Complicated dynamics from simple topological hypotheses

R.S.MacKay
Mathematics Institute
University of Warwick
Coventry CV4 7AL, U.K.

May 18, 2001

Abstract

There are many types of dynamical system for which quite simple topological hypotheses imply very complicated behaviour. A well known example is "Period 3 implies chaos" for continuous maps of an interval. A diverse range of more sophisticated examples has been produced by many researchers and is reviewed here. The relevance to physical applications, like mixing in fluid flows, chaotic motion of mechanical linkages, and fast dynamos, is explored.

1 Introduction

Fig 1 shows the action of a taffy pulling machine. The four prongs perform a periodic motion about each other in a plane, but they stretch any material connecting them. By forming the directed graph of Fig 2, indicating over which of the three regions – left arm L, right arm R, and in the middle M – the taffy in each region is forced to get stretched in one revolution, it can be shown that the total length of the material after n revolutions is at least $C\lambda^n$, for some $C > 0$, with $\lambda = 3 + \sqrt{8}$. Furthermore, if one were to examine in which of the three regions a particular particle lies at each time the machine returns to its initial configuration, one would find that most perform random-looking sequences. In fact, for each path in the graph of Fig 2, there is a particle (imagining the material to be infinitely divisible) which follows that sequence (multiple edges between two vertices correspond to different routes between the regions). Thus the taffy machine is a good mixer, at least on a macroscopic scale.

Figure 1: The action of a taffy pulling machine (following Rössler, 1983).

Figure 2: The directed graph of forced transitions for taffy particles.

The first aim of this article is to survey the growing body of mathematical results of this form, namely relatively simple topological hypotheses that imply complicated dynamics. We will understand "complicated dynamics" in the particular sense of existence of a trajectory for each path in some graph with a connected branched subgraph.

The second aim is to discuss the potential relevance of such results to physics. MacKay (1990) raised the question whether they could be useful to fluid mechanics or magnetohydrodynamics. There are now some suggestions that it can. Applications to mechanical linkages will also be given.

The paper begins by reviewing some simple examples of dynamical systems with complicated dynamics. Then some results are stated showing that any dynamical system possessing the same key topological features of these examples has at least as complicated dynamics. Next, the particular class of surface homeomorphisms is reviewed, including an explanation of the above claims for the taffy machine and some interesting work on periodic stirring protocols. Following this, results on geodesic flows are reviewed, and an application given to a mechanical linkage. Finally, the extent to which the results of Section 4 could generalise to 3D flows is addressed, and a few remarks made on dynamos.

There is also a large class of results (going back to Poincaré) that I will not review, where some more local feature implies complicated dynamics. This can occur in discrete-time dynamical systems with a non-invariant subset that maps over itself in a certain way, or in flows with a subset with certain disposition of entry and exit sets. The reader is referred for example to Burns & Weiss (1995), Conley (1978), Kennedy et al (2000), and Mischaikow & Mrozek (1994).

2 Some simple models with complicated dynamics

Let us begin by reviewing some simple examples of dynamical systems with complicated dynamics.

2.1 Doubling the angle on a circle

Think of the circle as $\mathbb{T} = \mathbb{R}/\mathbb{Z}$, and consider the map $g : \mathbb{T} \to \mathbb{T}$,

$$g(x) = 2x. \tag{1}$$

It is equivalent to the left shift on binary expansions $0.a_1a_2a_3\ldots \mapsto 0.a_2a_3\ldots$, where $a_j \in \{0, 1\}$, allowing for the usual ambiguities for dyadic rationals. Thus it has many periodic orbits (just choose periodic binary sequences) and

Figure 3: A tight interval map for the period-3 permutation (012).

dense orbits (just choose binary sequences containing every finite word). The dynamics is as random as a sequence of coin tosses: code the sequence as 0s and 1s, take the corresponding binary number as initial condition on the circle, and observe in which half of the circle – (0,1/2) or (1/2,1) – successive iterates lie, to recover the sequence of coin tosses.

2.2 Tight interval maps

Given a permutation π of $\{0, \dots, N\}$, define a map f of the interval $[0, N]$ to itself by joining the graph of the permutation with straight lines. Construct a graph with vertices $\{1, \dots, N\}$ representing the consecutive intervals between the integer points, and a (directed) edge from i to j if the interval j is in the image of the interval i. For every path in the graph there exists a point whose orbit under f follows the sequence of intervals specified by the path. If the graph has a connected branched subgraph then the map has complicated dynamics. An example is sketched in Figure 3.

Tight circle maps can be defined similarly. Take a permutation π of $\{0, 1/N, \dots, (N-1)/N\} + \mathbb{Z}$ which commutes with $x \mapsto x+1$. Then join its graph with straight lines and reduce modulo $\mathbb{Z}$, to obtain a degree-1 circle map. Construct the graph whose vertices represent the arcs between successive points n/N and edges allowed transitions. If the graph has a connected branched subgraph then the map has complicated dynamics.

