

## SOME OBSERVATIONS ON THE SIMPLE NEURON CIRCUIT

WALTER PITTS

THE UNIVERSITY OF CHICAGO

A new point of view in the theory of neuron networks is here adumbrated in its relation to the simple circuit: it is shown how these methods enable us to extend considerably and unify previous results for this case in a much simpler way.

### I. *Introduction and Definitions*

The sequent remarks are intended primarily as a prolegomenon and an illustration of a more general theory of neuron networks to be later redacted. Herein, as in the more general discussion, we shall adopt the linear model of excitation as employed by A. S. Householder in four previous papers in this *Bulletin*; (Householder 1941a, b, c, 1942) our point of view will be somewhat different, however, in that we shall be primarily concerned with non-steady-state activity and the conditions under which a steady-state may be attained, rather than in the ascertainment of equilibrium points per se. We shall nevertheless find that this point of view has interesting consequences for problems of network statics.

We may commence by defining the *total conduction time* of a fiber as the sum of its conduction time and the synaptic delay at the postliminary synapse: we shall suppose, as can be done without substantive loss of generality, that all the total conduction times of fibers of the circuit  $C$  in question are equal: and we shall measure time so that this quantity is unity. We shall consider explicitly only the case where all fibers of  $C$  have remained unstimulated up to  $t = 0$ , a constant stimulation  $\lambda_i + h_i$  is applied to each synapse  $s_i$  in the interval  $t \in (0, 1)$ , and the external stimulation assumes thereafter a steady value of  $s_i$ . We shall indicate desinently how all other cases are subsumable with inessential novations under this one.

We shall employ the same notation as in (Householder, 1941a, b, c, 1942) with 'A' for the product of the activity parameters of  $C$  and the abbreviation  $\mu_i = \sigma_i^{(n)}$ , where there are  $n$  fibers in  $C$ . In addition to these, we shall find the following notations of value.

The excitation-pattern of  $C$  may be described in a *matrix*  $E$ , of  $n$  rows and an infinite number of columns, each of whose elements

$e_{rs}$  represents the excitation at the synapse  $s_r$  during the interval  $(s, s + 1)$ . The successive entries in the excitation matrix  $E$  may be computed recursively from those in its first column—these are the quantities  $\lambda_r$ —by the following rule, whose validity is evident: Given the elements of the  $p$ -th column, compute those of the  $p + 1$ -st thus: if the element  $e_{ip}$  is negative or zero, place  $\sigma_{i+1}$  in the  $i + 1$ -st row and  $p + 1$ -st column, or in the first row of the  $p + 1$ -st column if  $i = n$ . Otherwise put  $\sigma_{i+1} + a_i e_{ip}$  in this place. We shall say that  $C$  is in a *steady-state* during a series of  $n$  intervals  $(s, s + 1), \dots, (s + n - 1, s + n)$  if, for every  $p$  between  $s$  and  $s + n$ , the  $p$ -th and  $p + n$ -th columns of  $E$  are identical. If  $s$  is the smallest integer for which this is the case, we shall say that the steady state begins at the interval  $(s, s + 1)$ .

It will be seen that the construction of the matrix  $E$  implies that its infinite diagonals—where we take a diagonal to start again at the top of the succedent column whenever it reaches the last row of  $E$ —are wholly independent of one another, so that if we know the starting point of a diagonal of  $E_s$ , we can calculate the entries along it uncognizant of any other values in the matrix. Physically, this of course means that the activity in  $C$  can be regarded as composed of wholly independent impulses, commencing originally at a synapse  $s_j$  with a value  $\lambda_j$ , and journeying around  $C$  in irrelatlon to the impulses beginning at other synapses. We shall find it convenient to adopt this standpoint, and consider only the case of a single impulse, so that the complete solution must be derived by combining the results of our subsequent procedures for the separate diagonals, and a steady-state for the whole circuit is attained only when one has accrued for each separate diagonal.

