# Duffing's equation in Brain Modelling.

By E.C. Zeeman

Symposium in honour of J.E. Littlewood's 90th Birthday,
9th June 1975.

To appear in Bull. Inst. Math. + Applications, 1976.

Mathematics Institute,
University of Warwick,
Coventry.

# Duffing's equation in Brain Modelling.

By E.C. Zeeman

Which branches of mathematics should be used to model the brain? I do not mean local models of the individual neuron, such as the Hodgkin-Huxley equations, but global models concerned with thinking and behaviour, that attempt to relate brain to mind. A brain model ought to be implicitly based on the underlying neurology, and explicitly capable of predicting some psychological behaviour. Models that only describe without predicting lay themselves open to the charge of being useless. And indeed under this criterion most published brain models to date are in fact useless, including one of my own [24]. Nevertheless in this paper I shall argue the case in favour of non-linear ordinary differential equations, and although there is still a considerable gap between the tentative predictions described here and those that would be required by the experimental scientist, at least the type of model proposed has qualities that would appear to be potentially useful in the design of experiments.

I am indebted to Larry Markus for many interesting discussions about differential equations. He once observed to me that the two prototype equations whose qualities characterise the non-linear theory are the classical oscillators of Van der Pol [21] and Duffing [4]. These particular equations are so rich that it is not surprising that they caught the attention of Littlewood in the

early 1940's, when he was already in his late 50's. They occupied much of his energy during the subsequent 25 years, and began his famous partnership with Mary Cartwright [3]. We shall not need to use the detailed analytic estimates found by Cartwright and Littlewood, but only some of the more elementary qualitative properties.

CONTENTS.


1. Classical oscillators
   - (i) Simple harmonic
   - (ii) Van der Pol
   - (iii) Duffing
   - (iv) Hopf bifurcation
   - (v) Van der Pol with large damping.
2. Neuro-psychological experiments.
3. What is an oscillator?
4. Summary of the modelling method.
5. Applications
   - (a) Sensory input
   - (b) Association
   - (c) Recall
   - (d) Mood
   - (e) Behaviour
   - (f) Anorexia nervosa
   - (g) Manic depression.




## 1. The classical oscillators.

We briefly recall their elementary properties.

### (i) The harmonic oscillator.

$$\ddot{x} + x = 0 \quad .$$

This induces a flow on the phase-plane $\mathbb{R}^2$, whose coordinates are $x, \dot{x}$. The orbits of this flow are concentric circles (Figure 1). The flow is structurally unstable because arbitrarily small damping changes the types of orbit.

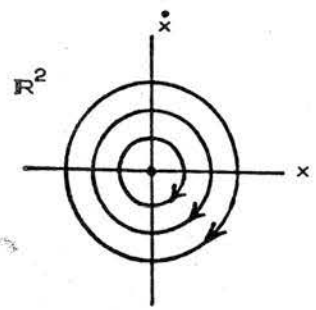



Figure 1 : Harmonic oscillator.

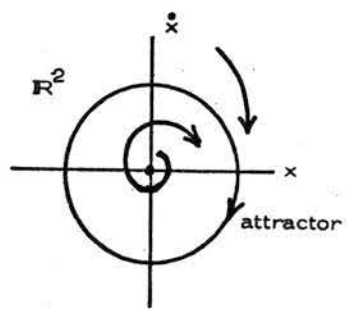



Figure 2. Van der Pol oscillator.

### (ii) The Van der Pol oscillator.

$$\ddot{x} + \varepsilon(x^2-1)\dot{x} + x = 0, \quad \varepsilon \text{ small constant} > 0.$$

This is a structurally stable perturbation of (i). The non-wandering set of the resulting flow on $\mathbb{R}^2$ consists of a repeller point at the origin, and an attractor of amplitude approximately 2, in other words an attracting closed orbit, or limit cycle, lying near the circle of radius 2 (see Figure 2). Here "near" means of order $\varepsilon$. This can be estimated by assuming $x = A\cos t$, computing $\dot{A}$, ignoring $\varepsilon^2$, and putting $\dot{A} = 0$. The rigorous proof follows by transversality (see [5,18]).

(iii) The Duffing oscillator.

$$\ddot{x} + \varepsilon k\dot{x} + x + \varepsilon\alpha x^3 = \varepsilon F \cos \Omega t \ ,$$

where $\varepsilon$ is a small constant > 0; k, F are constants > 0; $\Omega = 1 + \varepsilon\omega$, and $\alpha$, $\omega$ are parameters. This is another structurally stable perturbation of (i). It is the simplest non-linear forced damped oscillator. Being non-autonomous, it induces a flow not on the phase-plane $\mathbb{R}^2$, but on the solid torus $\mathbb{R}^2 \times T$, where T is a circle representing periodic time, with period $2\pi/\Omega$, where $\Omega$ is the frequency of the forcing term. For sufficiently small values of the parameter $\alpha$, $\omega$ the non-wandering set of the flow consists of either one attracting limit cycle, or else two attractors and one saddle-type limit cycle. The amplitude A and phase lag $\varphi$ of these limit cycles can be estimated (to order $\varepsilon$) by substituting

$$x = A\cos(\Omega t - \varphi)$$

into the equation, ignoring $\varepsilon^2$, equating the coefficients of $\cos \Omega t$ and $\sin \Omega t$, and solving for A and $\varphi$, giving

$$(*) \qquad A^2\left(\tfrac{3}{4}\alpha A^2 - 2\omega\right)^2 = F^2 - k^2A^2,$$

$$\tan \varphi = \frac{4k}{3\alpha A^2 - 8\omega} \ .$$

The equation (*) is called the Duffing amplitude relation. For rigorous treatments see [5,7,18]. (Do not see [26] because



it contains a mistake.) The equation (*) gives the graph of A as a function of the parameters $\alpha$, $\omega$ as shown in Figure 3. The graph has two cusp catastrophes [7, 20,27,28]. The cusp points