2.3 Hyperbolic toral automorphisms

Any 2×2 integer matrix A with determinant ± 1 induces a diffeomorphism of the 2-torus $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$: $x \mapsto Ax \mod \mathbb{Z}^2$. Regarding the torus as a group (under addition), the map is an automorphism. If A is hyperbolic (no eigenvalues on the unit circle) then the map has complicated dynamics in a strong sense. First one can choose a basis in which all components of A are non-negative. Then there is a finite cover of $\mathbb{T}^2$ by rectangles R_e with disjoint interiors, one for each edge e in the graph Γ with vertices 1 and 2 and A_{ij} edges from j to i, such that for every path $\mathbf{e} = (e_n)_{n\in\mathbb{Z}}$ in Γ there is a point whose orbit follows the sequence (R_{e_n}) (Manning, 1988). Fig 4 shows an example of such a partition for the map of $\mathbb{T}^2$ induced by

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}. \tag{2}$$

A finite cover of the state space X by subsets with disjoint interiors, each equal to the closure of its interior, such that for each path $\mathbf{e}$ in an associated graph there is a unique point $h(\mathbf{e})$ whose orbit follows that path, and every

Figure 4: Lift to the universal cover of a Markov partition for the map (2) on $\mathbb{T}^2$ (the dots represent the integer lattice), and associated graph.

Figure 5: Lift to the universal cover of a Markov partition for the map (2) on $\mathbb{S}^2_4$ (circles represent the integer lattice and dots the half-integer points).

point occurs as $h(\mathbf{e})$ for at least one path $\mathbf{e}$, is called a *Markov partition*. Each of the partitions discussed in this section is Markov.

Hyperbolic toral automorphisms are examples of *Anosov* systems (those for which the whole manifold is "uniformly hyperbolic").

2.4 Pseudo-Anosov maps on $\mathbb{S}^2_4$

Hyperbolic toral automorphisms have nice analogues on the sphere. Given 2×2 integer matrix A with determinant ± 1, let G be the group generated by $\mathbb{Z}^2$ and reflection through the origin (or more generally through some point z such that $(A - I)z \in \frac{1}{2}\mathbb{Z}^2$). Then $\mathbb{R}^2/G$ is a sphere with four "conical" points, which I denote $\mathbb{S}^2_4$, and A induces a homeomorphism of this sphere (mapping the set of conical points to itself). If A is hyperbolic then the induced map on the sphere again has complicated dynamics and is an example of a class of maps called "pseudo-Anosov". Fig 5 shows an example of a Markov partition for (2) and the associated graph.

3 Some sample results

Any degree-2 circle map $f : \mathbb{T} \to \mathbb{T}$ possesses all the dynamics of $g : x \mapsto 2x \mod \mathbb{Z}$. The precise statement is that f is semiconjugate to g. A continuous map $f : X \to X$ is *semiconjugate* to another one $g : Y \to Y$ if there exists a continuous surjection $h : X \to Y$ such that $h \circ f = g \circ h$ (they are *conjugate* if h has the stronger property of being a homeomorphism).

Here is a simple proof. Work with lifts to $\mathbb{R}$. Write $f(x) = 2x + \phi(x)$, with ϕ continuous and periodic. Look for a semiconjugacy in the form $h(x) = x + \psi(x)$ with ψ periodic. Then $h \circ f = g \circ h$ iff $2\psi(x) = \phi(x) + \psi(f(x))$. This is solved by $\psi(x) = \sum_{n=0}^{\infty} 2^{-(n+1)}\phi(f^n(x))$, which converges in C^0, making h a continuous surjection (actually, if f is C^1 with $f' > 1$ then h is also one-to-one and hence f and g are conjugate).

One might expect that the above would generalise to all continuous maps of closed orientable manifolds with degree greater than 1 in absolute value, but it does not, unless the map is assumed to be C^1 (see Ch 8 of Katok & Hasselblatt (1995) for a discussion). A similar result does hold, however, for hyperbolic mapping classes on $\mathbb{T}^2$ (and $\mathbb{T}^d$, $d \geq 2$): if $f : \mathbb{T}^d \to \mathbb{T}^d$ is isotopic to a hyperbolic automorphism A, then it is semiconjugate to A (Franks, 1970).

There is an important generalisation of how one dynamical system f may possess all the dynamics of another one, g. The map f may possess an invariant subset Λ such that $f|_\Lambda$ is semiconjugate to g. For example, if $f : I \to I$ is a continuous map of an interval I into itself, and possesses a periodic orbit of (least) period 3, then there is an invariant subset $\Lambda \subset I$ such that $f|_\Lambda$ is semiconjugate to the tight map of Figure 3. This is the basis for the "Period 3 implies chaos" result of Li & Yorke (1975) (note that the principle and many consequences had already been worked out by Sharkovskii (1964)).

4 Surface homeomorphisms

Work of Nielsen and Thurston has led to a beautiful and powerful classification of homeomorphisms of compact surfaces up to isotopy (for a clear review, see Boyland, 1994). Two homeomorphisms $f, g : M \to M$ are *isotopic* if there exists a continuous $F : [0,1] \times M \to M$ such that $F_0 = f$, $F_1 = g$ and for all $\lambda \in [0,1]$, F_λ is a homeomorphism of M. The theory also applies to compact surfaces with a finite set P of "marked" points, meaning that P is required to be invariant under each F_λ in an isotopy (including its end points).