We may define a set of intervals for every synapse  $s_i$ , which we shall call the ranges of succeeding synapses  $s_{i+j}$  from  $s_i$ , and denote by  $R_j(s_i)$ , by the following recursion:

$$(1). \quad \text{If } a_i > 0, R_1(s_i) = (\sigma_{i+1}, \infty).$$

$$\text{If } a_i < 0, R_1(s_i) = (\sigma_{i+1}, -\infty).$$

$$(2). \quad \text{If } R_j(s_i) = (m, n), R_{j+1}(s_i) = (\sigma_{i+j+1} + ma_{j+i},$$

$$\sigma_{i+j+1} + na_{j+i}), \text{ unless } m \text{ or } n \text{ or both are negative, in which case we replace } m \text{ or } n \text{ respectively in this expression by } 0.$$

If the range of  $s_{i+j}$  from  $s_i$  is  $(m, n)$ , the end-points  $m$  and  $n$  are denoted by  $L_j(s_i)$  and  $U_j(s_i)$ . If  $L_j(s_i) = U_j(s_i)$  for some  $j$ ,  $s_j$  is said to be *inaccessible* from  $s_i$ ; otherwise *accessible*. Clearly, if the suffixes  $i, j, h, 1$  are in cyclic order,  $R_{h-i}(s_i)$  is a proper subset of the

range of  $s_k$  from  $s_j$ ; it is also inaccessible from  $s_i$  or any preceding synapse. Physically, that the range of  $s_j$  from  $s_i$  is  $(m, n)$ , where say  $m < n$ , means that by properly varying  $y_i$ , we can cause  $y_j$  to assume any value within the interval  $(m, n)$ ; but we cannot from  $s_i$  cause  $y_j$  to diminish beyond  $m$ , nor to exceed  $n$ . If  $s_j$  is inaccessible from  $s_i$ , then nothing happening at  $s_i$  will have any effect upon the excitation at  $s_j$ , and consequently, if a circuit contain any synapse inaccessible from anywhere, it will attain a steady state immediately, in the second  $n$  columns of  $E$ , and this steady state will be wholly independent of the initial stimulation  $\lambda_i$ . Incidentally, if at any time there are two zeroes between  $s_i$  and  $s_j$ , between which is an odd number of inhibitory fibers,  $s_j$  is inaccessible from  $s_i$  (cf. Householder 1941b, lemma 1).

We shall now find it desirable to establish two lemmas of importance, the first of which will be necessary for our more general theory, the second whereof is a special case of a result later to be established for general networks.

#### LEMMA 1

*Let  $(m, n)$  be the range of a given synapse  $s_j$  from  $s_i$ . Suppose first that there are an even number of inhibitory fibers between  $s_i$  and  $s_j$ . Then we can find two numbers,  $\vartheta_0$  and  $\vartheta_1$ ,  $\vartheta_0 \leq \vartheta_1$ , such that for all values of  $y_i \leq \vartheta_0$  the consequent value of  $y_j$  is  $m$ ; for  $\vartheta_0 < y_i \leq \vartheta_1$ , the chain between  $s_i$  and  $s_j$  is completely active, and  $y_j$  is accordingly a linear function of  $y_i$ ; for  $y_i > \vartheta_1$ ,  $y_j$  is  $n$ . Second, suppose there is an odd number of inhibitory fibers between  $s_i$  and  $s_j$ . Then we can choose  $\vartheta_0$  and  $\vartheta_1$ ,  $\vartheta_0 \leq \vartheta_1$ , so that for  $y_i \leq \vartheta_0$ ,  $y_j = n$ ; for  $\vartheta_0 \leq y_i < \vartheta_1$ ,  $y_j$  is a linear function of  $y_i$ , and there are no zeroes between  $s_i$  and  $s_j$ ; for  $y_i > \vartheta_1$ ,  $y_i = m$ .*

