PRINCIPLES FOR THE QUANTITATIVE STUDY OF STABILITY IN A DYNAMIC WHOLE SYSTEM; WITH SOME APPLICATIONS TO THE NERVOUS SYSTEM.

By W. R. Ashby, M.A., M.D.,
Pathologist, St. Andrew's Hospital, Northampton;
Major, Specialist Pathologist, R.A.M.C.

[Received February 23, 1946.]

The discovery of the nervous impulse established that the nervous system was fundamentally mechanistic, and that it was highly dynamic; and all subsequent work has confirmed these earlier views. The stabilities of inanimate dynamic systems are frequently of high importance and rich in applications. It would seem likely that an investigation of the stabilities of the nervous system would yield interesting results, but it is only recently that the importance of stability in the nervous system has been unequivocally recognized. But it has now been established that such properties are of central importance in the processes of integration and co-ordination. Advance has probably been hampered by the difficulties of the subject.

It is proposed here to give an outline of the basic principles relating to the stability of complex dynamic systems, stressing particularly the quantitative aspect. Though the subject is extensive, this paper will be confined to essentials.

Complete Dynamic Systems.

Every physico-chemical system determinate in its behaviour has the property that its configuration at any instant specifies its subsequent configurations. In other words, given complete knowledge of its state at one instant, it is possible to predict what will happen subsequently. It may be shown that this is identical with saying that all such systems may be specified by equations of form—

$$\left.\begin{aligned} \frac{dx_1}{dt} &= f_1(x_1, x_2, \ldots, x_n) \\ \frac{dx_2}{dt} &= f_2(x_1, x_2, \ldots, x_n) \\ &\cdots\cdots \\ \frac{dx_n}{dt} &= f_n(x_1, x_2, \ldots, x_n) \end{aligned}\right\} \quad \ldots\ldots(1)$$

where $x_1, x_2, \ldots, x_n$ are the variables specifying the state of the system, t is the time, and the f's are functional forms. For example, the simple pendulum is specified by—

$$\begin{aligned} \frac{dx_1}{dt} &= x_2 \\ \frac{dx_2}{dt} &= -\frac{g}{l}\sin x_1, \end{aligned}$$

https://doi.org/10.1192/bjp.92.387.319 Published online by Cambridge University Press

where x_1 is the angle of deviation from the vertical, x_2 is its angular velocity, l is the length, and g is the acceleration of gravity (though the equation is not usually written in this form). The important point is that the fluxions of $x_1, x_2, \ldots, x_n$ can be specified in terms of $x_1, x_2, \ldots x_n$ themselves, *no other functions of the time being needed.* Such a system is called "complete." Only a complete system can properly be said to have stability, the concept having no meaning in other cases. Equations (1) are, then, the fundamental means of specifying all dynamic systems capable of having stability (or instability).

It is well known that such a set of differential equations has a unique set of solutions :

$$\left.\begin{aligned} x_1 &= F_1(x_1^\circ, x_2^\circ, \ldots, x_n^\circ ; t) \\ x_2 &= F_2(x_1^\circ, x_2^\circ, \ldots, x_n^\circ ; t) \\ &\cdots\cdots\cdots \\ x_n &= F_n(x_1^\circ, x_2^\circ, \ldots, x_n^\circ ; t) \end{aligned}\right\} \ldots\ldots (2)$$

where $x_1^\circ, x_2^\circ, \ldots, x_n^\circ$ is a set of n arbitrary constants. This set of solutions (2) may be interpreted as specifying the behaviour of the set of variables $x_1, x_2, \ldots, x_n$ if started at a configuration—

$$\begin{aligned} x_1 &= x_1^\circ \\ x_2 &= x_2^\circ \\ &\cdots\cdots \\ x_n &= x_n^\circ \end{aligned}$$

with t ranging from 0 onwards.

