Science Bulletin of NCTU, Vol. V, No. 2

SEMICONDUCTOR THEORY OF NERVE CURRENTS

LING Y. WEI

Biophysical Research Laboratory, Electrical Engineering Department
University of Waterloo, Waterloo, Ontario, Canada

(Received 12 July, 1971)

Abstract—*The semiconductor theory is promising to unite and correlate the essential features of electrodiffusion theory, ion exchange theory and dipole theory in dealing with nerve currents.*

I. INTRODUCTION

Ion current flow in nerve is one of the major problems in physiology. This electrical event in a biological material has presented stubborn resistance to attack on the theoretical front. As is well known, the Hodgkin-Huxley (H-H) theory is a phenomenological theory which gives a superb mathematical discription of the nerve currents in terms of g_{Na}, g_K, h, m and n[1]. However, the mechanisms underlying these FAKTOREN and their behaviour remain unknown. The electrodiffusion (ED) theory[2] is centered around the Nernst-Planck equation

$$J = qp\mu E - qD \text{ grad } p$$

To solve this equation, the electric field must be known a priori. In practical case, this is hard to obtain and one has to resort to sheer guess. The simplest (and hence naive) assumption is a constant field in the nerve membrane. By Poisson Equation, this is equivalent to assuming a "zero" charge density. That is probably near the case in the resting state of the membrane. However, under excitation the conductance of nerve membrane is increased by at least a hundred times and so is the charge density in it. The constant field assumption is then rather unjustified. This assumption can be dispensed with by some effort and is not a defect of the electrodiffusion theory. The inherent limitation of the theory in its present form is that it makes no room for chemical effects on ion flow, and that it gives no indication of other possible causes than current flow for the change of ion concentrations. In other words, the electrodiffusion theory centered only at the Nernst-Planck equation is incomplete and insufficient to deal with nerve phenomena[2].

In contrast, a chemical theory known as the ion exchange (IX) theory has recently been introduced to nerve problems[3-5]. In its elementary form, the theory considers only Na, K and Ca ions competing exchange with the anionic sites at the outer interface of the membrane. Though Ling's treatment is very basic[4], its application to nerve currents requires further elaboration. Adam's calculation[5] has made a good start but much more work needs to be done. In essence, the electro-

diffusion theory and the ion exchange theory are complementary and any theory which can combine the two will be definitely superior and hence should be sought after.

All the above theories (H-H, ED and IX) are ionic theories which deal directly with the moving ion species. Ions are slaves which move only under order. To understand the mechanism of nerve currents, one ought to find out who sitting behind really gives the order. Or what is the prime mover of the moving ions? The prime mover must sit in fixed positions and must have dynamic behaviour which can affect the conduct of moving charged species. The best candidate to meet both requirements is electric dipoles (polar molecules). Thus the dipole theory is bound to arise and is rising rapidly in recent years[6-15]. There are two great advantages of the dipole theory: it can establish a link between the membrane structure and functions; and it can deal with both electrical and non-electrical (optical and thermal) events[13,16]. Thus even though intense interest in the dipole theory came much later than the ionic theories, real understanding at the molecular level of nerve phenomena may rely more on the former than on the latter.

If a biological scientist looks deliberately into the physical sciences, he will probably find that the semiconductor theory developed in the past 25 years has contained many of the essential features of the electrodiffusion theory, ion exchange theory and dipole theory on nerve currents[8-11]. We shall show this from some elementary considerations.

II. FOUR BASIC QUESTIONS

Since electric current is determined by charge, potential, field and path conditions, we ought to ask the following basic questions before studying nerve currents: (1) What is the charge configuration in a nerve axon with three separate regions (external solution, membrane and axoplasm)? (2) What is the potential profile in the resting state of the axon? (3) What is the field distribution in the membrane under current flow? and (4) Are *all* the moving ions really going *through* the membrane without being stopped somewhere on the way?

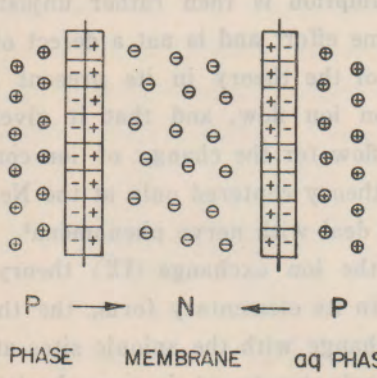


Fig. 1. The possible *pnp* charge configuration and the two dipole barriers of a nerve axon

The answer to the first question is depicted in Fig. 1 which was suggested in the Danielli-Davson model[17]. One sees from Fig. 1: (1) there are two electric dipole layers at the membrane interfaces, with the negative ends facing the aqueous phase, (2) within the membrane, there is an excess of mobile *negative* ions, and (3) in the immediate vicinity of the membrane in the aqueous phase, there is an *excess* layer of mobile *positive* ions whose concentration is greater than that in the bulk solution.