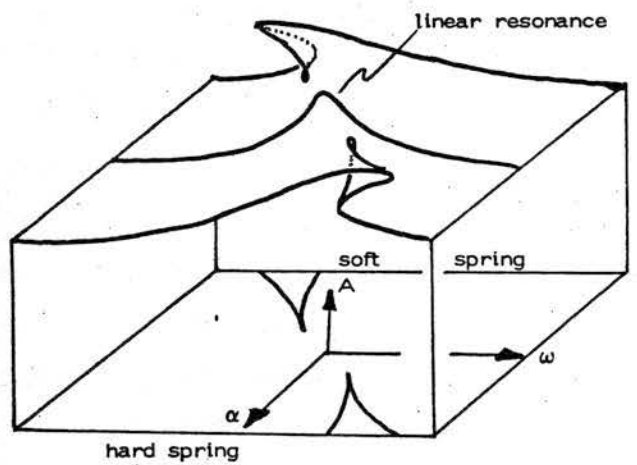



Figure 3. The Duffing amplitude relation has two cusp catastrophes.

are found by differentiating (*) twice with respect to $A^2$, and eliminating A, giving

$$(\alpha, \omega) = \pm \left(\frac{\sqrt{3}k}{2}, \frac{32k^3}{9\sqrt{3}F^2}\right) .$$

At each cusp the upper and lower sheets represent attractors, and middle sheet saddles. When $\alpha = 0$ the equation is linear, and there is always a unique attractor, whose amplitude reaches a maximum $A = F/k$ when $\omega = 0$, i.e. when the frequency of the forcing

term equals that of the original oscillator, $\Omega = 1$, causing resonance.

The case $\alpha > 0$ is called a hard spring. If, further, $\alpha > \frac{\sqrt{3}k}{2}$ then the graph becomes folded over because of the cusp catastrophe. If $\omega$ is slowly increased from negative to positive values, then A smoothly increases to the maximum $A = F/k$ at a point inside the cusp given by $\omega = \frac{3\alpha F^2}{8k^2}$ (which is given by the vanishing of the left hand side of (*)). If $\omega$ is increased further then at the right hand side of the cusp the larger attractor will coalesce with the saddle, and disappear, causing a catastrophic jump into the smaller attractor, a catastrophic drop in amplitude, and a catastrophic shift in phase. Imagine shaking a small tree in resonance, and then increasing the frequency until the tree suddenly "turns against" the shaker. Conversely a decrease in $\omega$ will cause a catastrophic increase in amplitude, and phase-shift, at the left hand side of the cusp. For the soft spring, $\alpha < 0$, events happen in the symmetrically opposite way.

If we add further non-linear terms, for example replace $\alpha x^3$ by $\alpha_1 x^3 + \alpha_2 x^5 + \ldots$, then the graph of A over the enlarged parameter space exhibits higher dimensional catastrophes such as the butterfly, etc. The important conclusions are

(1) Non-linear oscillators typically bifurcate according to the elementary catastrophes.

(2) Smooth changes in frequency (of the forcing term) can cause both smooth and catastrophic changes in amplitude and phase (of the oscillator).



(iv) The Hopf bifurcation.

An immediate word of warning is necessary, because not all stable bifurcations of oscillators are elementary catastrophes; it depends upon whether or not there exists a stably bifurcating Lyapunov function. As yet the non-elementary bifurcations are unclassified. The most famous counterexample is the Hopf bifurcation [8], which is the 1-dimensional bifurcation exhibited by the parametrised Van der Pol oscillator, with parameter b :

$$\ddot{x} + \varepsilon(x^2 - b)\dot{x} + x = 0 \quad .$$

When $b < 0$ the flow in the phase-plane $\mathbb{R}^2$ has only an attractor point at the origin. When $b > 0$ the origin turns into a repeller, and an attracting limit cycle appears of radius approximately $2\sqrt{b}$. Thus the non-wandering set, as b varies, consists of (or more precisely is differentially equivalent to and within $\varepsilon$ of) a paraboloid and its axis (Figure 4).

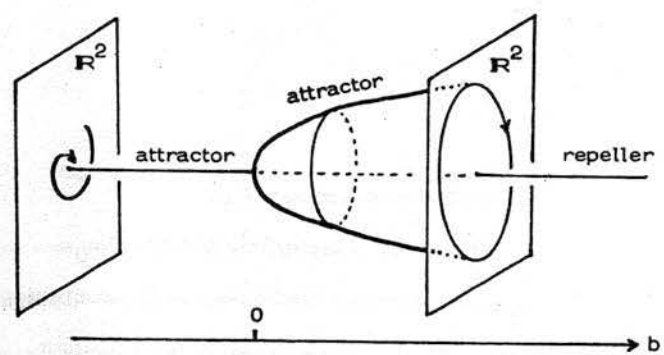



Figure 4. The Hopf bifurcation.

(v) Van der Pol with large damping.

Now consider what happens when the damping $\varepsilon$ becomes large. Let us replace $\varepsilon$ by K to indicate its largeness. The phase-plane with coordinates $x$, $\dot{x}$ is no longer a good geometrical way to represent the oscillator because even though $x$ may remain bounded the velocity $\dot{x}$ becomes very large. Therefore it is better to use the "dual" phase-plane with coordinates $x$, $\int x$, as follows. We begin with the Van der Pol oscillator

$$\ddot{x} + K(3x^2 - b)\dot{x} + x = 0$$

where K is a large constant, b a parameter, and the factor 3 is put in for convenience. Suppose that $x$, $\dot{x}$ take initial values $x_0, \dot{x}_0$. Let

$$a_0 = x_0^3 - bx_0 - \frac{1}{K}\dot{x}_0 \quad ,$$

$$a(t) = a_0 - \frac{1}{K}\int_0^t x(\tau)d\tau \quad .$$

Then $\dot{a} = -\frac{1}{K}x$ .

