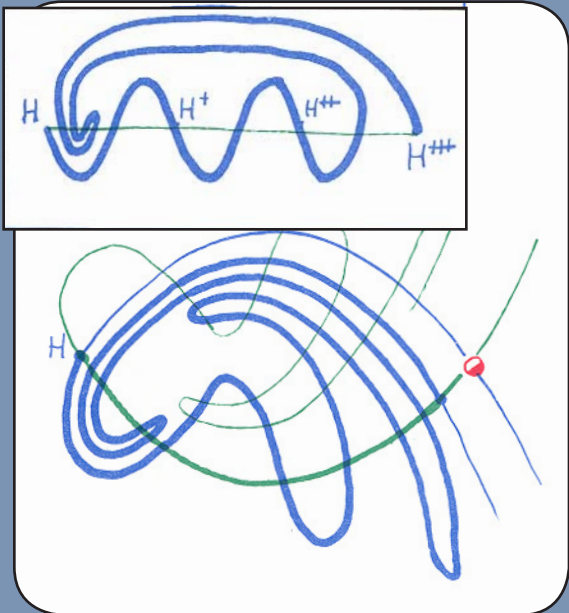


14.2.9.
Step 4. Repeat the single iteration step once more. This time, one of the elongating upper loops will make a new intersection with a segment of the inset curve belonging to the original Birkhoff signature, or one of its forward iterations. In this example, four new crossings have all appeared at once.

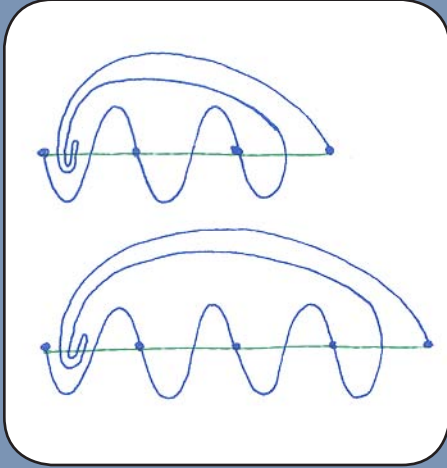
The Birkhoff signature is the first of our sequence. The figure in the inset above is the second. Let's try out these two on another example.



14.2.10.
 This is yet another tangle. It looks like Hayashi's, but is not. The Birkhoff signature (shown in the inset) is the same. Will our second signature reveal the difference?



14.2.11.
 The first iteration of the fundamental (Birkhoff) signature comes close to the fundamental, but does not cross it. The second image crosses the fundamental. The figure in the inset is the second signature of this tangle.



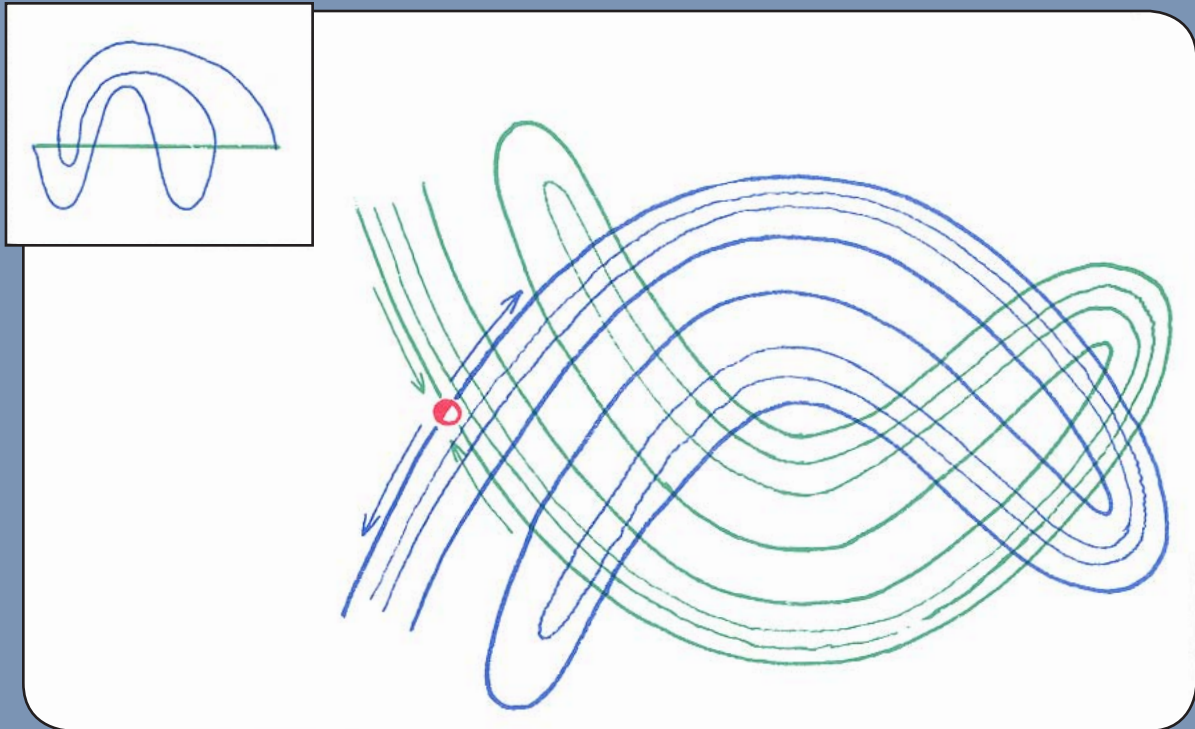
14.2.12.
Here, for comparison, is the second signature of the two tangles. The new tangle has two humps, while the Hayashi tangle has three, under the arching inset loop. They are topologically inequivalent tangles.

These two signatures are the first of an infinite sequence. See if you can draw the third signature in the examples above.

14.3.

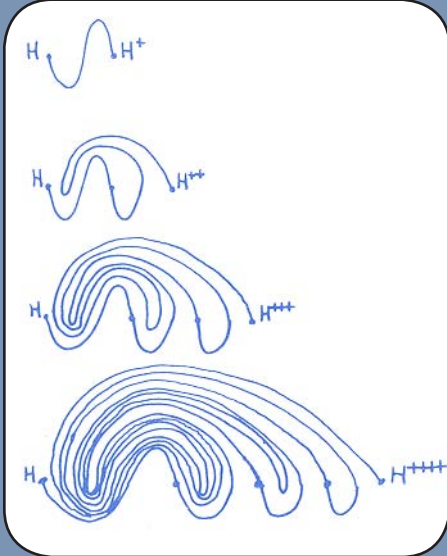
Horseshoes

In this section, we will tackle another tangle, called Smale's horseshoe. This third example originated as a geometric construction, but was subsequently observed in the forced Van der Pol system,⁴ and many others. Along the way, we will give an idea of the third signature of a tangle.

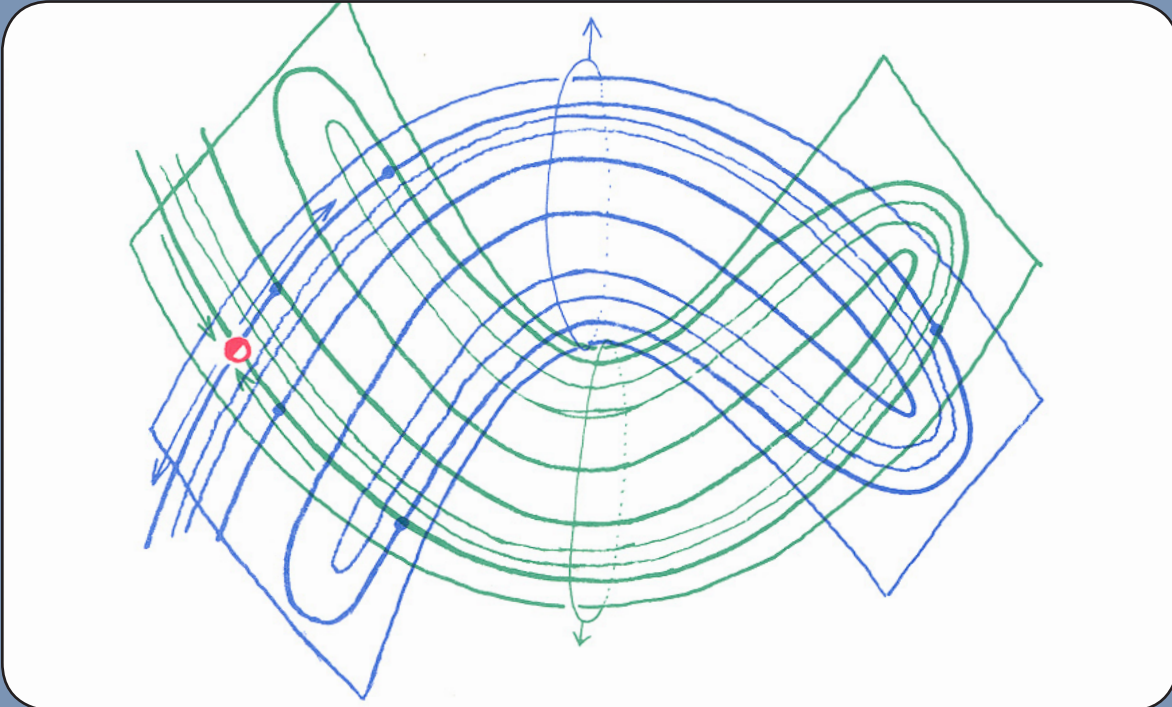


14.3.1.

