Initiation of the Propagated Disturbance

By W. A. H. Rushton

Research Fellow of Emmanuel College, Cambridge

(*Communicated by E. D. Adrian, F.R.S.—Received* 7 *July* 1937—
Revised 22 *September* 1937)

PART I—GENERAL THEORY

It has long been supposed that the mechanism of propagation is the stimulation of the inactive region just in front, by the advancing action-potential wave. Plausible as is this theory, it is only within the last year that evidence has been published which puts the matter beyond doubt.

Blair and Erlanger (1936*b*) obtained suggestive evidence by blocking a single impulse with calcium or anodal polarization. They found that this blocked impulse left a transitory state of enhanced conductivity which allowed a second impulse to be propagated. The nature of this enhancement was studied independently and in greater detail by Hodgkin (1937). He used cold or pressure to block the first impulse, and found that the lowering of threshold on the farther side of the block had the same spatial and temporal distribution as had the spread of the action current. Moreover, the action current spread had the same space-time distribution as that of a subthreshold current applied from an external circuit, and so arranged as to have the same time course as the action-potential wave.

Thus, whatever other factors may conceivably be involved, the action current certainly spreads forward into the inactive region of a nerve in the same way as does a current from the outside, and this spread will account for the lowering of threshold observed. But since this lowering may be as much as 90 % at some distance in front of the active region, there can be little doubt that close to the active region itself, the lowering will be far more than 100 %.

We may thus conclude that the action-potential wave normally propagated is of an intensity many times that sufficient to excite the neighbouring portion of nerve; it is quite a different matter, however, to suppose that this statement applies to the action-potential wave, as it is first initiated at the cathode of a stimulating circuit. For the normal propagated wave

[210]

consists of an active region some centimetres in length, each element of which may be supposed to contribute (though not equally) to the excitation of the region in front of the wave. On the other hand, the excitation wave at the moment of its inception will occupy only a very small length, and be correspondingly feeble and might in fact be insufficient to propagate unless fostered by the continued application of the external stimulus.

Now the classical treatment of the excitation process does not take this into consideration at all. It has been customary to assume not only that when a certain membrane in the neighbourhood of the applied cathode is depolarized to a critical extent the action-potential wave arises, but that this wave from the start is of sufficient intensity to propagate. This last conclusion, however, has never been justified, it is inconsistent with the above simple concept of the flow of action current, and moreover, the expectations to which it leads differ in certain respects from those where the feebleness of the initial response is considered. It therefore appeared worth while to investigate theoretically this concept of the initiation of the propagated disturbance.

Now, though a number of conclusions follow readily from intuitive considerations, one cannot obtain even approximately quantitative results without a detailed mathematical investigation, and this has the disadvantage that unless the assumptions are greatly over-simplified, the resulting equations are (to me at least) insoluble.

It has therefore seemed best to attempt the investigation as follows. Assumptions are made, in many cases grossly over-simple, but such as will allow of a relatively easy mathematical treatment over a very wide range of possible stimuli. A graphical analysis is described which will permit those not familiar with the underlying equations to obtain solutions in a routine manner for any given temporal form of stimulus, so that it will be relatively easy to compare experimental results with the expectations of theory over a wide field, and then it will often be easy to see how such deviations as will be found may be due to the over-simplification of the assumptions. Finally, if there is a general confirmation of the present concept, it may be worth while to develop the theory upon more accurate assumptions, but with incomparably greater difficulty.

The present treatment will consist of three sections. In the first the assumptions will be stated and discussed, the second will contain the mathematical treatment, and the resulting conclusions will be summarized in the third. Those, therefore, who do not wish to examine the details of the mathematics may omit the middle section and still appreciate the assumptions and the conclusions.

In Part II, the general conclusions of the first part will be applied to a number of specific cases to show that the present theory fits in a fairly satisfactory manner a wide range of known observation, and to point out a number of conditions where certain results should be expected which differ from the classical expectations. Mention will also be made of the principal observations which will not fit the present theory.

Assumptions

The nerve is supposed to have a cable structure, the axis cylinder and the interstitial fluid being relatively well conducting, and the myelin sheath being highly resistant. These resistances are supposed to be non-capacitative. The nodes of Ranvier are neglected. Closely applied to the myelin sheath is a fine membrane which is supposed to be capacitative, but so leaky that its resistance is negligible compared with that of the myelin sheath which is in series with it.

Each element of inactive nerve sheath may be represented by the circuit (A), fig. 1, where R is the ohmic resistance of the myelin, c the capacity of the fine membrane applied closely to it, ρ the leak resistance of the membrane, where ρ/R is negligible. Finally E is the resting potential across the sheath which is observed as the injury potential.

It is clear that the resting charge on the condenser will be altered by passage of current through the myelin, and we assume that when this alteration attains a critical value, the local condition of the membrane suddenly changes. The intimate nature of this change is not considered, but the result is that the resting potential E is abolished, so that the sheath in the active region is represented by (B), fig. 1.

Since any given element of sheath is connected with all other elements through the conducting interstitial fluid and axis cylinder, it follows that the change of potential arising where the given element changes from inactive to active state, will affect in some degree all points on the nerve, though principally the regions adjacent. The nature of this spread is illustrated in fig. 2.

Suppose that the whole nerve is inactive except for a stretch 5 mm. in length whose centre point is O. Then the potential of each point on the surface of the nerve due simply to the activity of this stretch is given by the lowest member of the family of continuous curves in fig. 2. The other members of the family correspond, in order, to the potential distribution when the active lengths are 10, 20, 30, 40, 50 mm.

It is seen that each curve is concave towards the horizontal axis in the central region, and convex on either side. The transition is at an abscissa corresponding exactly to the edge of the active region. The dotted curve passes through these inflexion points, hence the horizontal distance between
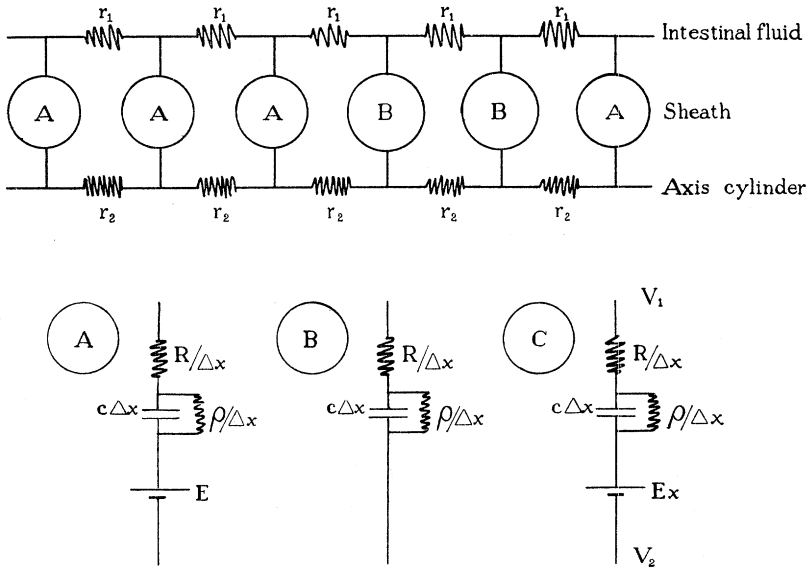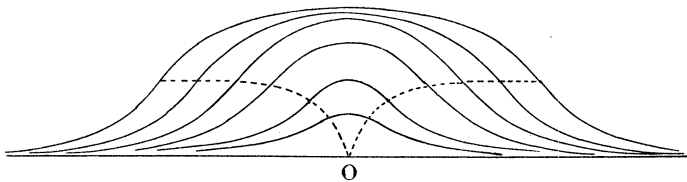

Fig. 1—Schematic circuit diagram of nerve.

Fig. 2—Abscissae: distances along nerve. Ordinates: potential of surface of nerve due to the activity of various lengths. O is the centre of the active stretch, and for each curve, the limits of the active region are the abscissae corresponding to the intersection of the curve with the dotted exponentials.

the two intersections for any member of the family gives the corresponding length of active nerve.

From fig. 2 it is easily appreciated that 5 mm. of active nerve only develops about one-quarter the action potential of a long active stretch, comparing the potentials at O, the mid-points of the two active regions;

and about one-half comparing the junctions of active and inactive regions in the two cases.

Shorter active lengths will clearly develop smaller potentials as given by the corresponding values of the dotted curve.

Mathematical Section

Let

V_1 = potential outside the sheath.

V_2 = potential inside.

i_1 = current flowing parallel to nerve outside.

i_2 = current flowing parallel inside.

$I = i_1 + i_2$.

r_1 = resistance per unit length of interstitial fluid.

r_2 = resistance per unit length of axis cylinder.

$1/R$ = conductivity per unit length of sheath for current passing radially.

$1/\rho$ = conductivity per unit length of membrane.

θ = charge per unit length on membrane.

θ_1 = critical value of θ at which nerve becomes active.

E = potential in series with sheath responsible for injury and action potential.

c = capacity of membrane per unit length.

$L^2 = R/(r_1 + r_2)$.

$\alpha = c\rho$ = finite time.

$\rho/R \to 0$.

Current Distribution in Nerve

The nerve is represented electrically by the scheme of fig. 1, where the elements are supposed to be infinitesimal and infinitely numerous. The element of sheath is represented either by (A) or (B), according as it is inactive or active at that point.