Substituting in the oscillator

$$\ddot{x} + K[3x^2\dot{x} - b\dot{x} - \dot{a}] = 0 \ .$$

Integrating

$$\dot{x} + K[x^3 - bx - a] = \text{constant}$$

$$= 0, \text{ initially, by choice of } a_0.$$

Hence in the dual phase-plane, with coordinates a, x, the oscillator is represented by the flow given by the two first order equations :

$$\begin{cases} \text{Fast equation} & \dot{x} = -K[x^3 - bx - a] \\ \text{Slow equation} & \dot{a} = -\frac{1}{K}x \ . \end{cases}$$

The qualitative difference between "fast " and "slow" is determined by the size of the damping K.



We call the curve in the (a,x)-plane given by

$$x^3 - bx - a = 0$$

the slow manifold. If we now allow b to vary then the same equation gives a surface in (a,b,x)-space which is none other than the canonical cusp catastrophe surface, with normal factor a and splitting factor b (see Figure 5)

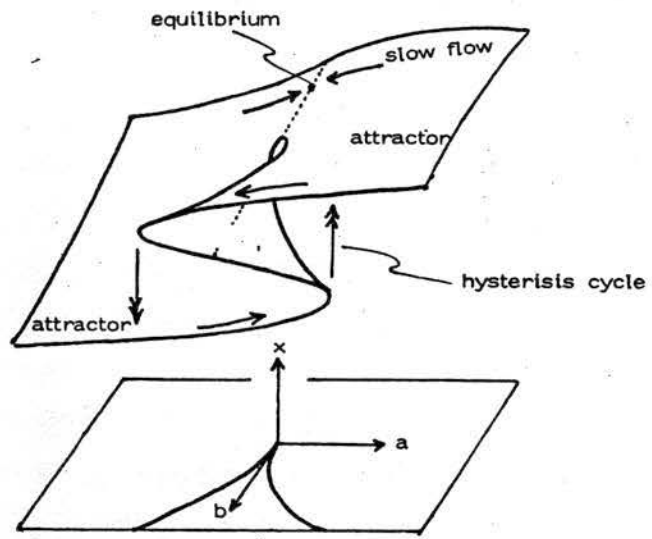



Figure 5. Van der Pol with large damping is a cusp catastrophe with feedback flow.

The slow manifolds are the sections of this surface given by b = constant. Off this surface the fast equation ensures that orbits are nearly parallel to the x-axis (where "near" means of order $\frac{1}{K}$) . Thus the fast equation acts as a catastrophe dynamic for the variable x, making the upper and lower sheets (given by $3x^2 > b$) into attractors, and the middle sheet (given by $3x^2 < b$) into a repeller.

This fast dynamic does three things : firstly it rapidly carries any point onto the attracting surface (or more precisely near the surface), secondly it holds the point on (or near) the attracting surface for as long as possible; thirdly when this becomes no longer possible, for example when the point crosses one of the fold curves (given by $3x^2 = b$) bounding the attracting surface, then the dynamic causes a catastrophic jump onto the other attracting sheet.

Once x lies on (or near) the surface, then $\dot{x} \neq 0$ (or is of order $\frac{1}{K}$), and so the slow equation comes into its own, causing the point to flow slowly along (or near) the slow manifold. In the language of catastrophe theory, the slow equation is a feedback flow of the fast "behaviour" variable x upon the slow "control" parameters a,b. When $b < 0$ the slow equation makes the origin $a = x = 0$ into a unique attractor point. When $b > 0$ the origin turns into a repeller point, and the slow equation creates a new hysteresis cycle, consisting of two portions of slow flow along the upper and lower attracting sheets of the slow manifold, alternating with two fast catastrophic jumps between the two sheets. (See [25] for more details and pictures).

Thus the Van der Pol oscillator with large damping, and its accompanying Hopf bifurcation, are both represented as additional structure superimposed upon the elementary cusp catastrophe, even though the Hopf bifurcation itself is non-elementary. The important conclusion is :



(3) Non-linear oscillators with large damping can sometimes be interpreted as slow feedback flows on elementary catastrophes.

Having extracted some general conclusions about the way that non-linear oscillators can behave and bifurcate, and having expressed them in the language of catastrophe theory, we now return to the brain. I like to approach the brain as a collection of strongly coupled oscillators, driving one another. The stability of our instincts, habits, and memories indicate strong stability of some of the oscillators, and the swiftness of our reactions indicates a coexisting instability due to strong coupling. But before we proceed further, it is a good idea to mention briefly a few neuro-psychological experiments to support this point of view.

## 2. Neuro-psychological experiments.

(i) Shutting the eyes and relaxing induces the α-rhythm. The α-rhythm is an observed frequency of about 10 cycles per second in the EEG (electroencephalograph) pattern. It probably indicates that large parts of the brain are oscillating in resonance.

(ii) I once asked a mathematician under EEG to calculate a complicated homotopy group. The recording needles stopped dead for a minute while he thought, and then started again once he had given me the answer. This probably indicated a lack of resonance during the specialised cortical activity.

(iii) During epilepsy operations Penfield [13] touched the surface of the brain with a small electric oscillator, and patients reported induced memories and a double reality. Moreover the same memories recurred when the experiment was repeated.

(iv) Adey [1] recorded oscillations from a number of electrodes implanted in a cat's brain, in the limbic system near the hippocampus. When the cat was relaxing the frequencies varied between 4 and 7 cycles per second, but when a trap-door opened leading to food, then all the oscillations locked onto one frequency (of about 6 cycles per second), with a specific phase ordering. If the cat made a wrong turning at a T-junction, then the phases hunted, until the cat turned back to the correct turning when they locked on to the correct ordering again.