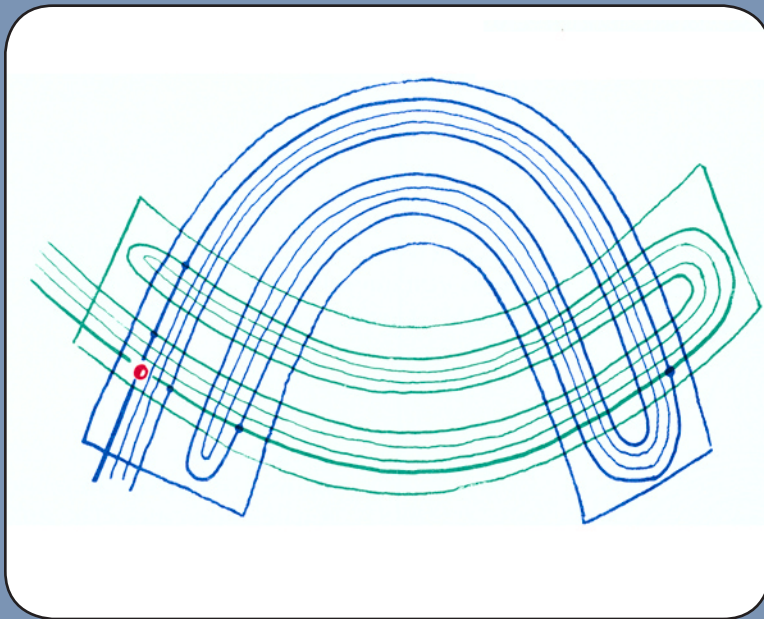
Here is yet another homoclinic tangle, the famous *horseshoe of Smale*. Note that the first signature is the familiar simplest one. But in the second signature, shown in the inset, the hump has been twisted back, creating two new intersections. To further characterize this tangle, we must draw the *third signature*.



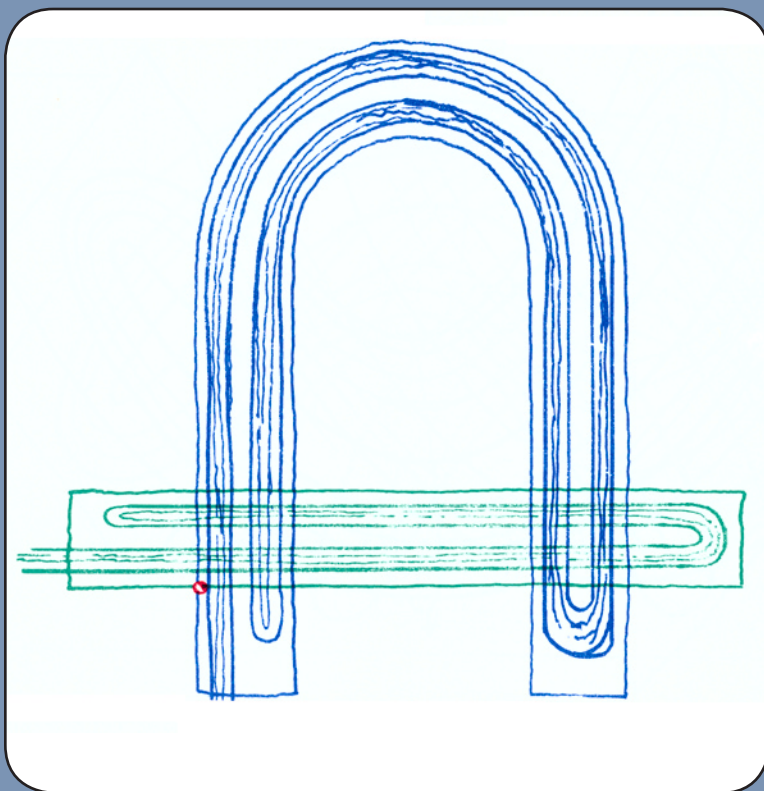
14.3.2.
Here are the first four signatures of our signature sequence, for Smale's horseshoe. The third signature is not identical to the third signature of the preceding example (try it and see).



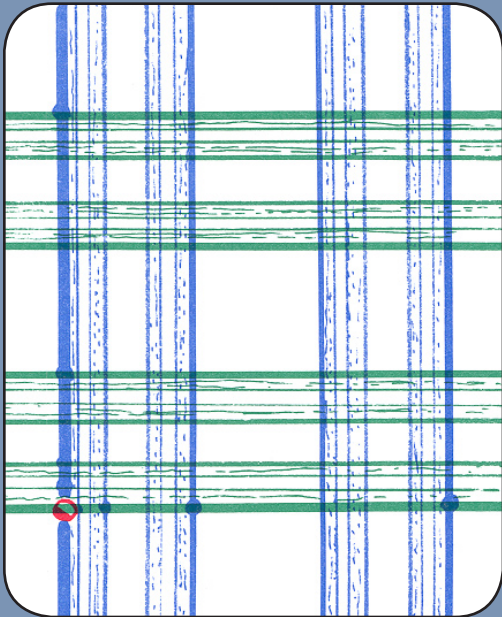
14.3.3.
The horseshoe has been untangled by Smale⁵ in a most ingenious way. Choosing a curved rectangular patch in the Poincaré section with some care, and applying the first return map yields another rectangular patch crossing the original patch at each end. Now, deform the whole picture by lassoing the two patches around the waist and pulling gently.



14.3.4.
Continue to pull the upper patch upwards by the waist, while pushing down on the ends. The idea is to straighten out the lower patch.



14.3.5.
There is the fully untangled tangle, the horseshoe of Smale. It is topologically equivalent to the messy original tangle, yet it admits a full analysis, as shown by Smale.



14.3.6.
The analysis is based upon a clever scheme for labeling all the points of intersection of the insets and outlets within the Poincaré section.



14.3.7.
Looking at a portion of the outset through a microscope, we see an infinite set of horizontal lines. Their intersection with a vertical line (such as the left edge of the box here) is much like Cantor's middle thirds set (see Figure 9.4.7).

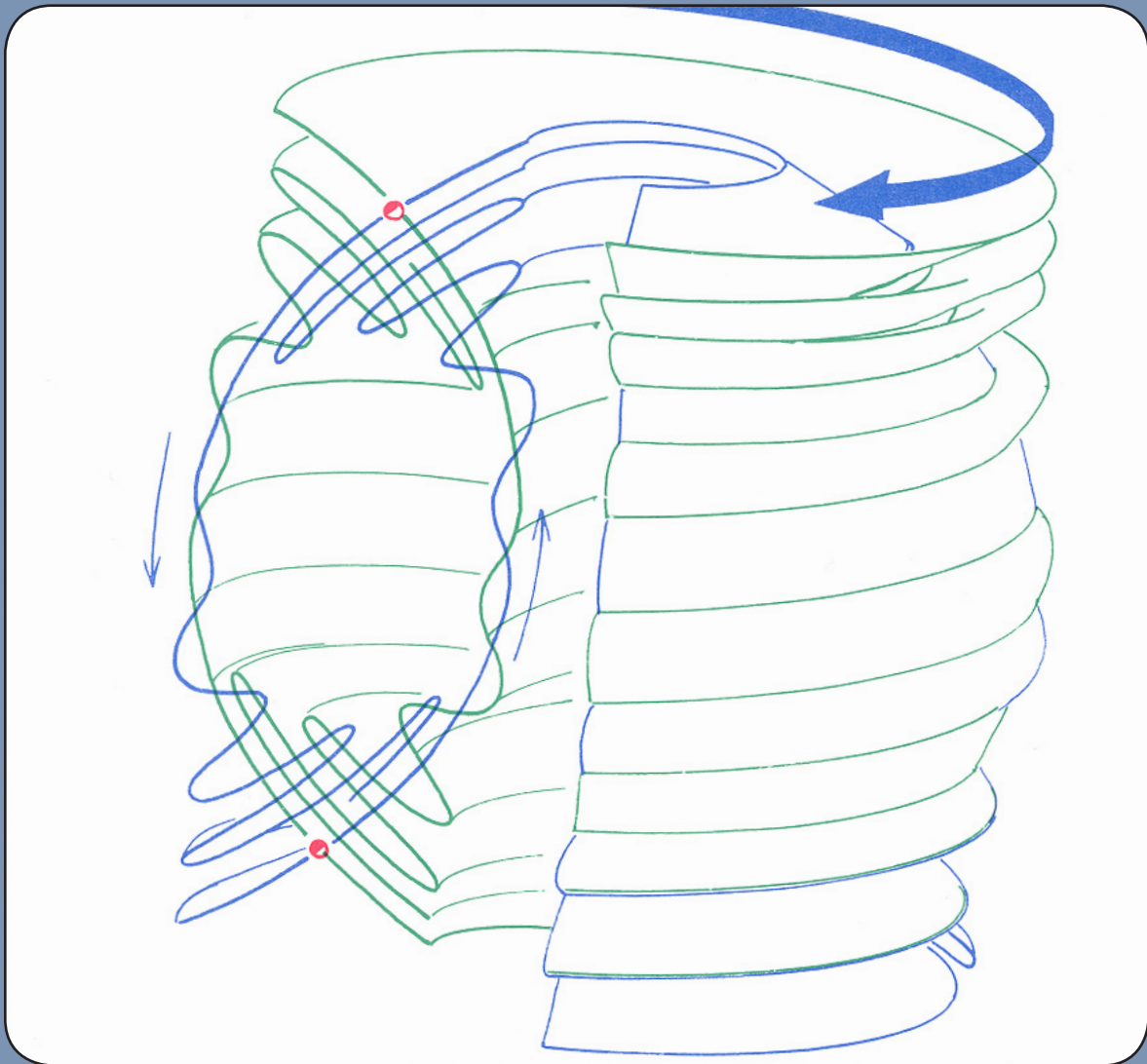
Smale's analysis of this particular tangle, based on combing it out and applying symbolic dynamics, might be applied to other homoclinic tangles, through careful use of the signature sequence.

The theory of homoclinic tangles is very important, and yet little known. Even in three dimensions, the lowest in which they occur generically, there are outstanding problems. In higher dimensions, little is known. Poincaré expressed the fear that they might defy analysis forever, but the theory of horseshoes, and the work of Zeeman, Newhouse, and others⁶ on more general shoes, gives hope.

14.4.