Currents from external circuits are led into the interstitial fluid according to any known time-space distribution, and it is required to find an expression for the charge on the membrane at any point and any time as a combined result of nerve activity and external stimulus.

To obviate the necessity of going over the working separately for the two types (A) and (B), let us assume that the sheath has the more general electrical structure (C), where E_x may be either E or zero as particular cases.

Then, applying Kirchoff's Laws to (C),

Since current through condenser element $= \dot{\theta}\Delta x$,

and current through shunt element $= \theta\Delta x/c\rho$,

$\therefore$ current through $R/\Delta x$ $= \left(\dot{\theta}+\dfrac{\theta}{c\rho}\right)\Delta x$,

$\therefore \qquad V_1 - V_2 = R\left(\dot{\theta}+\dfrac{\theta}{c\rho}\right) + \dfrac{\theta}{c} + E_x.$

But from the main network we have

$$-\frac{\partial V_1}{\partial x} = r_1 \dot{i}_1, \quad -\frac{\partial V_2}{\partial x} = r_2 \dot{i}_2, \quad \dot{i}_1 + \dot{i}_2 = I.$$

Also $\qquad \dfrac{\partial \dot{i}_2}{\partial x} = $ current density through sheath R

$$= \left(\dot{\theta}+\frac{\theta}{c\rho}\right).$$

$\therefore \quad \dfrac{\partial}{\partial x}\left[R\left(\dot{\theta}+\dfrac{\theta}{c\rho}\right) + \dfrac{\theta}{c} + E_x\right] = \dfrac{\partial}{\partial x}(V_1 - V_2)$

$$= r_2 \dot{i}_2 - r_1(I - \dot{i}_2).$$

$\therefore \quad \dfrac{\partial^2}{\partial x^2}\left[R\left(\dot{\theta}+\dfrac{\theta}{c\rho}\right) + \dfrac{\theta}{c} + E_x\right] = (r_1 + r_2)\left(\dot{\theta}+\dfrac{\theta}{c\rho}\right) - r_1\dfrac{\partial I}{\partial x}.$

Now putting $\qquad c\rho = \alpha, \quad R/(r_1 + r_2) = L^2,$

$$\frac{\partial^2}{\partial x^2}\left(\dot{\theta}+\frac{\theta}{\alpha}\right) - \frac{1}{L^2}\left(\dot{\theta}+\frac{\theta}{\alpha}\right) = -\frac{\theta}{\alpha}\cdot\frac{\rho}{R} - \frac{r_1}{R}\cdot\frac{\partial I}{\partial x} - \frac{\partial^2 E_x}{R\partial x^2},$$

and since $\rho/R \to 0$,

$$\left(\frac{\partial^2}{\partial x^2} - \frac{1}{L^2}\right)\left(\dot{\theta}+\frac{\theta}{\alpha}\right) = -\frac{r_1}{R}\cdot\frac{\partial I}{\partial x} - \frac{\partial^2 E_x}{R\partial x^2}.$$

To solve, we note that if

$$\left(\frac{L^2\partial^2}{\partial x^2} - 1\right)y = u,$$

$$2yL = e^{-\frac{x}{L}}\int_{-\infty}^{x} e^{\frac{x}{L}}u\,dx + e^{\frac{x}{L}}\int_{x}^{\infty} e^{\frac{x}{L}}u\,dx$$

is a solution, as may be verified by differentiation.

Thus putting
$$y = \left(\dot{\theta} + \frac{\theta}{\alpha}\right)\Big/ L^2,$$

$$u = -\frac{r_1}{R}\cdot\frac{\partial I}{\partial x} - \frac{\partial^2 E_x}{\partial x^2},$$

we obtain

$$-\frac{2R}{L}\left(\dot{\theta}+\frac{\theta}{\alpha}\right) = e^{-\frac{x}{L}}\int_{-\infty}^{x} e^{\frac{x}{L}}\left(r_1\frac{\partial I}{\partial x}+\frac{\partial^2 E_x}{\partial x^2}\right)dx + e^{\frac{x}{L}}\int_{x}^{\infty} e^{-\frac{x}{L}}\left(r_1\frac{\partial I}{\partial x}+\frac{\partial^2 E_x}{\partial x^2}\right)dx. \quad (1)$$

That this solution is the particular one to suit the boundary conditions · is seen by putting $x = \pm\infty$, when the second member of equation (1) vanishes, provided that $\partial I/\partial x$ and $\partial^2 E_x/\partial x^2$ are both zero for points far from the origin.

Since E_x is essentially a discontinuous function, being either E or zero, exception might be taken to the form $\partial^2 E_x/\partial x^2$ in (1), and the same in lesser degree applies to $\partial I/\partial x$. Since, however, equation (1) applies to any continuous function E_x, it will apply to $1 + \tanh qx$, which can be made to differ by less than any assignable quantity from the discontinuous function with which we have to deal, by making q large enough. Hence (1) applies to the discontinuous case with infinitesimal error.

Equation (1) may be put into a slightly more convenient form by integrating by parts the quantity $\partial^2 E_x/\partial x^2$, giving

$$-\frac{2R}{L}\left(\dot{\theta}+\frac{\theta}{\alpha}\right) = r_1\left\{e^{-\frac{x}{L}}\int_{-\infty}^{x}e^{\frac{x}{L}}\frac{\partial I}{\partial x}dx + e^{\frac{x}{L}}\int_{x}^{\infty}e^{-\frac{x}{L}}\frac{\partial I}{\partial x}dx\right\}$$

$$+\frac{1}{L}\left\{-e^{-\frac{x}{L}}\int_{-\infty}^{x}e^{\frac{x}{L}}\frac{\partial E_x}{\partial x}dx + e^{\frac{x}{L}}\int_{x}^{\infty}e^{-\frac{x}{L}}\frac{\partial E_x}{\partial x}dx\right\}, \quad (2)$$

where the coefficient of r_1 gives the contribution from the applied stimulus, and the coefficient of $1/L$ the contribution from the active portions of nerve.

It would be easy, integrating by parts again, to change the quantities under the integral sign from $\partial I/\partial x$, $\partial E_x/\partial x$ to I and E_x, but the above form is more convenient. For I and E_x in the following analysis will be constant over a certain range and then jump suddenly to a different constant value. This allows the integral in the form (2) to be written down at once. For, throughout the range in which I is constant, $\partial I/\partial x$ is zero and contributes nothing. If at x_n, I changes from I_n to I_{n+1}, then the value of

$$\int_{x_n-\Delta x}^{x_n+\Delta x} e^{-\frac{x}{L}}\frac{\partial I}{\partial x}dx = e^{-\frac{x_n}{L}}\int_{x_n-\Delta x}^{x_n+\Delta x} dI = e^{-\frac{x_n}{L}}(I_{n+1}-I_n).$$

The full value of the integral is thus simply

$$\sum^{n} e^{-\frac{x_n}{L}}(I_{n+1}-I_n). \quad (3)$$

BIPOLAR ELECTRODES

Suppose that the cathode is at $x = 0$, and the anode at $x = -s$. Then I vanishes except for values of x between the electrodes, since in the extrapolar regions all the current flowing away in the core must return in the interstitial fluid, and hence $i_1 + i_2 = I = 0$. Between the electrodes $I = I_1$, the current passed from the external circuit. These assumptions apply fairly accurately to the usual conditions of excitation. With regard to the potential E_x, we assume that until θ attains a value θ_1 it is E. But as soon as θ_1 has been attained E_x immediately assumes a fixed value zero which is maintained indefinitely.

At face value this is a poor approximation to the truth, since the action-potential wave appears to increase rather linearly than abruptly (Erlanger and Blair 1933), and certainly lasts 1 or 2 msecs. But it must be remembered that it has not been possible to record the onset of the action-potential wave in circumstances uncomplicated by propagation and capacitative spread, and all we know is that at the seat of generation of the action potential the time course must certainly be more rapid than any recorded wave.

With regard to the assumption of persistence of the action potential, this is chiefly erroneous for slowly rising stimuli where accommodation enters and complicates the matter in any case.

Suppose that at any moment t, the excited region extends from $-X_1$ to X_2, where X_2 is to the right of the cathode and $-X_1$ to the left. Consider the value of θ at a point x to the right of X_2.

Then in equation (2) the integrals from x to ∞ vanish, and applying formula (3) we immediately obtain

$$\frac{2R}{L}\left(\frac{\partial\theta}{\partial t}+\frac{\theta}{\alpha}\right) = e^{-\frac{x}{L}}(1 - e^{-\frac{s}{L}})I_1 r_1 + e^{-\frac{x}{L}}(e^{\frac{X_2}{L}} - e^{-\frac{X_1}{L}})\frac{E}{L}. \tag{4}$$

Now, if we make the substitution

$$\theta = \phi e^{-\frac{x}{L}}$$

in (4) and divide throughout by $e^{-\frac{x}{L}}$ we obtain

$$\frac{2R}{L}\left(\frac{\partial\phi}{\partial t}+\frac{\phi}{\alpha}\right) = (1 - e^{-\frac{s}{L}})I_1 r_1 + (e^{\frac{X_2}{L}} - e^{\frac{X_1}{L}})\frac{E}{L},$$

which shows that ϕ is independent of x.