## 3. What is an oscillator ?

What is the connection between the classical equations, and the behaviour of the complex biological systems that we have so glibly called oscillators? We cannot possibly measure all the important events taking place in the brain, or even those in one organ of the brain. Nevertheless it is possible to imagine a mathematical dynamical system of sufficient complexity to model those events. From general theory we now that such dynamical systems can have attractors, which can bifurcate. Moreover, although we cannot measure those attractors, nevertheless we can sometimes catch their bifurcations by means of artifacts. For



example the EEG pattern obtained from electrodes placed on the outside of the skull would appear to be an artifact, whose behaviour could not possibly give any significant information about the important dynamics within. Yet when the attractor inside makes a catastrophic jump, the artifact outside may well display a sudden qualitative change in behaviour. Such qualitative changes are easy to recognise in the resulting EEG pattern. Moreover when they occur at the same time as psychological behavioural changes, such as in the experiments described above, then it strongly suggests that the artifacts are echoing the important bifurcations and catastrophic jumps within. Extending this argument to more than one dimension, if an attractor is bifurcating according to a higher dimensional catastrophe, then the measurements of the artifact may reveal a diffeomorphic copy of the same catastrophe [29]. Hence, although the artifact may be but a pale shadow of the internal dynamics, yet its catastrophes may furnish a brilliant reflection of significant events. In this sense the artifact may provide a non-trivial qualitative model for the underlying neurology. Therefore the procedure of measuring and using neurological artifacts to predict psychological behavioural changes is scientifically tenable.

Let us now put the above discussion on a more concrete mathematical footing. Suppose that we model a biological system B by a mathematical dynamical system, that is a multidimensional manifold M, together with a vector-field X on M. In the case of the brain we are quite prepared for the dimension of M to be as high as $10^{10}$, the number of neurons, or $10^{14}$, the number of synapses.

The attractors of $X$ represent the homeostatic state of B. The $C^0$-density theorem of Smale [17], Shub [16], and d'Oliveira [11] says that, by making an arbitrarily small $C^0$-perturbation if necessary, we can assume that $X$ is structurally stable and that the only attractors of $X$ are points ( = stable equilibria) or closed orbits (= limit cycles).

Point attractors are easy to understand, and if B is parametrised, or driven, by another system then those points will bifurcate according to only the elementary catastrophes.

However there are two reasons why the closed orbit attractors are more important than the point attractors. Firstly the EEG evidence suggests that periodicity is the rule, and static equilibrium the exception. Secondly from the evolutionary point of view, the brain that can respond more swiftly than its neighbours to the environment has an evolutionary advantage. And if the dynamical system of the brain only had point attractors, it would remain stable when weakly coupled to any other stable system (representing some part the environment or some sensory input), and hence the brain could not respond. On the other hand a system with closed orbits can resonate with, or lock-on to, the attractors of the other system, however weakly coupled, thus enabling the brain to respond swiftly. (See [2,19] for the underlying theorems). Hence we would expect the brain to evolve non-gradient dynamics and limit cycles. By contrast the developing embryo does <u>not</u> want to be too perturbed by the environment during the crucial stages of development; hence we would expect it to evolve gradient dynamics and equilibrium states, as indeed it has.



We now bring this abstract multidimensional theory down to earth by relating it to ordinary differential equations, such as the classical oscillators that we started with.

<u>Lemma. If y is a measurement of a closed orbit of an arbitrary dynamical system, then there exists a second order differential equation having y as its unique attractor.</u>

<u>Proof.</u> Let C be the given closed orbit in M. By suitable choice of time unit, we can assume that C has period $2\pi$. Let T denote $\mathbb{R}$ modulo $2\pi$, representing periodic time, and let $T \to C$ be the diffeomorphism giving the timing round the orbit. Let $M \to \mathbb{R}$ be the given measurement. For instance a point in M might represent the brain state, and its image in $\mathbb{R}$ the resulting potential difference across two electrodes on the skull, measured by the EEG. Let y denote the composition

$$T \to C \overset{\subseteq}{\to} M \to \mathbb{R}$$

Then $y(t)$, $t \in T$, will be the periodic function of time recorded by the EEG. Let $\dot{y}(t)$, $\ddot{y}(t)$ denote the derivates with respect to t. Let $x$, $\dot{x}$ be coordinates in $\mathbb{R}^2$. Define

$$\phi : \mathbb{R}^2 \times T \to \mathbb{R}$$

by $\phi(x,\dot{x},t) = \ddot{y}(t) + 2(\dot{y}(t) - \dot{x}) + 2(y(t) - x)$.

Then

$$\ddot{x} = \phi(x,\dot{x},t)$$

is the required differential equation. It can be verified by substitution

that the general solution of the equation is

$$x(t) = y(t) + Ae^{-t}\cos(t-\varphi), \quad A,\varphi \text{ constants.}$$

Hence all the solutions decay to $y(t)$, which is therefore the unique attractor, as required.

## 4. Summary of the modelling method.

The main points of our discussion so far have been :

(a) We assume that we can model the activities of the brain by multidimensional dynamical systems.

(b) By the $C^0$-density theorem, we confine our attention to closed orbit attractors, and hence to second-order differential equations.

(c) Some of the bifurcations of the latter are modelled by elementary catastrophes.

(d) Neurological artifacts may exhibit diffeomorphic catastrophes, and hence provide measureable models for psychological prediction.

We conclude that some brain activities may be modellable by elementary catastrophes. It would be wrong to deduce any stronger statement, because

(a) It may not be possible to model the relevant brain activity by a dynamical system.

(b) The $C^0$-density theorem has not yet been generalised to parametrised systems.

(c) Some bifurcations are non-elementary.

(d) It may not be possible to measure anything that exhibits the relevant bifurcation.

For further discussion see [19,20,28,29].

We now give a number of applications to behaviour exhibiting catastrophic jumps. These examples are really not so much predictions as suggestions for designs of experiments.

(a) Sensory inputs.

Most sense organs convert amplitude into frequency. For example brighter lights cause the neurons in the optic tract to fire with the same action potential but more rapidly. Similarly louder noises, sharper pains, etc., all cause increased firing rates. When the frequency of firing reaches a certain threshold, then the brain will suddenly pay attention. Now the Duffing soft spring provides a simple model in which a frequency threshold causes a sudden jump in amplitude (see Figure 6). Here the forcing term represents the input message from the sense organ to the brain, and the oscillator represents the brain's response.

