A SIMPLE CASE OF THE WILSON-COWAN EQUATIONS

A. MUIR, *Department of Mathematics, The City University, London EC1 0HB England*

ABSTRACT The stability of equilibria, for Wilson-Cowan equations with piecewise-constant threshold functions, is investigated. It is shown that the characteristic equation has roots with positive real part, when the physiological parameters have reasonable magnitudes. The significance of this fact is discussed.

INTRODUCTION

Two important papers by Wilson and Cowan (1972, 1973) have presented equations governing the statistical dynamics of neural populations.

They suggest that the columnar organization of the primary sensory cortex provides reliability through redundancy along the radial dimension. The cortex is therefore functionally two dimensional, in this interpretation, and they argue for the same organizational principles in subcortical structures such as the thalamus.

For any such structure then, the appropriate functional variable will be some suitable average of the neural firing rates along the radial direction.

In this paper I want to examine one of the simplifying assumptions, which they employ to derive their equations, by comparing some of their predictions with those of an alternative simplification. I will take a critical look at the temporal smoothing procedure by which Wilson and Cowan reduce the complexity of their equations from an integro-difference system to an ordinary dynamical system.

It will be shown that the integro-difference equations may be tackled more directly by assuming constant threshold for the cells of the population. This alternative treatment yields oscillatory phenomena masked by the smoothing.

It is argued here that these oscillations must be reckoned with and not averaged away. More explicitly, it is suggested that the apparent redundancy of the columns of cortical tissue must be understood not just in terms of reliability but also as a time-sharing device for overcoming the limitations on information rate, which is imposed by the refractory period of the separate cells.

Spatial variation is not at issue here so we employ the simpler equations of the 1972 paper.

THE WILSON-COWAN EQUATIONS

To summarize the approach of Wilson and Cowan, they argue that the firing probabilities for the excitatory and inhibitory populations are sharply peaked about the mean and, further, that correlation between excitation and sensitivity of a cell is negligible. They are then able to

write their dynamical equation as:

$$E(t + \tau) = \left[1 - \int_{t-r}^{t} E(t')\,dt'\right] S_e(A), \tag{1}$$

where $A = \int_{-\infty}^{t} \alpha(t - t')\,[c_1 E(t') - c_2 I(t') + P(t')]\,dt'$.

Here, τ = synaptic delay, r = refractory period, α = membrane leakage time-course, and the coefficients c_i describe average connectivities between the excitatory and inhibitory populations; E and I are the mean firing rates of those populations, and P and Q are external inputs to them. S_e is a threshold function described more fully below.

There is an analogous equation for I that, together with Eq. 1, forms a pair of coupled integro-difference equations. The assumptions leading to these equations seem plausible enough to retain contact with the neural system they are intended to describe, though, as we shall see, the replacement of a discrete structure by a continuum leads to a problem of unboundedly large oscillatory frequencies.

However, the system is sufficiently complicated to demand further reduction and it is at this point that some rather ad hoc arguments are introduced to justify replacement of the variables by a moving time-average. The crux of the approach is to substitute for the integrals $\int_{t-r}^{t} E(t')\,dt'$ and $\int_{-\infty}^{t} \alpha(t - t')E(t')\,dt'$ the same "moving average" $\bar{E}$ (with similar handling of I) $\bar{E}(t + \tau)$ and $\bar{I}(t + \tau)$ are then approximated by their first-order Taylor expansion to yield an ordinary dynamical system in $\bar{E}$ and $\bar{I}$, amenable to phase-plane analysis.

The complexity of justifying this reduction on purely analytical grounds is great. Wilson and Cowan offer plausibility arguments backed by some computer check-out of the behavior of solutions of the unreduced against the reduced systems. In the end, however, they are mainly concerned to exhibit significant qualitative features of the dynamics, such as hysteresis and limit cycles. These phenomena show a good deal of resilience under alterations of the parameter magnitudes, and it might be guessed that the same would be true of structural alterations in the form of the equations, such as the replacement of integrals by moving averages.

Indeed, the latitude permitted is underscored by the different methods they employ in the two papers to effect that replacement. Comparing the spatially localized equations, we find that they are identical except that where the synaptic delay appears as a multiplier of the derivative in 1972, it is replaced, according to a different argument, by the membrane time-constant in 1973.

