

## The Essential Instability of Systems with Threshold, and some Possible Applications to Psychiatry

W. R. ASHBY, H. VON FOERSTER AND C. C. WALKER

*University of Illinois, Urbana, Ill. (U.S.A.)*

### INTRODUCTION

For half a century, the widespread occurrence of threshold in the nervous system, and the importance of threshold in the details of neuron activity, have been well known. There is less known, however, about how threshold would show in the large — in the behavior of the organism as a whole. Two studies (Beurle, 1956; Farley and Clark, 1961) have been made of the behavior of waves of activity traveling through a nerve net. Both studies have shown that such a net would have difficulty in maintaining a steady activity, for the wave of activity tends either to die out completely or to increase to saturation. Far from being tractable and steady, from the standpoint of biological usefulness such a network displays an essential instability. Not only does it tend rapidly to the extremes of inactivity or activity, but, once there, it can be moved away from the extreme only with difficulty.

This finding deserves emphasis because it is quite contrary to the plausible idea that threshold stabilizes a network. It also suggests that the actual brain must incorporate some mechanism that actively opposes the instability.

The studies cited are complex and do not allow the instability and the threshold to be related directly and simply. Here, we shall show that an extremely general and simple model still allows the relation to be displayed clearly. It also allows us to see more readily what is essential.

### I. ASSUMPTIONS AND DEFINITIONS

Consider reacting units in great numbers joined to form a network so large that we do not have to consider edge effects. Each reacting unit is affected by pulses (of some physical activity) that come to it, and it reacts by emitting (or not) pulses of the same physical activity. At any moment, these pulses will be in existence at various points in the net; by considering the actual number of pulses as a fraction of the possible number, we can speak of the 'density' of pulses over the net. We can also speak of the probability that a particular point has a pulse, provided we are careful to distinguish the two sample spaces: (1) various places at one moment, and (2) various

moments at one place. In this paper we shall be especially interested in how the density changes with time.

To be precise, let us assume that all the reacting units are identical, that each reacting unit has  $n$  input channels, and that it emits a pulse if and only if at least  $\theta$  of its inputs receive a pulse. We will assume here that time is advancing in discrete steps, and that the system's next state depends determinately on its state at the preceding instant of time. For this reason we assume that all pulses arrive at all inputs simultaneously. After  $\Delta t$ , the consequent pulses occur at the outputs, and thereby become the next set of pulses at the inputs.

The reacting unit's inputs are assumed to receive pulses independently, in that the occurrence of a pulse on one input is not to affect the probability of occurrence of a pulse on another input. The assumption implies that the net has no short loops, either by branches re-converging in feed-forward, or by short loops of feed-back.

*Change of density with time*

Given that the network is occupied by pulses to a density  $d$ , we can now find what will be its density  $d'$  at one step  $\Delta t$  later.

Since we are assuming that density and probability may be used interchangeably, the probability  $d_i$  that precisely  $i$  inputs on a particular reacting unit receive pulses is given by

$$d_i = \binom{n}{i} d^i (1 - d)^{n-i}. \tag{1}$$

Consequently, the probability that at least  $\theta$  inputs are active, *i.e.*, that the unit fires, is given by the cumulative binomial probability function (Ordnance Corps, 1952):

$$d' = \sum_{i=\theta}^n \binom{n}{i} d^i (1 - d)^{n-i}. \tag{2}$$

It is convenient to notice that  $d'$  is also given by

$$d' = \frac{n!}{(\theta - 1)!(n - \theta)!} \int_0^d x^{\theta-1} (1 - x)^{n-\theta} dx. \tag{3}$$

*Stability of the density*

These equations, by giving the next value ( $d'$ ) of  $d$ , as a function of its preceding value (and of the parameters  $n$  and  $\theta$ ) give, in effect, a difference equation: as  $d$  gives  $d'$ , so can  $d'$  give  $d''$ ,  $d''$  give  $d'''$ , and so on. In this way we can find what happens to the density as the time is continued indefinitely. When solved,  $d$  must be considered as a function of time, now introduced in the place of the mere difference  $\Delta t$ .

