IC/72/83

# INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS

## INFORMAL MEETING ON NEURAL NETWORKS

24 - 26 July 1972

(SUMMARIES)

INTERNATIONAL ATOMIC ENERGY AGENCY

UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION

1972 MIRAMARE-TRIESTE

IC/72/83

International Atomic Energy Agency
and
United Nations Educational Scientific and Cultural Organization

INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS

# INFORMAL MEETING
# ON
# NEURAL NETWORKS

24 - 26 July 1972

(SUMMARIES)

MIRAMARE - TRIESTE
August 1972

Please note that copies of papers referred to may be obtained direct from the authors and not from the ICTP.

# CATASTROPHE THEORY IN BRAIN MODELLING

E.C. Zeeman
University of Warwick, Coventry, UK.

If we implicitly assume that neurological activity can be modelled by generic differential equations (with arbitrarily many variables) then we can explicitly use catastrophe theory to explain and predict psychological phenomena. Mathematics is able to provide a link between neurology and psychology because of a deep theorem of Thom [1] classifying the elementary ways that differential equations can bifurcate.

We shall explain the idea by an example. But first consider Duffing's equation. I am grateful to Larry Markers for first introducing me to Duffing's equation, which is as follows:

$$\underbrace{\ddot{x} + x}_{\text{simple harmonic oscillator}} + \underbrace{\varepsilon k \dot{x}}_{\text{small damping term}} + \underbrace{\varepsilon x^3}_{\text{small Duffing term}} = \underbrace{F \cos \Omega t}_{\text{small forcing term}}$$

where

$t$ is time ,

$x$ is a real variable , $\dot{x} = dx/dt$ ,

$\varepsilon, k$ are small positive constants ,

$\Omega = 1 + \varepsilon\omega$

and

$\omega, F$ are control parameters .

Notice that the period of the forcing term is close to that of the oscillator. We assume for the moment that the controls $\omega, F$ are fixed.

Geometrically the manifold for Duffing's equation is the three-dimensional manifold $R^2 \times S^1$, where $R^2$ is the plane with co-ordinates $x$ and $\dot{x}$ , and $S^1$ is the circle with co-ordinate $t$ modulo $2\pi/\Omega$ (in other words, periodic time with period that of the forcing term). Duffing's equation gives a flow on this manifold. It turns out that this flow has 1, 2 or 3 closed orbits (depending upon the values of the controls) and all other orbits spiral towards them. The closed orbits are given by the following periodic solution (as can be seen by substitution):

$$x = A\cos\Omega t + \varepsilon\frac{A^3}{32}\cos 3\Omega t + O(\varepsilon^2) \qquad (2)$$

where

$$\frac{3}{4}A^3 - 2\omega A - F = 0 \quad . \qquad (3)$$

Here $A$ stands for amplitude because if we neglect $\varepsilon$ then (2) reduces to a simple oscillation of amplitude $A$. The value of $A$ is given by (3) which is called the Duffing amplitude relation. If we draw a graph of $A$ as a function of the controls $\omega, F$ then we obtain the characteristic cubic surface $S$ of the cusp catastrophe.

The surface $S$ has folds over the cusp:

$$128\,\omega^3 = 81\,F^2$$

in the control space $C$. This cusp is called the <u>bifurcation set</u>, because this is where the closed orbits bifurcate. If the control point $(\omega,F)$ lies <u>outside</u> the cusp, then there is a unique value of $A$ and therefore a unique closed orbit which is an attractor of the flow. If the control point lies <u>inside</u> the cusp then there are three values of $A$ and therefore three closed orbits of which two are attractors and the other is a saddle. The points of $S$ representing saddle-type orbits are shown dotted in the picture.

Intuitively the reason that two attractors can exist for the same control point is because we can either use the forcing term in-phase with a a large fast orbit, in order to slow it down to the correct frequency, or else use out-of-phase with a small slow orbit, in order to speed it up (here in-phase means $A, F$ have the same sign and out-of-phase means opposite sign). The points representing out-of-phase attactors are shown shaded in the picture. If we fix $\omega > 0$ and slowly and smoothly increase $F$ from negative to positive then the system will start in-phase and pass smoothly out-of-phase as $F$ changes sign, then at the second crossing of the cusp, the out-of-phase attractor vanishes by coalescing with the saddle, causing the system to jump into the other in-phase attractor. Here the word "jump" means that the speed of homing near the attractor is large compared with the rate of change of $F$, which we assume to be sufficiently slow. For instance, in electronic displays of Duffing's equation, the dot or circle on the cathode ray tube appears to the eye to jump.