PIECEWISE-CONSTANT EQUATIONS

In the dynamical Eq. 1 $S_e(x)$ is the proportion of excitatory cells that fire at excitation level x, so S_e can be considered as a cumulative distribution function for the thresholds of cells in the excitatory population (S_i is analogously defined for the inhibitory cells). We assume, following Wilson and Cowan, that these functions are sigmoid; this is equivalent to assuming a unimodal distribution of thresholds.

In this paper we explore the simplification that all cells in each population have the same threshold. Thus the threshold distribution is a delta function and S_e, S_i are piecewise constant.

$$\left.\begin{aligned} S_k(x) &= 0 \quad x < \theta_k \\ &= 1 \quad x > \theta_k \end{aligned}\right\} k = e \text{ or } i.$$

Of course, this is highly idealized, but the analytical tractability thereby achieved sheds some light on the moving averages approximation. Because no real neural system would exhibit such a markedly discontinuous feature, it is desirable to know how far such an assumption takes us away from more realistic situations. It would be desirable to know that the behavior of the constant threshold system stayed close to that of a similar system with a sharply peaked but continuous threshold distribution.

Some comments on this question are presented in the Appendix. Meanwhile, adopting the discontinuous S_k function, Eq. 1 becomes:

$$\begin{aligned} E(t + \tau) &= 1 - \int_{t}^{t} {}_{-r} E(t')\, dt' \quad \text{if } A > \theta_e \\ &= 0 \qquad\qquad\qquad\qquad\quad \text{if } A < \theta_e. \end{aligned} \tag{2}$$

There is again a similar equation for I, but to compare our analysis with the computationally based statements of Wilson and Cowan, we follow them in decoupling the two populations, writing $c_2 = 0$ in A and considering only the excitatory population.

In the absence of external stimulation, $P = 0$ and $A = \int_{-}^{t} {}_{\infty} \alpha(t - t') \cdot c_1 E(t') dt'$. Such an isolated population will have equilibrium states $E = E_0$ which satisfy

$$\begin{aligned} E_0 &= 1 - rE_0 \quad \text{if } c_1 \mu E_0 > \theta_e \\ &= 0 \qquad\quad \text{if } c_1 \mu E_0 < \theta_e, \end{aligned}$$

where $\mu = \int_0^\infty \alpha(t)\, dt$.

Thus $E_0 = 0$ is always an equilibrium, but there is also a nonzero equilibrium $E_0 = 1/(1 + r)$ if $c_1 \mu > \theta_e(1 + r)$.

STABILITY

To discuss the stability of these equilibria we write $E = E_0 + \Delta$ and examine conditions under which, for sufficiently small Δ, the system tends to return to $E = E_0$. Because of the integrals in the dynamical equations, the system's behavior depends on the whole of its past history. It is not, therefore, sufficient to consider the tendency to return to equilibrium after merely an initial displacement; rather we need to maintain the displacement over a time interval to permit the system's "memory" to accumulate.

For instance, the equilibrium $E_0 = 0$ yields $A = c_1 \int_{-}^{t} {}_{\infty} \alpha(t - t')\Delta(t') dt'$, and if $\Delta(t') < \epsilon$, we have $A < c_1 \mu \epsilon$ for all t. Then with any positive threshold it is possible to choose $\epsilon > 0$ such that $A < \theta_e$ for all t. This implies that holding the displacement at a magnitude $\Delta < \epsilon$ for such an ϵ, over any interval of time, will ensure, by Eq. 2, that the system returns to $E = 0$ after a further delay τ, after release.

Turning attention to the nonzero equilibrium, assuming it exists, the displacement equations are:

$$\Delta(t+\tau) = -\int_{t}^{t} {}_{-\tau}\Delta(t')\,dt' \quad \text{if } B > \phi$$
$$= -E_0 \quad \text{if } B < \phi, \tag{3}$$

where $B = \int_{-}^{t} {}_{\infty}\alpha(t-t')\Delta(t')\,dt'$ and $\phi = \theta_e/c_1 - E_0\mu < 0$.

