Biological Cybernetics

Volume 30 Number 4 1978

Indexed in Current Contents

Springer-Verlag Berlin Heidelberg New York

Biol. Cybernetics ISSN 0340-1200 BICYAF 30 (4) 187-248 (1978)

28. IX. 1978

Biological Cybernetics

Communication and control in organisms and automata.
Continuation of „Kybernetik" (Volume 1–16)
Biological Cybernetics appears about every month.

Subscription Information

Volumes 29–32 (4 issues each) will appear in 1978. The price of each volume is DM 158,– or $ 69.50. Information about obtaining back volumes available upon request.

North America. Subscription rate: $ 290.80 including postage and handling. Subscriptions are entered with prepayment only. Orders should be sent to your bookdealer or subscription agency or directly to: Springer-Verlag New York Inc., Service Center Secaucus, 44 Hartz Way, Secaucus, NJ 07094, USA, Tel. (201) 348-4033, Telex 125994

All Other Countries. Subscription rate: DM 632,– plus postage and handling. Orders can either be placed with your bookdealer or sent directly to: Springer-Verlag, Heidelberger Platz 3, D-1000 Berlin 33.

Manuscripts should be addressed to:
Prof. Dr. W. Reichardt
Max-Planck-Institut für biologische Kybernetik
Spemannstraße 38
D-7400 Tübingen, FRG

Springer-Verlag

Heidelberger Platz 3
D-1000 Berlin 33
Tel. (030) 822001
Telex 01-83319

Postfach 105280
D-6900 Heidelberg 1
Tel. (06221) 487-1
Telex 04-61690

Springer-Verlag New York Inc.
175 Fifth Avenue
New York, NY 10010
Tel. (212) 477-8200
Telex 232235

The concepts of transmission of information, processing of information and automatic control engineering originated within technology and physics. Today, however, these concepts have also proved useful in the biological sciences where analogous processes of communication and control are encountered. Despite the differences between nonliving and living systems, many of the logical procedures, the experimental and theoretical approaches and the mathematical techniques applicable to the physical sciences also find natural applications in the realm of the life sciences. In particular, by adopting this approach to sensory and neurophysiological problems new insight has been gained into the principles by means of which organisms handle and utilize information. Conversely, physicists and engineers have shown increasing interest in natural mechanisms of communication and control, including genetic communication and the control of reproduction.

The aim of "Biological Cybernetics" is to promote the exchange of experimental and theoretical information in the following fields: Quantitative analysis of behaviour, in both vertebrates and invertebrates; quantitative micro- and macro-physiological studies of information-processing in receptors, neural systems and effectors; mathematical models of communication and control processes in organisms, including reproductive mechanisms; biologically relevant aspects of information theory, network theory, theory of automata, theory of control systems.

Forthcoming Papers

D. S. Levine, C. D. Woody
Effects of Active Versus Passive Dendritic Membranes on the Transfer Properties of a Simulated Neuron

L. Bobrowski
Learning Processes in Multilayer Threshold Nets

H. W. Meyer
Phototaxis in the Walking Male and Female Fly *(Calliphora erythrocephala* Meig.*)* II. Water Balance and Phototactic Response

L. H. Zetterberg, L. Kristiansson, K. Mossberg
Performance of a Model for a Local Neuron Population

T. Torioka
Pattern Separability and the Effect of the Number of Connections in a Random Neural Net with Inhibitory Connections

S. Nagshineh, K. H. Ruddock
Properties of Length-Selective and Non-Length-Selective Adaptation Mechanisms in Human Vision

G. Gestri
Dynamics of a Model for the Variability of the Interspike Intervals in a Retinal Neuron

G. Palm
On Representation and Approximation on Nonlinear Systems

Biological Cybernetics

Communication and Control in Organisms and Automata

Nachrichtenübertragung, Nachrichtenverarbeitung, Steuerung und Regelung in Organismen und in Automaten

Springer-Verlag Berlin Heidelberg New York

Instructions to Authors

1. General

Manuscripts should be typewritten, double-spaced throughout, with wide margins. Papers should be as short and concise as possible. Generally 3 double-spaced manuscript pages correspond to about 1 printed page.

In general manuscripts should not exceed 25 manuscript pages including figures and tables.

To accelerate publication only **one set of proofs** is sent to authors. This shows the final layout of the paper as it will appear in the journal. It is therefore essential that manuscripts are submitted in their *final form,* ready for the printer, and that the approximate desired *positions of figures and tables are marked* on the margins. Proofreading should be *limited to the correction of typographical errors.* Any other changes involve time-consuming and expensive work, *and the costs will be charged to the author(s).* If absolutely necessary and space permits, additions may be made at the end of the paper in a "Note added in proof".

Each manuscript should include a separate sheet of "Instructions for the compositor" explaining markings and special types used.

2. Text

Abstract. All papers must have an **English abstract** and, if written in German or in French, an **English translation of the title** in addition.

Headings. Subheadings (Introduction, Materials and Methods, etc.) should be placed on separate lines. Nouns, adjectives, etc., should be capitalized.