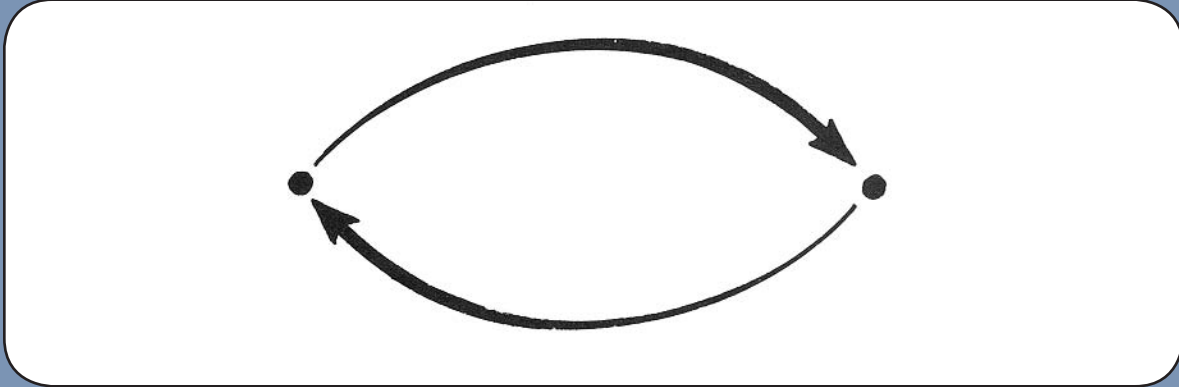
Hypercycles

An even more complicated situation occurs generically in dimension three or more. The insets and outlets of these may have transverse intersections, tangles, and heteroclinic trajectories in a daisy chain, called a *hypercycle*, or *heteroclinic cycle*.



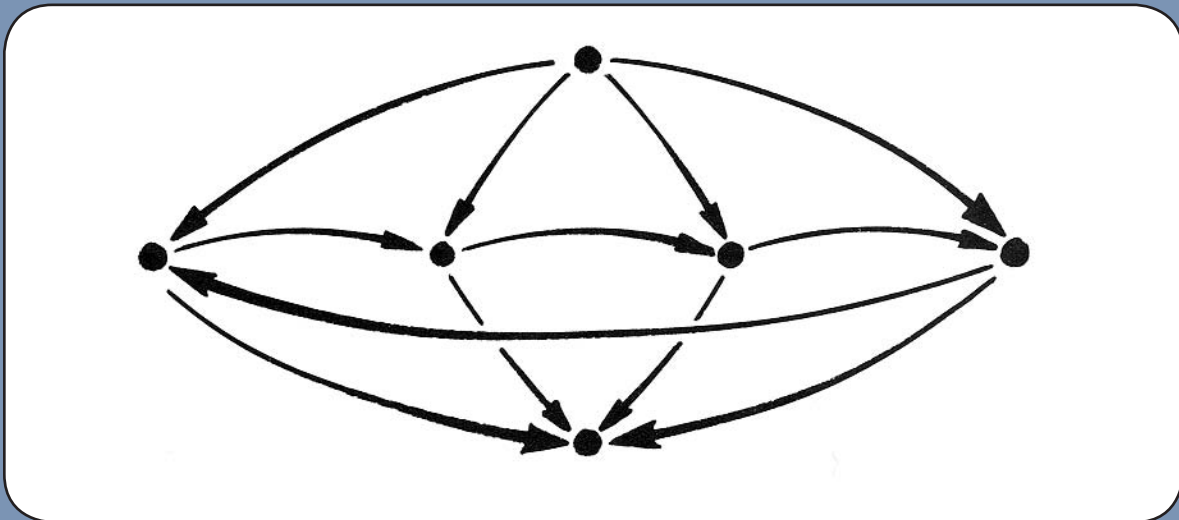
14.4.1.

Here is the simplest example of a cycle. In a three-dimensional state space, two closed orbits of saddle type (index 1) have heteroclinic trajectories, each to the other.



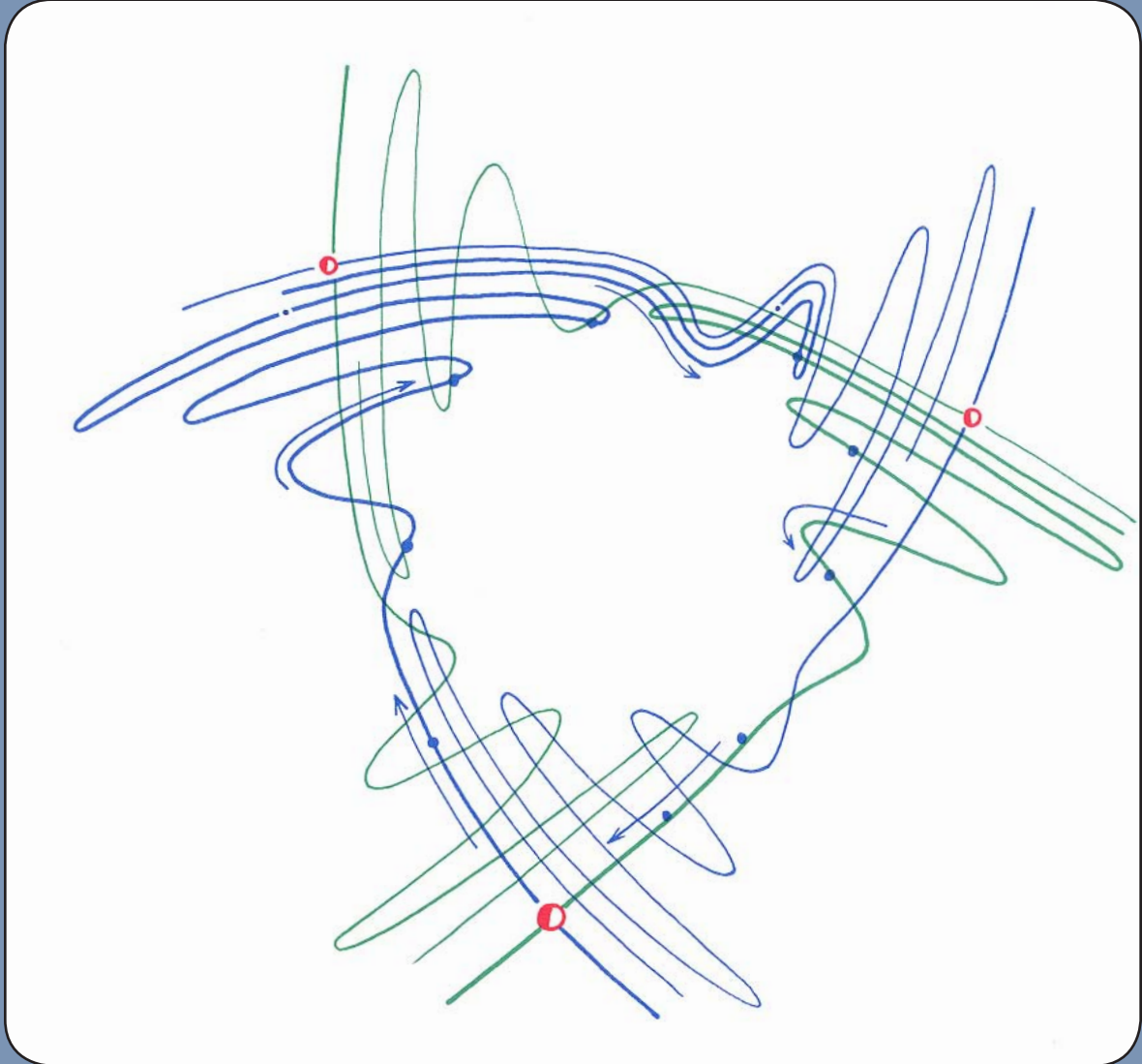
14.4.2.

This situation may be described by this diagram, called a *directed graph*, or *quiver*. This has a vertex for each of the limit cycles, and was introduced by Peixoto⁷ to describe generic systems in two-dimensional state spaces.



14.4.3.

More complicated cycles may involve larger sets of critical points, closed orbits, and even more complicated limit sets, in a daisy chain of saddle connections.

**14.4.4.**

Here is a hypercycle involving three limit cycles of saddle type. Each is heteroclinic to each of the others. In all of these situations, it can be proved, by topological analysis, that each of the limit sets involved is actually homoclinic. That is, membership in a heteroclinic cycle implies homoclinicity.

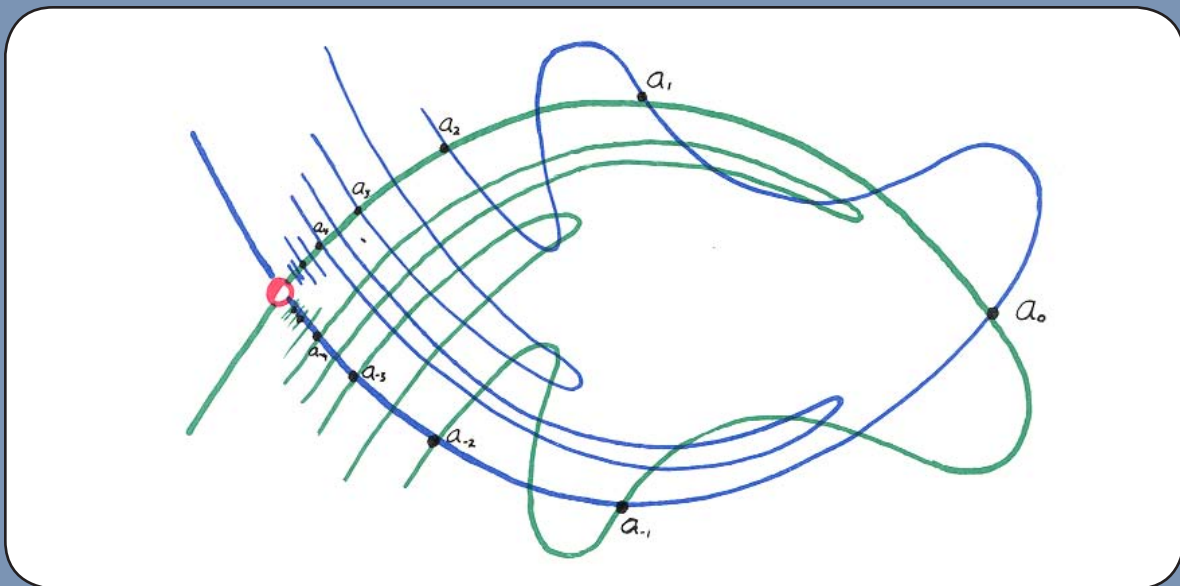
Cycles of heteroclinically related critical points are endemic in real dynamical systems, and are vitally involved in chaotic motions.