The question that concerns us here is: will the value of  $d$  tend to stay near its initial value, or will it tend to run away to an extreme? To answer this question, two distinct questions must be considered. First, do equilibril densities exist? Do densities exist which, once established, thereafter produce the same densities, densities for which  $d' = d$ ? (Such equilibril densities will be denoted by  $d^*$ .) Secondly, how do the densities change in the neighborhood of these  $d^*$ 's? If a perturbation from  $d^*$  is follow-

ed by a return to  $d^*$ , the equilibrium is 'stable'. Mathematically, a stable equilibrium exists at  $d^*$  when

$$\frac{\partial d'}{\partial d} < 1, \quad (4)$$

the derivative being given its value at  $d^*$ .

The manner in which  $d'$  depends on  $d$  can be seen by examination of the curve that shows their functional relation; variations in the parameters  $n$  and  $\theta$  will generate a family of such curves. They are best discussed by the separation off of the extreme cases.

Consider first the class in which  $\theta$  has those values other than its possible extreme values of 1 and  $n$ .

By differentiating equation (3) both once and twice with respect to  $d$  it can easily be seen that  $d'(d)$  has zero slope at both  $d = 0$  and  $d = 1$ , and exhibits a single inflection at the point  $d = (\theta - 1)/(n - 1)$ . Since  $d' = 1$  at  $d = 1$ , and  $d'(d)$  is continuous in the integral,  $d'(d)$  can cross the diagonal ( $d' = d$ ) only once. At that point (one of the values of  $d^*$ ) its slope must exceed one, for it starts below the diagonal and finishes above it. Thus when  $1 < \theta < n$ , the stability criterion (4) is not satisfied by any points between 0 and 1. The density is stable only at 0 and at 1.

At the extremes of  $\theta$ , when  $\theta$  equals 1 or  $n$ , the network is degenerate and may be dismissed as uninteresting. Its behavior would not be essentially different from that of the other case.

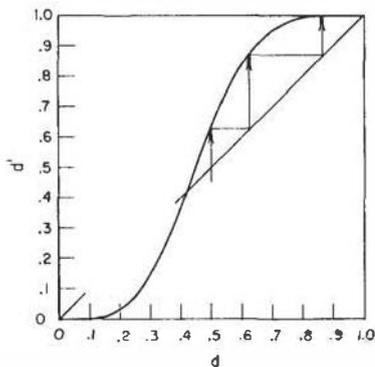


Fig. 1. Density at the next step ( $d'$ ) as a function of the previous density ( $d$ ), (threshold 5, 10 inputs).

Fig. 1 illustrates the situation in the simple case where  $n = 10$  and  $\theta = 5$ . The equilibria  $d^*$  are the three points where  $d'(d)$  intersects the diagonal: at 0, 0.42, and 1. Those at 0 and 1 are stable; that at 0.42 is unstable in the sense that the slightest perturbation from this value is followed by the density going to one of its extreme values. In Fig. 1 the changes in density from an initial value of 0.5 are indicated by the stairway. Fig. 2 shows how various perturbations provoke a runaway to an extreme value.

How fast does the density diverge? Fig. 2 (with the time marked in arbitrary steps) shows how rapid is the divergence in the particular case when  $n = 10$  and  $\theta = 5$ . Perturbations of more than one per cent result, in fewer than half a dozen steps, in the density being practically at 0 or 1. There is no reason to believe that our example is exceptional.