Subheadings should be classified according to the following system:

1. = first subheading
1.1. = second subheading
1.1.1. = third subheading
1.1.1.1. = fourth subheading.

Italics should be used for emphasis in text. Underline to indicate *italics.*

Small print. Materials and methods, and sections of lesser importance should be marked for small type. Because of higher composition costs, small type is no less expensive than regular type but serves to improve the organization of the text. Footnotes and tables are always printed in small type.

Footnotes to the text should be numbered consecutively. They should be placed at the foot of each page (not at the end of the article).

Measuring units. For *abbreviations, symbols,* and *units* the author should follow internationally agreed rules such as those defined by the International Organization of Standardisation (I.S.O.). The metric system of units should be used throughout.

3. Formulae

Equations should be typewritten whenever possible, showing special typefaces (alphabets) by underlining in color (see *Marking*). As there is no difference in size, extra special attention should be given to the placing of indices and exponents so that they are recognizable as such. Formulae should preferably be written in the "linear" form (see *Notations*). Equations should be numbered sequentially with arabic numerals in parentheses on the right-hand side of the page. In the text they are cited simply by the arabic numeral in parentheses, e.g., (7); the form Equation (7) or Equations (7) and (8) should be used only at the beginning of a sentence.

Marking. Letters in formulae are normally printed in italics, figures in ordinary typeface.
It will help the printer if in doubtful cases the position of indices and exponents is marked thus: $b\,\Uparrow,\ a\,\Downarrow$

Underlining for special type should be done according to the following code:

single underlining = Small letter
double underlining = Capital letter

yellow = Upright: abbreviations e.g. Re, im, grad, div, lim, NGC, and all elements

brown = Boldface:

A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S, T, U, V, W, X, Y, Z

a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q, r, s, t, u, v, w, x, y, z

red = Greek Lettering:

$\Gamma, \Delta, \Theta, \Lambda, \Xi, \Pi, \Sigma, \Phi, \Psi, \Omega$
$\alpha, \beta, \gamma, \delta, \varepsilon, \zeta, \eta, \theta, \vartheta, \iota, \varkappa, \lambda, \mu, \nu, \xi, o, \pi, \varrho, \sigma, \tau, \upsilon, \varphi, \phi, \chi, \psi, \omega$

green = Script:

$\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}, \mathcal{E}, \mathcal{F}, \mathcal{G}, \mathcal{H}, \mathcal{I}, \mathcal{J}, \mathcal{K}, \mathcal{L}, \mathcal{M}, \mathcal{N}, \mathcal{O}, \mathcal{P}, \mathcal{U}, \mathcal{R}, \mathcal{S}, \mathcal{T}, \mathcal{U}, \mathcal{V}, \mathcal{W}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}$
$a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q, r, s, t, u, v, w, x, y, z$

violet:
the single letters I and capital O (to distinguish them from the numerals 1 and zero)

The following are frequently confused:
$\cup$, **u, U**; *o, O,* 0; x, *x,* X, $\varkappa$; v, *v,* ν; ϕ, φ, Φ, $\varnothing$, Φ; *l,* l, ε, $\in$

also the handwritten Roman letters:
c, C; k, K; o, O; p, P; s, S; u, U; v, V; w, W; x, X; z, Z; e, l.

Please take care to distinguish them in some way.

Notations

preferred form	instead of
$7/8, (a+b)/c$	$\frac{7}{8}\ \frac{a+b}{c}$
$\exp(-(x^2+y^2)/a^2)$	$e^{-\frac{x^2+y^2}{a^2}}$
$\frac{\cos(1/x)}{(a+b/x)^{1/2}}$	$\frac{\cos\frac{1}{x}}{\sqrt{a+\frac{b}{x}}}$
$f: A \to B$	$A \xrightarrow{f} B$
sub lim, inf lim	$\underline{\lim}, \overline{\lim}$
inj lim, proj lim	$\varinjlim\ \varprojlim$

4. References

Literature citations in the text should be preferably by author(s) and year. Where there are more than two authors, only the first should be named, followed by "et al.".

The list of **References** should include only publications cited in the text. The references should be cited in alphabetical order under the first author's name, listing all authors (*surnames followed by initials throughout;* do *not* use "and") and the complete title (in English and French titles do *not* capitalize nouns, adjectives, etc.), according to the following rules and examples:

a) *Articles from journals and other serial publications:* author(s), title, series (abbreviated), volume followed by a comma, *inclusive* pages, year (in parentheses).

Example:
O'Leary, D.P., Segundo, J.P., Vidal, J.J.: Perturbation effects on stability of gravity receptors. Biol. Cybernetics **17**, 99—108 (1975)

b) *Articles from non-serial collective publications* (symposia volumes, encyclopedias, etc.); author(s), title followed by "In:", title of the volume and/or part (Vol., pt.) if appropriate, inclusive pages of article (pp.), name(s) of editor(s) followed by "ed(s).", publisher (place: name), year.

Example:
Hartline, H.K., Ratliff, F.: Inhibitory interactions in the retina of *Limulus.* In: Handbook of sensory physiology. Vol. VII/2, pp. 381—447. Fourtes, M.G.F., ed. Berlin-Heidelberg-New York: Springer 1972

c) *Books:* author(s), title, edition (edn.) if appropriate, publisher as under b), year.