15 Nontrivial Recurrence

In the history of dynamics, as in philosophy, the concept of *recurrence* frequently recurs. A periodic trajectory has the recurrence property: every one of its states will recur again and again. This is called *trivial recurrence*. The recurrence property also applies to more complicated (aperiodic) trajectories. This is called *nontrivial recurrence*. This concept already surfaced in the generic property G4, described in Section 11.3, and in the chaotic attractors of Part Two. In this chapter, more versions of this important phenomenon will be described.

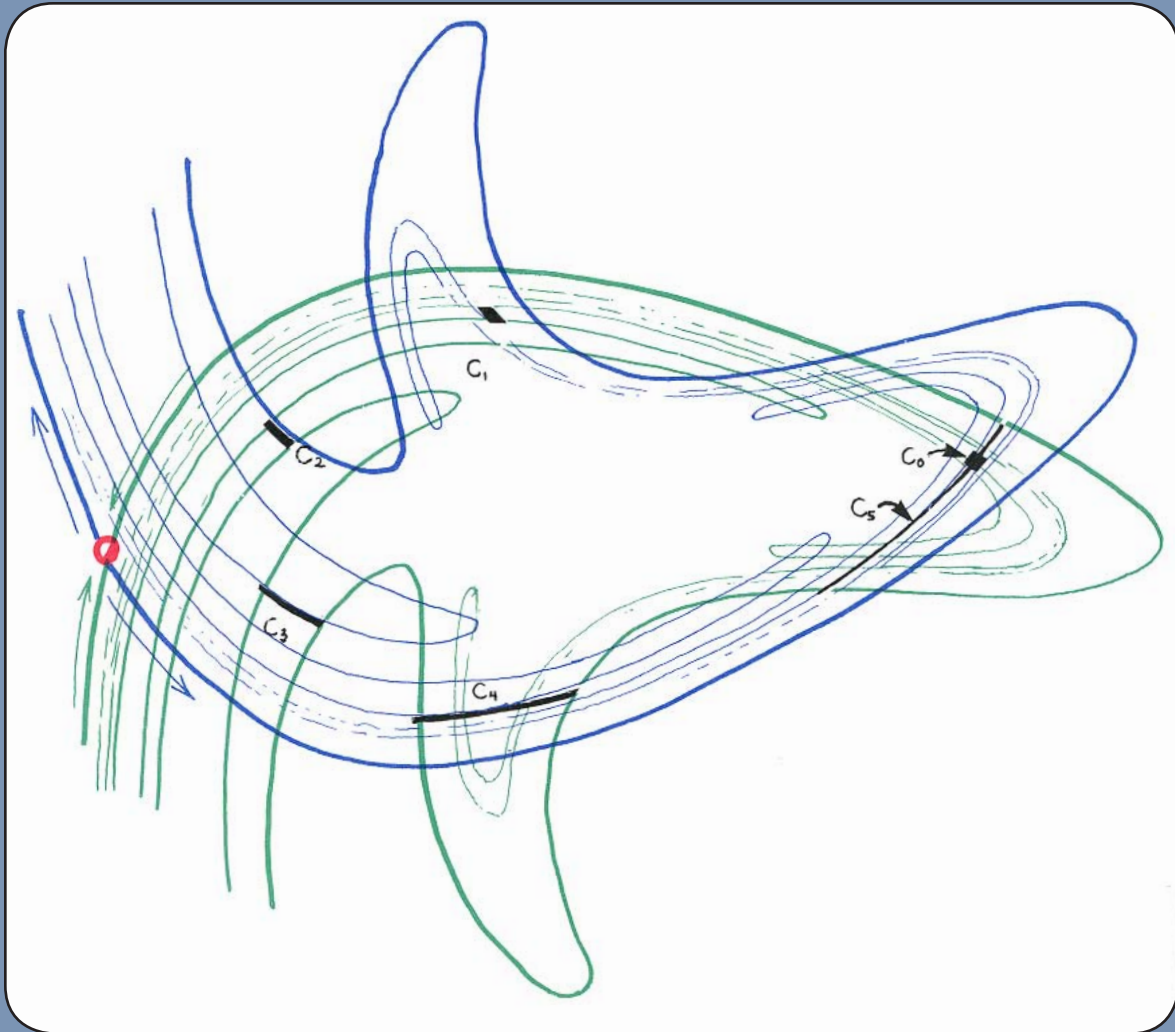
15.1. Nearly Periodic Orbits

Recall that generic property G4 limits the types of almost-periodic motions. Discovered by Peixoto in two dimensions, its genericity was established by Pugh in higher dimensions.¹ Suppose that we take a sequence of points in the state space, converging (approaching asymptotically closer and closer) to a point, and that each of the points belongs to a closed orbit (limit cycle, or periodic trajectory). Topological consequences of the generic condition G3 (transversality) force the periods of these periodic trajectories to get longer and longer. Thus, the oscillations they represent have frequencies that get lower and lower. The limit point of the original sequence lies on a trajectory that need not be periodic. But it is *nearly periodic*, in that observations cannot distinguish it from a low-frequency oscillation. We will denote the set of all nearly periodic points of the dynamical system by NP .



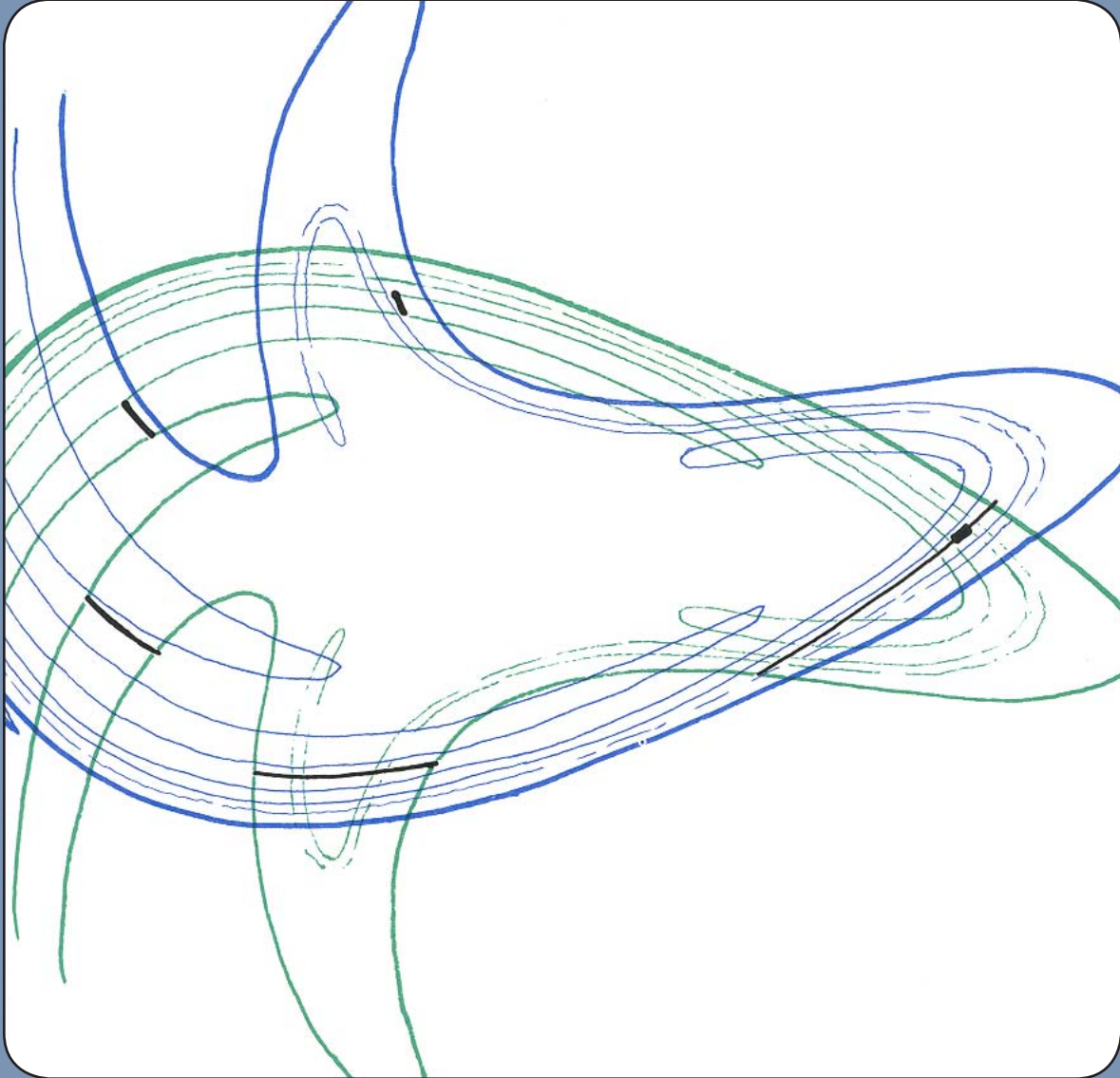
15.1.1. A homoclinic limit cycle provides good examples of nearly periodic points. Here is a Poincaré section of a homoclinic tangle. Look carefully at the trajectory of the point a_0 .