In n-dimensional space such a set of values of $x_1, x_2, \ldots, x_n$ depending on the single parameter t defines a line, and there is a one-to-one correspondence between such lines in the space and the modes of behaviour (i.e., changes of configuration) of the system. These lines are of fundamental importance because every property deducible by a direct study of the system's behaviour corresponds to a definite geometric property of the lines, and vice versa.

The fundamental definition of "stability" may now be stated : Given a complete system and its equations in form (2), and given $x_1^\circ, x_2^\circ, \ldots, x_n^\circ$ and some arbitrary value of t, say T ; if we can find finite numbers $x_1', x_1'', x_2', x_2'', \ldots, x_n', x_n''$ independent of the time so that for all values of t greater than T we have—

$$\begin{aligned} x_1' &< x_1 < x_1'' \\ x_2' &< x_2 < x_2'' \\ &\cdots\cdots\cdots \\ x_n' &< x_n < x_n'' \end{aligned}$$

then the line from $x_1^\circ, x_2^\circ, \ldots, x_n^\circ$ is *stable.*

In simpler but less accurate words, this means that, given the starting configuration, then however long the system is allowed to go on for, all the variables are to remain within some finite limits. It is easily shown that less general definitions are unsatisfactory. Most of the commonly used definitions apply only to systems of some restricted class. It will be noted that the

https://doi.org/10.1192/bjp.92.387.319 Published online by Cambridge University Press

definition is essentially quantitative; certain numbers either do, or do not, exceed certain other numbers.

The reader may be surprised that no mention is made of considerations of "energy" or of "force." These are in fact irrelevant in the general case, though often of use in more restricted investigations.

Perhaps more important from our point of view is the deduction that as one dynamic system can start from any one of many configurations, and as each resulting line of behaviour can be stable or not independently of the others, so *a single dynamic system may be stable to some disturbances and unstable to others*—a fact which will scarcely surprise a psychiatrist.

An important aspect of the line of approach is that it accords with one's common-sense approach to the subject. The differential equations are simply the most natural and direct way of writing down the way the individual variables affect one another, i.e., their physical relations, and this in turn is a direct consequence of the way in which the parts are assembled in relation to each other. The solution of these differential equations gives in direct form a description of how this same assembly of parts will behave when observed over some finite time. *The relation of the parts to the whole is therefore exactly the same as the relation of the differential equations to their solution.*

These theorems supply the basic requirements for a rigorous theory of stability in a dynamic system; the parts and their interactions specify the differential equations, the solutions of these give the observable behaviour of the system, and this can be tested for stability by a direct application of the definition. An interesting example is given in Ashby (1945*a*).

Ultrastable Systems.

Special types of dynamic systems will naturally have special types of stabilities. It has recently been shown (Ashby, 1945*b*) that dynamic systems which contain large numbers of step-functions have the peculiar and remarkable property of ultrastability—that they always undergo changes of internal organization until they are stable to any disturbances affecting them. (This property is quite different from the mere change of configuration after disturbance shown by every system which is stable: the "ultrastable" system will change from being inherently unstable to being inherently stable.) But this system is only of indirect interest to us now. It shows no quantitative gradation of stability as far as a given disturbance is concerned, since it is, in the limit, always just "stable."

If, however, the system is not only ultrastable but also contains large numbers of variables which are constant over finite intervals ("part-functions") then it may be shown that such a system, still ultrastable, can, and will, change instabilities to some disturbances into stabilities *without destroying previously established stabilities*. The only known examples of systems containing variables of these types (step- and part-functions) are the nervous system and certain types of social organization. It is possible, therefore, that it is to these particular peculiarities that the nervous system owes some, at least, of its remarkable properties.

https://doi.org/10.1192/bjp.92.387.319 Published online by Cambridge University Press

Applications.

That these stabilities are of fundamental biological importance to the organism has been noted repeatedly. Thus Pavlov (1927) writes: "(The animal organism) being a definite circumscribed material system, can only continue to exist so long as it is in continuous equilibrium with the forces external to it."