For convenience in statement, we shall assume an even number of inhibitory fibers between  $s_i$  and  $s_{j+i}$ ; the proof may be extended to the other case with inessential changes. It is clearly sufficient to show (1) that as we increase  $y_i$  from 0 to  $+\infty$ , then, unless  $L_j(s_i) = -\infty$ , when  $\vartheta_0 = -\infty$ , that intermediate zero between  $s_i$  and  $s_{j+i}$  which is nearest to  $s_{j+i}$  is removed at most once, to produce a state of complete activity; and (2), if we augment  $y_i$  still further, so that a new zero is formed—otherwise  $U_j(s_i) = +\infty$ ,  $\vartheta_1 = +\infty$ —this zero cannot be removed by further increase in  $y_i$ . (1) follows thus. At values of  $y_i$  sufficiently small so that if  $R_j(s_i)$  is finite  $y_{j+i}$  is equal to it, there will be zeroes at synapses between  $s_i$  and  $s_{j+i}$ , of which one, say  $s_k$  is closest to  $s_{j+i}$ . Now suppose that, as we increase  $y_i$ , the zero at  $s_k$  is removed for a least value of  $y_i$ , say  $y_{i_0}$ . Now if there were a zero between  $s_i$  and  $s_k$ , say at  $s_{k'}$ , which remained unchanged as  $y_i$  assumed

the value  $y_{i_0}$ ,  $s_k$  could, of course, not be affected. If all zeroes between  $s_i$  and  $s_k$  disappeared for  $y_i = y_{i_0}$ , a state of complete activity would result, in accordance with (1). The only other possibility is that there be no zero between  $s_i$  and  $s_k$  for  $y_i < y_{i_0}$ , but one is initiated, say at  $s_{k'}$ , for  $y_i = y_{i_0}$ . Now if there is to be no interval of complete activity at all, then, precisely when the zero at  $s_{k'}$  is formed, or  $y_{k'} = 0$ , and not at all before,  $s_k$  must cease to be a zero, so that  $y_k > 0$ . Now  $y_k$  is a linear function of  $y_{k'}$  for all values of the latter greater than or equal to zero, since there is then no zero between them: consequently if, as is clearly possible, we select a small  $y'_{k'}$  satisfying  $0 < y'_{k'} < y_k$ ,  $y'_{k'}$  will be the value of this function for some argument  $y'_{k'}$  which exceeds zero, by continuity and monotonicity; and since  $y'_{k'} > 0$ , the zero at  $s_k$  is broken down for this value of  $y_k$  also; and consequently for a  $y'_{k'} > 0$ , which is contrary to hypothesis. Hence (1) holds. (2) follows immediately: if we have complete activity between  $s_i$  and  $s_j$  for a given range of values of  $y_i$ , then all  $y_k$  for  $i < k < j$  vary linearly with  $y_i$ ; and it is only when one of these, say  $y_k$ , diminishes (linearly) to zero that the complete activity is interrupted. Now if complete activity were to be reestablished for a higher value of  $y_i$ , then  $y_k$ , being the same monotone function of  $y_i$ , would be exceeded by zero, which is not compatible with complete activity. By (1), the zero formed at  $s_k$  could not be removed in any other way.

It is clear, by the continuity of the value of  $y_{i+j}$  as a function of  $y_i$ , that the quantities  $\vartheta_{0i}$ ,  $\vartheta_i$ , and the ranges are connected by the important relations

$$L_j(s_i) = A_{i,i+j} \vartheta_0 + \sigma_{i+j}^{(j)},$$

$$U_j(s_i) = A_{i,i+j} \vartheta_1 + \sigma_{i+j}^{(j)}.$$

For the pair  $\vartheta_0$ ,  $\vartheta_1$  in the case where  $s_j = s_{i+n} = s_i$ , as determined thus, we shall use the permanent notation  $\vartheta_{0i}$ ,  $\vartheta_{1i}$ . In this case we have

$$L_n(s_i) = A \vartheta_{0i} + \mu_i,$$

$$U_n(s_i) = A \vartheta_{1i} + \mu_i.$$

#### LEMMA 2.

*If a circuit C has been in complete activity for a period  $qn$  time-units in length, and  $\xi < 1$ , then the excitation of  $s_i$  at the time  $qn + \xi$  is given by*

$$y_i(qn + \xi) = \mu_i \frac{1 - A^q}{1 - A} + \lambda_i A^q, \quad (1)$$

and that at the synapse  $s_{i+j}$ ,  $j < n$ , for any time  $qn + j + \xi$ ,  $\xi < 1$ , by

$$y_{i+j}(qn + j + \xi) = \sigma_{i+j}^{(j)} + A_{i,i+j} \mu_i \frac{1 - A^q}{1 - A} + \lambda_i A^q. \quad (2)$$

If  $C$  is in complete activity we may write a difference equation for excitation at  $t + n$  as a function of that at  $t$ ; this is

$$y_i(t + n) = \mu_i + A y_i(t).$$

One verifies that equation (1) is a solution of this by substitution; and equation (2) is an immediate consequence of equation (1).