Expansion of the tangle shows how the periodic orbits fit into this picture, from the cover of Hayashi's collected works.²



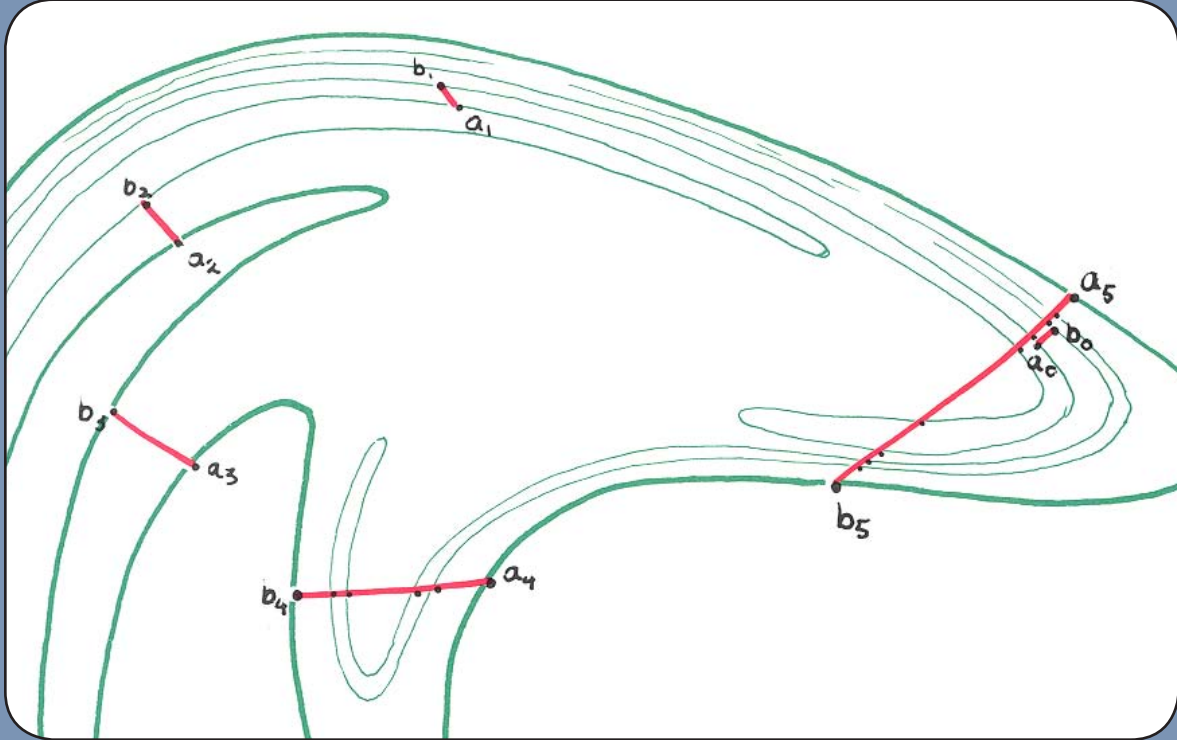
15.1.2.

Inside this tangle, there must be a periodic orbit.³ Let's follow the small red rectangle, marked c_0 . Its sides are segments of insets and outlets. After one revolution around the ring, its first return to the Poincaré section is again a small rectangle c_1 . Note that it is stretched in one direction and compressed in the other. Now follow its next five revolutions, noting that inset segments are stretched to longer inset segments, and outlet segments are compressed to shorter outlet segments. Note that c_5 intersects c_0 .

**15.1.3.**

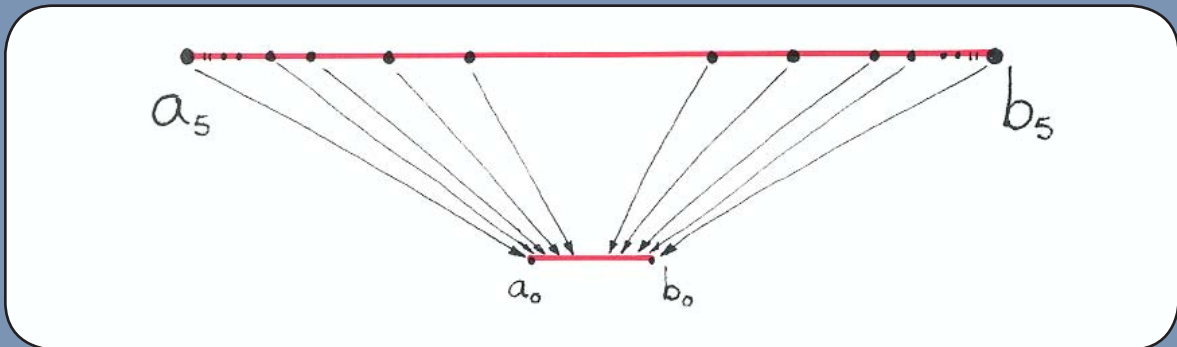
Now take the little piece of c_0 intersected by c_5 and follow it around five times. It will again pass through the initial rectangle. Continuing in this way, we obtain a sequence of nested boxes, which converge to a periodic point of period five, as predicted by the theorem of Birkhoff and Smith.⁴

We may use the expansion of the tangle as a magnifier, to zoom into the microstructure of the tangle.



15.1.4.

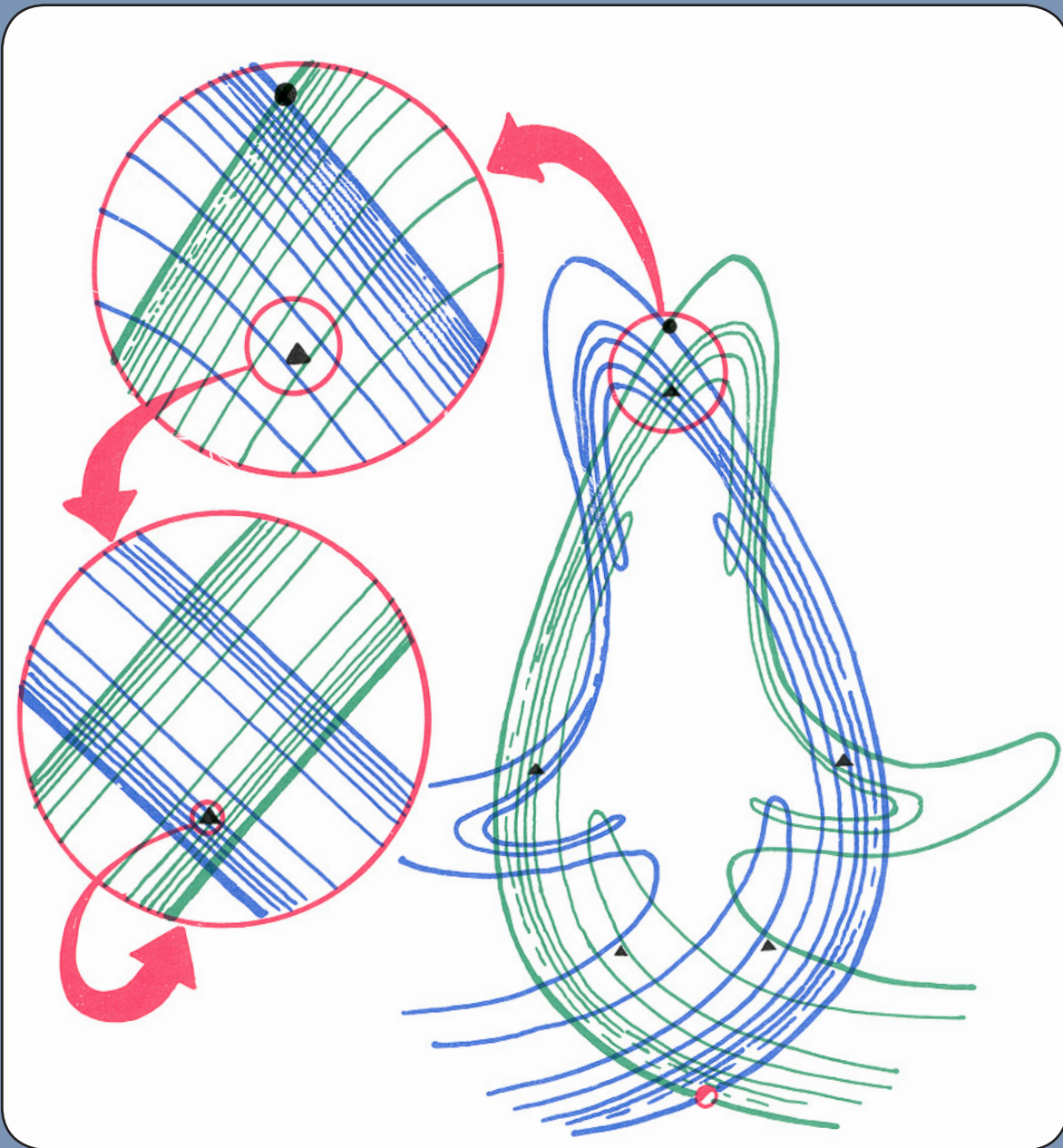
Now let's select two points a_0 and b_0 , and follow their fates. The line segment a_0b_0 becomes, after five revolutions, the segment a_5b_5 .



15.1.5.

All the intersections of the inset within this stretched segment ab , must also be found in the shorter segment ab , but they are *five generations smaller*.

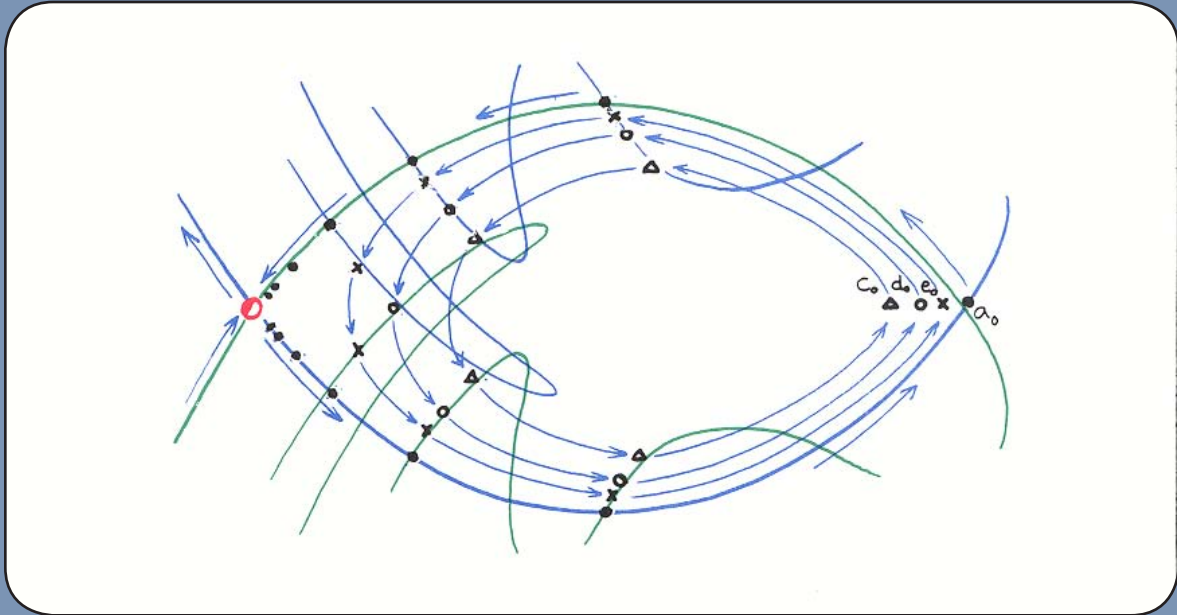
We may continue to zoom into this microscopic structure of the tangle.