Because Eq. 3 is linear it admits a Laplace transform approach. So long as $B > \phi$, we then have

$$e^{z\tau}\left[\bar{\Delta}(z) - \int_0^{\tau} \Delta(t)e^{-zt}\,dt\right] = \frac{(e^{-z\tau}-1)}{z}\left[\bar{\Delta}(z) + \int_{-}^{0} {}_{\tau}\Delta(t)e^{-zt}\,dt\right], \tag{4}$$

where $\bar{\Delta}(z)$ is the Laplace transform of $\Delta(t)$.

Let us assume that the initial conditions have been maintained for a sufficient length of time so that the two integrals in Eq. 4 give well-defined functions of z. Then the equation takes the form

$$G(z)\cdot\bar{\Delta}(z) = F(z),$$

where $F(z)$ contains the integrals of Eq. 4 and

$$G(z) = e^{z\tau} + (1 - e^{-\tau z})/z. \tag{5}$$

The inversion formula for the transform expresses Δ as a linear combination of exponentials e^{zt}, where z runs over poles of the transform. The poles that are intrinsic to the system and not dependent, through $F(z)$, on the initial conditions will then arise as zeros of the function $G(z)$.

The contribution of a term e^{zt} to the "memory" B is $\int_{-}^{t} {}_{\infty} e^{zt'}\,\alpha(t-t')dt'$. If, for instance, α has exponential form with time constant μ, this integral is $\mu[e^{zt} - e^{-t/\mu}]/(1+\mu z)$. Clearly we would then hope to exclude terms with $Re(z) > 0$, $IM(z) \neq 0$, because B would otherwise contain a term becoming arbitrarily large and negative, destroying the condition $B > \phi$. The system would then jump to $E = 0$ after a further time, τ.

Conversely, if $Re(z) < 0$ we can choose the coefficient of e^{zt} sufficiently small so that it will not destroy the condition $B > 0$ and its contribution to Δ will decay to zero.

This confirms analytically Wilson and Cowan's point that their unreduced equations can yield damped oscillations decaying asymptotically to the equilibrium of the reduced equations. However, their remark, based on computer solution of the unreduced equation, that "the period of this oscillation is dependent almost entirely on the length of the absolute refractory period" is not obvious. It is necessary to look more closely at the roots of the "characteristic equation," $G(z) = 0$.

THE CHARACTERISTIC EQUATION

Let us first note that there are no real solutions because when z is real $e^{\tau z}$ is positive, whereas $1 - e^{-\tau z}$ has the same sign as z.

Turning to the question of nonreal roots, we can remark that it is difficult enough to determine roots of a complex transcendental equation, let alone know whether one has

obtained them all, since they may be infinite in number. We are concerned here with whether values of τ and r, which are physiologically reasonable, ensure that the roots of the characteristic equation all have negative real part.

Unfortunately, this is not the case. For example, the equation with $r = 2, \tau = 3$ possesses a root $z = 0.148 + 0.796\,i$, that with $r = 3, \tau = 4$ possesses a root $z = 0.178 + 0.586\,i$, and so on. It is probably not very illuminating to map out in detail the location of such roots (Fig. 1); their very existence is enough to demand explanation, since as we saw in § 4 such a component will eventually cause the system to jump to the quiescent state, $E = 0$.

Two interpretations are immediate. Either the membrane leakage function decays faster than a merely passive exponential, which indicates an active process for membrane discharge. Or, we have to admit the inevitability of a transition to $E = 0$ in an isolated subpopulation, interpreted as a wiping clean of the past record, and rely on the external input or spatial interaction between populations to keep activity topped up.

However, both of these explanations could collapse if the characteristic equation admits roots with $Re(z)$ arbitrarily large, for then the exponential growth of the corresponding modes and the consequent jump to quiescence could occur within an arbitrarily short time. The problem of such roots has been too readily smoothed over by Wilson and Cowan's reduction.