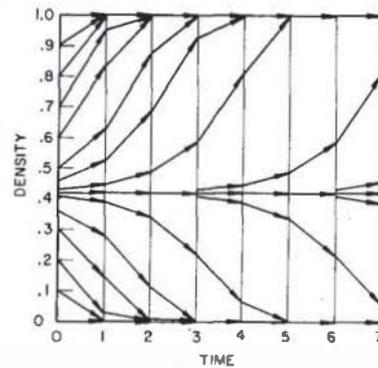


Fig. 2. How density changes with time, from various initial densities (threshold 5, 10 inputs).

## II. THRESHOLD VARIABILITY

The main point of our paper is to demonstrate that certain types of systems with threshold tend to be markedly unstable. As it is not impossible that the brain uses such systems, yet is not markedly unstable, it is of interest to glance at the question whether some simple extra factor could eliminate the instability.

One obvious possibility would occur if the threshold were not fixed but variable. Any density  $d$  (other than exactly 0 or 1) can be made to give a subsequent density  $d'$  (either more or less than  $d$ ) by a suitable value of  $\theta$ . So, if the threshold is varied quickly enough, the network's density can be driven in any desired direction; thus, a feedback from density to threshold could result in a network capable of operating indefinitely without  $d$  getting stuck on 0 or 1. Such a feedback could easily be provided by evolution and natural selection; there must be many ways by which an increase in pulse-density could lead, by changed physico-chemical conditions, to a change in threshold value.

Discussion of specific details would be out of place in this paper. We would, however, like to glance at the events that would occur if, once provided by natural selection, the feedback mechanism went wrong.

## III. FEEDBACK FAILURE AND PSYCHIATRY

Such a feedback would have to be provided by genetic control. What would happen in our network if the feedback were different — malformed or malfunctioning in some way? The question is not without psychiatric interest.

As a first point, we may notice that if the net is at all extensive, the corrective feedback would have to be applied all over it. As Fig. 2 showed, a very few steps uncorrected are sufficient to allow the density to go to an extreme value; and were the corrections fed back so that patches of neurons, or other sites of threshold, were more than a unit or two from a site of corrective action, then those patches would be unstable and would go rapidly to extreme values. Suppose that, for whatever reason, some patches failed to receive the corrective action; what would happen?

If the system generally were adaptive: very little. The uncorrected patches will

almost at once go to extreme values, and then one of two events will occur. If the values stay at the extreme, the values will become part of the *constant* conditions to which the adaptive parts are adapting, and the adaptation will occur in the usual way, the adaptation being merely forced to follow a different route. If the extreme value is not permanent but conditional, the value will fluctuate with the conditions around it, and the density in the patch will simply alternate between two values (equivalent to 0 or 1). The density in the patch has now simply become a variable of the system, and again adaptation will be in no way prevented, only forced to follow another route. In both cases failure of the corrective feedback would entail only that the patch, instead of having rich internal possibilities of variations, will either lose them completely, or will diminish them to the point of its becoming a mere binary. The performance of the whole system will tend simply to fall to that of a fractionally simpler system, the fractional fall being approximately proportional to the size of the patch.

So much for patchy failures. What about failures of the feedback generally? A point of interest is that every feedback, especially when its variation with time is continuous, readily develops oscillation — starts to ‘hunt’ — especially if there is an appreciable delay in the transmission of effects round the loop. Should such an oscillation occur, it would be the density that would show the rhythm. The changes would have to occur at a whole order of time slower than those that generated the density, for many pulses must occur, over much more than an average pulse-interval, before the density can be sensed as a density. Thus the hunting-oscillations would be expected to occur with a much longer period than that of the pulse-event. There is clearly a possibility that the Berger rhythm might owe its origin to this cause; more we can hardly say at the moment.

If the feedback becomes generally inactive, the whole system would soon go either to complete inactivity, or to maximal activity; and from these states it could be moved only with difficulty.