But we know that at $x = X_2$, $\theta = \theta_1$, thus at any greater value of x

$$\theta = \theta_1 e^{-\frac{x-X_2}{L}}. \tag{5}$$

Thus in the extrapolar region the charge diminishes exponentially with the distance away from the nearest point of the active region with length constant L. We may therefore learn whether the active region is advancing or not by finding whether the charge at some fixed point in front is increasing. This is obtained by combining (4) and (5), with the elimination of θ,

$$\frac{2R}{E} \cdot \frac{\partial\theta}{\partial t} \cdot e^{\frac{x}{L}} = -\frac{2R\theta_1}{\alpha E} e^{\frac{X_2}{L}} + \frac{LI_1 r_1}{E}(1 - e^{-\frac{s}{L}}) + e^{\frac{X_2}{L}} - e^{-\frac{X_1}{L}}, \tag{6}$$

from which we see that the second member must be positive if propagation is to occur at the instant considered.

Liminal Length—From equation (6) it is clear that the stimulus I_1 to be effective must produce a change of such a magnitude that when the stimulus is withdrawn $\partial\theta/\partial t$ will still be positive.

$$\therefore \quad (e^{\frac{X_2}{L}} - e^{-\frac{X_1}{L}}) > \frac{2R\theta_1}{\alpha E} e^{\frac{X_2}{L}}.$$

$$\therefore \quad X_2 + X_1 > -L \log_e\left(1 - \frac{2R\theta_1}{\alpha E}\right). \tag{7}$$

Now the second member is independent of the nature of the stimulus. Hence a necessary condition for stimulation is that a liminal length of nerve $X_2 + X_1$, defined as above, be excited by the time that the stimulus ceases.

Put
$$1 - \frac{2R\theta_1}{\alpha E} = h = e^{-\frac{X_2 + X_1}{L}} \tag{8}$$

at the liminal length. The constant h, which has no dimensions, plays an important part in all the present treatment and may be called the "propagation constant" of the tissue.

This condition of a liminal length has just been shown to be a necessary one, and it is easily seen to be sufficient. For, once the propagation has begun after cessation of the stimulus, the only change in conditions is that a greater and greater length of nerve becomes active, and hence a greater and greater stimulus is applied. The condition of the liminal length assures an initial velocity of propagation, positive though infinitesimal, and as a greater length becomes active the velocity will increase to a fixed value.

Velocity of Propagation and Safety Factor

At a great distance from the origin, equation (6) becomes

$$\frac{2R}{E} \cdot \frac{\partial \theta}{\partial t} = -\frac{2R\theta_1}{\alpha E} e^{\frac{X_2 - x}{L}} + e^{\frac{X_2 - x}{L}},$$

or, combining with (5) and (8),

$$\frac{\partial \theta}{\partial t} = \frac{h\theta}{\alpha(1-h)}.$$

But since in this region the impulse is travelling with constant velocity v, we have

$$\frac{\partial \theta}{\partial t} = -v\frac{\partial \theta}{\partial x} = \frac{v\theta}{L}.$$

Hence
$$v = \frac{Lh}{\alpha(1-h)}. \tag{9}$$

Now the safety factor may be defined as one less than the ratio of the existing value of E to the value which E must assume to reduce the velocity just to zero. But from (8) it is seen that E is inversely as $(1-h)$. Hence the safety factor is one less than the reciprocal of the ratio of the existing value of $(1-h)$ to the value that this must assume to abolish conduction velocity. But from (9) the latter condition is satisfied by $h = 0$ or $1 - h = 1$.

Hence
$$\text{Safety factor} = \frac{h}{1-h}. \tag{10}$$

The Rheobase

The rheobase must have the intensity to raise to the value θ_1 the charge immediately under the cathode, and it must furthermore be sufficient to allow propagation from this point indefinitely. We proceed to show that if the first condition is satisfied, the second will also be satisfied.

Let I_0 be the value of I_1 which satisfies the first condition, then from (4) we see that I_0 must have such a value that $\partial\theta/\partial t$ vanishes when

$$\theta = \theta_1 e^{-\frac{x}{L}},$$

X_2 and X_1 being zero. Thus

$$\frac{I_0 r_1 L}{E}\left(1 - e^{-\frac{s}{L}}\right) = \frac{2R\theta_1}{\alpha E} = 1 - h. \tag{11}$$

Putting $I = I_0$, equation (6) becomes

$$\frac{2R}{E} \cdot \frac{\partial \theta}{\partial t} \cdot e^{\frac{x}{L}} = -(1-h)\left(e^{\frac{X_2}{L}} - 1\right) + e^{\frac{X_2}{L}} - e^{-\frac{X_1}{L}}$$

$$= h\left(e^{\frac{X_2}{L}} - 1\right) + \left(1 - e^{-\frac{X_1}{L}}\right).$$

And since $X_2 > 0 > -X_1$, $\partial \theta / \partial t$ is always positive and I_0, the least current which will activate the cathode, will propagate indefinitely, and is hence the rheobase.

EXCITATION WITH SHORT INTERPOLAR LENGTH

We are now in a position to consider approximately the conditions of efficacy of a stimulus of any known time course. But since this efficacy depends upon the way in which propagation is initiated from the cathode, it will be convenient to consider specifically two cases of bipolar electrodes, namely, where the two electrodes are very close together, or very far apart, since in these two cases the problem is fairly simple.

When the electrodes are very close together, the excitation spreads at first entirely into the extrapolar region. For clearly the stimulus has no exciting effect at the point midway between the electrodes (which is very close to the cathode in the case considered), hence this point will not be raised to activity until there is already more than the liminal length of nerve excited in the extrapolar region. Thus the applied stimulus, by the time that it ceases, must have raised to activity the region from the origin to a point X, where from equation (8)

$$e^{-\frac{X}{L}} = h.$$

Consider then the value of θ at this point X. The condition that a stimulus may be adequate is that at the moment of its cessation θ has attained a value of at least θ_1. A stimulus is threshold when any diminution in strength abolishes the foregoing condition.*

Considering this value of θ in equation (4) we obtain

$$\frac{2R}{hE}\left(\frac{\partial \theta}{\partial t} + \frac{\theta}{\alpha}\right) = \frac{I_1 r_1 s}{E} + e^{\frac{X_2}{L}} - 1.$$

* It is not accurate to simplify the above statement and say that a threshold stimulus is one where $\theta = \theta_1$ at the cessation of the stimulus. Take, for example, a constant current which lasts for a time somewhat longer than the utilization time of the rheobase. If this current is below rheobasic intensity θ will never attain θ_1 (nor in fact $h\theta_1$). If on the other hand the current is just adequate, θ will attain θ_1 by the utilization period, and be much greater by the time that the current ceases.

Now there are two conditions in which X_2 is constant, i.e. in which the propagation is not advancing. The first is during that period when the charge under the cathode has not yet attained the value θ_1, in which case $X_2 = 0$, and the equation reduces to that for the charging of a shunted condenser. The second case is when, although X_2 has some positive value, yet, in the event of the removal or the diminution of the applied stimulus, the value of $\partial\theta/\partial t$ ceases to be positive. Since the active stretch continues to be active in such a case, X_2 remains constant, and propagation may be resumed if the stimulus increases again.

In both cases, therefore, we have an ordinary first order equation, which may be simplified by substitution from (8), and by introducing the rheobase value from (11), with s very small,

$$I_0 = \frac{E}{r_1 s}(1-h).$$

We thus obtain
$$\alpha\frac{\partial\theta}{\partial t} + \theta = h\theta_1\left(\frac{I_1}{I_0} + \frac{e^{\frac{X_2}{L}} - 1}{1-h}\right). \tag{12}$$

Now, when X_2 is increasing, the value of θ at X_2 is θ_1, hence at X

$$\theta = \theta_1 e^{\frac{X_2 - X}{L}} = h\theta_1 e^{\frac{X_2}{L}},$$

from which we can substitute for X_2 in equation (12) and obtain

$$\alpha\frac{\partial\theta}{\partial t} + \theta = \frac{\theta_1 h}{1-h}\left[\frac{I_1}{I_0}(1-h) + \frac{\theta}{h\theta_1} - 1\right],$$

or
$$\frac{\alpha(1-h)}{h}\cdot\frac{\partial\theta}{\partial t} - \theta = \theta_1\left[\frac{I_1}{I_0}(1-h) - 1\right]. \tag{13}$$

In passing between the two ranges governed by equations (12) and (13) account must be taken of the fact that, so long as the stimulus remains of finite intensity, θ is continuous. Thus, initially θ is zero, it increases to the value $\theta_1 h$ where it leaves the first range of integration, and this value is consequently the initial value in the second range of integration.

At the moment when the impulse is initiated, θ changes from one range of integration to the other. The value of θ when this occurs may be called the "transitional value".

Again, from equation (12) it appears that, so long as I_1 suffers no discontinuity at the moment of changing range, the value of $\partial\theta/\partial t$ will also be continuous. For, at the instant of changing, the three variables θ, I_1 and X_2 are all continuous and hence the expression for $\partial\theta/\partial t$ is continuous.

Thus, given a stimulus of any time course, the effect produced is obtained first by the application of equation (12) with X_2 and the initial value of θ both put zero. When θ attains the transitional value $h\theta_1$ we pass to (13) maintaining continuity of θ. A further change back to (12) will have to be made if, through diminution of stimulus, θ ceases to increase. A simple graphical method of working out these results will be described later in this paper.