15.1.6.

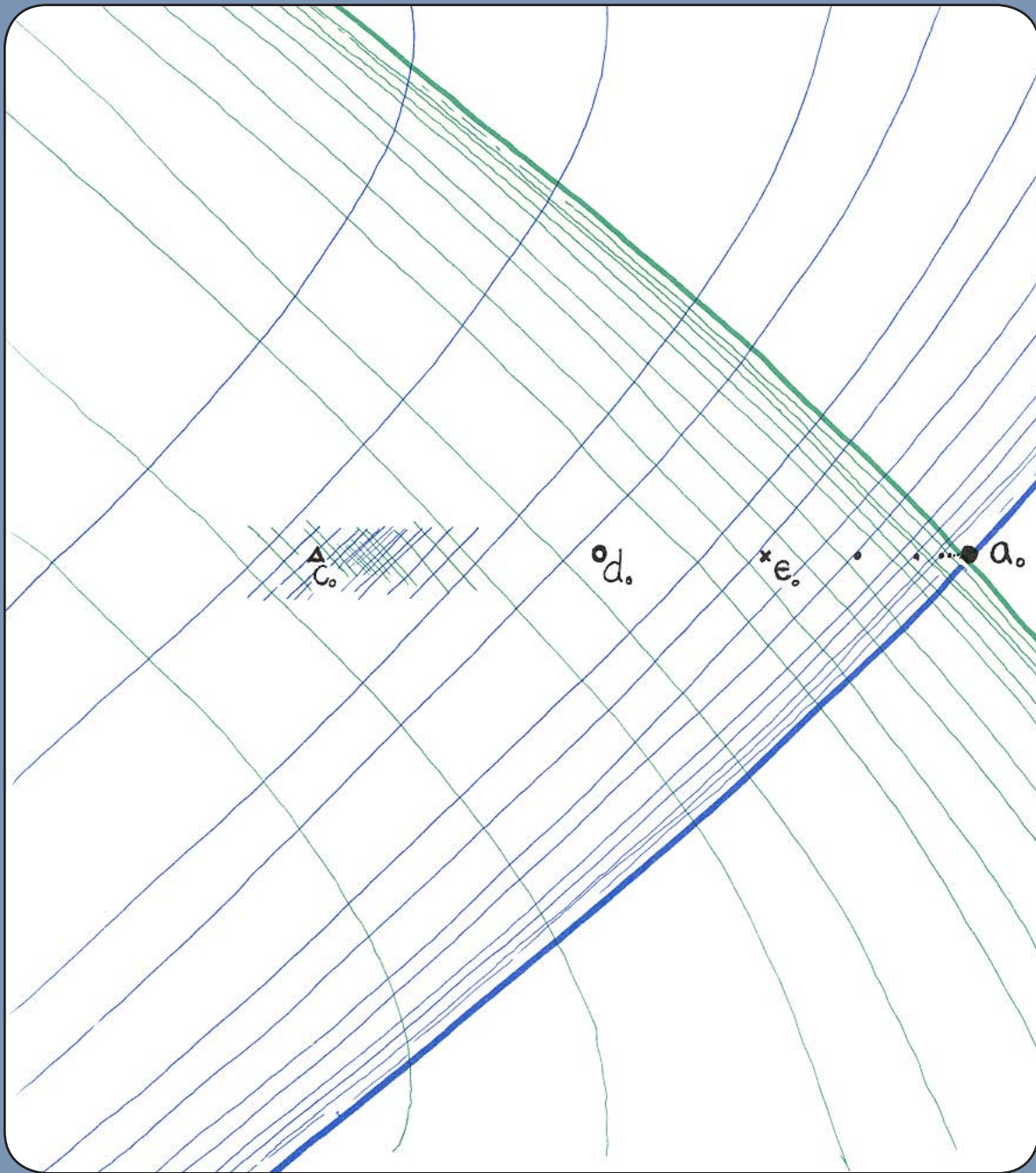
A few repetitions of the magnification method suffice to locate the periodic point as accurately as needed. It is within the small tangle.

By starting with other small rectangles and making judicious use of the zoom method, additional periodic points may be found.



15.1.7. Closer to the homoclinic point a_0 there must be another periodic point with a higher period, such as c_0 shown here. And even closer, another with an even higher period, such as d_0 . These may be located as accurately as needed by the zoom method described above.

Thus the original homoclinic point is the limit of a sequence of periodic points in the Poincaré section. In the three-dimensional state space, a sequence of closed orbits (periodic trajectories) asymptotically approach the homoclinic trajectory. Thus, every point on the homoclinic trajectory is *nearly periodic*, yet not periodic.



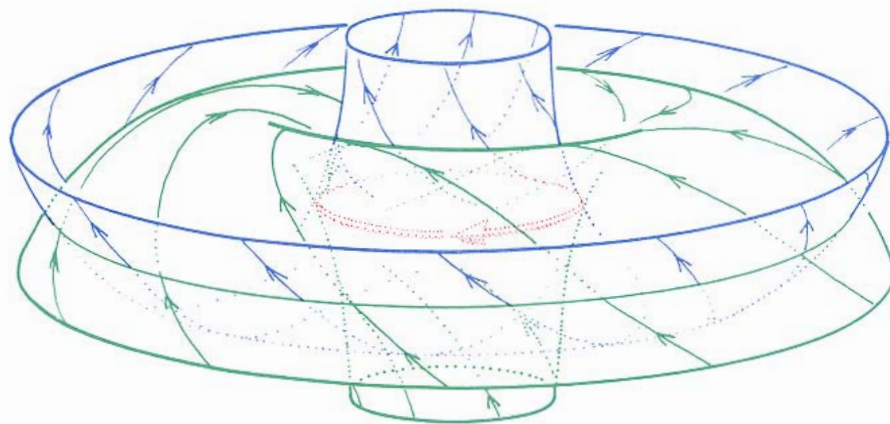
15.1.8.

Here, highly magnified, is a sequence of periodic points approaching closer and closer to a homoclinic point, which is nearly periodic, yet not actually periodic.

15.2.

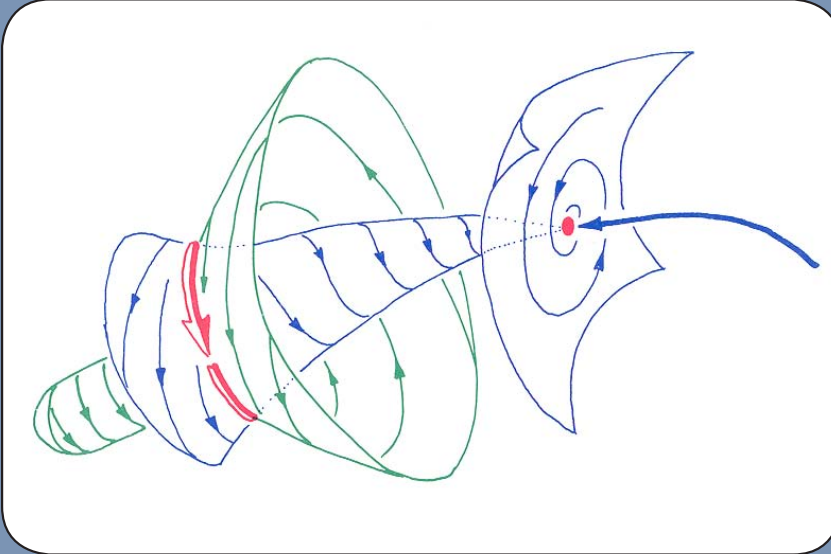
Why Peixoto's Theorem Failed in 3D

As described in Section 12.2, Peixoto's theory of structurally stable systems is restricted to the two-dimensional case. In the case of state spaces of three dimensions or more, it is still true that structurally stable systems must have the four generic properties: G1, G2, G3, and G4. But these conditions no longer ensure structural stability. In fact, structurally stable systems are rare (that is, hard to find) in higher dimensions. A complete characterization of structural stability in three-dimensional systems (having a global section) has been accomplished recently.⁵ This section describes the failure, and the remnants of Peixoto's theory that apply in higher dimensions.

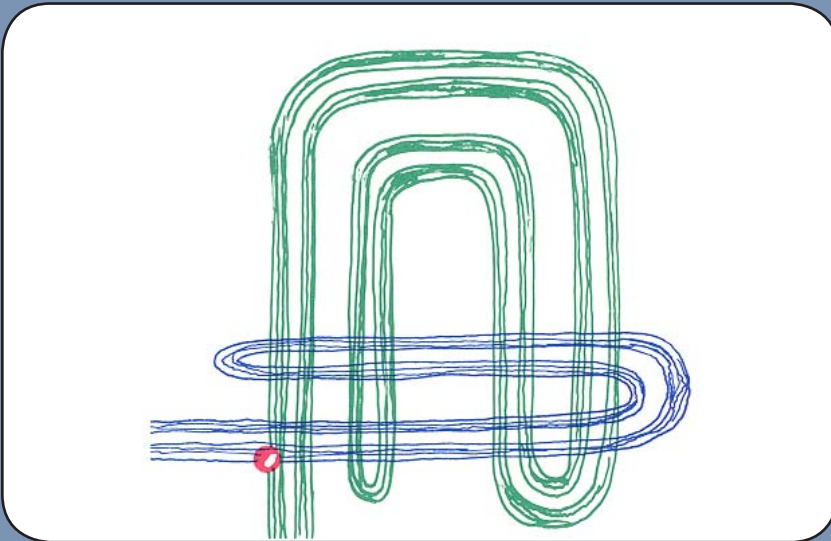


15.2.1.