The result is that every surface homeomorphism f is isotopic to one g which is either

finite-order: for some $n \geq 1$, $g^n = id$;

pseudo-Anosov: there is a $\lambda > 1$ and a transverse pair of invariant foliations with finitely many singularities modelled on the levels of the real and imaginary parts of $z^{k/2}$, for $z \in \mathbb{C}$, $k \in \mathbb{N} \setminus \{2\}$, $k = 1$ permitted only at marked points, or on the levels of the imaginary parts of (z, z^2) near a boundary $\Im z = 0$, with at least one pair per boundary component, carrying transverse measures which are respectively contracted/expanded by λ by g;

reducible: there is a finite invariant set of disjoint essential annuli ("essential" means not homotopic on $M \setminus P$ to a point or a boundary component), such that $g^n(A)$ is a Dehn twist ($y' = y \in [-1, +1], \theta' = \theta + \Omega(y) \in \mathbb{T}$, with $\Omega(\pm 1) \in \mathbb{Z}$) on each annulus A, where $n(A)$ is its period, and on each component C of the complement, $g^n(C)$ is finite-order or pseudo-Anosov.

Pseudo-Anosov maps have a Markov partition and the associated graph is connected and branched. A simple example is the map on $\mathbb{S}^2_4$ of Section 2.4. Actually, every pseudo-Anosov map of $\mathbb{S}^2_4$ is conjugate to one of

Figure 6: A Markov partition for the pseudo-Anosov representative of all taffy machine maps (thick lines indicate the partition element boundaries, some of which are internal), and part of the unstable foliation (thin curves).

Figure 7: Simplest representative for egg whisk maps (the annuli contain differential rotations interpolating between rest on the outside and period-4 motion inside).

the examples of Section 2.4 (Llibre & MacKay, 1992). The two foliations are those induced by the foliations of $\mathbb{R}^2$ by straight lines of slopes $1/\gamma$ and $-\gamma$, where $\gamma = (1+\sqrt{5})/2$. Furthermore, any map isotopic to one with a pseudo-Anosov component has exponential growth on the corresponding part of the fundamental group. Even better, there is an invariant subset Λ on which the map is semiconjugate to the pseudo-Anosov map.

For example, any map of $\mathbb{R}^2$ produced by one revolution of the taffy machine is isotopic to a certain pseudo-Anosov map on D^2_4, the disk relative to a set of four marked points corresponding to the prong positions at the initial time. The exponential growth on the fundamental group is easily understood by computing the matrix which gives the number of times that the images of the intervals L, M, R between the prongs cover each other. It has largest eigenvalue $3+\sqrt{8}$, hence all essential curves grow at least at this rate.

The pseudo-Anosov map turns out to be the map of $\mathbb{S}^2_4$ induced by $\begin{bmatrix} 3 & 2 \\ 4 & 3 \end{bmatrix}$ on $\mathbb{R}^2$, modulo the group generated by integer translations and reflection through the origin, with the point $(1/4, 0)$ blown up to the boundary of the disk. To find this, I used inspired guesswork (the Euler-Poincaré formula for singularities of foliations on D^2_4 allows just two options for a pseudo-Anosov map: either the boundary can be collapsed to a regular point, or the boundary has one prong and there is an interior 3-prong; in the first case, the map has a ramified cover by some toral automorphism, so I tried some until I found one which had the right action on essential curves), but there is an algorithm of Bestvina and Handel, which has been implemented efficiently on computer by T.Hall (Liverpool). Fig 6 shows a Markov partition for this map, chosen to match the L, M, R intervals. By the semiconjugacy result, for every path in the transition graph for the pseudo-Anosov, there is a point whose orbit follows that path.

It is instructive to contrast the case of (a two-dimensional idealisation of) an egg-whisk. It produces a map of a disk with eight marked points, forming two period-4 orbits. This map has reducible isotopy class with only finite-order components, however, sketched in Fig 7. So no great stretching or complicated dynamics can be deduced by these purely topological con-

Figure 8: Evolution of dye lines under stirring protocols with finite order (left column) and pseudo-Anosov (right column) isotopy class at times (a-g): 0-6, (h) 9 (from Boyland et al, 2000).

siderations (indeed, I think the function of an egg-whisk is not to mix but rather to entrain air).

Boyland, Aref & Stremler (2000) performed a very nice experiment, comparing two different but equally simple time-periodic 2D stirring protocols, with pseudo-Anosov and finite-order isotopy classes, respectively. They demonstrated experimentally that the pseudo-Anosov protocol is a much more efficient mixer than the finite-order one (Fig 8). Although the Thurston theory allows the complicated dynamics to occur on a set of measure zero, the experiment suggests that it occurs on one of positive measure for this Stokes flow and an intriguing project of Boyland's is to seek to explain this. Nevertheless, there are probably little elliptic islands embedded in the chaos, because the map is close to differential rotations near the boundaries.

As a final example of the results that can be obtained from Thurston-Nielsen theory, if a homeomorphism of $\mathbb{T}^2$ isotopic to the identity has three periodic orbits with non-collinear rotation vectors then it has orbits with all rotation vectors in their convex hull, because we showed that the map on the complement has pseudo-Anosov isotopy class (Llibre & MacKay, 1991).