Armed with these results, we shall easily be able to determine the possible types of activity in the circuit  $C$ . To this end we may divide the possible values of the activity  $y_i$  at  $s_i$  into three regions: first, the region  $\Gamma_1$ , comprising those which are less than  $\vartheta_{0i}$ , second the values  $\vartheta_{0i} < y_i \leq \vartheta_{1i}$ , making up  $\Gamma_2$ , and third, the  $y_i > \vartheta_{1i}$ , which constitutes  $\Gamma_3$ . Suppose the value of  $y_i$  at a given time belongs to  $\Gamma_1$ , i.e.,  $y_i \leq \vartheta_{0i}$ . By lemma 1, the consequent value of the impulse returning to  $s_i$  at  $t + n$  will be  $A \vartheta_{0i} + \mu_i$ . If this is also less than  $\vartheta_{0i}$ , the same value will recur at  $t + 2n$ ; and we shall have a steady state of  $C$  determined by  $y_i = A \vartheta_{0i} + \mu_i$ ; if, however,  $A \vartheta_{0i} + \mu_i \geq \vartheta_{0i}$ , the circuit will, by lemma 1, go into complete activity, and there will be no steady state with  $y_i = A \vartheta_{0i} + \mu_i$ .

Suppose that a given  $y_i$  belongs to  $\Gamma_2$ , i.e., that  $\vartheta_{0i} \leq y_i < \vartheta_{1i}$ . Then  $C$  will be in complete activity, and, by lemma (2), the course of activity will be described by

$$y_i = \mu_i \frac{1 - A^t}{1 - A} + A^{t+1} \lambda_i.$$

Consider the value of this expression as  $t$  becomes very large. If  $|A| > 1$ , then, unless both  $\mu_i$  and  $\lambda_i$  vanish, in which case also  $C$  is not in complete activity,  $y_i$  will increase indefinitely, so that a steady state with complete activity is not possible. If  $A = \pm 1$ , we arrive at a special case to be treated later. If  $|A| < 1$ ,  $y_i$  will approach an asymptotic value

$$y_i = \frac{\mu_i}{1 - A},$$

and if this value for  $y_i$  puts  $C$  into complete activity, it will determine a possible steady state of complete activity: i.e., this will be the case

if and only if

$$\vartheta_{0i} \leq \frac{\mu_i}{1-A} < \vartheta_{1i}.$$

The possibility of a steady state for a  $y_i$  in  $\Gamma_3$  may be treated analogously to that of  $\Gamma_1$ : we shall have, as necessary and sufficient condition for the existence of such a value  $y_i = A \vartheta_{1i} + \mu_i$  that  $A \vartheta_{1i} + \mu_i \geq \vartheta_{1i}$ . Collecting these conditions, we shall have the

THEOREM.

*A circuit C may have at most three possible steady states:*

A). One given by  $y_i = A \vartheta_{0i} + \mu_i$

B). One given by  $y_i = \frac{\mu_i}{1-A}$

C). One given by  $y_i = A \vartheta_{1i} + \mu_i$ .

*Necessary and sufficient conditions for the actual existence of each of these are given respectively by the requirements:*

A).  $\mu_i \leq (1-A) \vartheta_{0i}$ .

B).  $\vartheta_{0i} < \frac{\mu_i}{1-A} \leq \vartheta_{1i}, \quad |A| < 1.$

C).  $\mu_i > (1-A) \vartheta_{1i}$ .