In Figure 6 the frequency threshold occurs at $\omega_2$. If the input frequency is now reduced again, then the model makes a prediction about hysteresis, which could be tested experimentally : the response does not switch-off again at $\omega_2$, but at a lower frequency $\omega_1$. Indeed between $\omega_1$ and $\omega_2$ the amplitude of the response is in fact slightly enhanced. The hysteresis $\omega_2 - \omega_1$ could be measured.

Another question that could be tested experimentally is the hardening or softening effect due to changes in arousal, fatigue or practice. A second prediction would be an anti-correlation between threshold and hysteresis : if the threshold $\omega_2$ drops (implying softening) then the hysteresis $\omega_2 - \omega_1$ would increase, and conversely if the threshold increases (implying hardening) then the hysteresis would drop.

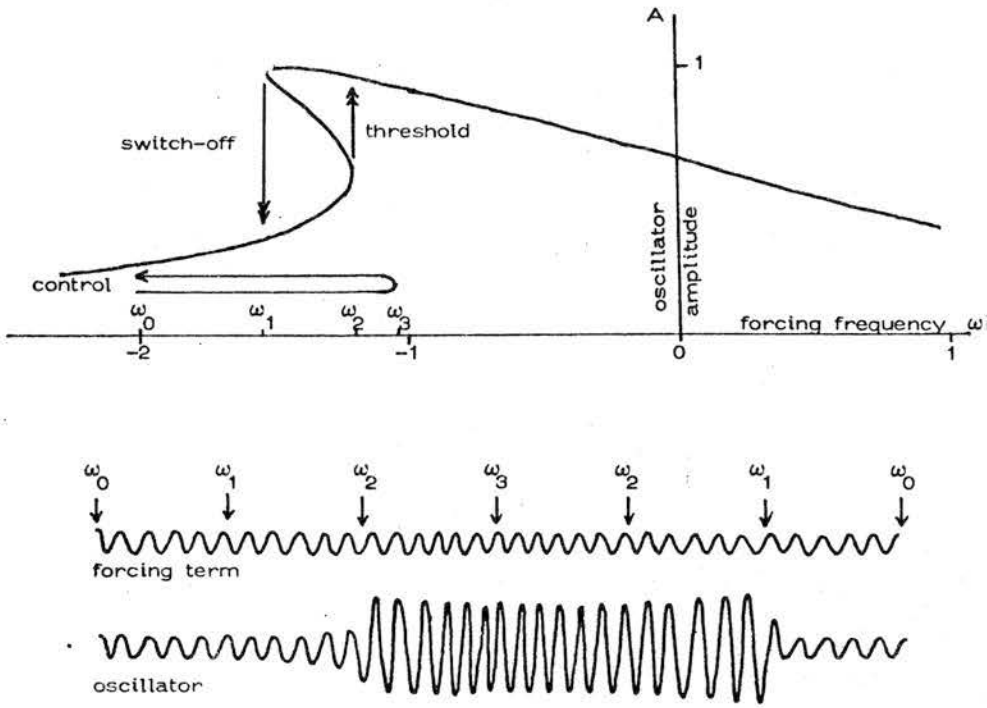



Figure 6. The Duffing soft spring sketched for $k = F = 1$, $\alpha = -4$. The input frequency is increased to up to $\omega_3$ and down again. The threshold occurs at $\omega_2$, and switch-off at $\omega_1$.

(b) Association.

We begin with an idea of Thom [19]. Suppose two memories, as yet unassociated, are each represented by an attracting closed orbit of a dynamical system. The words "as yet unassociated" mean that we can represent both together by the product of the two systems. Therefore when the two memories are stimulated together they are represented by the product of the two



attractors, which is a linear flow on an attracting torus. However the latter is structurally unstable, and an arbitrarily small generic perturbation will, by Peixoto's Theorem [12], furnish a new stable system with a new attracting closed orbit on the torus. (This process is analogous to the locking-on phenomemon in the forced Van der Pol oscillator.) The new attractor represents the new associative memory, associating the two previous memories.

The beauty of this model is that it needs only arbitrarily small random synaptic perturbations in order to work. By contrast most models of associative memory, particularly those used in the design of artificial intelligence machines, need to assume a feedback system of synaptic changes in order to work. For example a typical machine "learns" by turning up the gain on all those synapses it has just used every time it gives the correct answer; therefore it needs a feedback system from the "answer" to the synapses. And it is unlikely that a real brain could contain the abundance of such feedback systems, that would be necessary to explain the apparent ease with which we make all manner of associations all the time.

(c) Recall.

Consider a stimulus recalling a memory. Represent this by an oscillator driving another oscillator. The simplest model is Duffing's equation, with the forcing term representing the stimulus, and the oscillator the memory. Let us further suppose that when the

the memory is first laid down the equation is linear, $\alpha = 0$; therefore when the stimulus hits the resonant frequency the memory resonates accordingly (not unlike Thom's model above ).

Now suppose that the memory is allowed to lie dormant for some period, before the stimulus is given again. There are two possibilities according as to whether or not the mind has been thinking about closely related thoughts during the intervening period. If it has, then facilitation of the nearby synaptic pathways may make the oscillator easier to drive, in the sense that its amplitude and frequency become correlated, because of the proliferation of short-cuts available in the neural pathways. In other words the oscillator turns into a hard spring, $\alpha > 0$. Therefore when the stimulus is given, represented by increasing the forcing frequency to the original resonant frequency $\omega = 0$, then the amplitude of the oscillator will increase smoothly to a value lower than the original resonance. Hence the memory will <u>flow smoothly to mind</u>, quite unobtrusively.

By contrast, if the mind has not been thinking about any closely related thoughts, then the oscillator may become more difficult to drive, in the sense that amplitude and frequency become anticorrelated. In other words it has become a soft spring, $\alpha < 0$ (like a simple pendulum). In this case when the stimulus is given then the amplitude will make a sudden jump just before the forcing frequency hits the original resonant frequency (see Figure 6). In other words the memory will <u>suddenly spring to mind</u>.