Example:
Ledbetter, M.C., Porter, K.R.: Introduction to the fine structure of plant cells. Berlin-Heidelberg-New York: Springer 1970

5. Tables

Tables should be prepared consulting a recent issue of the journal and should be numbered consecutively with Arabic numerals. Footnotes in tables should be indicated by lower-case suffix letters, beginning with [a] in each table.

6. Illustrations

Illustrations should be limited to materials essential for the text, and line drawings should be used wherever possible. Double documentation of the same point in figures *and* tables is not acceptable.
Requests for color illustrations cannot be approved unless the author(s) agree to bear the costs.
Figures, tables and *diagrams* should be submitted on separate sheets and not incorporated into the text.
The figures should not extend beyond the column width (8.1 cm) or page width (16.8 cm). Print area: 16.8 cm × 22 cm. Several figures should be grouped into a plate on one page. For *line drawings,* sharp glossy prints in the desired final size are preferred. The inscriptions should be clearly legible. Letters 2 mm high are recommended. For *half-tone illustrations,* well-contrasted photographic prints, trimmed at right angles and in the desired printing size are essential. Inscriptions should be about 3 mm high. Parts of figures should be indicated by lower-case boldface letters (a, b, c). All figures must have legends, which should be submitted on a separate sheet. Illustrations taken from other publications should be accompanied by full information as to their source.

7. Offprints

Fifty (50) offprints of each paper will be supplied free of charge. Additional copies may be ordered at cost price when the proofs are returned.
Manuscripts must be accompanied by the full address of the author. Should the author change his address before the manuscript appears in print, he should in his own interest notify the publisher immediately.

Biol. Cybernetics 30, 187—193 (1978)

Biological Cybernetics

A Model for Light Adaptation: Producing Weber's Law with Bleaching-Type Kinetics

S. M. Dawis

Laboratory of Neurophysiology, University of Minnesota, Minneapolis, Minnesota, USA

Abstract. An "adaptation model" having two stages is introduced and its mathematical properties are examined. The two stages are the "adaptive process" (parameter K_b), which has bleaching-type kinetics, and the "response function" (parameters K_r and n), which incorporates response saturation. In order to study the increment threshold functions generated by the "adaptation model" the concept of a "detector" is required. It is demonstrated that without an adaptive process the compression hypothesis, in the form of the "difference equation", produces increment threshold functions which saturate and do not obey Weber's law. It is then shown that an adaptive process with bleaching-type kinetics can prevent saturation and produce Weber's law behavior provided that the "adaptive strength" of the system exceeds the "detector sensitivity".

1. Introduction

In recent years there has been much interest in the description of light adapted responses in vertebrate photoreceptors. In 1970 Boynton and Whitten demonstrated that an increment threshold function (ITF) generated from data on monkey foveal late receptor potentials was very similar to the psychophysical ITF for cone mechanisms obtained from human subjects. Specifically, the generated function obeyed Weber's law. This finding suggested to Boynton and Whitten that "a major part of the adaptive machinery is in the cones themselves". The idea that photoreceptors are a major site of visual adaptation is now shared by many researchers who have since then studied photoreceptors by intracellular recordings (Grabowski et al., 1972; Norman and Werblin, 1974; Kleinschmidt and Dowling, 1975)[1].

In the same paper Boynton and Whitten (1970) proposed a model to explain the observed responses to increment flashes superimposed on adapting background fields. Recently, Williams and Gale (1977) have questioned whether this model will generate Weber's law; they demonstrated that what they termed the additive equation certainly will not. However, it can be shown that the crucial component in the Boynton-Whitten model which prevents saturation and produces Weber's law behavior is photopigment bleaching. The Boynton-Whitten model has, therefore, the very attractive capability of producing Weber's law behavior provided that the effect of photopigment bleaching is strong enough. As Boynton and Whitten have shown, this may be the case for the cynomolgus macaque visual system.

Since the initial transduction process in all eyes involves the absorption of light quanta by photopigment molecules, the Boynton-Whitten model is, theoretically, applicable to all vertebrate retinae. However, photopigment bleaching and regeneration is not the only process of light and dark adaptation[2]. An adaptation which is more rapid and more powerful than photochemical adaptation has been observed to occur in photoreceptors (Frank, 1971; Grabowski et al., 1972; Dowling and Ripps, 1972; Witkovsky et al., 1973; Normann and Werblin, 1974; Kleinschmidt and Dowling, 1975). This rapid, powerful adaptation has often been called neural adaptation; however, in order to differentiate it from other forms of neural adap-

1 See MacLeod (1978) for a discussion on adaptation in receptors

2 A very common observation is that Weber's law begins at an intensity thousands of times less intense than that required to bleach photopigment (Dodt and Echte, 1961; Dowling, 1967; Weinstein et al., 1967; Frank, 1971; Dowling and Ripps, 1972; Witkovsky et al., 1973; Normann and Werblin, 1974). The beginning of Weber's law can probably be associated with the onset of light adaptation. These studies, therefore, demonstrate that the vertebrate eye possesses an adaptive process other than that provided by photopigment bleaching. See also Riggs (1971) and Dowling (1977)