EXCITATION WITH GREAT INTERPOLAR LENGTH

In this case all the conditions of excitation are symmetrical on either side of the cathode, consequently propagation will spread equally in both directions, and we have

$$X_1 = X_2,$$

and hence the liminal length is excited when the impulse has travelled only half as far from the cathode as in the former case with short interpolar length. Thus if X is the distance from the cathode in the present case

$$e^{\frac{X}{L}} = \sqrt{h}.$$

Putting these values in equation (4), substituting from (8) and introducing from (11) the value of the rheobase I_0, with s infinity,

$$I_0 = \frac{E}{r_1 L}(1 - h),$$

we obtain an equation analogous to (12) for the value of θ at X.

$$\alpha \frac{\partial \theta}{\partial t} + \theta = \theta_1 \sqrt{h}\left[\frac{I_1}{I_0} + \frac{e^{\frac{X_2}{L}} - e^{-\frac{X_2}{L}}}{1 - h}\right]. \tag{14}$$

Now when X_2 is increasing

$$\theta = \theta_1 e^{\frac{X_2}{L}}\sqrt{h}.$$

Hence
$$\alpha \frac{\partial \theta}{\partial t} = \theta_1 \sqrt{h}\left[\frac{I_1}{I_0} + \frac{\sqrt{h}}{1 - h}\left(\frac{\theta}{\theta_1} - \frac{\theta_1}{\theta}\right)\right]. \tag{15}$$

Now θ/θ_1 in the range of equation (15) lies between $\sqrt{h}$ and unity, hence if h is not less than 0·75, we may with an error less than 2·5 % put

$$\frac{\theta_1\sqrt{h}}{\theta} = 2 - \frac{\theta}{\theta_1\sqrt{h}}.$$

Introducing this into equation (15) we obtain

$$\frac{\alpha(1-h)}{1+h} \cdot \frac{\partial \theta}{\partial t} - \theta = \frac{\theta_1 \sqrt{h}}{1+h} \left[\frac{I_1}{I_0}(1-h) - 2 \right]. \tag{16}$$

SOLUTION BY GRAPHICAL ANALYSIS

In a recent publication (Rushton 1937) I have described how by means of a simple appliance made from a piece of card, and by a routine manipulation, it is easy to solve the equation

$$a \frac{\partial y}{\partial t} + y = f(t),$$

where a is a constant, and $f(t)$ any given function of t.

Now all the equations (12), (13), 14), (16) are of this type and can therefore be solved in this way, which is indeed usually the easiest treatment. Moreover, since the matter is a perfectly routine one, it may be carried out by those unfamiliar with the mathematics underlying the equations to be solved, and they can in this way readily obtain an accurate graphical solution for any given form of stimulus.

We shall first consider the case where the electrodes are close together, and describe the procedure of the analysis without digressing to justify the statements made. Then the modifications required for the case of great interpolar distances will be given, and finally the validity of the process will be demonstrated.

Two analysers are required, one with subtangent α,. and the other $\alpha(1-h)/h$. We know the rheobase I_0 and the stimulus I_1 at any instant, hence we may plot I_1/I_0 on a convenient vertical scale against time measured horizontally, and so scaled that α occupies about 2·5 cm. An example is shown in fig. 3 where the stimulus is a current which increases exponentially from zero to a constant value, and is represented by the continuous curve I_1/I_0 passing through the origin.

Two horizontal lines are drawn, one dotted at ordinate value unity, and one continuous at $1/h$. The curve I_1/I_0 is now redrawn, "reflected" as it were in the dotted horizontal, but with a different ordinate scale. We shall refer to this curve as the "reflected curve". It may be obtained first by reflecting I_1/I_0 in the ordinary way, and then reducing the ordinate distances between the curve and the dotted horizontal in the ratio $(1-h)/h$, i.e. the ratio of the two subtangents to be used. The reflected curve thus represents the quantity

$$\left[1 - (1-h)\frac{I_1}{I_0} \right] \frac{1}{h}.$$

Now using the analyser with subtangent α and starting at the origin, a curve is obtained which represents the development of θ at a point distant the liminal length from the cathode. Ordinates represent a linear scale in θ; the horizontal dotted line is $h\theta_1$, the transitional value of θ, and consequently is the level at which we change from equation (12) to (13); the upper horizontal line is θ_1, and indicates the earliest point at which the current may be cut off and still allow propagation. When the curve of θ
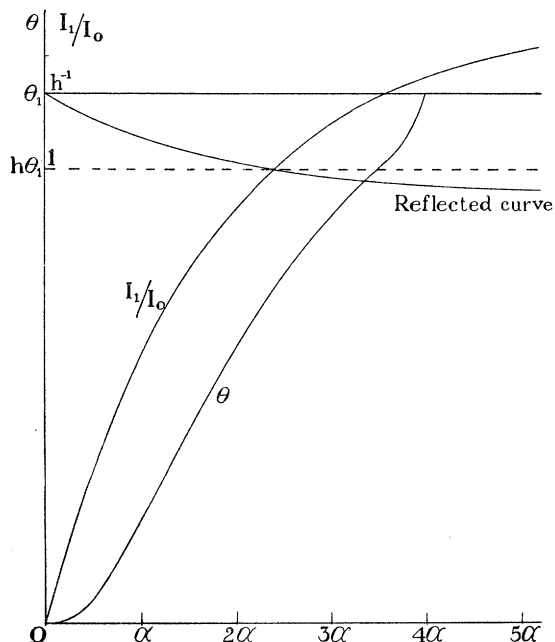

FIG. 3—Graphical analysis to obtain the curve of θ as function of time, given I_1, a stimulus of any known time course (see text).

reaches the dotted horizontal, the other analyser with subtangent $\alpha(1-h)/h$ is taken, and the curve continued without a break. This analyser must be used with the pivot to the left of the tracing pencil, since the subtangent in this case is negative; the curve to be analysed now is not I_1/I_0 but the reflected curve

$$\left[1 - (1-h)\frac{I_1}{I_0} \right]\frac{1}{h}.$$

The curve of θ is seen to suffer no discontinuity of gradient (as already noted), but (in the present instance) to experience a point of inflexion at the transitional value. Though the impulse is initiated when the curve

cuts the dotted horizontal, it is not until the upper horizontal is reached that the stimulus may be withdrawn without abolishing propagation.

If now the stimulating electrodes are far apart instead of being close together, some of the quantities in the foregoing analysis must be altered. The curve of I_1/I_0 is plotted as before, also the horizontal at ordinate unity, but this now corresponds to $\theta_1\sqrt{h}$, the transitional value for this case, hence the upper horizontal corresponding to θ_1 is drawn at ordinate $h^{-\frac{1}{2}}$. The first analyser is the same as before, but the second has a subtangent

$$\alpha\frac{1-h}{1+h}.$$

The reflected curve is drawn with distances from the dotted horizontal reduced in the ratio

$$\frac{1-h}{1+h},$$

which is the ratio of the subtangents in the present case, and therefore the curve represents the expression

$$\frac{2-(1-h)I_1/I_0}{1+h}.$$

As before, when the dotted horizontal is reached, the second analyser is used (pivot to the left), and the reflected curve is the one now treated. At the moment when θ meets the upper horizontal, the stimulus may be removed without abolishing conduction.

JUSTIFICATION

If in fig. 3 we adopt the scale which gives the value $\theta_1 h$ to the dotted horizontal, then the curve marked I_1/I_0 becomes the curve of $h\theta_1 I_1/I_0$, and this, analysed with subtangent α, gives a solution of θ which satisfies

$$\alpha\frac{\partial\theta}{\partial t}+\theta = h\theta_1\frac{I_1}{I_0},$$

which is equation (12) with X_2 zero.

The solution also satisfies the conditions that initially θ is zero and that it only applies to values below the transitional value $h\theta_1$. It is therefore the required solution.

In the same way the second range of the analysis satisfies equation (13), taking into account that the subtangent in this case is negative, being $-\alpha(1-h)/h$. Similarly the procedure for the case of great interpolar length satisfies the analogous equations (14) and (15).

TRIPOLAR ELECTRODES

If the current is led into the nerve by more than two electrodes, certain modifications must be made in some of the foregoing treatment, but in general the results are much the same, and do not merit a detailed consideration here. One case, however, is of particular interest, namely when the cathode of a tripolar system has two equal anodes symmetrically situated on either side of it, for unless s, the distance between two adjacent electrodes, exceeds a certain value, excitation is impossible however strong the stimulus.
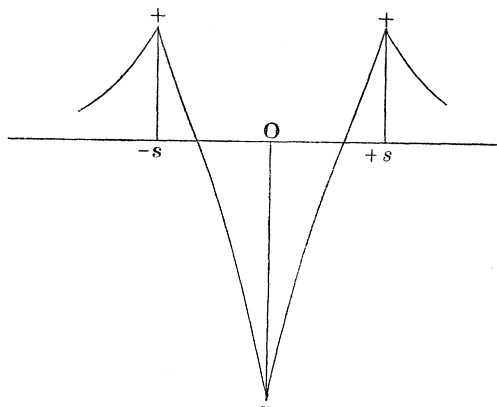

FIG. 4—Symmetrical tripolar electrodes. The density of current (ordinates) crossing the sheath is shown for each point on the nerve (abscissa).