The difference between memories that flow-to-mind and spring-to-mind is a well known phenomenon; for instance consider the ease with which we remember our friends' names, compared with the difficulty of putting a name to a familiar face that has not been seen for sometime. It is remarkable how that missing name can sometimes suddenly spring to mind. It would be interesting to try and devise psychological experiments to test this difference between memories that flow-to-mind and spring-to-mind, and to measure the size, s, of the catastrophic jump in the latter case. If this were possible then the model gives a quantitative prediction, as follows.

During the period while the memory is lying dormant and while the oscillator is softening, there is a unique critical moment, at time $t_0$ say, when $\alpha$ crosses the cusp-point. This is the precise moment when the latent memory is "forgotten", in other words is switched from being a flow-to-mind into a spring-to-mind type. In the neighbourhood of the cusp point, the size of the latent catastrophe increases parabolically with time. In other words we have the quantitative prediction

$$s = \lambda\sqrt{t-t_0} + O(t-t_0), \quad \lambda \text{ constant} > 0 .$$

(d) <u>Mood.</u>

The influence of environment upon mood and emotion has features that strongly suggest the use of catastrophe models. For example the persistence of mood, the sudden changes of mood, the delays before those chang the possibility of different moods under similar circumstances, and the inaccessibility of other intermediate moods, all these five properties are typical

of the bifurcation of oscillators [28]. How do we measure mood? According to MacLean [10,22] emotions are probably generated in the limbic system (roughly the middle third of the brain), and so the oscillators involved would be modelling limbic organs, notably the hypothalamus. It is difficult to record directly from the limbic system, but the direct connections from the hypothalamus to the frontal lobes suggest that the latter might provide artifacts, that would echo limbic catastrophes. Similarly physiological indicators of autonomic nervous activity can provide artifacts, such as the facial expressions of a dog indicating the levels of fear and rage [9].

Contrary to what might be suspected at first sight, mood may in fact be one of the simplest brain activities to model. Possibly much simpler than, for example, the old favourites of visual perception, aural perception, language and problem solving. For emotionally we tend to be in one mood at a time, whereas intellectually we are able to grasp many things at a time. Perhaps this is because the limbic system tends to oscillate in resonance, due to its 3-dimensional interconnectivity, whereas the cortex is able to oscillate out of resonance, differently is different parts, due to its 2-dimensionality and its lateral inhibition. If mood <u>can</u> be represented by a single attractor, then although the infinite variety of nuances of mood would be difficult to measure because they would be represented by smooth variations of this attractor, nevertheless the noticeable changes of mood might be relatively easy to model using catastrophe models of the attractor's bifurcations, and might be relatively easy to predict using artifactual measurements.



(e) <u>Behaviour.</u>

The influence of mood and emotions upon behaviour is the next step. Again, the persistence of behaviour, the sudden behavioural changes, the delays before making those changes, the possibility of divergent behaviour under similar emotions, and the inaccessibility of other intermediate behavioural patterns, all suggest catastrophe models. A simple example is the cusp catastrophe model of fear and rage as conflicting factors influencing aggression [28]. There are many such psychological models, for both man and animals, waiting to be tested experimentally against neurological and physiological measurements.

(f) <u>Anoresia nervosa.</u>

Anorexia is a psychological disorder, in which dieting degenerates into obsessive fasting, leading to severe malnutrition and possibly death. It sometimes develops a second phase of alternately fasting and gorging. The psychotherapist J. Hevesi and I conjectured that the cause might be an elementary bifurcation of the brain oscillator underlying eating and satiety. This gave rise to a 5-dimensional butterfly catastrophe model [28] of both the disorder and its cure under Hevesi's successful technique of trance therapy. This model was effective in permitting a coherent synthesis of a large number of observations that would otherwise appear disconnected. It also made sense of some of the victims' bizarre descriptions of their own disorder. Furthermore it gave insight into what might be the key operative suggestions of the therapy, which is psychiatrically useful

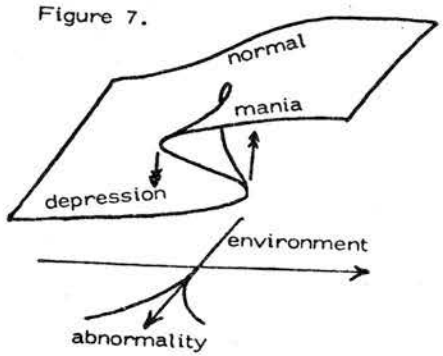
in helping to explain the technique to other therapists.

One of our projects is to extend the model to include sleep, because the disorder interferes with the natural catastrophic jumps of falling asleep and waking up. It is likely that the enlarged model could be a section of the 10-dimensional double-cusp catastrophe, of which the mathematics is as yet poorly understood [30]. Even so, the 5-dimensional geometry of the butterfly catastrophe is already sufficiently rich to have made some qualitative predictions that have been confirmed by observation. What is now needed, parallel to the theory, is a programme of quantitative testing of the model, for instance the monitoring of EEG and physiological changes in patients during the different states of fasting, gorging, sleeping, dreaming, therapy, etc., and the development of numerical techniques to convert these readings into geometric form, in which higher dimensional catastrophes can be recognised, and verified.

(g) Manic-depression.

I am indebted to T.C. Dunn for introducing me both to his patients and to the literature on the subject. In [27] I suggested briefly that normality, mania and depression might fit into a cusp catastrophe, as shown in Figure 7 with some measure of compatibility with the environment as

Figure 7.






a normal factor, and some clinical measurement of the abnormality as splitting factor. The changes between the two pathological states do seem to be catastrophic in the sense that they take place relatively quickly compared with the time spent in either state. I had in mind a bifurcation of some attractor in the limbic brain, similar to the anorexia model, of frequency say a few cycles per second. However M. Schmocker of Tübingen University Nervenklinik suggested that an oscillator with a 24 hour period might be more appropriate. Her studies on sleep deprevation, with results similar to those of Pflug [14], had shown many indications of the disturbance of circadian rhythms amongst depressive and manic-depressive patients.