0340-1200/78/0030/0187/$01.40

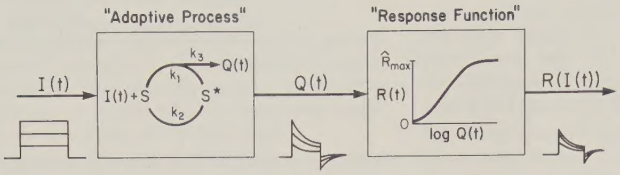

Fig. 1. A schematic diagram of the "adaptation model". The "adaptive process" has bleaching-type kinetics, and the "response function" exhibits response saturation. The important parameter for the "adaptive process" is K_b which is defined to be the ratio k_2/k_1; for the "response function" the important parameters are K_r and n. See the text for a detailed discussion about the parameters, variables, and behavior of the model

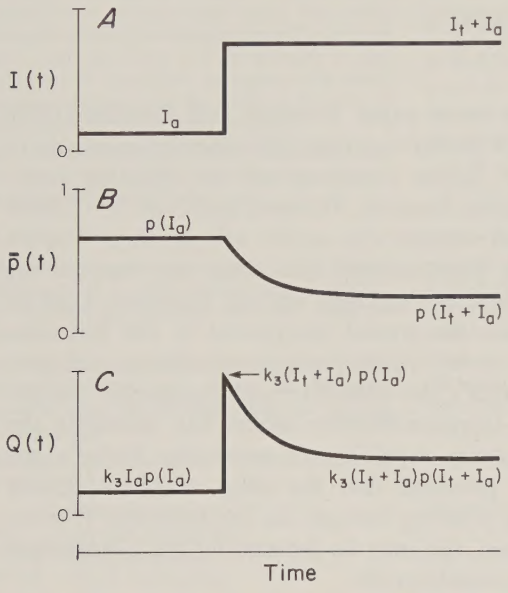

Fig. 2A–C. The effect of an increment step of intensity I_t superimposed on an adapting background of intensity I_a. **A** The time course of the light intensity $I(t)$. **B** The time course of the fraction of subsystems, $\bar{p}(t)$, in state S. Note that $\bar{p}(t)$ has a constant value of $p(I_a)$ before the stimulus onset; this is because it is assumed that the adaptation to I_a has reached equilibrium. After the onset of the step increment, $\bar{p}(t)$ declines to a new steady state value of $p(I_t+I_a)$. **C** The time course of the effect $Q(t)$ of the "adaptive process" on the "response function". $Q(t)$ exhibits an overshoot and adaptation when the increment step of light is applied. Therefore the bleaching-type process in the first stage of the "adaptation model" is an adaptive process

tation, it will be referred to as "photoreceptor adaptation"[3].

In this paper it will be shown that the capacity to avoid saturation and exhibit Weber's law behavior is a property of a general "adaptation model" in which the adaptive process has bleaching-type kinetics. Although one specific realization of this general model is the Boynton-Whitten model, it is *not* necessary to identify the bleaching-type kinetics with photopigment bleaching. The "adaptation model" can also be shown to describe some properties of "photoreceptor adaptation" in a vertebrate cone system (Dawis, in preparation). In such a case Weber's law behavior can still be anticipated from the properties of the "adaptation model". Given this general applicability of the "adaptation model" it becomes important to understand how it behaves.

2. The "Adaptation Model"

The "adaptation model" considered in this paper has two stages: 1) the "adaptive process" which has bleaching-type kinetics, and 2) the "response function" which displays response saturation. The model is shown schematically in Figure 1.

2.1. The "*Adaptive Process*"

Consider a system composed of many identical subsystems. Suppose that each subsystem can exist in two states, S and S^*. Defining $\bar{p}(t)$ as the fraction of subsystems which are in state S at time t, assume that the differential equation which governs the distribution of the subsystems between states S and S^* is given by

$$\frac{d\bar{p}(t)}{dt} = -[k_1 I(t) + k_2]\,\bar{p}(t) + k_2, \tag{1}$$

where k_1 and k_2 are rate constants. Let us also assume that the effect, $Q(t)$, of the "adaptive process" on the response function is proportional to the rate at which subsystems are transformed from state S to state S^*;

$$Q(t) = k_3 I(t)\,\bar{p}(t), \tag{2}$$

where k_3 is a constant of proportionality. The "adaptive process" is defined by (1) and (2). The reader may have recognized that (1) is identical in form to the bleaching equation (cf., Rushton and Henry, 1968) and that $Q(t)$ is analogous to quantal capture rate. As shown in Figure 2 this bleaching-type kinetics does provide adaptation [4].