It is Step 1 in Peixoto's proof which is specifically two-dimensional. That step established that there are only a finite number of closed orbits (limit cycles) in the two-dimensional case. Here is an example of a *generic* portrait in three-dimensions. The homoclinic tangle forces the occurrence of an infinite number of limit cycles, as described in the preceding section. This example makes Step 2 wrong as well, as Step 2 is a simple consequence of Step 1.



15.2.2. Step 3 remains true in higher dimensions. It assumes *property F*: the limit sets consist of a finite number of limit points and limit cycles only, as well as the four generic conditions. These are sufficient to ensure structural stability. This is a difficult result, due to Palis and Smale.⁶ Here is an example of such a portrait, in three dimensions.



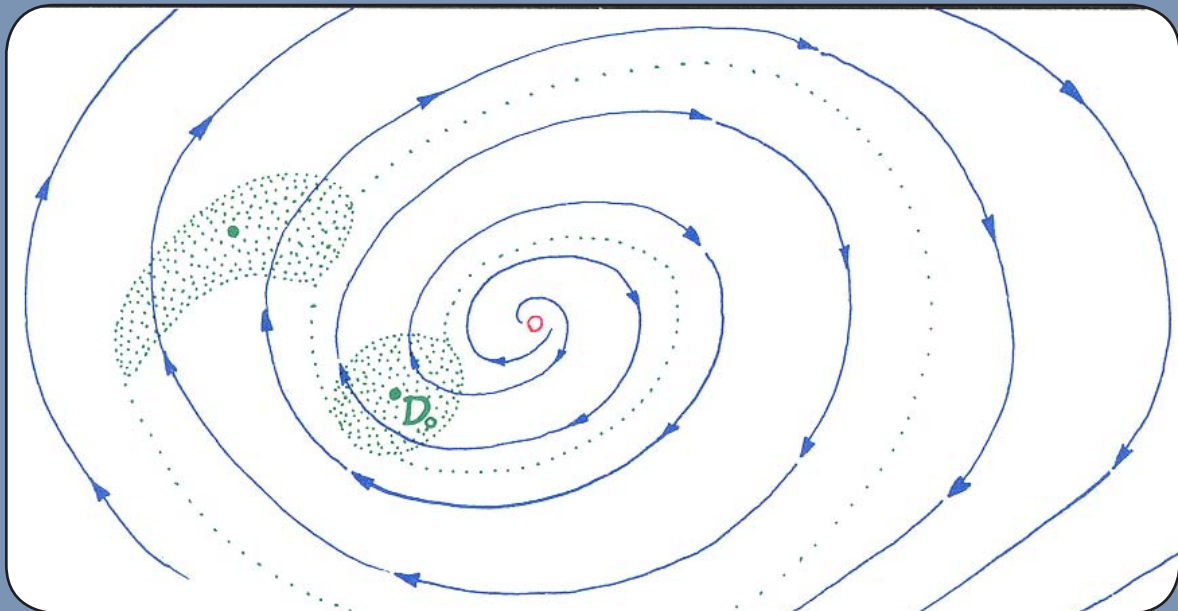
15.2.3. Step 4 fails in higher dimensions. Structural stability does ensure the four generic conditions. This is a relatively easy result, due to Markus and Robinson.⁷ But structural stability does not ensure property F, the finiteness of the limit sets. The generic homoclinic tangles can be structurally stable, as Smale has shown for the example shown here.⁸

The progress of dynamical systems theory stalled briefly at this point, until it occurred to Smale to regard a homoclinic tangle as a generalized limit cycle and propose generic properties for it as a unit. He called this a *basic set*. The main example is the horseshoe, dissected in the preceding chapter. This was a prototype for the chaotic attractors, described in Part Two. One of the fundamental properties of a basic set is nonwanderingness, described in the next section.

15.3.

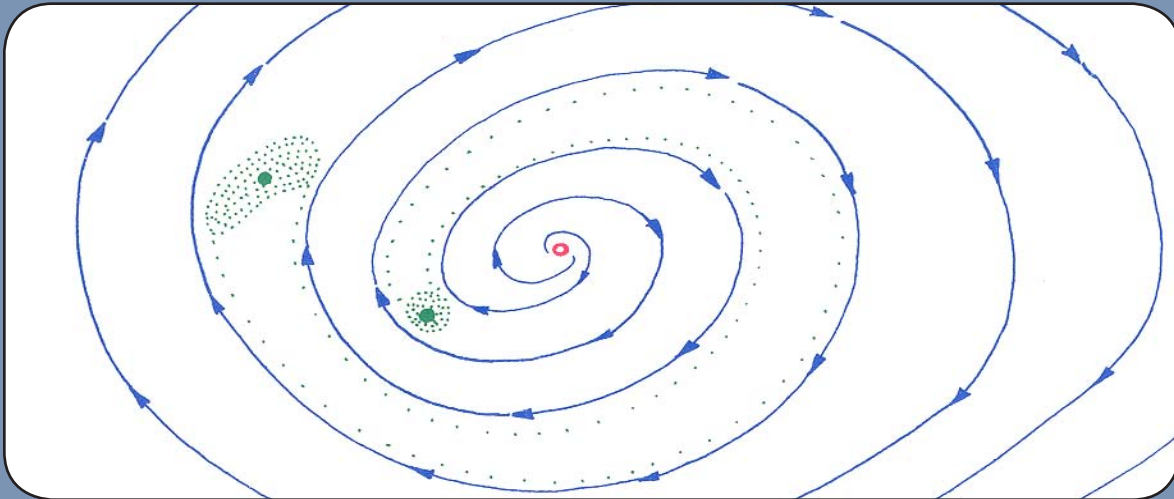
Nonwandering Points

One of the most restrictive versions of the recurrence property is near-periodicity, defined above, in Section 15.1. And one of the least restrictive versions is the property of *nonwandering*, defined in this section.



15.3.1.

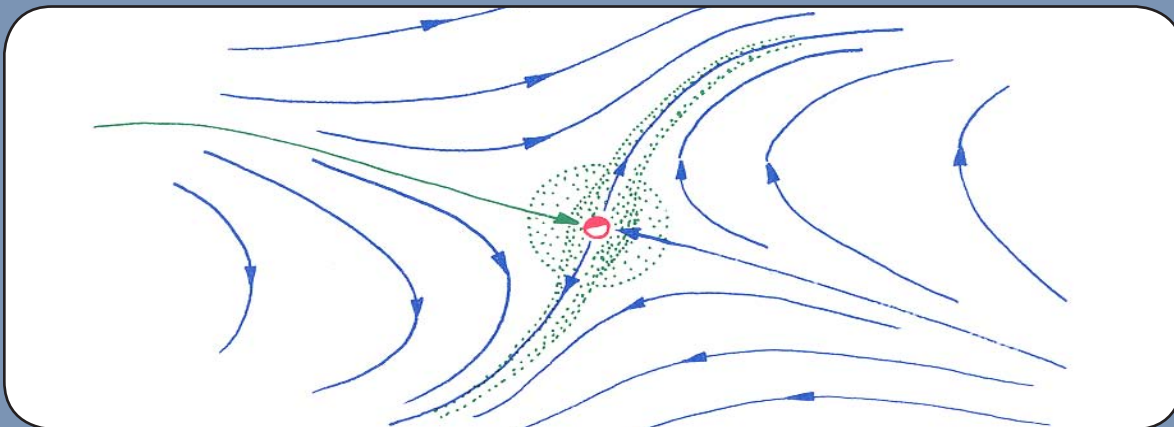
Suppose, having picked out a point in the state space and a little disk centered on it, that we follow the future meandering of the entire disk. If wide enough, it may meet up with itself along its meander.



15.3.2.

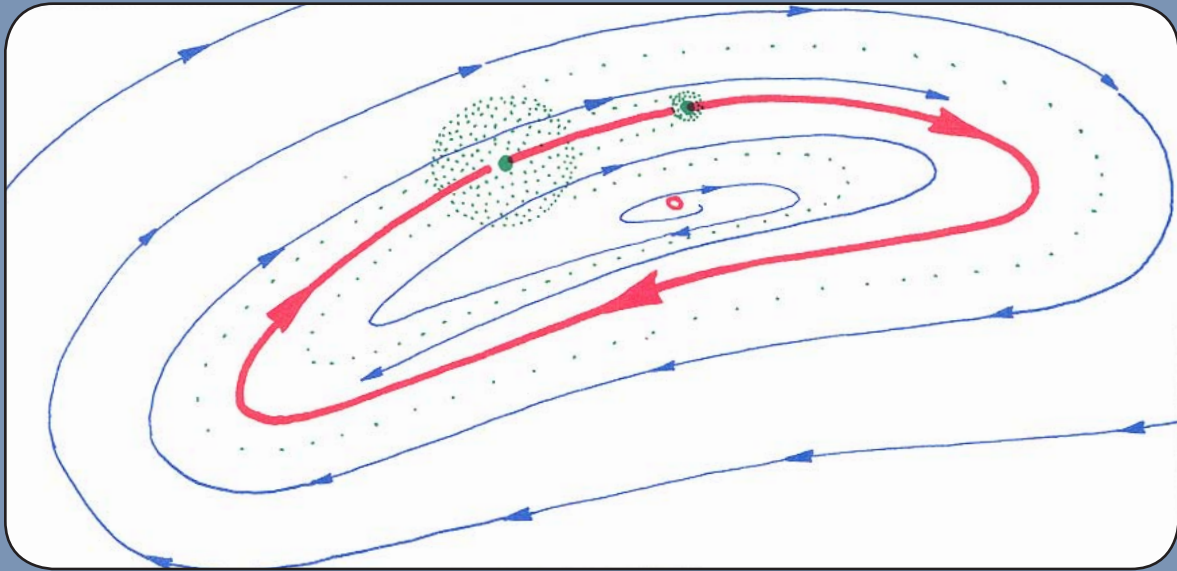
If so, start with a smaller disk, and repeat the construction. If now the meandering disk leaves its original position, wanders away, and never returns to overlap its original position, then the original point at the center of the disk is called a *wandering point*.