By way of corollary from these conditions—which differ from those of (Householder, 1941b and c) in giving explicit expressions for the equilibrium points—we may derive the general results of (Householder, 1941b and c) for the simple circuit in a very much easier way. Suppose first that  $0 < A < 1$ . Then (A), (B), and (C) become

A').  $\frac{\mu_i}{1-A} \leq \vartheta_{0i}.$

B').  $\vartheta_{0i} < \frac{\mu_i}{1-A} \leq \vartheta_{1i}, \quad |A| < 1.$

C').  $\frac{\mu_i}{1-A} < \vartheta_{1i}.$

Clearly one, and only one of these conditions can be fulfilled by a given circuit. We shall consequently have the

COROLLARY.

*When  $0 < A < 1$ , the circuit C has a unique steady-state activity*

pattern. This will also be the case when  $A < -1$ , except that here no steady state of complete activity is possible.

Suppose now that  $A > 1$ . Here again, condition (B) is excluded, and there is no steady state of complete activity; the other possibilities may be expressed as

$$A''). \quad \frac{\mu_i}{1-A} > \vartheta_{0i}$$

$$C''). \quad \frac{\mu_i}{1-A} \leq \vartheta_{1i},$$

which are not incompatible, so that we may well have two possible steady states, given by  $A \vartheta_{0i} + \mu_i$  and  $A \vartheta_{1i} + \mu_i$  respectively, both of which contain at least one zero.

It may be desired to trace the course of activity in  $C$  from the initial value  $\lambda_i$  to whatever steady state, if any, is finally reached. This may be done without difficulty by considering which region contains  $\lambda_i$ , the initial value of  $y_i$ . If this region contain a possible steady-state value of  $y_i$ , then, if it is  $\Gamma_1$  or  $\Gamma_3$  the steady state will obviously be attained immediately, while if it is  $\Gamma_2$ , it will be approached asymptotically, in accordance with (1). If the region containing  $\lambda_i$  do not contain a steady state then, if it be  $\Gamma_1$  or  $\Gamma_3$  the value of  $y_i$  will enter immediately and move through the adjacent regions until the first one containing a steady state is entered, whereupon that steady state is either attained immediately or asymptotically accordingly as this final region is  $\Gamma_1$  or  $\Gamma_3$ , or is  $\Gamma_2$ ; while if the initial region of  $y_i$  be  $\Gamma_2$  then, if only one of  $\Gamma_1$ ,  $\Gamma_2$  contain a possible steady state value, then that one will be entered, at a time determinable by solving (1) for  $t$  with  $y_i = \vartheta_{0i}$  or  $\vartheta_{1i}$  accordingly as the final region be  $\Gamma_1$  or  $\Gamma_3$ ; but if both  $\Gamma_1$  and  $\Gamma_3$  contain possible steady states,  $y_i$  will be described by (1) until that value  $y_i = \vartheta_{0i}$  or  $\vartheta_{1i}$  which occurs first befalls, whereupon it goes into the steady state of  $\Gamma_1$  or  $\Gamma_3$  respectively.

It will be worth while to conclude our discussion with a consideration of the interesting and important case where  $A = \pm 1$ . We notice first that if  $A = \pm 1$ , then the conditions (A) and (C) of the theorem become

$$A'''). \quad \mu_i \leq 0$$

$$C'''). \quad \mu_i \geq 0,$$

at least one of which must hold, so that there always exists a steady state for  $C$  with at least one zero in this case. The interesting case

arises when we have complete activity, however, at least initially, so that we shall suppose  $\vartheta_{0i} < \lambda_i \leq \vartheta_{1i}$ . Equation (1) becomes in this case

$$y_i = \mu_i \frac{1 - (\pm 1)^t}{1 - (\pm 1)} + (\pm 1)^{t+1} \lambda_i,$$

for the values of  $t$  specified in lemma 2, which is indeterminate; evaluating the quotient by limits we obtain

$$y_i = t \mu_i (\pm 1)^t + (\pm 1)^{t+1} \lambda_i. \quad (3)$$

If  $\mu_i \neq 0$ ,  $y_i$  becomes indefinitely large with  $t$ , so that a steady state of complete activity is not possible; and the case has no special interest. When  $\mu_i = 0$ , however, we have