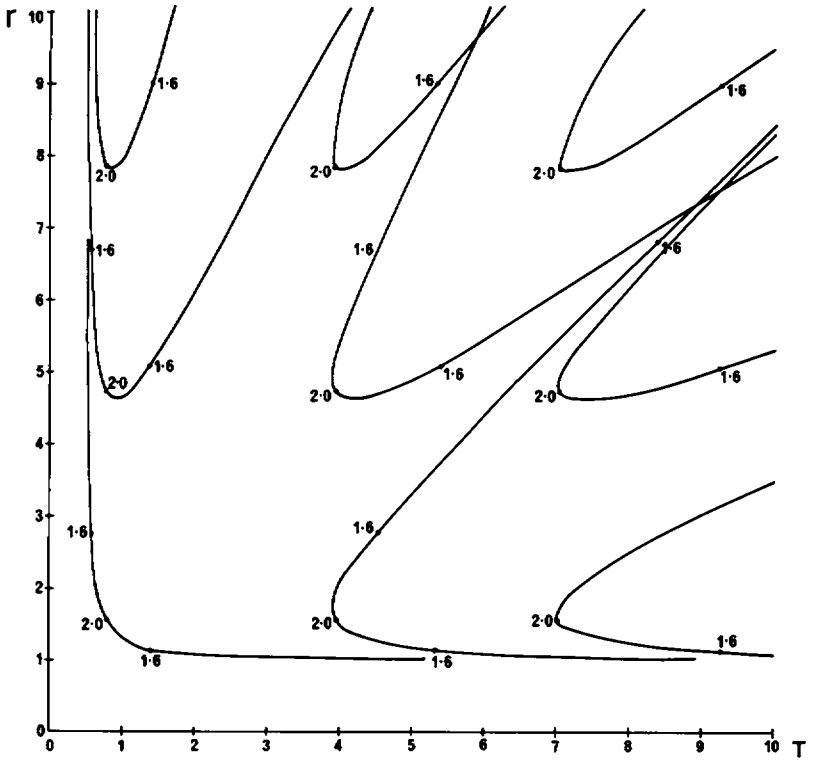

FIGURE 1 Graph of (τ, r) for which there exists roots of the characteristic equation with zero real part. The numbers on the curves show the imaginary part of the root—the angular frequency of the oscillation—at the indicated points.

Turning to $Im(z)$, the angular frequency of a mode of oscillation, we can get some idea of how this is distributed over the poles of G by considering those roots for which $Re(z) = 0$.

Writing $z = i\beta$, the characteristic equation becomes:

$$\begin{aligned} \cos \beta\tau &= -\sin \beta r/\beta \\ \sin \beta\tau &= (1 - \cos \beta r)/\beta. \end{aligned} \tag{6}$$

Treating these as parametric equations for (τ, r) in terms of β, we can solve them explicitly as:

$$\begin{aligned} &\beta\tau = M + 2k\pi, \beta r = 2\pi - 2M + 2\ell\pi \\ \text{or} \quad &\beta\tau = \pi - M + 2k\pi, \beta r = 2M + 2\ell\pi, \end{aligned} \tag{7}$$

where $M = \sin^{-1}(\beta/2)$.

Thus, there is a doubly infinite set of curves in the $\tau - r$ plane, consisting of those (τ, r) for which a solution exists with $Re(z) = 0$, each branch being defined by a choice of integer pair (k, ℓ). A sketch graph is presented from which a number of properties of the solutions may be inferred. There appear to be (τ, r) values yielding any designated finite number of roots. For instance, $r = 9.2$, $\tau = 5.6$ lies near the points on two distinct curves corresponding to $z = 1.2i$ and $z = 1.6i$. Continuity of the roots with respect to variation of τ and r indicates that exact roots of the characteristic equation for the stated r, τ will be close to these z-values. Extension of the diagram to larger values of τ, r shows similarly that $r \approx 14, \tau \approx 15$ lies close to 5 curves, yielding an expected 5 solutions, at least.

It is fortunate that physiologically reasonable magnitudes for τ and r, say < 10 ms, place us in a region with few branches, indicating the possibility of approximating G by a low degree polynomial, a point to which we return.

POLYNOMIAL APPROXIMATION

A striking feature of the curves is their almost linear appearance, except near $\beta = 2.0$; this is most conspicuous in the basic L-shaped solution $k = \ell = 0$. If we express M in Eq. 7 by a Maclaurin series in β, then to achieve an accuracy of ≈ 0.02 we need only employ the constant approximation for $\beta < 0.7$ and the next (quadratic) term for the remainder of the domain.