2.2. The "*Response Function*"

The second stage of the "adaptation model" is the "response function". Notice in Figure 2C that the peak

3 It is thought that adaptation (or summation) pools in the retina are an important stage of visual adaptation (Rushton, 1965; Tong and Green, 1977). This adaptation, being independent of photochemical kinetics, is another form of neural adaptation. Since this neural adaptation occurs at a site proximal to the receptors, Dowling (1977) suggests using the term "network" adaptation for this process

4 The function p is defined by $p(x) \triangleq K_b/(x+K_b)$ where $K_b \triangleq k_2/k_1$

change of the effect $Q(t)$ is given by $k_3 I_t p(I_a)$; that is, the peak change of the effect of the "adaptive process" is proportional to the test stimulus I_t. However, what is usually observed is a saturation of the response. Therefore let us suppose that the response, $R(I(t))$, of the model is given by

$$R(I(t)) = \frac{Q(t)^n}{Q(t)^n + K_r^n} \hat{R}_{\max}, \tag{3}$$

where n and K_r are positive valued parameters and $\hat{R}_{\max}$ can be positive or negative depending on whether the response is one of depolarization or hyperpolarization, respectively. The "adaptation model" is defined by (1), (2), and (3).

2.3. The "*Adaptation Model*" and the Experiment

The experimental protocol commonly used to study light adapted responses is to adapt the system to light of intensity I_a and then to present a test flash of intensity I_t. The parameter commonly measured is the peak response to the incremental flash. Under these conditions one can see from Figure 2C and (3) that the normalized peak response is given by

$$\frac{\hat{R}(I_t|I_a)}{\hat{R}_{\max}} = \frac{(I_t + I_a)^n p_a^n}{(I_t + I_a)^n p_a^n + K_r^n} - \frac{I_a^n p_a^n}{I_a^n p_a^n + K_r^n}, \tag{4}$$

where $p_a \triangleq p(I_a)$, the units are chosen such that $k_3 = 1$, and $\hat{R}(I_t|I_a)$ is the peak response to a test flash of intensity I_t presented on an adapting background of intensity I_a[5]. One should realize that (4) is formally equivalent to the model proposed by Boynton and Whitten (1970). However, it is important to realize that the present model does not identify the "adaptive process" with photopigment bleaching; all that is required is that the "adaptive process" have bleaching-type kinetics.

2.4. The Dark Adapted Response Function

Let us examine the special case when the system is dark adapted, that is, when $I_a = 0$. Equation (4) then becomes

$$\frac{\hat{R}(I_t|0)}{\hat{R}_{\max}} = \frac{I_t^n}{I_t^n + K_r^n}. \tag{5}$$

This generalized Naka and Rushton (1966) response function is frequently used to describe dark adapted responses. The parameters $\hat{R}_{\max}$ and K_r now take on familiar meanings: $\hat{R}_{\max}$ is the maximal peak response under dark adapted conditions, and the value of K_r is equal to the test intensity which will evoke a half maximal response under dark adapted conditions[6]. The response described by (5) exhibits response saturation.

2.5. The Compression Hypothesis

The primary purpose of the "adaptation model" is to describe light adapted responses. To emphasize the importance of the "adaptive process" in this description let us consider the case when adaptation is absent. We can effect this by setting $k_1 = 0$ in (1). The result is that $\bar{p}(t)$ will always have a value of 1; consequently, (4) becomes the "difference equation"

$$\left.\frac{\hat{R}(I_t|I_a)}{\hat{R}_{\max}}\right|_{\text{no adaptation}} = \frac{(I_t + I_a)^n}{(I_t + I_a)^n + K_r^n} - \frac{I_a^n}{I_a^n + K_r^n}. \tag{6}$$

By substituting (5) into (6) we obtain

$$\hat{R}(I_t|I_a)|_{\text{no adaptation}} = \hat{R}(I_t + I_a|0) - \hat{R}(I_a|0).$$

This statement has been referred to as the "compression" hypothesis (Boynton and Whitten, 1970; Dowling and Ripps, 1972; Normann and Werblin, 1974).

It is important to realize that the ITFs generated by the "difference equation" saturate and do not obey Weber's law. For the case when $n = 1$ Williams and Gale (1977) prove that regions are produced with slopes of 2 rather than slopes, as prescribed by Weber's law, of 1. A more rigorous proof reveals that the region of slope 2 will increase if the criterion response for threshold detection is decreased[7]. This is shown in Figure 3. For the general case, $n > 0$, it can be shown that if the criterion response is chosen small enough then the ITF will approach a slope of $n+1$ over a certain region. This is illustrated in Figure 4. The compression hypothesis by itself will not prevent saturation nor will it provide Weber's law behavior[8]; what is required is an adaptive process.