5 Geodesic flows and mechanics

The special class of geodesic flows has a long history and many results. A *geodesic flow* is the flow on the unit tangent bundle SM of a Riemannian manifold M whose trajectories have unit speed and zero curvature (equivalently, every short enough piece of trajectory minimises the length between its end points).

Every compact orientable surface M of negative Euler characteristic supports a Riemannian metric of constant negative curvature. The geodesic flow of such a metric is chaotic in a clear sense: there is a choice of generators and relations for the fundamental group $\pi_1(M)$ and choice of fundamental domain in the universal cover, such that for every path in the Cayley graph of $\pi_1(M)$ with no back-tracking nor words corresponding to more than half a relation, nor certain words corresponding to the concatentation of parts of relations starting and ending with a half relation, there exists a unique trajectory which makes that sequence of moves of fundamental domain in the universal cover (for a review, see Series, 1991). Fig 9 shows such a fundamental domain for a closed surface of genus two. All constant negative curvature metrics on the same surface give equivalent geodesic flows.

Morse (1924) proved a result that implies (after a little more work, in

Figure 9: Fundamental domain for a closed surface of genus two (each edge is matched to its "prime" by an orientation-preserving isometry of the Poincaré disk).

Denvir & MacKay, 1998) that for any Riemannian metric on a compact orientable closed surface M with negative Euler characteristic, there is an invariant subset of SM with a semi-equivalence to the constant negative curvature case. A (continuous) flow $\phi : \mathbb{R} \times M \to M$ is (topologically) *semi-equivalent* to another one $\psi : \mathbb{R} \times N \to N$ if there exists a continuous surjection $h : M \to N$ and a continuous map $T : \mathbb{R} \times N \to \mathbb{R}$ with $T(.,y)$ an orientation-preserving homeomorphism of $\mathbb{R}$ for each $y \in N$, such that

$$h(\phi_t(x)) = \psi_{T(t,h(x))}(h(x)), T(s+t,y) = T(s,y) + T(t, \psi_{T(s,y)}(y)) \qquad (3)$$

for all $t \in \mathbb{R}, x \in M$ (the two flows are (topologically) *equivalent* if h can be chosen to be a homeomorphism).

Morse's invariant subset consists of the unit tangents to the *class A* geodesics, those which minimise the length in the universal cover between any two of their points. In passing, it should be remarked that Hedlund subsequently did the same for the torus, the only difference being that the constant curvature metrics for the torus are flat and the dynamics of their geodesic flows are not complicated.

Denvir & MacKay (1998) provided a slightly cleaner proof of Morse's result, in which more precise control of the deviation of the semiconjugacy from the identity on the universal cover was obtained, and extended it to the case with geodesic (or geodesically convex) boundary. The extension allows one to deduce nice results such as the existence of complicated dynamics for the geodesic flow on a torus with a contractible closed geodesic, by removing the disk bounded by the contractible closed geodesic. Similarly, three closed geodesics on a sphere bounding disjoint disks implies chaos.

Such results (though without quite as precise control as ours) can also be obtained as simple consequences of Gromov's theory of quasi-geodesics (e.g. Coonaert et al, 1990), which generalises to arbitrary dimension. A Riemannian manifold is *Gromov-hyperbolic* if the universal cover has the *thin triangle property*, namely there exists $\delta > 0$ such that for every triangle formed from class-A geodesics, each side is within δ of the union of the other two. For a compact manifold, this is a property of the fundamental group only, viz. given a choice of generators, all triangles in the Cayley graph are uniformly thin. Given a metric h on a manifold (or even a metric space) M, a parametrised curve $x : \mathbb{R} \to M$ is called a (C,D)-*quasigeodesic* if for all $t, t' \in \mathbb{R}, C^{-1}|t'-t| - D \leq d_h(x(t), x(t')) \leq C|t'-t| + D$ on the universal cover. In particular, given any other metric g (for M compact) there exist C, D such that class-A geodesics of g are (C,D)-quasigeodesics of

Figure 10: The triple linkage. There are two types of pivot: the circles are fixed, while the dots are free to move in the plane.

h. Gromov's theory shows that there exists R depending only on C, D such that every quasigeodesic has a class-A h-geodesic within R of it. Applying this to a hyperbolic metric h (for which every geodesic is class A) and in the opposite direction, one deduces semi-equivalence to the h-flow from the g-flow on the set of unit tangents to its class-A geodesics.

The results of this section can be applied to mechanical systems with a Gromov-hyperbolic configuration manifold M and Lagrangian $L(q,\dot{q}) = T(q,\dot{q}) - V(q)$ on TM, where $ds^2 = 2T(q,dq)$ is a Riemannian metric (twice the kinetic energy) and V a function (potential) on M. The system conserves energy $H(q,\dot{q}) = T(q,\dot{q})+V(q)$ and the dynamics on an energy level $H^{-1}(E)$ with $E > V_{\max} = \sup_{q\in M} V(q)$ is equivalent up to time-reparametrisation to the geodesic flow of the *Jacobi metric* $ds^2 = 4(E - V(q))T(q,dq)$. For $E \leq V_{\max}$, the Jacobi metric is degenerate on the boundary of $M_E = \{q \in M : V(q) \leq E\}$, but we can still apply our results if each boundary component of M_E is surrounded by something geodesically convex and the remaining part of M_E supports a hyperbolic metric.