Thus, to this accuracy, the basic solution becomes:

$$(\tau, r) = \left(\frac{1}{2}, \frac{2\pi}{\beta} - 1\right) \text{ or } \left(\frac{\pi}{\beta} - \frac{1}{2}, 1\right) \quad \text{for } \beta < 0.7$$

$$= \left(\frac{1}{2} + \frac{1}{48}\beta^2, \frac{2\pi}{\beta} - 1 - \frac{1}{24}\beta^2\right) \text{ or } \left(\frac{\pi}{\beta} - \frac{1}{2} - \frac{1}{48}\beta^2, 1 + \frac{1}{24}\beta^2\right) \quad \text{for } 0.7 < \beta < 2.0.$$

All other curves are then obtainable by displacement in both coordinate directions through multiples of $2\pi/\beta$.

We have here a partial confirmation of Wilson and Cowan's point that the period of oscillation, $2\pi/\beta$, has the same order of magnitude as the refractory period. This is certainly true for the vertical arm of the basic solution and its corresponding part on other branches, but it neglects the horizontal arm where we find oscillatory periods tied to the synaptic delay, τ.

Wilson and Cowan state that oscillations of period less than that of the refractory period cannot propagate information.

The argument, however, is unconvincing. The frequencies in question are properties of the neural aggregate that does not function as a synchronous entity. Therefore we are not requiring that each individual neuron be able to register the fluctuations in firing rate of the aggregate. All that is needed is that a sufficient number of neurons should become nonrefractory within the period of collective oscillation, for the excitation to be maintained.

A perhaps more satisfactory reason for ignoring high frequencies lies in the available experimental methods, rather than in the interpretation of their neural significance. Gross recording focuses attention on components below 50 Hz, filtering out higher frequencies along with random fluctuations. Thus β, in the millisecond units we are using for τ and r, is only detected below approximately $\pi/10 \simeq 0.3$. Assuming that oscillations decaying with a time constant less than a typical refractory period are also ignorable, we further require $Re(z)$, the reciprocal of that time constant, to be $< 1/r$.

With reasonable values for r, such considerations indicate that although the characteristic equation might have solutions z with large modulus, these will be experimentally neglected unless $|z|$ is $< \approx 0.5$. Now, $G(z)$ is analytic throughout the complex plane, so for any z and given error bounds, its Maclaurin series can be truncated suitably so as to approximate $G(z)$ within those bounds. If z is a zero of $G(z)$ it will appear as an approximate root of an appropriately truncated polynomial equation.

Thus, if we are seeking zeros of $G(z)$ whose moduli are less than a designated size we can solve instead a suitable approximating polynomial. Zeros of that polynomial not satisfying the modulus inequality need not approximate to zeros of $G(z)$ but will anyway be discarded.

Consider, for example, the fourth degree approximation:

$$(1 + r) + (\tau - r^2/2)z + (\tau^2/2 + r^3/6)z^2 + (\tau^3/6 - r^4/24)z^3 + (\tau^4/24 + r^5/120)z^4 = 0.$$

An acceptable condition for this approximation to be valid is that the first neglected term, evaluated for maximum observable $|z|$, should be less than the allowable error, δ. Thus,

$$(\tau^5/120 - r^6/720) \times [\max |z|]^5 < \delta.$$

For instance, with $|z| < 0.5$ and, say, $r = 2$, $\tau = 3$, the quartic can replace $G(z)$ to an accuracy ≈ 0.06. The roots of the quartic are then $0.1617 \pm 0.7245i$ and $-0.6880 \pm 1.0108i$, none of which have $|z| < 0.5$. Thus there will be no observable roots for the characteristic equation with $r = 2, \tau = 3$.

CONCLUSION

Repeating such calculations, for various r and $\tau < \approx 5$, yields the same result—no roots of modulus < 0.5. I have suggested that Wilson and Cowan's equations become dubious when they effect their moving averages reduction. The earlier stages of their derivation seem more firmly based but lead, as indicated above, to significant neural phenomena not detectable by usual recording methods.

Whether or not the high frequency oscillations are meaningful in terms of information

propagation between different neural aggregates, those with $Re(z) > 0$ will lead to a switching off of local excitation. This can occur, of course, through a high frequency oscillation which is unobservable. It must be remembered, however, that we have excluded inhibition from the model and it is well known that this is needed to control the runaway behavior of a purely excitatory population.

An immediate question that now arises is whether the multitude of oscillatory modes is merely an artifact of the mathematical model. That is, whether the neural systems it is intended to describe could exhibit the same features.