It is clear from Fig. 1 that the charge configuration from inside to outside the membrane looks like *exactly pnp* where *p* stands for excess positive ions and *n*, for excess negative ions. The potential profile across the membrane based on the charge distribution and the two dipole layers is shown in Fig. 2. The resting membrane potential is given by $E_R = E_i + E_M + E_0$. Where E_0 and E_i are the barrier potentials at the outer and inner interfaces and E_M, the true membrane potential. The outer barrier potential E_0 is in the order of 20–30 mV, as estimated by several authors[5,9,15,18]. A simple calculation by Wei[9] has indicated that an E_0 of this magnitude will keep the Na ions to stay out without the aid of a sodium pump.

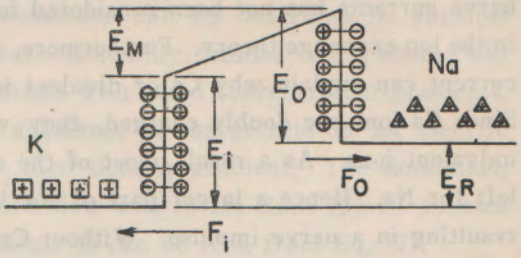
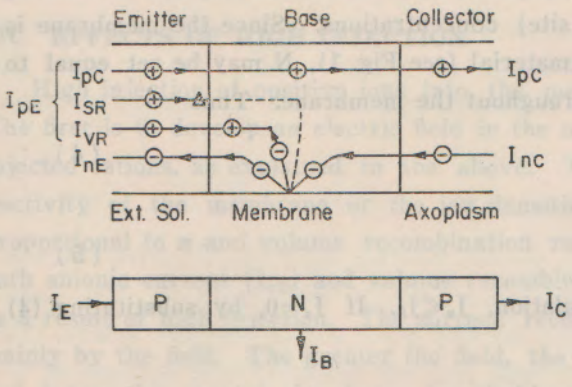


Fig. 2. Potential profile in the resting state of a nerve axon

The *pnp* charge configuration and the potential profile in a nerve axon are similar to those characterizing a *pnp* transistor[19]. Even though the detailed structures of the two are quite different, the physical theory on electrical currents should be applicable to both equally well[8]. Let us see how the semiconductor theory would answer the third and the fourth questions given in the above and then exercise our judgement.

For convenience, let us consider the fourth question first. Fig. 3 shows all possible currents in the three regions of a nerve axon treated as a *pnp* transistor. One sees that the positive ion current I_{pE} starting from the emitter (external solution) tumbles all the way along: a part I_{SR} (surface recombination current) stopping at the outer interface by recombination with surface traps (say, anionic binding sites), a second part I_{VR} (volume recombination current) resulting from recombination with anions at the recombination centers in the membrane and a third part I_{pc} finally reaching the collector (axoplasm). This shows that the membrane serves not only as a pathway but also as a *terminal* of current flow. Thus a nerve axon is not a two-terminal but a *three-terminal* device. This is a great departure from the conventional idea. The "base current" is the sum of all current flows, anionic and cationic, which terminate at the surface and in the bulk of the membrane. This part of the



$$I_E = I_{pE} + I_{nE} \approx I_{pE}$$
$$I_C = I_{pC} + I_{nC} \approx I_{pC}$$
$$I_B = I_{SR} + I_{VR} + I_{nE} - I_{nC}$$

Fig. 3. Possible currents in a nerve axon taken as a transistor

where p_E=the positive ion density at the outer interface (emitter junction), N_t= the number of surface traps, A_t=the effective trapping cross-sectional area of each trap and s=the surface recombination velocity. We shall consider the change in I_{SR} with respect to the change in I_{pE}, the positive ion current at the outer interface.

$$\frac{\partial I_{SR}}{\partial I_{pE}}=sqN_tA_t\frac{\partial P_E}{\partial I_{pE}}=sqN_tA_t\frac{\partial P_E}{\partial Z}\cdot\frac{\partial Z}{\partial I_{pE}}=KsqN_tA_t\partial p_E/\partial Z \tag{12}$$

where $Z=KI_{pE}$ as given before. p_E can be obtained from Eq. (8) by setting $x=0$,

$$2p_E/P-\ln(1+p_E/P)=Z$$

$$\partial p_E/\partial Z=P(1+p_E/P)/(1+2p_E/P)=Pg(Z) \tag{13}$$

where $g(Z)$ is a function which decreases sharply from unity near $Z=0$ and very gradually to 0.5 after $Z=5$. Substitution of (13) into (12) gives

$$\frac{\partial I_{SR}}{\partial I_{pE}}=KsqN_tA_tPg(Z)=Ag(Z) \tag{14}$$

This shows that the surface recombination current drops sharply as the ion injection increases from zero. From the above discussion, we can conclude that high injection would be to decrease I_{SR} and to increase both I_{VR} and I_{nE}. The reverse would be true at low injection. Hence too small and too large a stimulating current would not excity the nerve. The first effect implies a threshold for excitation and the second effect is known as the Wedensky's paradox[21]. Webster had obtained[20],

$$\frac{1}{\alpha}=\frac{\partial I_B}{\partial I_E}\cong\frac{\partial I_B}{\partial I_{pE}}=\frac{\partial I_{SR}}{\partial I_{pE}}+\frac{\partial I_{VR}}{\partial I_{pE}}+\frac{\partial I_{nE}}{\partial I_{pE}}\simeq Ag(Z)+(B+C)(1+Z) \tag{15}$$

where A, B and C are constants related to conductivities, diffusion constant and dimensions, and

$$Z=KI_E$$

In a previous paper[8], the above expression for α had been used to interpret a wide range of nerve phenomena including those found in perfused squid axon. This is really a new adventure of the semiconductor theory outside its home base. We shall next attempt to explain sodium and potassium currents and active transport with this theory.