Boyland & Golé (1999) took up these ideas, to prove existence of orbits of time-periodic Lagrangian systems on TM with M supporting a hyperbolic geodesic flow, similar to those of the geodesic flow, with arbitrarily large average speeds. Furthermore, Angenent remarked to me during a conference at S'Agaro in June 1995 that one can construct minimax and higher index orbits for examples like our torus with a contractible closed geodesic, by considering paths that cross the disc. This idea has been taken further in Ghrist et al (2001).

I close this section with a nice mechanical example to which the results apply. It is the triple linkage of Thurston & Weeks (1984), which consists of three rods A_iB_i, $i \in \{1,2,3\}$, with the A_i fixed in a triangle and the B_i free to rotate in a common plane, linked by rods B_iC to a common freely moving pivot C in the same plane, sketched in Fig 10. Depending on the choice of the triangle $A_1A_2A_3$ and the lengths of the rods, the configuration space for the linkage can take different forms, but it is generically a smooth orientable closed surface (with perhaps more than one component), and there is a region of parameter space (including that of Fig 10) for which it is a (connected) surface of genus 3 (so Euler characteristic -4). Take any smooth component M of the configuration space with negative Euler characteristic; such a surface carries a metric of constant negative curvature and so a flow with complicated dynamics. The free dynamics with energy $E > 0$ is up to time-scale the geodesic flow of the kinetic energy metric (which depends on the masses, moments of inertia, centre of masses and lengths of the rods). Thus it possesses at least all the dynamics of the constant negative curvature case.

Actually, by an extensive computational search, Hunt (2000) found a region of parameter space (for the lengths and masses) where the curvature appears to be everywhere negative. This led to a proof of existence of a region where the geodesic flow for the triple linkage is Anosov (Hunt & MacKay, 2001). It gives a simple mechanical example of a system which is Anosov on every energy level. A real triple linkage suffers from friction, but one could imagine designing a controller which restores energy in such a way that a given energy level (or a submanifold of TM close to an energy level) is invariant; then for weak friction and control, by structural stability of Anosov systems, the dynamics would still be Anosov. A general weak energy restoration mechanism, however, is unlikely to have this property, as it would be rare that it creates an invariant submanifold near an energy level (the energy levels for the frictionless system are very far from being normally hyperbolic).

The triple linkage can also be subjected to a potential $V : M \to \mathbb{R}$. Using the Jacobi metric, one deduce the existence of chaotic dynamics for each energy $E > V_{\max}$. Furthermore, using a result of Berger (1971), one can choose V and $E_1 > V_{\max}$ such that the curvature of the Jacobi metric is constant for $E = E_1$, and hence the dynamics is Anosov for all nearby E. Finally, the triple linkage could be submitted to a time-periodic potential, providing an illustration of Boyland and Golé's result.

The results of this section can probably also be extended to mechanical systems with a magnetic (or Coriolis) field in addition to a potential. This means a Lagrangian of the form $L(q, \dot{q}) = T(q, \dot{q}) + \langle \omega(q), \dot{q} \rangle - V(q)$. In the autonomous case, the energy $H(q, \dot{q})$ is again conserved and the dynamics on an energy level $H^{-1}(E)$ with $E > V_{\max}$ is equivalent up to time-reparametrisation to the geodesic flow of the generalised Jacobi metric $ds = 2\sqrt{(E - V(q))T(q, dq)} + \langle \omega(q), dq \rangle$. This is not Riemannian if ω is non-zero, but still defines a geodesic flow via the variational characterisation. I suspect that the results for quasigeodesics can be generalised to this non-symmetric case; this would be a nice research project.

6 More general 3D flows

The domains of applicability of the previous two sections can be regarded as special cases of 3D flows: a surface homeomorphism can be suspended to produce a 3D flow and a geodesic flow on a surface is a 3D flow on its unit tangent bundle. But 3D flows merit more general treatment. A class of dynamical system of particular physical interest is the motion of "particles" along a 3D volume-preserving vector field, but non-volume-preserving ones are also of interest as systems of ODEs. I will restrict attention to time-independent vector fields. I also suppose the domain to be compact, and the vector field to be C^1 and tangent to any boundaries. The vector field could represent an incompressible fluid velocity field or a vorticity field or a magnetic field or an electric current. Compressible fluid flows can be included, by taking the vector field to be the mass flux or the preserved "volume" to be the mass. Flows with smooth no-slip boundaries can be extended to more mathematically generic-looking ones, by scaling by a smooth function which looks locally like distance to the boundary, thus replacing the identically zero velocity field on the boundary by the "skin-friction" field.

Are there simple topological hypotheses on such 3D flows which force existence of complicated trajectories?

For 3D vector fields in a domain that fibres over a circle and has positive component along the circle, or more generally a cross-section homotopic to a fibre, the answer is yes. A *cross-section* to a flow is a codimension-1 submanifold transverse to the flow such that the orbit of every point intersects it (by compactness, this happens in a bounded time both backwards and forwards: Fried, 1982). Take the return map to the cross-section, which is a surface diffeomorphism, and apply the results of Section 4.