## Generalization to the brain

We now generalize Duffing's equation to a model of the brain. Suppose some part X of the brain is influencing another part Y. The three-dimensional manifold $R^2 \times S^1$ is replaced by an n-dimensional manifold M implicitly describing the states of Y. Here n could be large, like $10^{10}$, and M could represent states of all neurons and synapses in Y, or chemical concentrations, or spike densities, etc. The differential equation is replaced by a generic flow $\varphi$ on M, representing activity of Y. Generic implies that $\varphi$ has attractors. Meanwhile the control parameters are replaced by another high-dimensional manifold Q representing states of X. As the control varies in Q it causes the flow $\varphi$ to vary and bifurcate. The bifurcation points form a subset B of Q which is in general stratified. In particular the strata of codimension 2 determine elementary bifurcations of a single attractor of $\varphi$ into two attractors. In such a case the bifurcation is characterized by taking any two-dimensional disc C in Q transversal to the stratum.

Meanwhile, if the attractor is a submanifold of M (of arbitrary co-dimension k) we can construct a two-dimensional surface S in M by plotting the intersections of the attractor with a fixed transverse k-dimensional disc, as the control varies in C. We now appeal to Thom's theorem: the map $S \to C$ is differentiably equivalent to the above picture. Therefore this three-dimensional graph is an accurate description of bifurcation of brain activity.

Now the attractors are the stable global dynamic states of X and if we assume implicity that remembered concepts or stored behaviour patterns are represented by attractors then the way that the mind jumps from one concept to another or the way an animal switches from one behavioural pattern to another must be represented by the bifurcation of attractors. Therefore the above graph is an accurate model for elementary jumps of mind or behaviour.

Example (i)

Rage and fear are conflicting factors influencing aggression (see Refs. 2 and 3).

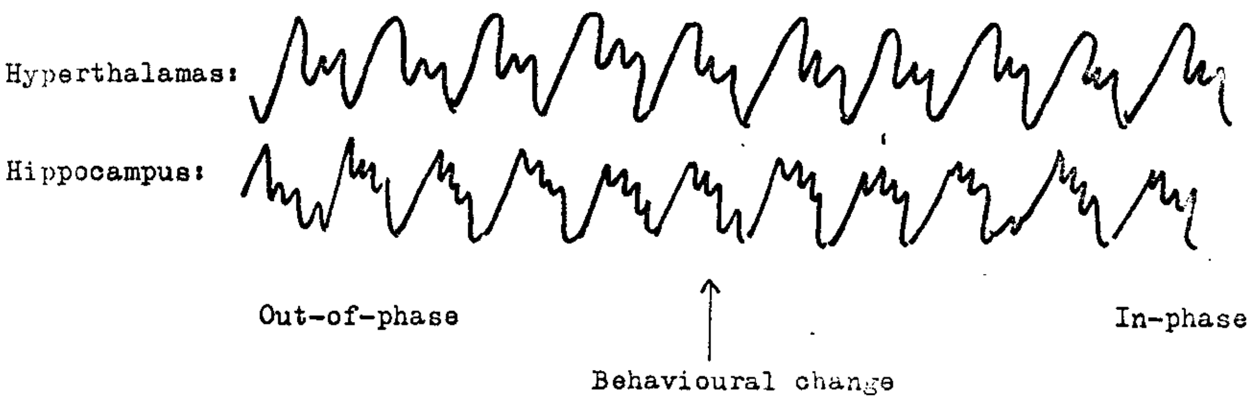
Example (ii)

Fix $\omega < 0$. A smooth increase of F causes a smooth increase of A. This might represent the recall of a continuous memory. Suppose now that this memory is not recalled for several days or years. One might expect a slight disorganization of the main synaptic pathways involved in the attractor representing this particular memory, due to the fact that all pathways are multipurpose and nearly all pathways are used nearly all the time in greater or lesser degree for nearly all brain activities. As a result one might expect a marginal slowing down of the attractor and therefore a relative speeding up of the forcing term. Therefore, in the notation of Duffing's equation, the frequency $\omega$ would change from being a negative constant to a positive constant. Notice that we are making assumptions here that the controls in part X of the brain have qualities of amplitude F and frequency $\omega$ similar to those of the forcing term in Duffing's equation.

If the recall-instruction F is now stimulated, in other words if there is a smooth increase of F, then the recalled memory A will jump and omit the middle section. Perhaps this is why an immediate recall of yesterday's events is full of gaps. However, if we work at trying to remember what was in those gaps, then the synaptic pathways may be marginally reorganized again, the attractor speeded up, the forcing term relatively slowed until $\omega < 0$, and consequently the intermediate continuous memory A recovered. Perhaps this is why with a little thought we seem to be able to remember things.