Fig. 4 shows the charge produced at each point on the nerve between the points $-s$ and s by a current too weak to excite the cathode. It is easily seen from equations (2) and (3) that if at the cathode the charge is $2(1 - e^{-\frac{s}{L}})$ then at any point x lying between 0 and s the charge θ_0 is given by

$$\theta_0 = 2e^{-\frac{x}{L}} - e^{-\frac{s-x}{L}} - e^{-\frac{s+x}{L}}, \tag{17}$$

hence the charge is zero when

$$e^{\frac{2x}{L}} = 2e^{\frac{s}{L}} - 1.$$

Suppose that the distance between the two points of zero charge is the liminal length, then

$$e^{-\frac{2x}{L}} = h$$

or

$$e^{-\frac{s}{L}} = \frac{2h}{1+h}. \tag{18}$$

Now it is clear that if s has a value less than this, propagation from the cathode will be impossible, for we have seen that a region no longer excited by external stimulus cannot be rendered active by less than the liminal length, and the region beyond the points of zero charge is negatively excited. Thus equation (18) represents the minimum value which s can have if any stimulus is to be effective. If s exceeds this value, however, we shall show that a rheobasic current merely sufficient to excite the cathodal point will propagate.

For from equation (4) we may write down the condition that charge shall increase at any instant, taking into account, that the contribution to excitation of the active region is independent of the nature of the electrodes but that the contribution of the applied stimulus is not the coefficient of I_1 in (4) but some multiple of θ_0 in (17).

We thus obtain

$$\frac{2R}{E}\left(\frac{\partial\theta}{\partial t}+\frac{\theta}{\alpha}\right) = e^{-\frac{x}{L}}(e^{\frac{X_2}{L}}-e^{-\frac{X_2}{L}})+A\{2e^{-\frac{x}{L}}-e^{-\frac{s}{L}}(e^{-\frac{x}{L}}+e^{\frac{x}{L}})\}.$$

Introducing the value that the current is rheobasic when $\partial\theta/\partial t$ vanishes for $\theta=\theta_1$ and $x=X_2=0$ we obtain

$$\frac{2R}{E}\frac{\partial\theta}{\partial t}+\frac{\theta}{\theta_1}(1-h) = e^{-\frac{x}{L}}(e^{\frac{X_2}{L}}-e^{-\frac{X_2}{L}})+\frac{1-h}{2(1-e^{-\frac{s}{L}})}\{2e^{-\frac{x}{L}}-e^{-\frac{s}{L}}(e^{-\frac{x}{L}}+e^{\frac{x}{L}})\}.$$

Now introducing condition (18), and putting $x=X_2$, which also means $\theta=\theta_1$,

$$\frac{2R}{E}\frac{\partial\theta}{\partial t} = -1+h+1-e^{-\frac{2X_2}{L}}+(1+h)e^{-\frac{X_2}{L}}-h(e^{-\frac{X_2}{L}}+e^{\frac{X_2}{L}})$$

$$= e^{\frac{X_2}{L}}(1-e^{-\frac{X_2}{L}})(e^{-\frac{2X_2}{L}}-h),$$

and since $\qquad\qquad 1 > e^{-\frac{X_2}{L}} \quad\text{and}\quad e^{-\frac{2X_2}{L}} > h,$

$\partial\theta/\partial t$ is always positive, and propagation will occur as far as the liminal length, but no farther. If, however, at this point the stimulating current ceases, propagation will continue indefinitely.

CONCLUSIONS FROM THE MATHEMATICS

1—In the extrapolar region the charge at the limit of the excited active portion has the critical value θ_1; the charge at other points at this instant

diminish exponentially with their distance from this limit. The exponential constant is L, the constant of length for the cable (5).*

2—When an applied stimulus stops, the necessary and sufficient condition that propagation should continue is that a liminal length of nerve should have been excited. This length is given by expression (7).

3—A constant of no dimensions h is defined in (8) and called the "propagation constant". It appears in most of the mathematical expressions which have been derived.

4—It is clear that an adequate brief shock must excite the whole liminal length immediately. An enduring current on the other hand need only excite one point, whence the impulse, aided by the continued stimulus, will propagate indefinitely (11).

5—Expressions are derived for the final velocity of propagation (9) and for the safety factor (10).

6—Under the heading "Graphical Analysis" a method is described whereby the time course of excitation may be plotted in a routine fashion by those unfamiliar with differential equations, for the case of a stimulus of any given time course.

7—In the case of symmetrical tripolar electrodes, with the cathode in the middle, excitation is impossible unless the distance s between adjacent electrodes exceeds the value given by expression (18).

PART II—APPLICATION OF THE THEORY

1—TEMPORAL CONSIDERATIONS

a—Brief Shocks

One of the chief difficulties in understanding the excitation process is that the only thing which can be observed about excitation itself, is that a given stimulus is or is not effective. All that we believe about the process is inference from the properties of various stimuli which either are or are not effective, together with a little collateral evidence derived from such physical and chemical concomitants as we imagine to be relevant.

In Part I of this paper a method has been described which will enable us to obtain a curve representing the way in which the excitation process is supposed to develop under the action of any given stimulus. At first

* The numbers refer to the equations.

sight it would seem that this whole curve is of no significance except as a means of determining the fact that in one case propagation will occur, and in another it will not; the intermediate part of the curve would seem to rest as highly speculative and be outside the range of experimental investigation. It becomes, therefore, a matter of the highest importance to the practical application of the theory that these curves may in fact be experimentally obtained by superposing brief shocks upon the stimulus to be analysed.

One Shock.

The action of a single instantaneous shock (hereafter called merely a "shock") is obviously to charge the nerve instantly to a value proportional to the strength of shock and distributed in space as is the coefficient of r_1 in equation (2). If θ is not raised to the critical value of θ_1 over a region equal to at least the liminal length, propagation cannot occur. A threshold shock will raise instantly to activity exactly this length. If we apply to a single shock the general graphical method described in Part I, it is clear that the effect of a threshold shock is to raise θ instantly to θ_1. Any other strength of shock will similarly increase θ by a value proportional to this strength.

Now turning to the more general case, we have seen that, from the given time curve of the stimulus, we can derive the curve of θ. We shall know θ at any instant if we know what we must add to θ to bring the value to θ_1. But it is just this information which is given by superposing a shock, and if at any instant this is done, and simultaneously the stimulus is cut off, it is clear that the threshold value of this shock is that necessary to increase θ to θ_1. By superposing threshold shocks, and simultaneously cutting off the stimulus at various instants, therefore, we may obtain experimentally a curve for comparison with that predicted.

Two Shocks.

The simplest case to consider is the curve due to one shock investigated by adding a second after a varying interval. Fig. 5 gives the results to be expected from the theory in the case where the electrodes are far apart. The curves fall into three classes. In the first the shock raises θ to an initial value θ_0 less than $\theta_1\sqrt{h}$, and the curve, after rising instantly, returns to zero along an exponential of subtangent α. In the second class, θ_0 has a value between $\theta_1\sqrt{h}$ and θ_1, which signifies that a region of nerve from $-X_2$ to $+X_2$ is made active where

$$e^{-\frac{X_2}{L}} = \sqrt{h}\,\theta_1/\theta_0.$$

W. A. H. Rushton

The subsequent curve of θ is given by (14), where $I_1 = 0$, and

$$e^{-\frac{X_1}{L}} = e^{-\frac{X_2}{L}} = \sqrt{h}\,\theta_1/\theta_0.$$

$$\therefore \;\; \alpha\frac{\partial\theta}{\partial t} + \theta = \frac{\theta_1\sqrt{h}}{1-h}\left(\frac{\theta_0}{\theta_1\sqrt{h}} - \frac{\theta_1\sqrt{h}}{\theta_0}\right).$$
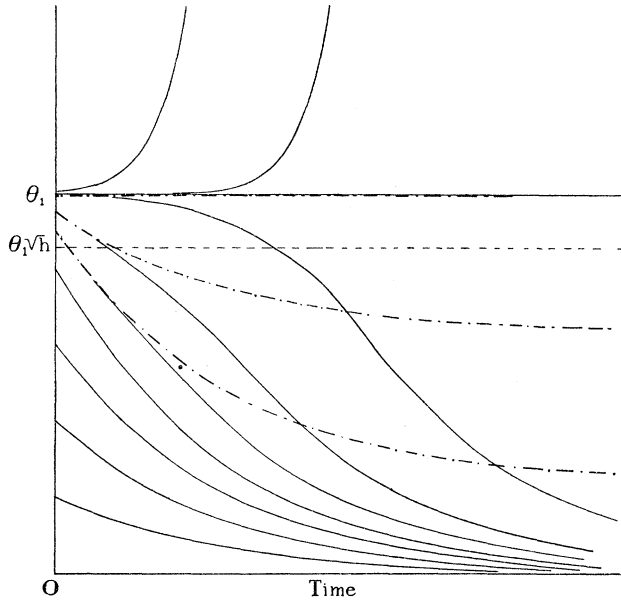

FIG. 5—The time course of θ following a single instantaneous shock applied at time O. The shape of the curve depends upon the strength of the shock (see text). For comparison with the experimental results of Katz (1937).