IV. SODIUM AND POTASSIUM CURRENTS

In a transistor, there are only two species, positive holes and negative electrons. In a nerve axon, there are three major ion species, Na^+, K^+ and Cl^-. The names of Na and K currents are based on the starting place and on the time course. The inward Na flow starts from the external solution (emitter) and appears in the earlier time of the total current. The outward K flow starts from the axoplasm (collector) and has a delayed rise in the time course. Thus from time and place of origin, we have little question to make I_K equivalent to I_c, the collector current. However, there is some ambiguity about the equivalence of I_{Na} since the base current I_b is a part of the emitter current I_E (see Fig. 3). From the circuit

point of view (lower diagram of Fig. 3), $I_E = I_b + I_c$. If I_b is taken as I_{Na}, then $I_E = I_{Na} + I_K$ which is the total current just as what it should be. In the previous discussion and from Fig. 3, the base current is a part of the current flow which terminates at the surface and in the bulk of the membrane (base). When Na ions start to flow inward, they will first meet the surface traps, next the anions in the membrane and finally get into the interior. Thus in time sequence, the base current which consists largely of the surface and volume recombination currents should appear earlier, just as I_{Na} in the total nerve current. This equivalence of I_{Na} to I_b in the time course gives us a new understanding on the fate of inward going Na ions. A large part of them do not get into the interior but rather are arrested (or resting) at the surface and in the bulk of the membrane.

Granted the equivalence of I_{Na} and I_K to I_b and I_c in a *pnp* transistor, what then would be their forms in the time course under stimulation? Let us first assume the stimulus as a step voltage across the outer interface (emitter-to-base junction). Analysis by means of Laplace transform[22,23] has yielded I_b and I_c shown in Fig. 4. These forms look like I_{Na} and I_K as observed but not quite. The rising portions of I_b and I_c are steeper than actually observed I_{Na} and I_K, and I_c should have terminated shortly after reaching saturation if it were equivalent to I_K. The non-conformities arise from the assumed *step* stimulus across the outer interface. The fact may not be so. According to the dipole theory[13], under stimulation, the dipoles at the outer interface will rotate and the barrier height will change accordingly. Calculation shows that the barrier potential will rise more gradually than a step voltage but will decline after the peak like an exponential[13]. When the I_b and I_c in Fig. 4 are each multiplied by this form of barrier potential, one will find that they will more closely resemble to the actually observed I_{Na} and I_K.

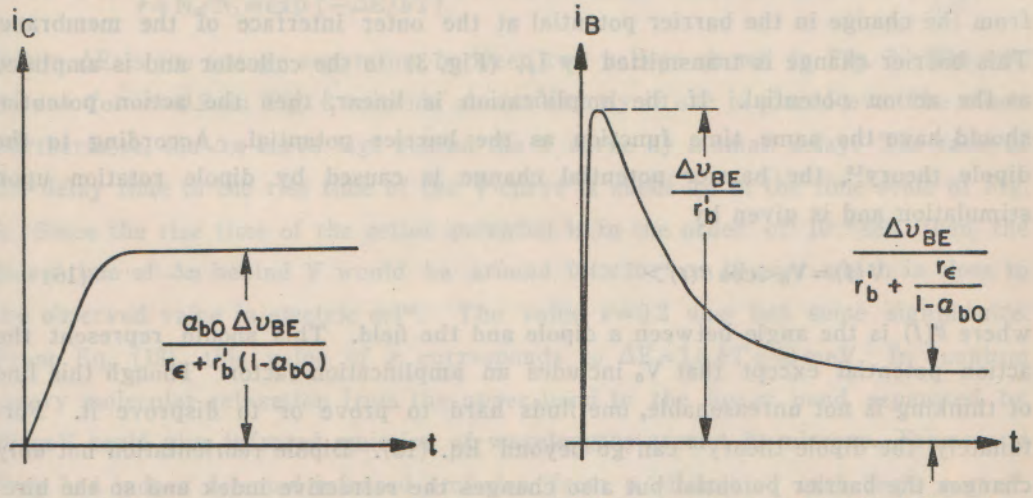


Fig. 4. The base current I_b and collector current I_c in a transistor under stimulation of a step voltage across the base-to-emitter junction