This has some applications. For example, flow down a pipe formed by joining curved sections in a periodic manner (e.g. Ottino, 1989) can be considered as being in a solid torus, and if there is no backflow (more formally, every trajectory of the flow, extended to the skin-friction field in the no-slip case, goes from minus infinity to plus infinity) then a result of Fried (1982) shows that there is a cross-section. It follows that if the flow possesses a pseudo-Anosov braid then it has complicated dynamics. As for the 2D time-dependent case, a good strategy for efficient mixing could be to design the pipe to have such a braid in the flow. The braid could even be forced by inserting a set of wires to realise it, though it is not clear how to hold them in place!

The same considerations apply whenever a flow ϕ has a *surface of section*: an immersed surface Σ whose interior is embedded and transverse to the flow, whose boundary consists of closed orbits, and for which there is a $T > 0$ such that for all x, $\phi_{[0,T]}(x) \cap \Sigma \neq \emptyset$. Then the first return map to the interior of Σ is defined and is a homeomorphism. If this map has pseudo-Anosov isotopy class then the flow contains complicated dynamics.

For example, the figure of eight knot can be spanned by a torus with a disk removed and any flow for which it is a surface of section has return map isotopic to the pseudo-Anosov map which is obtained from (2) on the torus by blowing up the origin to its circle of directions of approach (Thurston, 1982), and hence at least as much dynamics. Every 3D transitive Anosov flow has a surface of section with pseudo-Anosov return map (Fried, 1983).

What about more general 3D flows? Is there an analogue of Thurston's classification of surface homeomorphisms up to isotopy and simplest representative and does everything in the homotopy class have at least the dynamics of the simplest representative?

There is a good analogue of isotopy for *non-singular* vector fields on 3-manifolds M, i.e. those with no equilibria (it is unfortunate terminology but well established), viz. homotopy via non-singular vector fields. There are good candidates for simplest representatives of some homotopy classes. For example, the Hopf flow on $\mathbb{S}^3$, Seifert flows on solid tori minus a torus knot, and all Anosov flows and pseudo-Anosov flows (a flow analogue of pseudo-Anosov maps; for a recent paper on them, see Fenley, 1999).

There are good candidates for reducible homotopy classes too. In the theory of 3-manifolds, it is usual first to cut along any 2-sphere which doesn't bound a ball and then glue in a ball to each resulting boundary sphere (the *prime decomposition* of 3-manifolds); to extend this to simplest vector fields, I think one would wish the glued-in ball to carry a parallel flow, so that it plays no role, thus the flow on the sphere would be inward on one hemisphere, outward on the other, and tangent along the equator (to avoid creating equilibria, one could make the equator a periodic orbit). The second usual step is to cut along some incompressible 2-tori (i.e. with injective map of fundamental groups); it would be quite natural to require the simplest vector field to leave them invariant (and carry a Poincaré flow) (though the "round handle decomposition", e.g. Morgan, 1978, requires other options).

Assuming these reductions, Thurston's ambitious programme to classify all 3-manifolds (Thurston, 1982, 1983) proposes simplest representatives as Riemannian metrics with particular homogeneity properties. There are 8 classes of these, the most interesting of which are the hyperbolic manifolds. Examples of hyperbolic manifolds include the complement of many knots or links in $\mathbb{S}^3$, for example the figure-eight knot, the Whitehead link, and the Borromean rings. If the link contains an unknot, by fattening it to a solid torus and removing its interior, one obtains hyperbolic knot or link complements in the complementary solid torus, which are realisable fluid flow domains. It is tempting to surmise that every non-singular vector field on a 3-manifold with a hyperbolic component in its reduction has complicated dynamics based on some graph of allowed transitions in the fundamental group.

Note one can not expect results for the other geometries. For example, the suspension manifold of a hyperbolic toral automorphism has "solve" geometry and the vector field for the (Anosov) suspension flow can be homotoped to that for the holonomy flow along the strong unstable manifolds, which is not complicated. Similarly, the unit tangent bundle to a surface of negative curvature has $\tilde{SL}(2,\mathbb{R})$ geometry, and the vector field for geodesic flow on it can be deformed to that for the (stable or unstable) "horocycle" flow and from there to ones for which every orbit is periodic with the same period: simply apply a constant magnetic field (Guillemin & Uribe, 1989).

There is even an apparent problem with hoping that everything in a homotopy class has at least the dynamics of some simplest representative, namely, the possibility of *flow plugs*. These are deformations of a flow that cut one periodic orbit without introducing any new ones (nor equilibria) (see Ghys, 1995, for a survey). For example, K.Kuperberg designed a C^∞ flow plug and hence constructed a C^∞ vector field on $\mathbb{S}^3$ with no periodic orbits nor equilibria (made analytic by G&K.Kuperberg). So one can not expect to keep all the periodic orbits of a simplest representative. K.Kuperberg's flow plug is not volume-preserving, however, and although G.Kuperberg made a C^1 volume-preserving flow plug, a smoother than C^1 volume-preserving flow plug is not known. If there are none, restricting to smoother volume-preserving flows would avoid this problem.