5 The objective of the "adaptation model", therefore, is not to reproduce the precise temporal structure of the response to light but to describe the effect of adapting light on an experimentally measured parameter, the peak response. Since the "adaptation model" restricts its view to response peaks it is incapable of dealing with such phenomena as integration time (see MacLeod, 1978)

6 The value for n is species dependent, usually falling in a range from 0.7 to 1.0. The value also depends on the response which is measured (van Norren and Baron, 1977). Normann and Werblin (1974) have suggested that the duration of the test flash may also be a factor in determining n

7 All mathematical proofs mentioned in this paper are available upon request and will be contained in the doctoral thesis of the author

8 Dowling and Ripps (1972), for example, constructed an ITF for the skate rod photoreceptor which was unsaturated and obeyed Weber's law. On the basis of this, one would predict that the compression hypothesis cannot describe the observed light adapted responses. Dowling and Ripps demonstrated in their Figure 12 that this was indeed the case

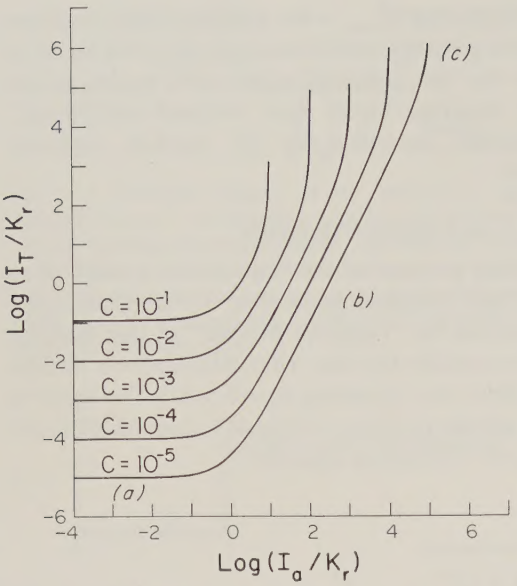

Fig. 3. ITFs produced by the "difference equation" when $n=1$. Although there are two parameters, K_r and C, the shapes of the ITFs are determined by the parameter C alone. The ITFs exhibit plateaus in segment (a) where I_T approaches a constant value of $CK_r/(1-C)$ and saturate at $I_a=(1-C)K_r/C$ [segment (c)]. Furthermore, as C is chosen smaller and smaller an intermediate segment (b), where the slope approaches 2, expands

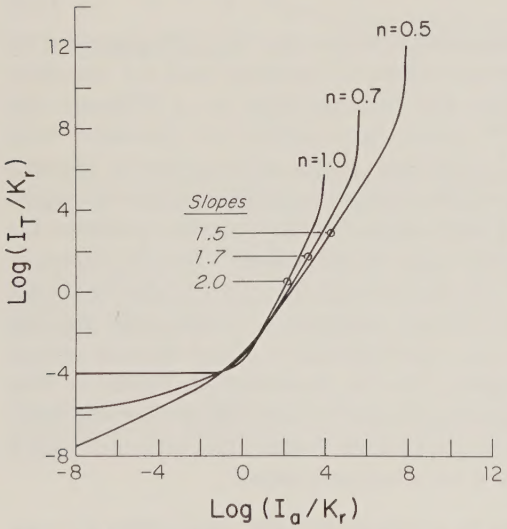

Fig. 4. ITFs generated by the "difference equation" with $n=0.5$, 0.7, and 1.0. A small value (10^{-4}) was selected for the criterion response, C, in order to provide regions having slopes of $n+1$

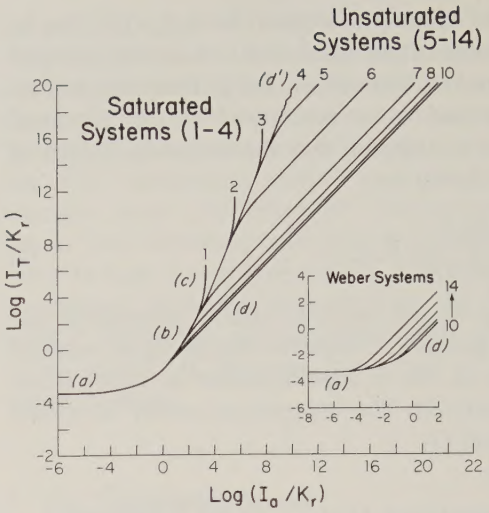

Fig. 5. A "(C,n) family" of ITFs generated by the "adaptation model". Values $C=0.005$ and $n=0.7$ are chosen, and ITFs corresponding to different K_b/K_r ratios are shown: (Curve *1*; $K_b/K_r=\infty$), (*2*; 1933), (*3*; 1923.5), (*4*; 1923.4274), (*5*; 1923.42), (*6*; 1920), (*7*; 1000), (*8*; 100), (*9*; 10), (*10*; 0.98581572), (*11*; 0.1), (*12*; 0.01), (*13*; 0.001), (*14*; 0.0001). The shapes of the ITFs are determined by C, n, and the ratio K_b/K_r. Curve *4*, a saturated system, is very close to the boundary between saturated and unsaturated systems

2.6. Properties of ITFs *Generated by the "Adaptation Model"*

An adaptive process with bleaching-type kinetics can prevent the ITF from saturating and provide Weber's law behavior. There are two possible types of systems which result from the "adaptation model". There are unsaturated systems which produce ITFs that have plateaus at low adapting intensities and approach slopes of 1 at higher adapting intensities. In Figure 5 the plateau and Weber's law behavior of an unsaturated system are shown in segments (a) and (d), respectively. There also are saturated systems. The ITFs produced by saturated systems have plateaus and regions of saturation; these two characteristics of saturated systems are illustrated in Figure 5 by segments (a) and (c), respectively. As we will see, it is the "adaptive process" which makes Weber's law behavior possible.