$$y_i = (+1)^{t+1} \lambda_i. \quad (4)$$

In the case  $\lambda_i = +1$ , this means that the circuit  $C$  has an *infinitude of possible steady states*, one for each value of  $\lambda_i$  satisfying  $\vartheta_0 \leq \lambda_i < \vartheta_{1i}$ , so that the parameters of  $C$  do not determine the steady state conditions even among a finite set of values. Another such case arises when  $\lambda_i = -1$ ; here, however, we have no steady state at all, but instead continual oscillations in  $C$  of amplitude  $|\lambda_i|$ . It may be remarked that this possibility might have been determined by the methods of (Householder, 1941a, b, c) except for an oversight: the application of Cramér's rule on page 68 of (Householder, 1941a) is invalid when  $\Delta = 0$ , which is equivalent to supposing  $A = 1$  ( $\Delta = 1 - A$ ); on account of this the principal result of (3) is not in general valid. If the  $\leq$  sign be struck out of the statements of this result, however, and replaced by  $<$ , it becomes correct. A similar correction can generally be made at a few other places in (Householder, 1941a, b, c, 1942) where this oversight enters. The two possible cases in respect of the equations (7) of (Householder, 1941a), namely inconsistency and indeterminacy, correspond respectively to the case where  $\mu_i$  does and does not vanish for all  $i$ . In the first case, as we saw above, equation (2) gives an indefinitely great  $y_i$  for sufficiently late  $t$ , and no steady state of complete activity is indeed possible; in the second we have the determination of the substantial one of an infinite number of possible steady states by the initial values  $\lambda_i$ , as we should infer from the equations (7) of (Householder, 1941a) in this case.

We may perhaps conclude the present discussion with two remarks. First, with regard to the relationship of our present analysis to the purely static one of (Householder, 1941a, b, and c); we have been able to obtain all the results there found with a rather simpler procedure. We are, moreover, enabled to solve explicitly a problem



treated less directly in (Householder, 1941a, b, and c). A theorem there enables us to calculate all the stimulus patterns consistent with one or two specified activity patterns; since there are only a finite number of activity patterns, we can always enumerate them until we find all the pairs consistent with any given stimulus pattern, which is what we really want. This method may be very laborious, however: there are  $2^n$  distinct *AP*'s, and  $2^n(2^n - 1)$  pairs to be tested in general. Theorem 1 above on the contrary enables us to determine at once explicitly how many steady states there are in any given circuit, and what values of  $y_i$  determine them.

Secondly, the above results may be extended (1) to the case where the initial stimulation between  $t = 0$  and  $t = 1$  is not a constant  $\lambda_i$ , but a bounded function  $y_i(t)$ , by simply dividing the interval  $(0, 1)$  into sufficiently small segments so that all points of a given segment, when taken as the  $\lambda_i$  above, permit the circuit to approach the same steady state in the same way; the predictions thus obtained for each segment are to be applied only throughout those later time-intervals which are congruent to it modulo  $n$ ; (2), to cases where the initial variation in stimulation has lasted arbitrarily long before becoming steady, by waiting until the last synapse of  $C$  receives a steady stimulation, counting that interval as the first, and applying (1).

Desinently, I should like to express my appreciation of the counsel and criticisms of Dr. A. S. Householder, who suggested the problem of this and later papers on the theory of neuron networks.

This work was aided in part by a grant from the Dr. Wallace C. and Clara A. Abbott Memorial Fund of the University of Chicago.

#### LITERATURE

- Householder, A. S. 1941a. "A Theory of Steady-State Activity in Nerve-Fiber Networks: I. Definitions and Preliminary Lemmas." *Bull. Math. Biophysics*, 3, 63-69.
- Householder, A. S. 1941b. "A Theory of Steady-State Activity in Nerve-Fiber Networks: II. The Simple Circuit." *Bull. Math. Biophysics*, 3, 105-112.
- Householder, A. S. 1941c. "A Theory of Steady-State Activity in Nerve-Fiber Networks: III. The Simple Circuit in Complete Activity." *Bull. Math. Biophysics*, 3, 137-140.
- Householder, A. S. 1942. "A Theory of Steady-State Activity in Nerve-Fiber Networks: IV.  $N$  Circuits with a Common Synapse." *Bull. Math. Biophysics*, 4, 7-14.