If the density went to 0, and remained there, we would see a system that would remain inactive if undisturbed, that would respond only to a minimal degree if disturbed, and that would return to complete inactivity as soon as the disturbance ceased. Its behavior would be much like that of a patient in melancholia. On the other hand, if the system had gone to maximal excitation, and was kept there by its own instability, one would see a system active at all times, even when no disturbances come from the outside. It would be active because it was driven by its own activity, and would give an appearance much like that of a patient in acute mania. These two correspondences, between dynamic theory and clinical observation, are obviously of interest, but much more work would be necessary before we can take the correspondence further.

One last possibility remains to be considered. How might drugs and other biochemical factors affect the system with such an instability, corrected by a feedback-varied threshold? If the drug acts simply to cause a general and uniform shift in the threshold, any change it induces will be at once opposed by changes caused by the supposed correcting feedback; so the drug, in the presence of the correcting mechanism, would simply have a reduced effect. If however, the drug, instead of affecting the threshold, weakens the feedback itself, the system would become one in which

localized runaways could occur. The system would break, by its own dynamics, into a patchwork of high, low, and moderate densities. The behavior shown by such a system would probably be characterized chiefly by its bizarreness. Today a number of drugs are known — the so-called ataractics — characterized by their power of inducing behaviors of extremely mixed type. Again, whether the mechanism we have considered will be found to be relevant to these questions will have to be elicited by further study.

#### ACKNOWLEDGEMENT

Work supported by Contract NSFG 17414 of the National Science Foundation.

#### SUMMARY

Any large system that uses threshold as the criterion for whether transmission is to occur at each node is fundamentally unstable in its density of active points. Any increase in density, by increasing the chance that other stimuli will be successfully transmitted, tends to cause yet further increases in density. Calculation of the exact effects confirms the expectation.

Were the threshold fixed, and were the conditions as assumed, any network using threshold would be erratic in action, tending continually to run away either to complete inactivity or to maximal activity. Since many processes in the nervous system are not normally seen to exhibit such runaways, there must be some factors opposing the primary instability. One is described briefly, showing that stabilization of a thresholded network is not difficult.

The application of these facts to various symptoms seen in psychiatry is considered briefly.

#### REFERENCES

- BEURLE, R. L., (1956); Properties of a mass of cells capable of regenerating pulses. *Philosophical Transactions of the Royal Society of London*, **B**, **240**, 55-94.
- FARLEY, B. G., AND CLARK, W. A., (1961); Activity in networks of neuron-like elements. *Proceedings of the Fourth Symposium on Information Theory*. C. Cherry, Editor. London, Butterworth, Ltd.
- ORDNANCE CORPS, (1952); *Tables of the Cumulative Binomial Probabilities*. **ORDP 20-11**. Washington, D.C. Office of the Chief of Ordnance.

#### DISCUSSION

WIENER: A system where you can easily get instability is a system of alternators. If you have switched in an alternator of a very different frequency or an appreciable different frequency or which has been wrongly phased, the system will blow up. On the other hand, if they are all locked in nothing will happen. How is this engineering problem of phasing a system which may exist over the whole country handled? In the old days human powers were used. A man watched two needles moving ahead and representing the rotation, run by synchronous motors, from different generators.

When the needles were nearly at the same speed, the apparatus was switched in. Today this is done automatically. An automatic mechanism switches the generator into the circuit only when the phase and frequency of the generator is correct. An unloaded generator wanders through the phases and when it comes to the right phase it is pulled in. If by some catastrophe one generator is overloaded a circuit breaker is automatically blown out and the generator is again freewheeling. In other words, in systems of the sort where instabilities are not only possible but very likely, the instabilities are limited by highly non-linear apparatus which will close or open a circuit at the right time. If you stay in a narrow range, there is also a tendency, and this is most important, for the different generators to pull one another into frequency. In studying the nervous system it will be most useful to study analogies of this sort which exist in much smaller (operationally) similar complexes of machinery.