Nevertheless, I feel that "possess at least the dynamics of" needs to be to be interpreted slightly more loosely than topological semi-equivalence on a subset. I propose to say ϕ is *weakly topologically semi-equivalent* to ψ if (3) holds for some continuous surjection h and continuous map T with $T(.,y)$ non-decreasing for each $y \in N$. This allows an orbit of ϕ to be mapped to just a piece of an orbit of ψ, and so h^{-1} of a periodic orbit of ψ could for example be the union of a flow plug and a homoclinic orbit to it. Allowing to break some orbits into equilibria and connecting orbits in this way is quite natural in other contexts. For example, a vector field on a torus possessing a rational winding ratio (also known as homology direction or asymptotic cycle) does not necessarily possess a periodic orbit of that winding ratio but does possess an invariant subset semi-equivalent to one.

Restricting to non-singular vector fields is quite important, else no topological conditions force complicated dynamics (one can choose the vector field to be zero outside a neighbourhood of prescribed dynamics), though some mild restrictions on the set of singularities, e.g. contractible, might restore some interest. Allowing flux through the boundary looks a promising idea for how to incorporate prime decomposition without introducing equilibria. Some assumptions on the set of equilibria and a set of transverse sections allow to deduce a lot (Ghrist et al, 1997). Some other ideas on how to cope with equilibria might be gained from the theory of flows on surfaces, where equilibria are forced by non-zero Euler characteristic, e.g. Aronson et al (1996). Although there is not room for chaos, the bifurcation diagrams can be quite complicated (e.g. Ch.4 of Baesens et al, 1991), and there are subtle issues about the deviations of orbits from a straight line in the universal Abelian cover (see review by Zorich, 1999), a problem arising in the motion of a semiclassical electron on a Fermi surface in a magnetic field.

Restricting to volume-preserving flows might lead to some progress. The trajectories are critical points of the integral of a 1-form (at least on the universal Abelian cover), which suggests links with geodesic flows. Equilibria can be included leading to construction of a skeleton which can often imply complicated behaviour (though not necessarily with a simplest representative) (MacKay, 1994). A further restriction, viz. to Reeb flows (Hofer, 1993) – these correspond to force-free magnetic fields (Etnyre & Ghrist, 2000) – is producing interesting results, in particular existence of certain periodic orbits.

7 Dynamos

A particular area of physics where topological methods have made an impact and can be expected to do more is dynamo theory. The main interest focusses on *fast dynamos*: conducting fluid flows for which the limit infimum of the growth rate of magnetic energy for the linearised problem as magnetic diffusivity goes to zero is positive. Following work of Oseledets on the torus, Kozlovski (results presented in Arnol'd & Khesin, 1998) proved that an area-preserving diffeomorphism of a surface is a fast dynamo iff its induced action on first homology has spectral radius greater than 1 (the limiting growth rate is its logarithm). By the Thurston classification, such a diffeomorphism must have a pseudo-Anosov component in its simplest representative. A pseudo-Anosov component is not sufficient for fast dynamo action, however, e.g. (2) on $\mathbb{S}^2_4$ or D^2_3 acts as just a permutation on the (three-dimensional) first homology (the expansion on free homotopy classes cancels out in homology).

Some Anosov flows have been proved to be fast dynamos (with respect to particular choices of Riemannian metric) (Arnol'd & Khesin, 1998). One might hope that, as for the 2D discrete-time case, the fast dynamo property would be independent of the metric and even of the representative of the homotopy class (though the growth rate would need normalising to some average flow speed), but this remains open.

8 Conclusion

Simple topological hypotheses on a dynamical system can have profound effects. Some implications for fluid flows and mechanical linkages have been drawn.

Acknowledgements

I am indebted to Philip Boyland for introducing me to the Nielsen-Thurston theory and the works of Fried and Fenley, to Thierry Barbot for explaining Gromov's theory of quasigeodesics to me, and to Rob Ghrist and Caroline Series for useful comments.

References

[1] Arnol'd VI, Khesin BA, 1998, Topological methods in hydrodynamics (Springer).

[2] Aronson S KH, Belitsky GR, Zhuzhoma EV, 1996, Introduction to the qualitative theory of dynamical systems on surfaces (Am Math Soc).

[3] Baesens C, Guckenheimer J, Kim S, MacKay RS, 1991, Three coupled oscillators: mode-locking, global bifurcations and toroidal chaos, Physica D 49, 387–475.

[4] Berger MS, 1971, On Riemannian structures of prescribed Gaussian curvature for compact 2-manifolds, J Diff Geom 5, 325–332.

[5] Boyland PL, Topological methods in surface dynamics, 1994, Topology and Appl 58, 223–298.

[6] Boyland PL, Aref H, Stremler MA, 2000, Topological fluid mechanics of stirring, J Fluid Mech 403, 277–304.

[7] Boyland P, Golé C, 1999, Dynamical stability in Lagrangian systems, in: Hamiltonian systems with three or more degrees of freedom, ed Simo C (Kluwer) 90–114.