We shall now come to the intricate question of active transport. According to the conventional idea, extrusion of Na from the interior is against its electrochemical

potential hill and hence would require help from some active device such as "sodium pump". The semiconductor theory has clarified two points. First, a large part of the inward flowing Na are residing at the surface and in the bulk of the membrane and only a small part gets into the interior. Secondly, there is the "Self-Aid Principle", that is, at any time the instantaneous field in the membrane is always to aid the flow of *main stream* of positive ions regardless of its direction. This principle is dictated obviously by Eq. (6). In the falling phase of action potential, the main stream of flow is the outward flow of potassium ions. The instantaneous field in the membrane is then increasing in the outward direction. The Na ions, be they in the interior or in the membrane will be *automatically* aided by this field to extrude out and no sodium pump is needed. This is also true for the return trip of K ions left outside. They will be helped by the membrane field in return to their home base during the rising phase of the action potential. Thus the inward and outward flows should consist of both Na and K ions though in unequal proportions. In the perfused experiment, there may be no Na ions outside and/or no K ions inside and hence the conventional inward Na flow and outward K flow would be completely misleading. For these reasons, we think that the terms of emitter, base and collector currents borrowed from transistor terminology may be a better description. Particularly, the base current I_b which gives us a deeper insight into the real (but hidden) situation of the nerve currents should be given due recognition.

V. ORIGIN OF ACTION POTENTIAL

From the viewpoint of transistor theory, the action potential in nerve is a collector-to-emitter potential variation which should be an amplification of a base-to-emitter potential variation. Thus the origin of action potential should have come from the change in the barrier potential at the outer interface of the membrane. This barrier change is transmitted by I_{bc} (Fig. 3) to the collector and is amplified as the action potential. If the amplification is linear, then the action potential should have the same time function as the barrier potential. According to the dipole theory[13], the barrier potential change is caused by dipole rotation upon stimulation and is given by

$$V(t) = V_0 \langle \cos \theta(t) \rangle \tag{16}$$

where $\theta(t)$ is the angle between a dipole and the field. This should represent the action potential except that V_0 includes an amplification factor. Though this line of thinking is not unreasonable, one finds hard to prove or to disprove it. Fortunately, the dipole theory[13] can go beyond Eq. (16). Dipole reorientation not only changes the barrier potential but also changes the refractive index and so the birefringence. It has been found in recent years that in nerve axons under stimulation, the birefringence change almost coincides with the action potential except by a small delay[24-27]. If the dipole theory can show the near coincidence of the barrier potential change with the birefringence change and a small delay of the latter

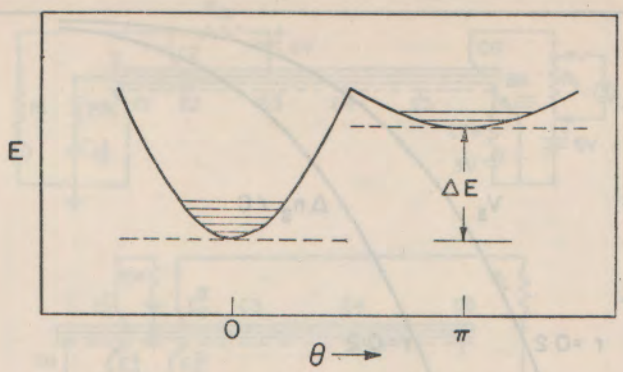


Fig. 5. Energy distribution of dipoles versus orientation

behind the former, then there will be a stronger reason in favor of the origin of the action potential coming from the barrier potential change or from the dipole reorientation under stimulation. We have just done this calculation[13]. In the calculation, we need the energy distribution of dipoles. According to Frohlich[28], ordered dipoles generally have two stable states at the valleys of two energy bands and separated by 180°. We approximated this energy distribution by the diagram in Fig. 5 in which the contour of each band is taken as a cosine curve. Then we calculated the barrier potential given by Eq. (16) and the birefringence change by

$$\Delta n = C < (3\cos^2\theta(t) - 1)/2 > \tag{17}$$

The result is shown in Fig. 6. The r in that figure is the ratio of the dipole populations in the upper and in the lower bands,

$$r = N_2/N_1 = \exp(-\Delta E/kT) \tag{18}$$

where ΔE is the energy separation between two valleys shown in Fig. 5. The two curves for $r=0.2$ in Fig. 6 coincide almost exactly if one is placed over the other. Furthermore, the Δn curve lags behind the V curve by a small delay. The ratio of the delay time to the rise time of the V-curve is about 0.4 in the time scale of Fig. 6. Since the rise time of the action potential is in the order of 10^{-4} sec, then the delay time of Δn behind V would be around 0.4×10^{-4} or 40 μsec which is close to the observed value in electric eel[25]. The value $r=0.2$ also has some significance. From Eq. (18), this value of r corresponds to $\Delta E = 1.6\,kT = 40$ meV. In quantum theory molecular relaxation from the upper band to the lower band separated by 40 meV could give infrared emission of wavelengths around 30 microns. Fraser and Frey had indeed detected infrared emission from a stimulated crab leg's nerve[29]. Thus the semiconductor theory and the dipole theory seem to be able to shed some light on the origin of action potential, birefringence change and infrared emission; the origin would be one and the same, namely dipole reorientation at the outer interface of the nerve membrane under stimulation.