Since the second member of the equation does not vary with time, the curve of θ is an exponential of time constant α, as shown by the dotted curves in fig. 5. The initial value of each curve is θ_0, and the final value, obtained by putting $\partial\theta/\partial t = 0$, is

$$\theta_{t=\infty} = \frac{\theta_1}{1-h}\left(\frac{\theta_0}{\theta_1} - h\frac{\theta_1}{\theta_0}\right).$$

In particular it will be seen that when $\theta_0 = \theta_1$, both initial and final values are θ_1, and hence the curve coincides with the horizontal at that level.

Now the curves are calculated on the assumption that the action potential of the active stretch persists indefinitely, but we know that this is not the

case, and that in fact the curves of θ instead of continuing horizontally at various levels must soon fall below these levels and return to the axis. The course then, that we should actually expect to find, is one which would coincide with the dotted curves of fig. 5 in their earlier course, and then fall to the axis as shown by the continuous curves in fig. 5. The third class of curves is when θ_0 is greater than θ_1, in which case propagation will occur unless an *anodic* shock is applied to reduce θ below θ_1 again. The curve in this case is obtained from the graphical analysis of the second range, which solves equation (16). The curves are shown in fig. 5, rising steeply above the level θ_1. They are exponentials asymptotic to the level θ_1, with subtangent $\alpha \dfrac{1-h}{1+h}$.

The special significance of these curves is that qualitatively at least they resemble very closely those recently obtained by Katz (1937) in his experimental investigation of this relation. His results show curves of the three classes mentioned above, each exhibiting the appropriate characteristics.

It may be of interest to remark that the present theory and Katz's experiments were independent. Hodgkin and I had worked out these expectations, and sketched some experimental details of the modification considered below, when Katz came to discuss with us his first evidence of deviation from the simple exponential decay.

Three Shocks

If a shock is quickly followed by one that is equal but in the opposite direction, the subsequent excitability, as measured by a third shock in the same direction as the first, should show hardly any change from the resting condition, provided that the initial shock was less than $\sqrt{h}$ threshold intensity. If this value is exceeded, however, the resulting excitability curve should be that due to the local action potential alone. If the third shock is sent in the same direction as the second it will test the action potential (if any) which arises from the anode.

<div align="center">*b—Constant Currents*</div>

Fig. 6 shows the curves obtained by the graphical analysis of constant currents whose intensity in multiples of the rheobase is indicated by the number on each curve. It is seen that subrheobasic currents cause an exponential rise of θ to a final level proportional to the current but never rising higher than the transitional value $h\theta_1$ for short interpolar length, or $\sqrt{h}\,\theta_1$ for great interpolar length. Currents greater than the rheobase give rise to curves which, below the transitional value, are the same as the foregoing except that the ordinates are proportionally greater, but above this value the curve changes from concave to convex.

To establish the exact form of this upper portion consider the following operation. Suppose that we perform the graphical analysis of the first part of the curve again, but this time on a piece of tracing paper, and with every thing reduced in scale in the ratio $(1-h)/h$ (for the case of short interpolar lengths). The curve obtained will naturally be the same as before, but similarly reduced. The analyser used will have subtangent $\frac{\alpha}{h}(1-h)$ and hence will be the analyser used for the second portion of the curves in fig. 6. Now, suppose that we take the tracing paper and put it upside down
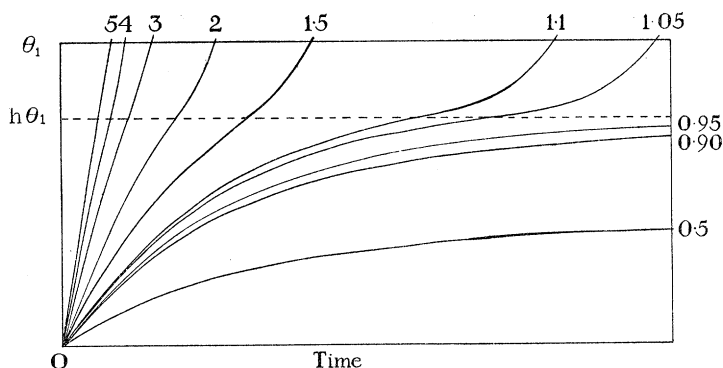

FIG. 6—The time course of θ following the making of a constant current at time O. For comparison with the experimental results of Rushton (1932) which are plotted the other way up.

on fig. 6, with the horizontal axis coincident with the upper horizontal of fig. 6, then it is easy to see that the two dotted horizontals must coincide, and the line I_1/I_0 on the tracing paper will coincide with the "reflected curve".

It follows that the second part of the curve in fig. 6 must be the same as the curve on the tracing paper, since they are both derived by analysing the same curve by the same analyser between the same limits. And thus the portion of the curve above the dotted horizontal is similar to the portion below, and similarly situated with respect to the point of intersection with this horizontal; they stand in the proportion of the two subtangents.

The same result holds for great interpolar lengths, allowing for the approximation made in obtaining equation (16).

It follows from the geometrical symmetry of the two parts of the curve,

that the time taken for θ to attain θ_1 is a constant multiple of the time taken to attain the transitional value, this constant being

$$1/h \quad \text{or} \quad 2/(1+h) \quad \text{for short or great lengths.}$$

Consequently the strength duration curve has the form

$$\frac{I_0}{I} = 1 - e^{-\frac{ht}{\alpha}} \quad \text{or} \quad \frac{I_0}{I} = 1 - e^{-\frac{1+h}{2\alpha}t}, \tag{19}$$

where I is the current and I_0 the rheobase.

This is the condenser formula of excitation which has been proposed countless times. Evidence is rather conflicting as to its accuracy, but on the whole it is probably a good approximation to the experimental data. Where the present theory differs from the classical condenser concept is that the time constant of the physical process α is significantly smaller than that of the excitation process.

A more striking difference is seen in the curves of the way that θ develops under the action of a constant current. According to the condenser theory all curves should run as do those in fig. 6, below the transitional value, and excitation should occur when this value is attained. Exactly this has been described by Bishop (1928), who was the first to investigate the matter by the method of superposed shocks. Unfortunately he restricted himself to subrheobasic currents, which do not exhibit the characteristic inflexion, and differ from the condenser theory expectations only in that a just subrheobasic current requires a moderate shock to be superposed, and not an infinitesimal one. This Bishop found, but explained it away by an alleged property of his circuit which appears to be inconsistent with Kirchoff's laws. Erlanger and Blair (1931b) repeated the observation and disproved, for their circuit, Bishop's explanation. They were chiefly interested in questions of accommodation to a polarizing current and also confined themselves to subrheobasic currents. Their curves which plot, for the most part, action-potential heights, are not strictly comparable with the curves of the present analysis, but in a general way they agree over short durations where accommodation is not significant.

At the same time, but unaware at first of their investigations, I made a number of similar observations, much cruder in method, but extending into the range of super-rheobasic currents. The curves obtained showed the inflexion as in fig. 6, and were obviously inexplicable on the simple classical theories. I was unwilling to publish these without being able further to explain their significance, and merely mentioned the matter

in a short note (Rushton 1932) wherein the curves are seen to be extremely similar to those of fig. 6 plotted the reverse way up. More detailed discussion of these and related results will be presented in a separate paper.

c—Condenser Discharges

The condition that a condenser discharge should be just threshold is easiest established from the geometry of fig. 7, where A and B are two exponential curves asymptotic to RS whose ordinates RP and RQ are in a constant ratio. OPS is the tangent at P, hence from the property of the exponential RS the subtangent is constant. Thus from similar triangles OQ is also constant and independent of the position of P.
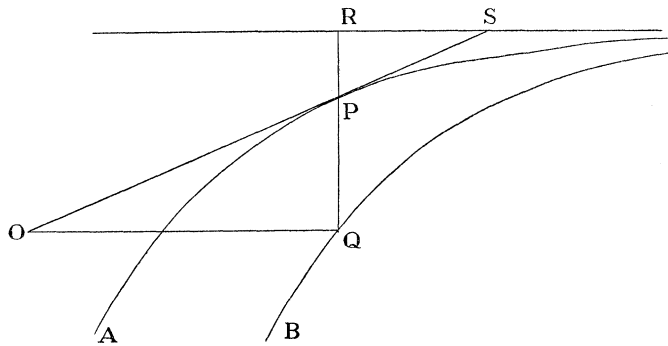

FIG. 7

Now the "reflected curve" in the graphical analysis of a condenser discharge is an exponential of the form B, and hence it is clear from the construction, that if the curve of θ in the second range of analysis starts by lying on A, so situated with regard to B that $OQ =$ subtangent of the second analyser, it will continue to trace out A indefinitely. Obviously this is the threshold condition since the slightest deviation on either side of B will cause the curve of θ quickly to diverge much farther.

We thus conclude that the utilization period of a threshold condenser discharge is the full period of the discharge. This is not experimentally the case. There is uncertainty as to the exact experimental relation but agreement that an appreciable fraction of a slow discharge may be cut off without affecting the efficacy. Three factors at least may enter into this divergence from the theory. Conditions are nearly always slightly super-threshold and in this case the utilization period is very considerably diminished. Measurements are not usually made with electrodes very

close together, and so the theoretical results are invalidated by the approximation introduced into equation (16). Accommodation will diminish the utilization period. According to the recent results of Fabre (1933) a condenser continues to act until its voltage has fallen well below rheobasic voltage, but not to zero, it therefore fits the present theory as modified by accommodation, etc. It will not, however, fit the classical condenser theory, according to which a subrheobasic intensity can never be present at the instant preceding excitation.