Experiment

Examples (i) and (ii) suggest a large variety of possible psychological experiments involving easily observable jumps of mind or behaviour. What would be even more interesting would be to correlate the psychology with the neurology. The mathematics has found a bridge between the two and if the bifurcation involved is Duffing-like, then the mental or behavioural jump might coincide with observable phase shifts in EEG recordings. For example, recordings from electrodes embedded in hyperthalamus and hippocampus or frontal lobes might exhibit a phase shift of the following type at a behavioural jump:



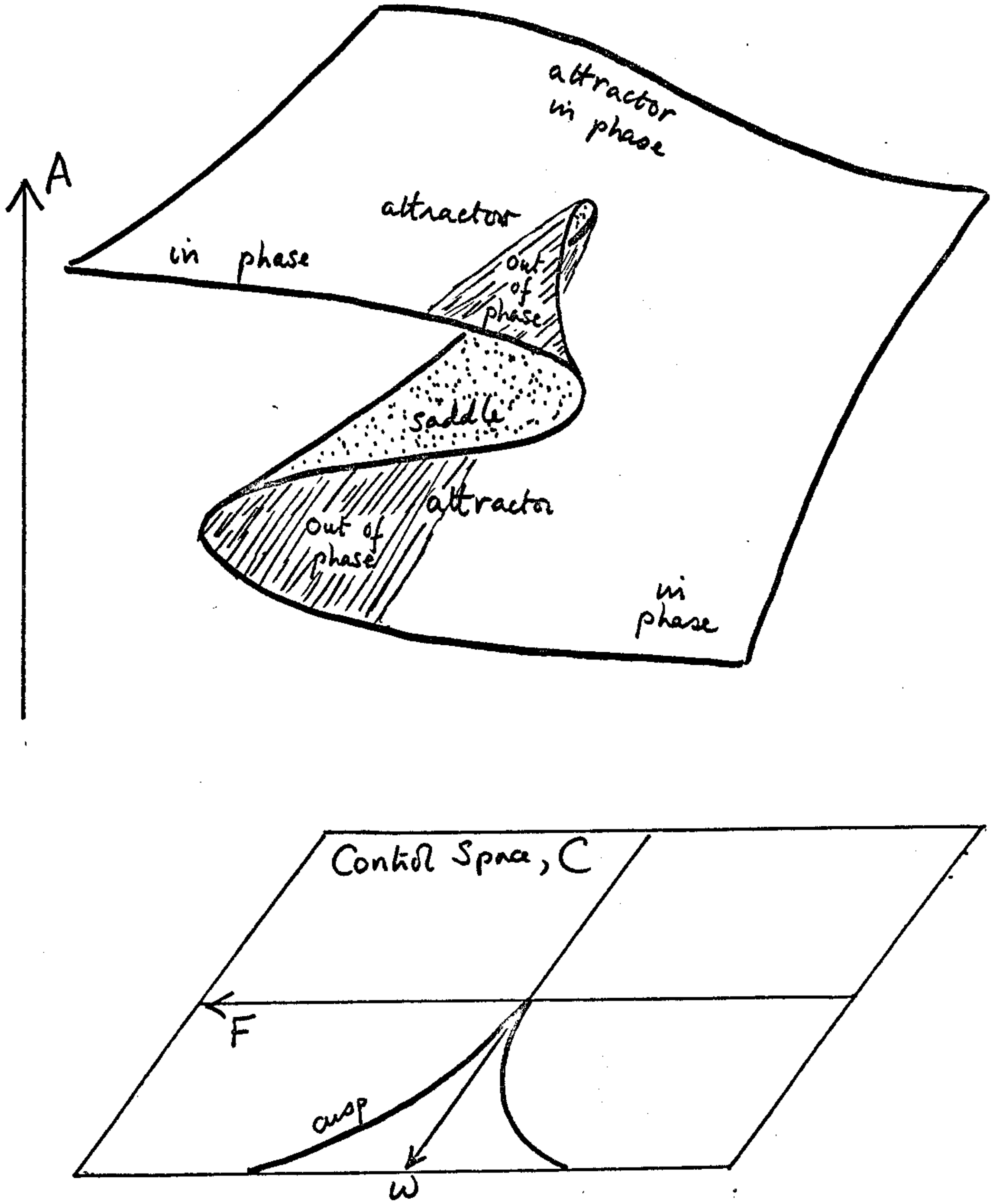



Fig.1

The behaviour of the amplitude A of the solution of Duffing's equation as a function of the control variables ω,F .

The observer might not notice such a shift unless looking for it. Phase shifts have been observed by Adey and others[4] in cats when the cat makes a wrong turning in a maze.

## REFERENCES


1) R. Thom, Topological Models in Biology, Topology 8, 313-335 (1969).

2) E.C. Zeeman, Geometry of Catastrophe, Times Lit. Suppl. (10 December 1971).

3) E.C. Zeeman, Applications of Catastrophe Theory, Bull. Lond. Math. Soc. (to appear).

4) W.R. Adey, C.W. Dunlop and C.E. Hendrix, Arch. Neurol. 3, 74-90 (1960).