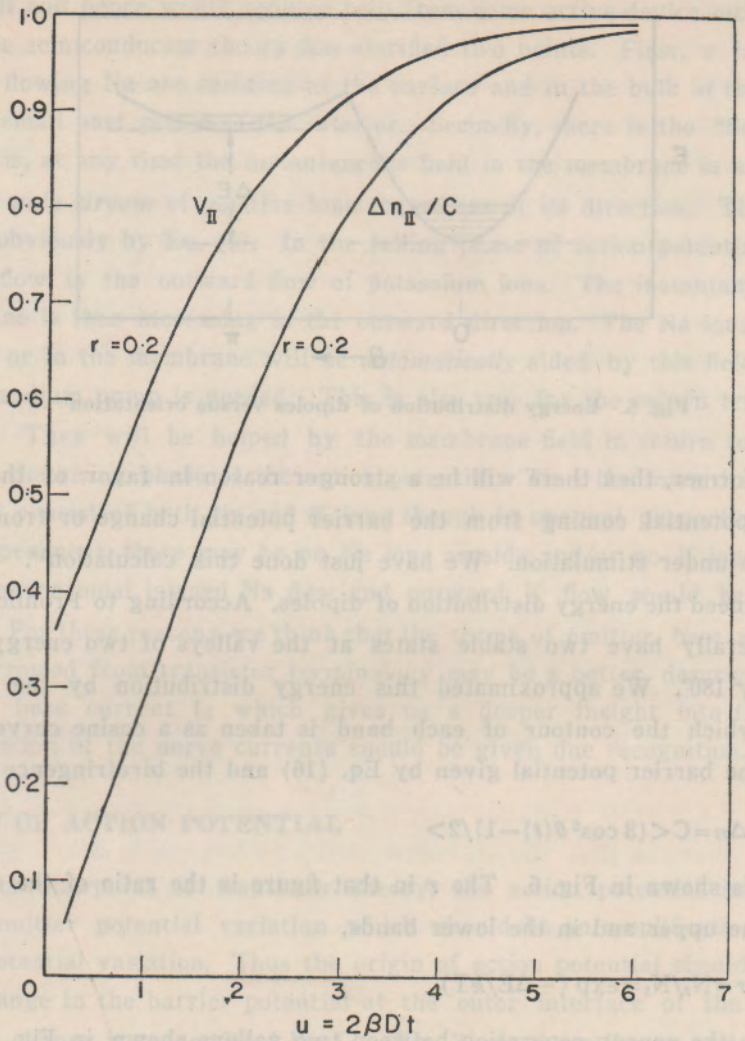


Fig. 6. The calculated barrier potential V and birefringence change Δn in a nerve axon based on the energy distribution of Fig. 5

VI. HARDWARE *PNP* MODEL

So much for the theory. A question which has been put to me since I first proposed the *pnp* transistor theory for nerve conduction in 1966[8] is, can a *pnp* semiconductor structure be fabricated to demonstrate some nervelike properties before naked eyes? The answer to this question would be interesting both the biologist and the electrical engineer. If the answer were yes, then the biologist would at least have some trust in the capability of a new theoretical tool—the semiconductor theory—which could be used to deal with nerve problems, and the electrical engineer would see a possibility even remote of making a zero-loss pulse transmission line, a pressing demand for handling today's heavy traffic of information. With the latest "Integrated Circuit" (IC) technique, we have been able to fabricate hardware *pnp* semiconductor active line or "stretched transistor" which did give the answer desired. The longest line we made was 2 cm long. The line

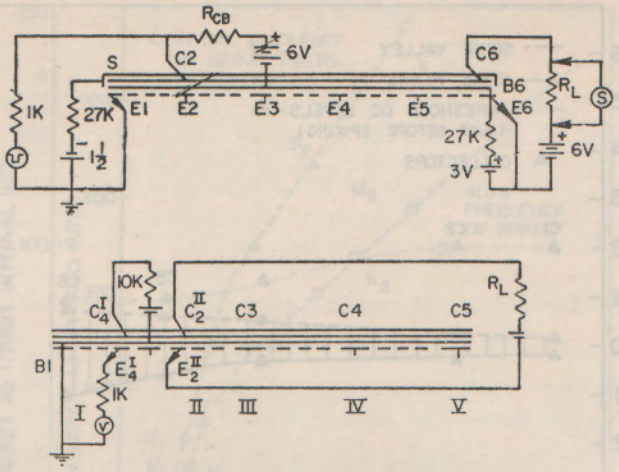


Fig. 7. Displaced collector-driven (upper) and emitter-driven (lower) *pnp* active lines for spike generation and transmision