An other matter in which I am very much interested is the question of the one stage and the large scale change that you get by the action of the system. I am working on this in some statistical mechanical fields. I think it is possible to set up a statistical theory of the system to make the small changes of the system an infinitesimal transformation. Under certain circumstances this is what I am doing now with some statistical mechanical work. I hope to show that if this small change does not alter probabilities and the system is in equilibrium, neither will a large change. In other words, I feel that statistical methods and those involving integration in function space are powerful tools available for this work.

COWAN: I must confess to being rather sceptical of this line of attack. Noise is used to stabilize outputs of a system but then there is the problem of controlling the noise and making sure that the output is not a random function of the inputs. Suppose there is no noise, then, if for one element the result is what Von Neumann termed restoration, the result is trivial. If we are not dealing with one linear threshold element as this appears to be with a network, then the result is also not very interesting because the classic functions which can be obtained from such a network are trivial. You can only get  $2^n$  plus a small number of functions out of the  $2^{2^n}$  functions that one generally deals with in neural nets.

I have only considered systems in which you have excitation and what we call linear thresholds elements. There is fairly good evidence, although I hesitate to make the jump from neural nets to the nervous system, that there is a lot of inhibitions in the nervous system even at the single unit level. There are a lot of interactions of afferents in the nervous system and this gives a much richer set of functions to play with. In fact, you can get any kind of functional behavior out of a system. The nervous system does not appear to have this kind of simple linear elements, but in fact has a large collection of different kinds of elements: some monotonic, some anti-monotonic, all which are noisy.

ROSS ASHBY: The effect of inhibitory contributions has not yet been considered by us.

WIENER: I think that one of the major problems is that the nervous system is not working near a level of zero activity. Fluctuations about random activity must be considered which means that you cannot neglect the background activity.

ROSS ASHBY: One point I would like to make is that if the pulses are not to be grossly inefficient as carriers of information, their distributions and arrivals must be largely independent statistically. But independence of arrival generates the binomial distribution in the numbers arriving in a given time, and it is the binomial shape, with a hump in the middle rather than a hollow, that causes the instability. So there is, in the case we are considering, a fundamental conflict between the need for efficiency as a channel of communication and the need for stability.

COWAN: Surely you are not interested in inputs which are carrying independence but in fact are carrying patterns.

ROSS ASHBY: If the nervous system is to be efficient, it must re-code the patterns, getting rid of the redundancy implied by the 'pattern', until the information is compressed into a more efficient form, which must be one whose variables vary independently.

WIENER: I do not see any need of the assumption that the nervous system is working near maximum information carrying efficiency. It probably takes tremendous waste to make it work at all. I think this assumption, that the nervous system is governed by maximum possible information carrying efficiency, is wrong from the beginning.

ROSS ASHBY: I must disagree with you, unless we are talking about different things. I am not assuming and do not believe that the nervous system works with the efficiencies achieved by communication engineers: I was considering 'efficiency' only in the sense of avoiding the grossest redundancy.

WIENER: What reason have you to believe that there is no great redundancy?

ROSS ASHBY: In that case it would be highly efficient. It would be carrying up to its maximal capacity.

COWAN: The trouble with the informational analysis of the nervous system is a communication channel. It completely neglects what is very basic, the engineering aspect of cost function. If the nervous system is regarded as a communication channel supplying energy to the system it may turn out that what is informationally a good thing to do, in fact, is an extremely bad thing to do. The important consideration is the energetic one rather than the informational one.

ROSS ASHBY: I agree that both are important. When a rabbit, for instance, has to run for its life, the weight of one extra gram of brain may be an appreciable handicap. If the species develops brains with an extra gram of substance, that gram must give at least a compensating advantage of better muscular coordination, or maneuverability, or learning the predator's ways, etc. Because weight alone is a handicap, natural selection will incessantly press for the avoidance of gross inefficiencies in information processing.

WIENER: It is more or less a teleological argument that we must be near the maximum efficiency informationally. We may be near the maximum we can do as an engineering job, and likely we are, but that involves the use of every bit of information.