The voltage-capacity relation for threshold condenser discharges has been investigated by countless workers; the recent experiments of Hill (1936 b) are probably the most accurate. He found that the relation fitted precisely the theoretical curve obtained on the classical condenser theory, hence it is important to find what curve is to be expected upon the present theory. Unfortunately the working, though straightforward, is rather long, and I shall therefore merely mention the method and quote the result.

Taking the case of large interpolar distance, as in Hill's experiments, we introduce into (14) V_0 for I_0 and $Ve^{-t/\beta}$ for I_1, where β is the time constant of the condenser discharge. Putting X_2 zero we obtain the value t_0 when $\theta = \theta_1 \sqrt{h}$. Now introducing the threshold conditions of fig. 7, we may eliminate t_0 and obtain

$$\log\left(\frac{V}{V_0}\right) = \frac{1}{1 - \frac{\beta}{\alpha}} \left[\log\left(\frac{2\frac{\alpha}{\beta}}{1+h}\right) - \frac{\beta}{\alpha} \log\left(1 + \frac{\alpha}{\beta} \cdot \frac{1-h}{1+h}\right) \right]. \qquad (20)$$

When $h = 1$ this reduces to Hill's formula.

Table I shows the result of computing the value of $\log_{10}\left(\frac{V}{V_0}\right)$ for various values of $\log_{10}\left(\frac{\beta}{\alpha}\right)$ where $h = 0 \cdot 75$. The third line gives the value of $\log_{10}\left(\frac{V}{V_0}\right)$ according to Hill's formula, where V_0 is the same in the two cases, but in the second α has been replaced by $2\alpha/(1+h)$. It is seen that the two curves never differ by more than $0 \cdot 026$ and thus the results of Hill's experiments fit the present theory as accurately as they do the simpler formula which

TABLE I

$\mathrm{Log}_{10}(\beta/\alpha)$	2	1	0	-1	-2	-3
$\mathrm{Log}_{10}(V/V_0)$ (present theory)	$0 \cdot 02$	$0 \cdot 11$	$0 \cdot 44$	$1 \cdot 134$	$2 \cdot 07$	$3 \cdot 06$
$\mathrm{Log}_{10}(V/V_0)$ (Hill's theory)	$0 \cdot 02$	$0 \cdot 12$	$0 \cdot 46$	$1 \cdot 160$	$2 \cdot 07$	$3 \cdot 06$

he applied. Fabre's results, on the other hand, certainly favour the present theory.

The time constant $2\alpha/(1+h)$ of the voltage-capacity curve (Hill's k) is interpreted differently on the two theories. According to the condenser theory this is the constant of the charging process in the nerve, but on the present theory α is the charging constant. The same difference was found above (19) in the case of constant current stimulation. According to the present theory, therefore, a nerve should have the same form of curve both for physical charging and for excitation, but the latter should have a somewhat longer time constant. Exactly this result has been found by Harris and Rosenberg (1935) if the development of a given electrotonic potential be regarded as a measure of the physical charging. Since, however, the assumptions of the present theory require that the charging of the membrane should affect infinitesimally the electrotonic potential, it is perhaps unjustifiable to claim support from Harris and Rosenberg's evidence. However, Katz (1937) has found that the charging rate, as measured by the exponential decay in the excitation remainder following a single weak shock, corresponds to the rate found by Harris and Rosenberg and is considerably smaller than the constant derived from voltage-capacity excitation curves.

d—Evaluation of Constants

We are now in a position to assign numerical values to the constants of the foregoing equations. From equation (9) we have

$$v = \frac{hL}{\alpha(1-h)},$$

where v = velocity of conduction = 30 mm./msec. at 20° C.

L = length constant for nerve in air = 3 mm. (Rushton 1934).

$\dfrac{2\alpha}{1+h}$ = Hill's k = 0·34 msec. at 20° C. (Hill 1936b).

Substituting the above values we obtain

$$h = 0\cdot75, \quad \alpha = 0\cdot3 \text{ msec.}$$

This is the value which has been assumed for h above in making approximations.

The Safety Factor (10) = $\dfrac{h}{1-h}$ = 3.

The Liminal Length (8) = $-L\log_e h$ = 0·86 mm.

As it happens this is about the distance between one Node of Ranvier and the next.

2—Spatial Considerations

If the current is led to and from a nerve by conductors which are not small compared to the nerve trunk, we may assume that the current is distributed around each fibre with radial symmetry. The current then passing across the sheath is given by equation (2) where E is put zero. This formula is essentially the same as the one I put forward (1928) for nerve in fluid, and showed to satisfy certain variations in the spatial distribution of stimulus. The constant L, deduced in the first place from excitability measurement, should have the physical significance $\sqrt{\{R/(r_1 + r_2)\}}$. This comparison has been made (Rushton 1934), and, though the accuracy of the experiments was not very great, there was no detectable difference between L as found by resistance and by excitability measurement. The present theory, therefore, appears to account for the general spatial as well as for the general temporal aspects of excitation. Three qualifications, however, must be made.

First, the value of L is assumed to be the ratio of purely ohmic resistances. But we know that nerve has capacity, and R is presumably less for brief than for longer shocks. I have in fact found that L is less with briefer stimuli, and several workers (Cardot and Laugier 1914; Bouckaert and Katz 1935, discussed by Hill (1936b)) have found that the time constant of the strength-duration curve is less with short than with long interpolar lengths, which is another way of expressing the same thing. As we have seen (19), even excluding capacity, some difference in time constant is to be expected with great or small interpolar lengths, from the different equations applicable in these two cases. But the effect thus deduced is small, and the sense opposite to that observed.

The second qualification is due to the connective tissue sheath which deflects the current in a manner distinctly appreciable at short interelectrode distances.

The third refers to tripolar electrodes. As has been shown in (18), in the symmetrical case, if the distance between adjacent electrodes is less than

$$L \log_e\left(\frac{1+h}{2h}\right) = 0\cdot 5 \text{ mm. (approx.)},$$

propagation cannot take place. This result may be related to the well-known observation of Pratt (1930) and Gelfan (1930) that a micro-electrode intimately applied to a single muscle fibre gives rise to a localized contraction which will not propagate. Conditions increasing the length excited

allow of propagation. Now the current leaving the fibre at the micro-electrode must enter it elsewhere and in fact it enters through the regions just either side of the electrode, thus forming a system not unlike the tripolar arrangement from which (18) was deduced.

Two other characteristics of the localized contraction in the basihyoid membrane, mentioned by Kato (1934), support the present interpretation. First, whenever the localized contraction spread more than 5 mm., it propagated throughout the whole fibre. Second, the localized effect was less prominent in fresh muscles than in those kept some time in Ringer's fluid, or especially those narcotized. Thus, conditions lowering the Safety Factor, as we shall deduce below in (21), increase the liminal length, and the phenomena localized within it.

According to the assumption of an action potential persisting indefinitely, it is obviously impossible for the impulse to be propagated in one direction only from the point of stimulation, for, as the active length becomes great enough, the region on the other side of the electrode will propagate. This does not follow, however, if we take into account the brief duration of the action potential. If the cathode is far from another electrode, conditions will be symmetrical on either side, but if the anode is near, the initial propagation (as we have seen) will be entirely into the extrapolar region. At the moment when the stimulus is removed, propagation will continue into the extrapolar region which is still cathodally charged, but hardly in the other direction where the charge is still anodal. Thus, in a short while, when the first stretch has become inactive, the new active stretch in the extrapolar region will be considerably longer than the new stretch in the interpolar region. And if the former was barely sufficient to continue propagation (which is the threshold condition) the latter will be insufficient, and propagation will occur in the extrapolar direction only. This will only happen with short interpolar distance and threshold strength (unless we consider very strong currents with the classical anodal block).

3—ELECTRIC RESPONSE

Any satisfactory theory of nerve activity should describe and if possible explain the form of the action-potential wave. In the present exposition this whole matter has been sacrificed for the sake of a relatively simple treatment. If the theory turns out to be of value, this is the direction in which it will most need development.

One further matter in connexion with the electric response, however, needs consideration. Should an active stretch just less than the liminal

length give rise to an appreciable action potential? Many workers have looked for and failed to find an action potential resulting from a just subthreshold stimulus. Blair and Erlanger (1936*a*), considering just such a possibility as is worked out in this paper, conclude that the action potential associated with an active length less than liminal cannot exceed 2 % of the full spike.

To find the value to be expected on the present theory we note that, in the absence of external stimulus, the stimulation at X_2 due to the active region from $-X_1$ to X_2 is

$$K_1(1 - e^{-\frac{X_2+X_1}{L}}),$$

where K_1 is the stimulation when $X_2 + X_1$ is great, i.e. the stimulation due to a full-grown impulse. But if F is the Safety Factor, then by definition $K_1/(F+1)$ is a stimulation which will barely propagate, and is thus equal to the stimulation produced by the liminal length X.