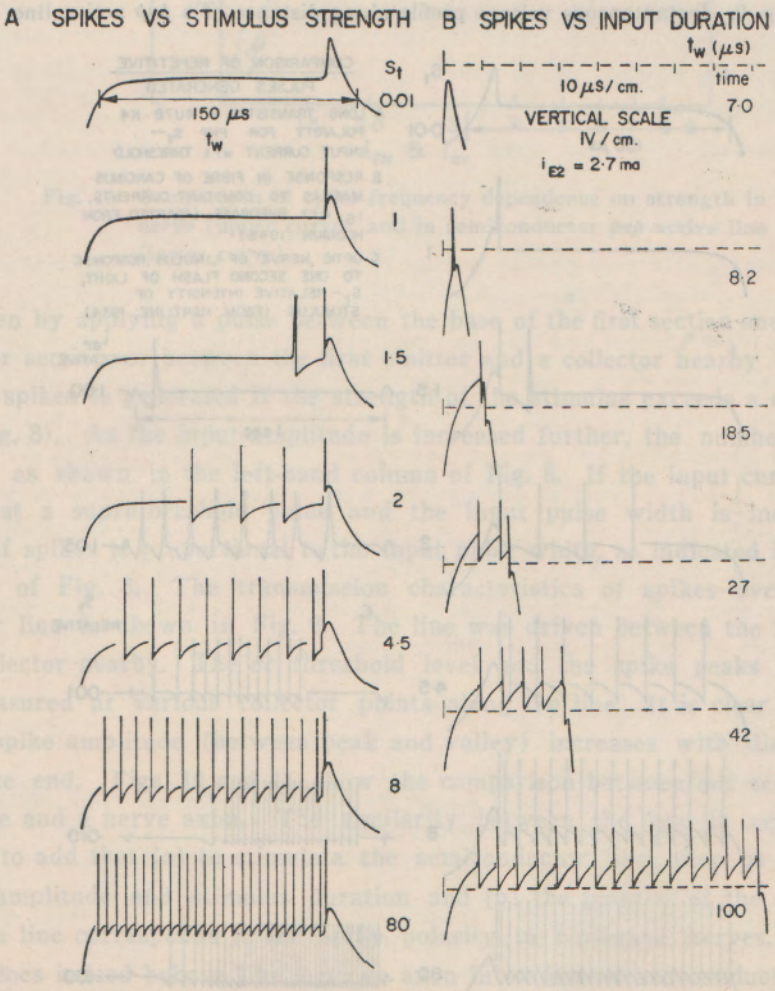


Fig. 8. Spike generation for various input pulse amplitudes (left column) and for various input pulse widths (right hand column) in a *pnp* active line