The principal thing to understand is how arbitrarily high oscillatory frequencies might be achieved. This is best invisaged by imagining, for simplicity, groups of neurons firing in volleys and entering their refractory period, with the excitation being taken over by another group that has just become nonrefractory. In real systems the frequencies would be bounded above by the actually finite nature of the cortex, in contrast to the continuum model employed above.

Redundancy in the cortical columns might well be for reliability, as Wilson and Cowan suggest, but the number of cells per column which they quote, ranging up to 2×10^5 in humans, seems to overdo such caution.

The alternative view, which seems to be implied by the results above, is that the column is indeed a functional unit, but one whose quantitative size is in part employed to permit faster information handling by time-sharing between the units.

The relative status of these two possibilities depends on assessing the familiar trade-off between information rate and reliability.

I wish to thank one of the referees for detailed critical comments that have enabled me to strengthen the original account.

Received for publication 3 October 1978 and in revised form 21 April 1979.

APPENDIX

Let two excitatory neural populations differ only in their threshold functions S and S_1, the parameters τ, r, α being the same.

Denote the excitation levels by E and E_1 and arguments of the threshold functions by A and A_1.

Then,

$$E(t + \tau) = \left[1 - \int_{t}^{t} {}_{-r} E(t')\, dt'\right] S(A) \tag{1A}$$

with a similar equation, mutatis mutandis for the other system.

Assume that E, E_1 differ by less than some positive quantity δ up to time t.

Then,

$$|A - A_1| \leq \int_{-\infty}^{t} \alpha(t - t')\, |E(t') - E_1(t')|\, dt' \leq \delta\mu.$$

Now by Eq. 1A,

$$
\begin{aligned}
&|E(t+\tau) - E_1(t+\tau)| \\
&\left|\left[1 - \int_{t}^{t} {}_{-r} E(t')\,dt'\right] S(A) - \left[1 - \int_{t}^{t} {}_{-r} E_1(t')\,dt'\right] S_1(A_1)\right| \\
&\leqslant \int_{t}^{t} {}_{-r} |E_1(t') - E(t')|\,dt' \cdot S(A) + \left[1 - \int_{t}^{t} {}_{-r} E_1(t')\,dt'\right] |S(A) - S_1(A_1)| \\
&\leqslant r\delta + |S(A) - S_1(A_1)|,
\end{aligned}
$$

because the other factors are between 0 and 1.

If S_1 has continuous first derivative with maximum modulus M on [0,1] we have:

$$
\begin{aligned}
|S(A) - S_1(A_1)| &\leqslant |S(A) - S_1(A)| + |S_1(A) - S_1(A_1)| \\
&\leqslant |S(A) - S_1(A)| + M \cdot \sigma\mu
\end{aligned}
$$

Finally, if S and S_1 differ by at most γ, we have $|E(t+\tau) - E_1(t+\tau)| \leqslant (r + M\delta)\mu + \gamma$, which we denote by k.

We are concerned here with whether the two trajectories which have remained close up to time t will continue to do so. Given $\epsilon > 0$ we want to argue that δ and γ, the former characterizing the nearness of the initial trajectories, the latter the nearness of S and S_1, can be chosen sufficiently small to make $k < \epsilon$.

There is no problem if both S and S_1 can be continuous and we can therefore state the following. Theorem: Let S, the threshold function for a neural system satisfying Eq. 1A, be continuous. Then for any $\epsilon > 0$, there exist $\delta, \gamma > 0$, such that a similar system having threshold function S_1 within γ of S and having initial conditions within δ of the given ones, has a solution trajectory staying within ϵ of the given system for a further time interval τ.

Actually, by a finite number of repetitions of this result, δ, γ can be chosen to ensure the desired closeness over any given finite time interval.

Unfortunately, our discontinuous S does not approximate any continuous function in the desired pointwise manner. The best we can hope for with a sigmoid S_1 is that it should approximate 0 and 1 to the left and right, respectively, of some small interval around the jump-point θ.

The theorem will then still hold so long as the value of A does not enter that interval. For the purposes of this paper, where we are analyzing small perturbations around the nonzero equilibrium, this will be satisfied.

REFERENCES

WILSON, H. R., and COWAN, J. D. 1972. Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys. J.* **12**:1–24.

WILSON, H. R., and COWAN, J. D. 1973. A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik.* **13**:55–80.