Therefore let us take the Duffing oscillator as a tentative model, with the forcing term representing the external rhythm of day and night, and the oscillator representing the diurnal variation in the blood content, measured for example by the level of plasma cortisol [15,23]. The limbic brain is directly involved because cortisol production is controlled by the pathway : hypothalamus → pituitary → ACTH secretion → adrenal cortex → cortisol secretion → blood (see [23]). With this type of control a correlation between frequency and amplitude of the oscillator is plausible, since overstimulation of hypothalamus is liable to increase both. Therefore let us assume the oscillator behaves like a hard spring, $\alpha > 0$.

What we have said so far applies to the normal person. We now turn to the manic-depressive, and take as our main hypothesis that the underlying cause of the abnormality is a speeding up of the

internal circadian rhythm. Mathematically this is equivalent to decreasing the relative frequency $\omega$ of the forcing term. Figure 8

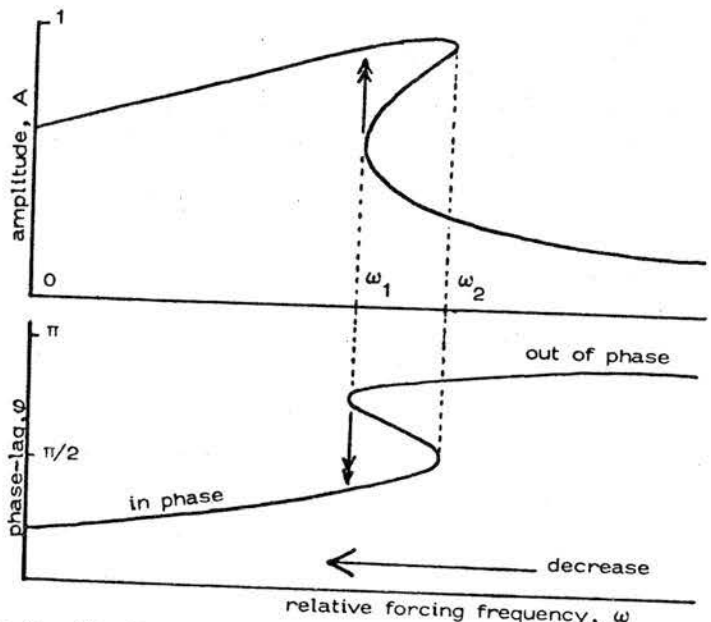



Figure 8. The Duffing hard spring drawn for k = F = 1, $\alpha$ = 4. A decrease in the relative forcing frequency $\omega$ causes a catastrophic increase in amplitude and forward phase-shift.

illustrates the consequences : when $\omega$ reaches the threshold $\omega_1$ this causes a catastrophic rise in the amplitude A and a catastrophic drop in the phase-lag $\varphi$. This is exactly what is observed in some depressives, a substantial increase in plasma cortisol and a forward phase-shift [15]. The fact that the oscillator is now nearly in phase with the forcing term could mean that cortisol secretion is now being paradoxically triggered by the presence, rather than the absence, of cortisol already in the plasma. Moreover the hysteresis effect locks



the system in its abnormal state, even if the internal circadian rhythm slows down again : $\omega$ has to reach $\omega_2$ before recovery is possible.

We now turn to the psychological effects. It is known in Cushing's Syndrome [23], that an overproduction of cortisol due to other causes, for instance an adrenal tumour, is liable to cause mental changes, such as changes of mood ranging from depression to mania. Therefore we can expect these moods to arise from an increase in amplitude due to the Duffing catastrophe. Meanwhile phase-shift symptoms are also observed : sleep patterns are disrupted, manics often enjoy a vigorous night-life, and depressives are often unable to face the day. If half the body metabolism is out of phase, and sending conflicting messages to the limbic brain, no wonder the latter is liable to generate abnormal moods.

If the underlying cause of the abnormality is so simple, there should be an equally simple cure : just remove the forcing term. Then the abnormal internal circadian rhythm will reassert itself, causing a catastrophic renormalisation of both cortisol level and phase. This is exactly what happens in the sleep deprivation treatment [14]. A depressed patient is kept vigorously awake and active all one night, and in the morning is cured! One 30-year old manic-depressive patient reported :

> "After this she felt like a changed person, she could enjoy her breakfast and experience pleasure again. She wanted to make things and read. She said she had not felt so well for 6 months."

The snag is that the cure generally lasts only 2 or 3 days. Some fortunate patients stay well (protected by the hysteresis effect) but the more severe depressives revert as soon as the forcing term reasserts itself, i.e. as soon as the patient begins to keep regular hours again.

A simple test of the model would be to remove the forcing term for a longer period, by providing an artifical environment without day and night, or by living in the continuous sunlight of the artic summer for instance, so that the patient could revert to his own circadian rhythm. Once the period had been established, then he might be able to devise a life-style to suit; admittedly it can be awkward living a non-24-hour day, but worthwhile if it gives freedom from depression. For example Schulte [14] reports that some victims had independently discovered a method of relieving their own depression by taking sleepless nights periodically.

Another possible cure might be to place a tiny alternating electric field across the patient's bedroom; for Wever [22] has shown that for subjects isolated for a month in an underground electromagnetically shielded bunker, a background field as small as 2.5 volts per metre at 10 cycles per second can have the power of synchronizing the endogenous oscillators, and entraining the circadian rhythm. The usual treatment for manic depression is the drug lithium [6], which in terms of the model, may interfere with the oscillator sufficiently to cancel the Duffing catastrophe. However the trouble with lithium is that it can have unpleasant side effects, and an over-accumulation can cause lithium poisoning, possibly resulting in tremor, slurred speech, even coma or epileptic siezures.