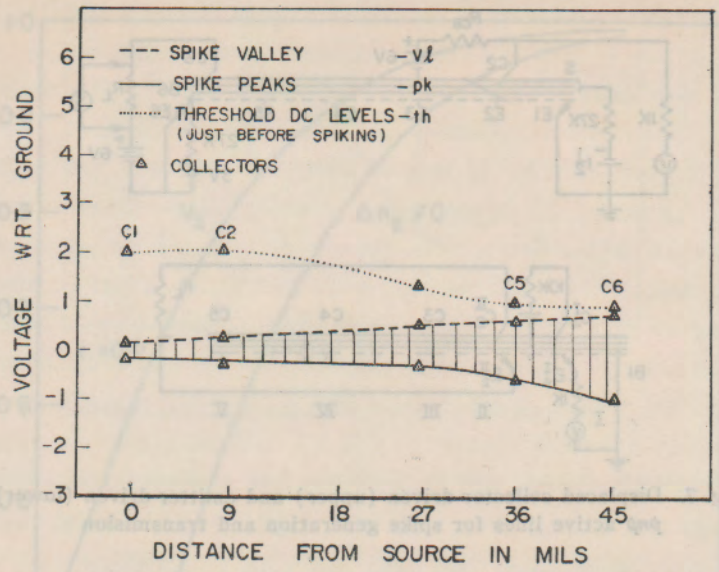


Fig. 9. Instantaneous voltage profile along distance of a *pnp* active line

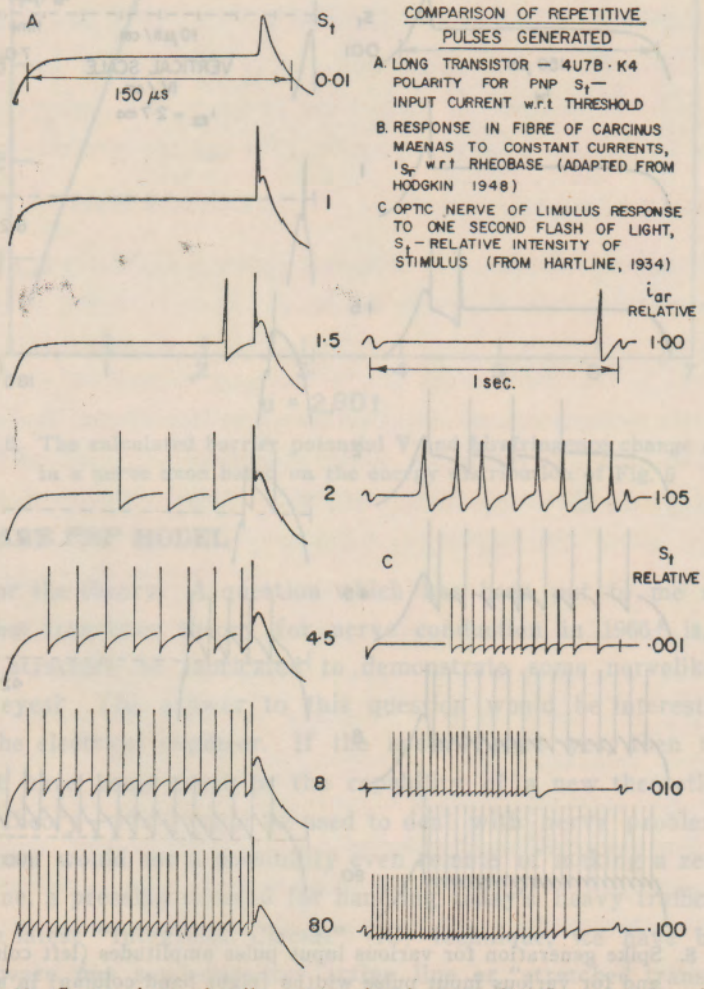


Fig. 10. Comparison of spike generation in a semiconductor *pnp* active line (left side) and in nerves (right side)

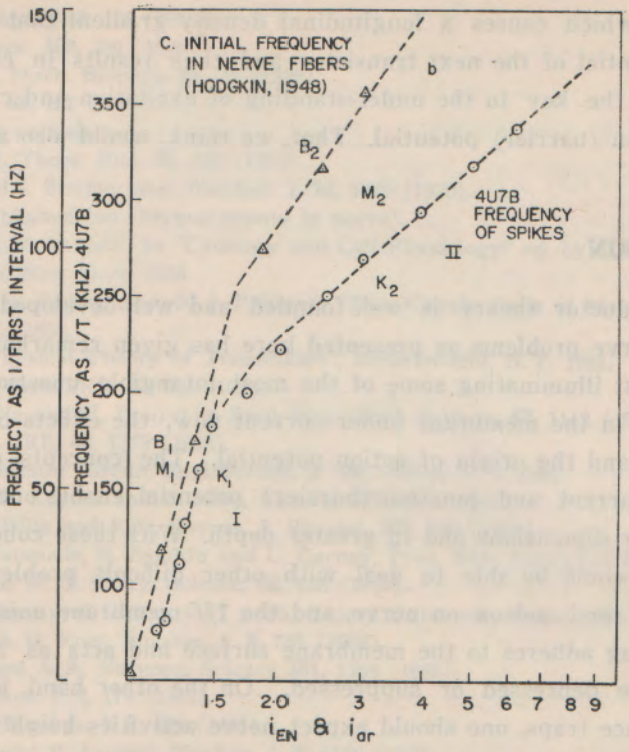


Fig. 11. Comparison of spike frequency dependence on strength in nerve (upper curve) and in semiconductor *pnp* active line (lower curve)

was driven by applying a pulse between the base of the first section and the emitter of another section, or between the first emitter and a collector nearby (Fig. 7). A train of spikes is generated if the strength of the stimulus exceeds a certain threshold (Fig. 8). As the input amplitude is increased further, the number of spikes increases as shown in the left-hand column of Fig. 8. If the input current is held constant at a suprathreshold value and the input pulse width is increased, the number of spikes is proportional to the input pulse width, as indicated in the right-hand-side of Fig. 8. The transmission characteristics of spikes over the semiconductor line is shown in Fig. 9. The line was driven between the first emitter and a collector nearby. The *dc* threshold level and the spike peaks and valleys were measured at various collector points along the line. It is clear from Fig. 9 that the spike amplitude (between peak and valley) increases with distance from the source end. Figs. 10 and 11 show the comparison between our semiconductor active line and a nerve axon. The similarity between the two is very striking. We wish to add that (a) to stimulate the semiconductor line, there is a threshold, both in amplitude and stimulus duration and (b) the polarity of the spikes for a *pnp* active line corresponds to the spike polarity in biological nerves. Thus our *pnp* line does indeed behave like a nerve axon in excitation and conduction. Every point of the *pnp* line is actually a tiny *pnp* transistor. When the emitter barrier of that tiny transistor is suddenly lowered by stimulation, a large transient current

passes through which causes a longitudinal density gradient that in turn changes the junction potential of the next transistor and thus results in conduction along the line. Here, the key to the under-standing of excitation and conduction is the change of junction (barrier) potential. That, we think, would also apply to a nerve axon.

VII. CONCLUSION

The semiconductor theory is well-founded and well-developed in physics. Its application to nerve problems as presented here has given remarkable results. The theory sheds light illuminating some of the most intangible questions such as the field distribution in the membrane under current flow, the effects of high injection, active transport and the origin of action potential. The concepts of traps, recombination, base current and junction (barrier) potential enable one view the nerve currents in wider dimensions and in greater depth. With these concepts thoroughly understood, one would be able to deal with other difficult problems such as the effects of drugs and poison on nerve, and the $1/f$ membrane noise voltage[30]. For example, if a drug adheres to the membrane surface and acts as Na—traps, nerve activities will be depressed or suppressed. On the other hand, if drug molecules can fill the surface traps, one should expect nerve activities heightened.