The theme has been discussed in more detail by Ashby (1940). The nervous system, like the other organs of the body, has been fashioned by natural selection; and survival implies that the organism can maintain constant (or within "physiological" limits) a number of important variables, e.g., body temperature, oxygen supply, food supply. But this "maintaining constant" conforms to the definition of stability given above. There is therefore little doubt of the basic importance of stability to the organism and therefore to the nervous system.

The principles outlined above apply equally to systems living or dead, and the distinction is, in this connection, irrelevant. That they will apply to the nervous system is a prime hypothesis of the physiologist, since they apply to all systems whose behaviour is determined at every moment by the state of the system at that moment. If it is granted that the nervous system is of this type, then the other conclusions follow, not of physiological, but of mathematical necessity.

The simple applications of these principles will be seen only in the simpler activities of the nervous system. In the reflexes are seen various types of mechanism whose sole activity is to restore a *status quo* after disturbance. No further discussion can be given without going into individual details.

More interesting, however, are the activities involving the cerebral cortex. It is here that the nervous system possesses both the step-functions which enable it to show the *self*-stabilizing features of an ultrastable system, and also the part-functions which enable it to *accumulate* stable neuronic arcs. The two aspects will be discussed in order.

(1) That the cerebral cortex must contain many variables behaving as step-functions may be shown decisively by the proof that certain forms of behaviour necessarily imply the existence of such variables. This proof, though in a sense indirect, is a direct mathematical deduction from the observed facts. To the existence of these step-functions must be attributed the recognized property of the higher nervous system that although it may at first produce inappropriate behaviour (i.e., a child trying to grasp at dancing flames), yet it invariably tends, after a time, to change to a mode more compatible with survival ("the burned child dreads the fire"). This change is a direct consequence of its ultrastability (Ashby, 1945*b*).

(2) The existence of part-functions in the nervous system is well known from more direct evidence. The facts that threshold values exist and that summation is usually necessary to excite a cell (e.g., Sherrington, 1931) are themselves merely examples of part-functions, as may be seen at once by comparison with the definition. These physiological features necessarily

https://doi.org/10.1192/bjp.92.387.319 Published online by Cambridge University Press

make the nervous system multistable, since it has been proved that any ultrastable system tends to the multistable form as the number of variables which are part-functions increases. The most important characteristic of a multistable system is that it will progressively improve its stabilization against disturbance, the essential point being that the later reorganizations necessary to improve stability against later disturbances can be made without disorganizing the stabilities previously established against previous disturbances. This is obviously of prime importance in the development of any really complex adaptation. It may therefore be suspected that the properties of " threshold " and " summation " are really of fundamental importance in neural adaptation, and that, far from being produced as by-products of other requirements, are really of importance in themselves, and that they have been established by natural selection, nervous systems without them being defective in powers of adaptation. Whether changes in these physiological features, made perhaps by disease, would produce the obvious maladaptations of mental disorder is a question which involves too much to be discussed here.

Summary.

Stability of the nervous system is necessary both for biological adaptation and survival, and also for the production of normal behaviour. Investigation of the subject has been hampered partly by its complexity and partly by lack of suitable technical methods.

An account is given of the fundamental principles for handling such complex stabilities, the subject being treated by an extension of the methods of mathematical physics.

Reference is made to recent developments, and some applications to the nervous system are discussed.

It is a pleasure to acknowledge my thanks to Prof. J. A. Carroll for his interest and encouragement.

References.

Ashby, W. R. (1940), *J. Ment. Sci.*, **86**, 478.
Idem (1945*a*), *Nature*, **155**, 242.
Idem (1945*b*), *J. Gen. Psychol.*, **32**, 13.
Pavlov, I. P. (1927), *Conditioned Reflexes*. Oxford.
Sherrington, C. S. (1931), *Brain*, **54**, 1.

https://doi.org/10.1192/bjp.92.387.319 Published online by Cambridge University Press