We now return to our main hypothesis, and ask what happens if the circadian rhythm slows down instead of speeding up. The prediction from Figure 8 would be a slight decrease in amplitude, and slight phase-shift the opposite way, but since no catastrophes would be involved, this would hardly be detectable, except by statistical analysis of a large number of normal people. However if the internal rhythm were to slow down exceptionally to a period of more than 36 hours, then the oscillator instead of locking-on to the forcing term would lock-on to its second harmonic with a similar catastrophe as before. This is exactly what happens in certain rare cases of manic-depressives, who develop a 48-hour cycle [6]. The patient described in [6] regularly suffered from one day of depression alternating with one day of hypomania for thirteen years, with the change of state occuring each night during sleep, usually between 2 and 3 a.m. When the patient was put in an artificial 22-hour environment, then, as would be expected from the model, he locked-on to a 44-hour psychotic cycle.

We conclude by emphasising the tentative nature of the model, and pointing out some reservations. Firstly it does not explain the difference between mania and depression; perhaps it could be combined with Figure 7 into a higher dimensional catastrophe model, compatible with the 48-hour syndrome. Secondly, a closer analysis of cortisol [15] reveals that it is secreted in a series of 7 to 9 major episodes during the 24 hours. Therefore superimposed

upon the circadian rhythm are physiological subrhythms, and any effective model should not only be compatible with the latter, but also perhaps offer an explanation for, and prediction of, the psychological subrhythms of eating/satiety and sleeping/dreaming. One thing is clear :to develop and test such models will need long-term colaboration between mathematicians interested in dynamical systems and physiologists and psychiatrists.



References.


1. Adey, W.R., Dunlop, C.W. & Hendrix, C.E., Hippocampal slow waves in learning, Arch. Neurol. 3 (1960) 74-90.

2. Arrowsmith, D., Products of flows, Proc. Camb. Phil. Soc. 73 (1973) 301-306.

3. Cartwright, M.L. & Littlewood, J.E., On non-linear differential equations of the second order, Jour. Lond. Math. Soc. 20 (1945) 180-189, and Ann. Math. (2) 48 (1947) 472-494.

4. Duffing, G., Erzwungene Schwingungen bei veränderlicher Eigenfrequenz, Braunschweig, 1918.

5. Hale, J.K., Ordinary differential equations, Wiley Interscience, 1969.

6. Hanna, S.M., Jenner, F.A., Pearson, I.B., Sampson, G.A., & Thompson, E.A., The therapeutic effect of lithium carbonate on a patient with a forty-eight hour periodic psychosis, Br. Jour. Psychiatry 121 (1972) 271-280.

7. Holmes, P.J. & Rand, D.A., The bifurcation of Duffing's equation : an application of catastrophe theory, Jour. Sound and Vibration 44, 2 (1976) 237-253.

8. Hopf, E., Abzweigung einer periodischen Lösung von einer stationären Losung einer Differentialsystems, Ber. Verh. Sächs, Akad. Wiss. Leipzig. Math. Phys., 95 (1943) 3-22.

9. Lorenz, K., On agression, Methuen, London, 1966.

10. MacLean, P.D., The limbic brain in relation to the psychoses, Physiological correlates of emotion Academic Press, N.Y., 1970.

11. de Oliveira, M.M.C., $C^0$-density of structurally stable vector fields, Thesis, Warwick University, 1976.

12. Peixoto, M.M., Structural stability on two-dimensional manifolds Topology 1 (1962) 101-110.

13. Penfield, W. & Roberts, L., Speech and brain mechanisms, Princeton, 1959.

14. Pflug, B., Therapeutic aspects of sleep deprivation, 1st Europ. Congr. Sleep Res., Basel (1972) 185-1

15. Sachar, E.J., Hellman, L., Roffwary, H.P., Halpern, F.S., Fukushima, D.K. & Gallagher, T.E., Disrupted 24-hour patterns of cortisol secretion in psychotic depression, Arch. Gen. Psychiat. 28 (1973) 19-24.

16. Shub, M., Structurally stable diffeomorphisms are dense, Bull. A.M.S. 78 (1972) 817-818.

17. Smale, S., Stability and isotopy in discrete dynamical systems, Dynamical systems (Ed. Peixoto, M.M.) Academic Press, N.Y., (1973) 527-530.

18. Stoker, J.J., Non-linear vibrations, Interscience, 1950.

19. Thom, R., Topologie et signification, L'Age de la Science, 4, (1968) 219-242.

20. Thom, R., Stabilité structurelle et morphogénèse, Benjamin, N.Y. 1972, Eng. translation by Fowler, D.H., Benjamin-Addison Wesley, N.Y. 1975.

21. Van der Pol, B., On oscillation hysterisis in a triode generator with two degrees of freedom, Phil. Mag. (6) 43 (1922) 700-719.

22. Wever, R., Human circadian rhythms under the influence of weak electric fields and the different aspects of these studies; Int. J. Biometeor 17 (1973) 227-232.

23. Wright, Sampson, Applied physiology, 12th Edition (Rev : Keele, C.A. & Neil, E) Oxford Univ. Press, 1971.

24. Zeeman E.C. Topology of the brain, Mathematics and computer science in biology and medicine, Med. Res. Council Publ., 1965.

25. ---------- Differential equations for the heartbeat and nerve impulse, Towards a theoretical biology, 4 (Ed. C.H. Waddington) Edinburgh Univ. Press (1972) 8-67.

26. ---------- Catastrophe theory in brain modelling, Inter. J. Neuroscience 6 (1973) 39-41.

27. ---------- Applications of catastrophe theory, Manifolds Tokyo 1973, Tokyo Univ. Press., 11-23.

28. ---------- Catastrophe theory, Scientific American, 234, 4 (1976) 65-83.

29. ---------- Brain modelling, Symp. on Catastrophe Theory, Seattle 1975, Springer Lecture Notes in Maths., (to appear).

30. ---------- The umbilic bracelet and double-cusp catastrophe, ibid.