Mathematically, the semiconductor theory of transport employs two master equations, the continuity equation and the Nernst-Planck equation, and thus can handle a wider variety of problems than either of electrodiffusion theory and ion-exchange theory. If the concepts of lifetime, scattering probability and energy states of ions were adopted and become familiarized with, then the analysis of nerve problems would move to an even higher level. On the other hand, intelligent use of a proper equivalent circuit of transistor could make a complex analysis of nerve currents and potentials much simpler and easier. In fact, the trend is already seen and is rising in recent years in taking the semiconductor approach to the studies of fixed charged membrane[31], Nitella[32], lipid membranes[33–36] and nerve membranes[5,8,37]. It is hoped that the wealth of the semiconductor theory accumulated over the past 25 years can benefit the nerve studies in general and can lead to the unveiling of the mechanism of nerve currents in particular.

REFERENCES

1. A. L. Hodgkin and A. F. Huxley, J. Physiol. **117**, 500 (1952).
2. K. S. Cole, "Menbranes, Ions and Impulses", University of California Press, Berkeley, 1968.
3. G. N. Ling, "A Physical Theory of the Living State", Blaisdell Publ. Co., N. Y. 1962.
4. I. Tasaki, In "Biological Membranes" ed. by D. Chapman, p. 347-410, Academic Press, N. Y. 1968.
5. G. Adam, In "Physical Principle of Biological Membrane", edited by F. Snell *et al.*, p. 35-64, Gordon and Brench Science Publ. N. Y., 1970.
6. D. E. Goldman, Biophys. J. **4**, 167 (1964).
7. D. E. Goldman, In "The Molecular Basis of Membrane Eunction" ed. by D. C. Tosteson, Prentice-Hall, Englewood Cliffs. N. Y., 259-280 (1969).
8. L. Y. Wei, IEEE (Inst. Elec. Electron. Engr.) Spectrum, **3**, 123 (1966).

9. L. Y. Wei, Biophys. J. **8**, 396 (1968).
10. L. Y. Wei, Science, **163**, 280 (1969).
11. L. Y. Wei, Bull. Math. Biophys. **31**, 39 (1969).
12. L. Y. Wei, Ibid, **33**, 187 (1971).
13. L. Y. Wei, to be publisheh.
14. D. Wodschall, J. Theor. Biol. **21**, 439 (1968).
15. B. B. Hamel and I. Zimmerman, Biophys. J. **10**, 1029 (1970).
16. L. Y. Wei, unpublished (on thermal events in nerve).
17. F. Brown and J. F. Danielli, In "Cytology and Cell Physiology" ed. by G. H. Bourne, 249-250, Academic Press, New York, 1964.
18. F. H. Johnson, H. Eyring and M. J. Polissar, "The Kinetic Basis of Molecular Biology", p. 564, Wiley, N. Y., 1954.
19. L. B. Valdes, "Physical Theory of Transistors", McGraw-Hill, N. Y. 1961.
20. W. M. Wedster, Proc. IRE, **42**, 914 (1954).
21. R. Lorento de No and V. Honrubia, Proc. Nat. Acad. Science, **52**, 1142 (1964).
22. J. L. Moll, Proc. IRE, **42**, 1773 (1954).
23. R. F. Shea, Transistor Circuit Engineering, p. 306, Wiley, N. Y. 1958.
24. L. B. Cohen, R. D. Keynes and B. Hille, Nature, **218**, 438 (1968).
25. L. B. Cohen, B. Hille and R. D. Keynes, J. Physiol. **203**, 489 (1969).
26. I. Tasaki, R. Watanade, R. Sandlin and L. Carnay, Proc. Nat. Acad. Sci. **61**, 883 (1968).
27. L. D. Carnay and W. H. Barry, Science, **65**, 608 (1969).
28. H. Frohlich, "Theory of Dielectrics", p. 55, Oxford University Press, 1958.
29. A. Fraser and A. H. Frey, Biophys. J. **8**, 731 (1968).
30. H. E. Derksen and A. A. Verveen, Science **151**, 1388 (1966).
31. A. Mauro, Biophys. J. 2, 179 (1962).
32. H. G. L. Coster, Biophys. J. **5**, 669 (1965).
33. W. E. Bamberg and P. Lauger, Biophys. J. **9**, 1150 (1969).
34. B. Neumcke, D. Waltz and P. Lauger, Biophys. J. **10**, 172 (1970).
35. B. Rosenberg and H. C. Pant, Chem. Phys. Lipids, **4**, 203 (1970).
36. D. R. Kalkware, D. L. Frasco and W. H. Brattain, In "Physical Principles of Biological Membranes" ed. by F. Snell *et al.*, p. 165-174, Gordon and Breach Science Publishers, N. Y. 1970.
37. F. W. Cope, Proc. Nat. Acad. Sci. **61**, 905 (1968).
38. The work was supported by the National Research Council of Canada under Grant No. A1252.