On the other hand, it may happen that, no matter how small you draw the original disk, it always comes back to overlap itself. Or, it may never cease overlapping itself, no matter how long you wait. In these cases, the original point is a *nonwandering point*. The set of all nonwandering points of a given dynamical system will be denoted by NW .



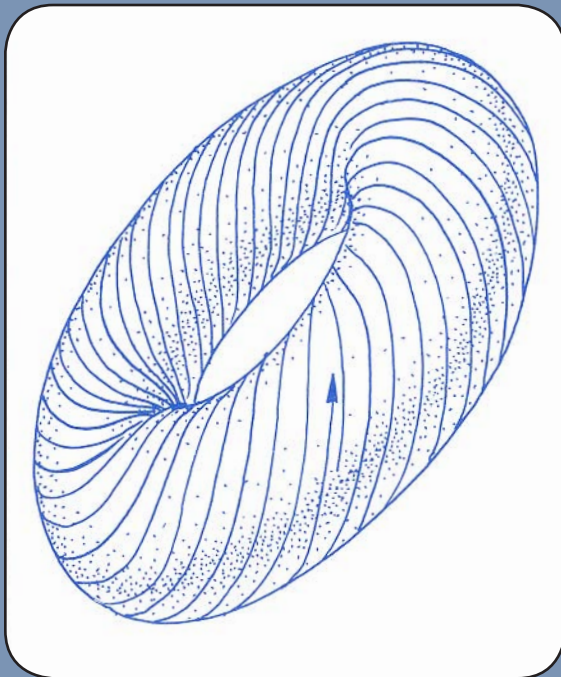
15.3.3.

For example, a limit point (equilibrium) is nonwandering. The little disk is tied down at the center.



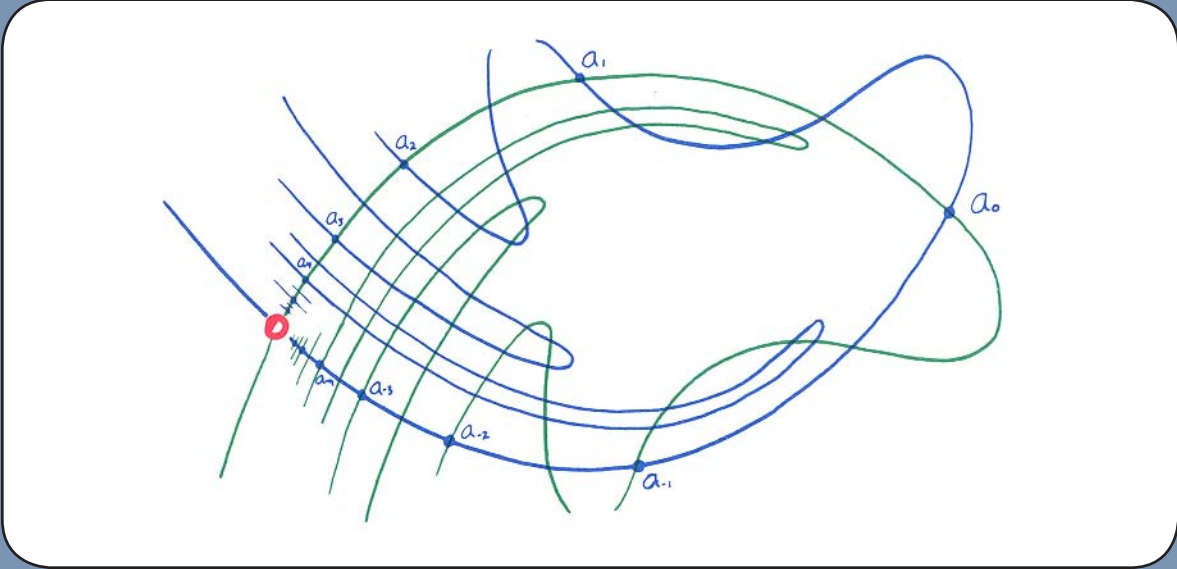
15.3.4.

Similarly, a closed orbit (limit cycle) is nonwandering. The center of the little disk keeps passing through the initial point, again and again. In fact, the set of nearly-periodic points, NP , is contained in the set of nonwandering points, NW , for topological reasons.



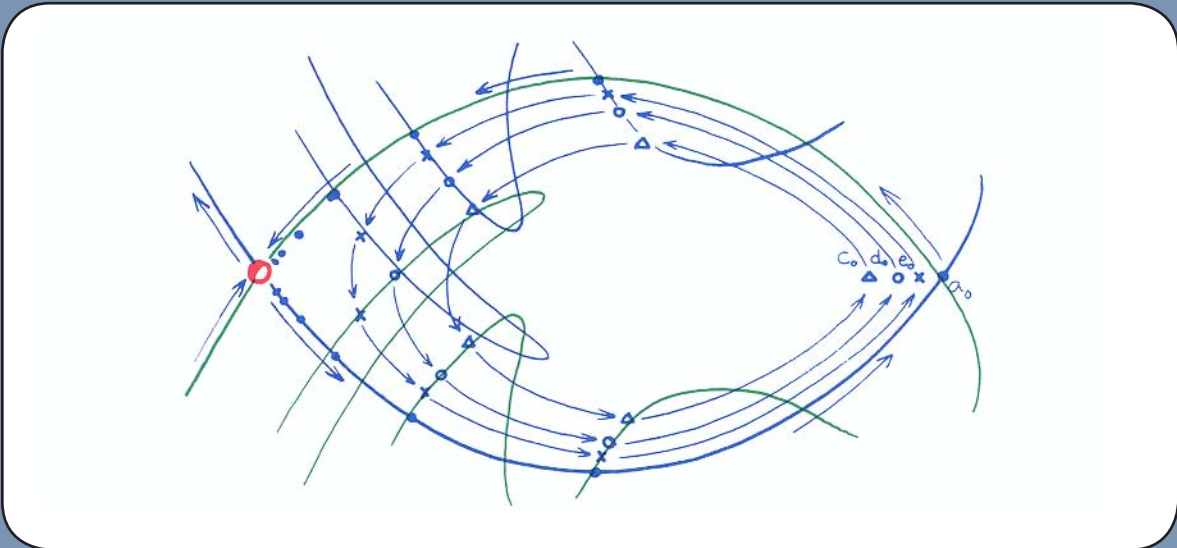
15.3.5.

Here is an outstanding example of a nonwandering point which is *not* nearly periodic. In this solenoidal flow on the torus, called a *Kronecker irrational flow*, every point is nonwandering, yet no point is periodic, or even nearly periodic.



15.3.6.

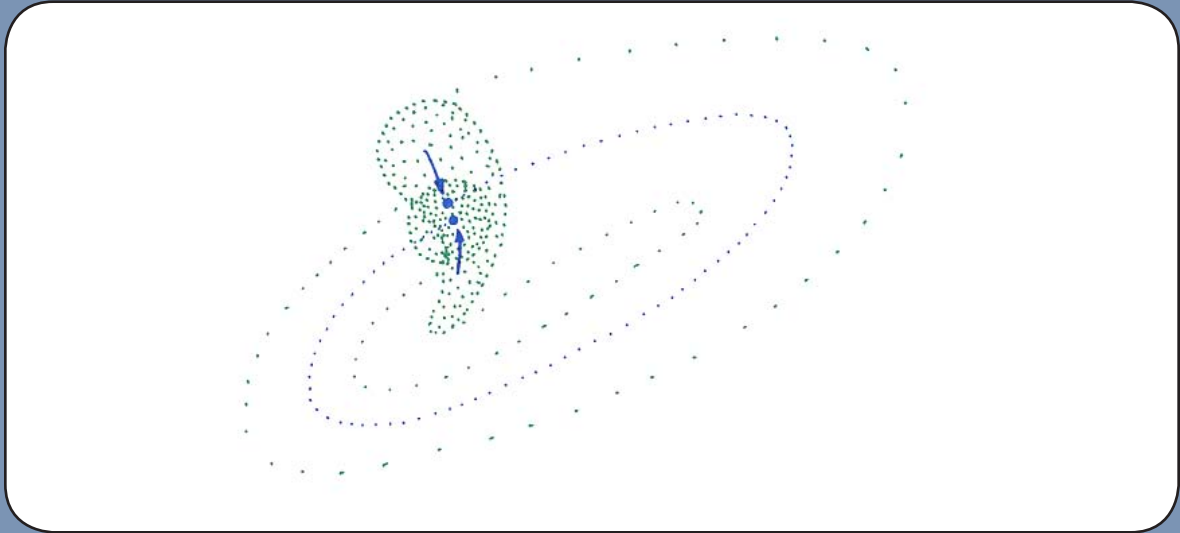
This is an example of a nonwandering point which is not itself recurrent in any sense. The flow has a limit cycle of saddle (index 1) type, which is homoclinic, and satisfies G3 (transversal intersection). The heteroclinic trajectories within this tangle are nonwandering.



15.3.7.

The theorem of Birkhoff and Smith, later generalized to higher dimensions by Smale, shows that these trajectories are nearly periodic. That is, they are approximated by limit cycles of very low frequencies. The heteroclinic trajectories belonging to a heteroclinic cycle of tangles are also nearly periodic.

Generic property G4, discussed previously in Section 11.4, can now be simply stated: $NP = NW$. That is, a dynamical system has property G4 if its every nonwandering point can be approximated by periodic points (points belonging to limit cycles). This property is generic, as proved by Peixoto (in 2D) and Pugh (in higher dimensions).

**15.3.8.**

The proof of the genericity of this property is intuitively simple, yet it is one of the most difficult in the whole literature of mathematical dynamics to carry out in detail. The key step, called the *Closing Lemma*, makes small changes in the vectorfield, so that a closed orbit is found in the disk that meets itself.

Warning: As described briefly in Chapter 12, this property is generic only in a very weak sense. The reason is that the violation of G4 by persistent solenoidal flows (equivalent to irrational Kronecker flows on invariant tori) occurs with positive expectation. Thus, in the sense of probability, G4 violation is also generic. We may call this the G4 paradox. It will be explained further in Part Four.