[8] Burns K, Weiss H, 1995, A geometric criterion for positive topological entropy, Commun Math Phys 172, 95–118.

[9] Conley C, 1978, Isolated invariant sets and the Morse index, Conf Board Math Sci 38 (Am Math Soc).

[10] Coonaert M, Delzant T, Papadopoulos A, 1990, Géométrie et théorie des groups, les groupes hyperboliques de Gromov, Lect Notes Math 1441 (Springer).

[11] Denvir J, MacKay RS, 1998, Consequences of contractible geodesics, Trans Am Math Soc 350, 4558–4568.

[12] Etnyre JB, Ghrist RW, 2000, Contact topology and hydrodynamics III: knotted orbits, preprint.

[13] Fenley SR, 1999, Surfaces transverse to pseudo-Anosov flows and virtual fibres in 3-manifolds, Topology 38, 823–859.

[14] Franks J, 1970, Anosov diffeomorphisms, in: Global Analysis, Proc Symp Pure Math 14, 61–93.

[15] Fried D, 1982, The geometry of cross sections to flows, Topology 21, 353–371.

[16] Fried D, 1983, Transitive Anosov flows and pseudo-Anosov maps, Topology 22, 299–303.

[17] Ghrist RW, Holmes PJ, Sullivan M, 1997, Knots and links in three-dimensional flows, Lect Notes in Math 1654 (Springer).

[18] Ghrist RW, Vandenberg JB, Vandenvorst RC, 2001, Closed characteristics of fourth-order twist systems via braids, Comptes Rendus Acad Sci Paris Math, to appear.

[19] Ghys E, 1995, Construction de champs de vecteurs sans orbite périodique, Astérisque 227.

[20] Guillemin V, Uribe A, 1989, Circular symmetry and the trace formula, Invent Math 96, 385–423.

[21] Hofer H, 1993, Pseudoholomorphic curves in symplectizations with applications to the Weinstein conjecture in dimension three, Invent Math 114, 515–563.

[22] Hunt TJ, 2000, Low dimensional dynamics: bifurcations of cantori and realisations of uniform hyperbolicity (PhD thesis, Univ of Cambridge)

[23] Hunt TJ, MacKay RS, 2001, Anosov parameter values for the triple linkage, in preparation.

[24] Katok A, Hasselblatt B, 1995, Introduction to the modern theory of dynamical systems (Cambridge Univ Press).

[25] Kennedy J, Kocak S, Yorke JA, 2001, A chaos lemma, Amer. Math. Monthly 108, 411- 423.

[26] Li T-Y, Yorke JA, 1975, Period three implies chaos, Am Math Monthly 82, 985–992.

[27] Llibre J, MacKay RS, 1991, Rotation vectors and entropy for homeomorphisms of the torus isotopic to the identity, Ergod Th Dyn Sys 11, 115–128.

[28] Llibre J, MacKay RS, 1992, Pseudo-Anosov homeomorphisms on a sphere with four punctures have all periods, Math Proc Camb Phil Soc 112, 539–549.

[29] MacKay RS, 1990, Knots for Three-dimensional vector fields, in: Topological fluid mechanics, eds Moffatt HK, Tsinober A (Cambridge Univ Press) p.787

[30] MacKay RS, 1994, Transport in 3D volume-preserving flows, J Nonlin Sci 4, 329–354.

[31] Manning AK, 1988, A Markov partition which displays the homology of a toral automorphism, in: Papers presented to Christopher Zeeman (Maths Inst, Univ Warwick) 137–141.

[32] Mischaikow K, Mrozek M, 1994, Isolating neighbourhoods and chaos, Jap J, Ind Appl Math 11, 185–202.

[33] Morgan JW, 1978, Non-singular Morse-Smale flows on 3-dimensional manifolds, Topology 18, 41–53.

[34] Morse HM, 1924, A fundamental class of geodesics on any closed surface of genus greater than one, Trans Am Math Soc 26, 25–60.

[35] Ottino JM, 1989, The kinematics of mixing: stretching, chaos and transport (Cambridge Univ Press).

[36] Rössler OE, 1983, The chaotic hierarchy, Z Naturforsch 38a, 788–801.

[37] Series C, 1991, Geometrical methods of symbolic coding, in: Ergodic theory, Symbolic dynamics and Hyperbolic spaces, eds Bedford T, Keane M, Series C (Oxford Univ Press), 125–151.

[38] Sharkovskii AN, 1964, Coexistence of cycles of a continuous map of the line into itself, Ukr Math Z 16, 61–71 (an English translation appears in Int J Bif Chaos 5 (1995) 1263–73).

[39] Thurston WP, 1982, Three dimensional manifolds, Kleinian groups and hyperbolic geometry, Bull Am Math Soc 6, 357–381.

[40] Thurston WP, 1997, Three-dimensional geometry and topology (Princeton).

[41] Thurston WP, Weeks JR, 1984, The mathematics of three-dimensional manifolds, Sci Am 251 (July) 94–106.

[42] Zorich A, 1999, How do the leaves of a closed 1-form wind around a surface?, in: Pseudoperiodic topology, eds Arnold VI, Kontsevich M, Zorich A, Am Math Soc Transl 197, 135–178.