Thus
$$K_1(1 - e^{-\frac{X}{L}}) = \frac{K_1}{F+1},$$

$$e^{\frac{X}{L}} = 1 + \frac{1}{F}. \tag{21}$$

But the action potential from the mid-point of an active region of length X is

$$K_2(1 - e^{-\frac{X}{2L}}),$$

where K_2 is the potential from an active region extending far in each direction. Hence, when X is the liminal length, the ratio of the "liminal action potential" to the maximum action potential is

$$1 - e^{-\frac{X}{2L}} = 1 - \left(1 + \frac{1}{F}\right)^{-\frac{1}{2}}. \tag{22}$$

Now in the section "Evaluation of Constants" the safety factor came to 3, which when substituted in (22) gives the liminal action potential as 13·5 % maximal—a value far too great to have escaped observation by Blair and Erlanger. On the other hand, the safety factor is certainly greater than 3. Conduction early in the refractory period and in anodal block suggest that 10 is a better approximation (Hodgkin 1937) and substituting this in (22) gives 4·7 % maximal for the liminal action potential. Moreover, it can easily be shown that if this value is approximately true and the stimulus is n % below threshold, then the resulting action

potential expressed in terms of maximal will also be n % less than in the threshold case. Thus a 97 % threshold shock in the case above would cause an action potential only 1·7 % maximal, and be within Blair and Erlanger's limits of error.

Since the excitation of the liminal length appears to give rise to an action potential too small to be recorded, and since the velocity of propagation in these circumstances is practically zero, a finite latent time will elapse before the action potential reaches a visible size. This latency has been found by Blair and Erlanger (1936a), but the capacity of the tissue must also play a large part, as they suggest.

The present theory will never stand upon a firm basis until a sub-threshold shock has been shown to leave a localized action potential, and to be followed by a refractory period. The treatment in this paper is too simplified for quantitative estimates to carry much weight, but it does appear possible that the liminal action potential may lie below the limits so far recorded, but that a small extension of those limits will reveal it.

4—ACCOMMODATION

Hill (1936a) and Solandt (1936) have shown that the relative inefficacy of currents of long duration is accurately accounted for by assuming that the threshold for excitation is not constant, but rises according to a first order differential equation. It is possible to interpret this by supposing that there is some polarizable structure in series with the excitable element, and that the observed inefficacy of slow currents is due to their partial neutralization by the polarization currents.

This concept may be applied equally to the present theory. I have recently shown (Rushton 1937) how to use a graphical analysis to solve Hill's equation for the effect of accommodation, and this same method may be applied in the two ranges of graphical analysis in the present paper, assuming that a nerve accommodates to its own action potential in the same way as to an applied stimulus. It is obvious that in this way accommodation may be brought within the scope of the present theory. This has not been included in the Mathematical Section, however, partly for simplicity, but partly because there is an altogether different phenomenon which must also be considered.

If the action potential does not continue indefinitely, but only for some 2 msec., an adequate stimulus must not only be able to raise to activity the liminal length, but it must do this rapidly, otherwise the region first excited will already have become refractory before propagation has reached

the liminal length. Now the classical idea of accommodation was that unless the critical polarization was accomplished within a brief time, the strength of stimulus had to be greatly increased, so we may ask whether accommodation can be explained entirely on the basis of the refractory period of the first excited portion of tissue.

That refractoriness and accommodation are connected is strongly suggested by the comparisons of Erlanger and Blair (1931 *a*) between the relative refractory period and the post-cathodal depression following a single subthreshold shock. But refractoriness can hardly explain the anodal excitation at the opening of a constant current, and refractoriness is, I suppose, still present in a citrated nerve where accommodation is practically absent.

It therefore appears that two distinct phenomena may enter into accommodation, first, the refractoriness of the portion of nerve first excited, and second, the polarization of some structure by both the stimulus and the subthreshold action currents. The relation between relative refractoriness and post-cathodal depression may thus have a twofold basis, for the latter may be partly caused by the *localized* refractory period, while the *propagated* refractory period may be due in part to polarization from the action current. We should thus expect the post-anodal enhancement following a single just subthreshold shock to differ in shape from the post-cathodal depression, assuming the former to be uncomplicated by a localized activity followed by a refractory period. The effect at the cathode should therefore show a greater and more prolonged enhanced period followed by a more marked depressed period. This difference is seen in fig. 7 of Erlanger and Blair's paper.

Discussion

The present theory has, for simplicity, made two very serious errors in assumption. First, the nerve is taken as non-capacitative, and thus all the temporal aspects of electrotonus cannot be considered. Second, the action potential is assumed to rise instantly to its full value and persist indefinitely, consequently considerations relating to the form of the observed wave cannot apply. In addition these two erroneous assumptions affect to some extent the quantitative validity of all the conclusions of this paper, as has been pointed out in many places above. The results must also be modified by considerations of accommodation and localized refractory period (as just discussed).

Finally, it must be stated that there are other and quite different sets of assumptions which will give, in a general way, most of the results of this

paper. The value of the present assumptions is that they are necessary deductions from the generally accepted theories of excitation and propagation.

In spite, however, of its errors and limitations, the theory has something to its credit. It yields, in a fairly manageable form, the results of stimulation with currents of any temporal form whatever, applied in any spatial distribution whatever. It describes in a quantitative and verifiable way the process both of excitation and of conduction. If we are given the velocity of conduction and length constant L from physical measurements, and also in given circumstances the rheobase I_0 and time constant (Hill's k), we have all the constants required for the entire working of the theory. Of these four constants only I_0 and k are arbitrary, the other two are physical measurements, thus, in spite of its scope the theory has no more arbitrariness than was formerly required to describe the strength-duration curve alone. One may even hope that, with the introduction of assumptions as to capacity into the present theory, the constant k may be deducible from physical measurements upon charging rate.

With regard to excitability relations, the theory will describe the strength-duration and voltage-capacity curves, with an accuracy as great as that of the experiments, and will account for the results of superposing shocks in a way that invites attention.

The chief claim of any theory, however, lies in its suggestions for further investigation. To this end the present ideas have been put forward with as definite a physical significance as possible. In this way the concepts become sharper, more provocative to those inclined to attempt disproof, more clearly verifiable for those attracted. The theory offers suggestions for experimental correlation over a wide field, and if this is carried out, our knowledge should be appreciably extended.

The ideas developed in this paper have arisen largely out of informal discussions with Mr. A. L. Hodgkin, for whose criticisms and suggestions it is a pleasure to express my indebtedness.

SUMMARY

1—We assume that a certain depolarization of a nerve at a point is the necessary and sufficient condition that an element of action potential be generated at that point, and that propagation results from the stimulation of the inactive region in front of the propagated wave by the spread of the action current. It follows that excitation is inadequate unless it

activates a sufficient length of nerve to give an action current big enough to propagate.

2—In Part I of this paper the matter is worked out on an over-simplified scheme. The "Mathematical Section" is separated from the "Assumptions" and the "Conclusions from the Mathematics", so that those not interested in the technical treatment may yet see in summary form both what is assumed (p. 212) and what is concluded (p. 227).

3—In Part II the theory is applied to a selection of the known observations upon excitability. The strength-duration curve, the voltage-capacity curve, and the effect of two successive shocks, are satisfactorily described, including certain aspects which will not fit the classical theories. The spatial relations of excitation are also satisfactorily described, including an explanation of the non-propagated contraction arising from micro-stimulation of single muscle fibres.

4—The theory, though too simplified in the present form, suggests a large number of experiments, in the light of which further development should be possible.

REFERENCES

Bishop, G. H. 1928 *Amer. J. Physiol.* **84**, 417.
Blair, E. A. and Erlanger, J. 1936*a* *Amer. J. Physiol.* **114**, 309.
— — 1936*b* *Amer. J. Physiol.* **117**, 355.
Bouckaert, J. P. and Katz, B. 1935 *J. Physiol.* **85**, 10*P*.
Cardot, H. and Laugier, H. 1914 *C.R. Soc. Biol., Paris*, **77**, 276.
Erlanger, J. and Blair, E. A. 1931*a* *Amer. J. Physiol.* **99**, 108.
— — 1931*b* *Amer. J. Physiol.* **99**, 129.
— — 1933 *Amer. J. Physiol.* **106**, 565.
Fabre, P. 1933 *C.R. Soc. Biol., Paris*, **112**, 1425.
Gelfan, S. 1930 *Amer. J. Physiol.* **93**, 1.
Harris, D. T. and Rosenberg, H. 1935 *J. Physiol.* **84**, 11*P*.
Hill, A. V. 1936*a* *Proc. Roy. Soc.* B, **119**, 305.
— 1936*b* *Proc. Roy. Soc.* B, **119**, 440.
Hodgkin, A. L. 1937 *J. Physiol.* (in the Press).
Kato, G. 1934 "Microphysiology of Nerve." Tokyo.
Katz, B. 1937 *J. Physiol.* **90**, 6*P*.
Pratt, F. H. 1930 *Amer. J. Physiol.* **93**, 680.
Rushton, W. A. H. 1928 *J. Physiol.* **65**, 173.
— 1932 *J. Physiol.* **75**, 16*P*.
— 1934 *J. Physiol.* **82**, 332.
— 1937 *Proc. Roy. Soc.* B, **123**, 382.
Solandt, D. Y. 1936 *Proc. Roy. Soc.* B, **119**, 355.