In order to examine the mathematical properties of the ITFs generated by the "adaptation model" we must first determine how the threshold increment I_T depends on the adapting intensity I_a. We begin with (4) since it describes the behavior of the "adaptation model" under the conditions used in obtaining an ITF. To begin, we must introduce the concept of a "detector". Assume that for threshold detection $\hat{R}(I_T|I_a)/\hat{R}_{\max}=C$, where C is a constant such that $0<C<1$. The constant C can be conveniently interpreted as the criterion response (normalized by $\hat{R}_{\max}$). Substituting (4) into the above equation and then solving for I_T gives

$$I_T=\frac{K_r}{p_a}\left[\frac{C+f(I_a p_a)}{1-C-f(I_a p_a)}\right]^{1/n}-I_a, \tag{7}$$

where the function f is defined by $f(x) \triangleq x^n/(x^n+K_r^n)$ with $x \geqq 0$ and $n>0$. The ITFs generated by the "adaptation model" are obtained by plotting $\log(I_T)$ as a function of $\log(I_a)$ using (7). The family of ITFs shown in Figure 5 was so generated.

2.6.1. Saturated and Unsaturated Systems. The Effect of "Adaptive Strength".

From a consideration of limits the following result can be obtained:

The "adaptation model" can produce two types of systems:

Unsaturated systems, when parameters are chosen such that

$$(K_r/K_b)^n > C/(1-C), \tag{8}$$

or saturated systems, when parameters are chosen such that

$$(K_r/K_b)^n < C/(1-C). \tag{9}$$

In the regime of the "adaptation model", let us define the "adaptive strength" of a system to be the value $(K_r/K_b)^n$. This definition will help us to verbalize the mathematical findings. For example, one might define the "detector sensitivity" as the value $C/(1-C)$; in that case, (8) states that the "adaptive strength" of the system must exceed the "detector sensitivity" in order to prevent the ITF from saturating.

For a saturated system the saturation point is given by

$$I_a = \frac{K_b K_r((1-C)/C)^{1/n}}{K_b - K_r((1-C)/C)^{1/n}}. \tag{10}$$

2.6.2. The Plateau Region of the ITF. Both saturated and unsaturated systems have plateaus at low levels of adapting intensities. This can be seen from examining I_T for small I_a. From (7) we obtain for the plateau value

$$\lim_{I_a \to 0} I_T = K_r(C/(1-C))^{1/n}. \tag{11}$$

It is interesting to note that the parameter K_b is not involved in the plateau value. This means that the plateau value is unaffected by the "adaptive strength" of the system (see Fig. 5). It is not necessary with the "adaptation model" to invoke the concept of dark light (Barlow, 1972) in order to account for the plateau region of the ITF (Boynton and Whitten, 1970).

2.6.3. Weber's Law Behavior in the ITF. Let us examine under what conditions Weber's law is approximated. For an unsaturated system, we have from (7)

$$\lim_{I_a \to \infty} (I_T/I_a) = k,$$

where the constant k is defined as

$$k \triangleq \frac{K_r}{K_b}\left[\frac{C+f(K_b)}{1-C-f(K_b)}\right]^{1/n} - 1. \tag{12}$$

Therefore as I_a becomes large we have

$$\log(I_T) \approx \log(I_a) + \log(k),$$

and the slope of the ITF approaches 1. Thus we see that at high adapting intensities an unsaturated system approximates Weber's law[9]. Those familiar with Weber's law will recognize that the constant k is for all practical purposes the Weber constant. This gives an importance to (12) since this equation relates all the parameters of the "adaptation model" to the Weber constant.

For saturated systems one can prove that Weber's law can be satisfied, but only under circumstances so restrictive that it is unlikely to occur[10]. For the saturated systems considered in Figure 5 the only ITF displaying Weber's law behavior is Curve *4*, and even then only the faintest hint of Weber's law appears in segment (*d'*).

3. Discussion

The behavior of the ITFs generated by the "adaptation model" results from an interaction of two underlying characters. One character is due to the compression hypothesis which produces saturation and, if the criterion for threshold detection is small enough, a region of slope $n+1$. The other character is provided by the "adaptive process" which tends to prevent saturation

9 As a specific realization of the "adaptation model" the model proposed by Boynton and Whitten (1970) has the same general properties. Therefore their model can produce an unsaturated ITF which obeys Weber's law. Boynton and Whitten (1970) measured the values $n=0.07$ and $K_r^n = 10000^{0.7}$ (there is a typographical error concerning K_r^n throughout the article; however, by examining their Figure 2, one can determine what they had intended). In addition, they decided that a reasonable value for K_b was 10000 trolands and for C, 0.005. With these parameter values $(K_r/K_b)^n = 1$ and $C/(1-C) = 0.00502512$; so condition (8) for an unsaturated system is well satisfied. The ITF is essentially given by Curve *10* in Figure 5

10 Because of this behavior the "adaptation model" has to be modified if it is to describe the rod ITF which exhibits both Weber's law behavior and saturation (Aquilar and Stiles, 1954; Daw and Pearlman, 1971; Green, 1973; Kleinschmidt and Dowling, 1975). A modification may be justified since in nature it appears that the rod ITF is somehow modified by the presence of a cone system (Daw and Pearlman, 1971; Dowling and Ripps, 1972)

and produce a slope of 1[11]. The balance between the two characters depends on the "adaptive strength" of the system. When the "adaptive strength" of a system falls short of the "detector sensitivity" [condition (9)] the character provided by the compression hypothesis dominates; when it exceeds the "detector sensitivity" [condition (8)] the character provided by the "adaptive process" dominates.

One can see in Figure 5 the transition occurring in the character of the ITF as the "adaptive strength" is varied. Curve *1* is the ITF generated by an "adaptive strength" of zero and therefore displays the behavior expected from the compression hypothesis. Since n was chosen to be 0.7 and the criterion response is small, a slope of 1.7 occurs in segment (b). Re-emphasizing an earlier point, we see that without an adaptive process the compression hypothesis produces an ITF which saturates and does not obey Weber's law. As the "adaptive strength" is increased (K_b/K_r decreased) we see from Curves *2* and *3* that the saturation point moves to higher levels of adapting intensities. Eventually a slope of 1 appears (Curve *4*); however, virtually as soon as this occurs the saturation disappears. As the "adaptive strength" is increased further, the region of slope 1 grows and eventually supplants the region of slope 1.7. This is shown in Curves *7* through *10* (see Footnote[12] about the behavior exhibited by Curves *5* and *6*). Then a point is reached when an ITF is produced which has a very simple form. For the parameters chosen this occurs when the K_b/K_r ratio attains a value of 0.98581572 (approximately 1); the corresponding ITF in Figure 5 is Curve *10*. This ITF exhibits a plateau at low adapting intensities and Weber's law behavior at high intensities. Let us refer to any system which produces an ITF of this simple form as a "Weber system". This simple form in the ITF is retained with further increases in the "adaptive strength". This is illustrated by Curves *10* through *14* shown in the inset of Figure 5. By the above terminology the systems which generated these curves are "Weber systems".

It should be clear from the preceding analysis and discussion that it is the "adaptive process" which prevents the ITFs generated by the "adaptation model" from saturating. It is also the "adaptive process" which makes it possible for the "adaptation model" to obey Weber's law. In the regime of the "adaptation model" a system must have a strong "adaptive process" in order to be unsaturated and exhibit Weber's law behavior.

In the "adaptation model" it is assumed that the "adaptive process" has bleaching-type kinetics. Since it remains, otherwise, unspecified the "adaptation model" is a descriptive model. When it is possible to specify more precisely the "adaptive process" the model becomes an explanatory model. A case in point is the Boynton-Whitten model in which the "adaptive process" is identified with photopigment bleaching. This flexibility is very important in light of the observations in many vertebrate photoreceptors of a strong adaptive process which could not have been due to photopigment bleaching. Although the Boynton-Whitten model cannot explain the light adapted responses observed in these photoreceptors the more general "adaptation model" might still be applicable. If such is the case then an adaptive process with bleaching-type kinetics can be used to describe the observed light adapted behavior. There are indications that the "adaptation model" is more than just a curve fitting model. It can be shown that for the model there is an optimal "adaptive strength"; I have found that the visual systems which can be described by the "adaptation model" are very near to the theoretical optimum, as might be expected if selective pressures for intensity discrimination have been important in the evolution of vertebrate visual systems.

11 The "adaptation model" can produce regions in ITFs having slopes of 0 (the plateau region), 1 (Weber's law) and $n+1$ where $n>0$ (compressive behavior). The model, however, has difficulty reproducing the slopes of 1/2 (deVries-Rose law) in human psychophysical ITFs obtained when small diameter test flashes are presented under "high" scotopic to mesopic conditions. If receptoral mechanisms are to be invoked to explain this psychophysical phenomenon, the "adaptation model" must be modified. Using a different model Barlow (1957) found that the deVries-Rose law could be anticipated when the effects of quantum fluctuations and background noise ("dark light") were considered. Furthermore, by varying the quantum/spike gain Barlow (1965) was able to generate Weber's law. On the other hand, "network" mechanisms rather than receptoral mechanisms might be involved. MacLeod (1978) gives an argument that responses evoked by small diameter stimuli may reflect effects of lateral interactions as well as receptor function, and Dowling (1977) provides evidence that in the later stages of the visual system it is not until high background intensities are attained that receptoral mechanisms determine ITF behavior

12 There is a significant region above segment (b) in Figure 5 which has a slope of 2.4. This region is gradually replaced in Curves 5 through 7 with a Weber's law region. I was unable to determine a mathematical formula for this slope; it seems that the "adaptation model" may also be capable of producing regions in the ITF having slopes of $2n+1$

Acknowledgments. I wish to thank Dr. Richard Purple for reading the many versions of this paper and for his helpful discussions. I also thank Dr. Dwight Burkhardt for providing me with a better perspective of the "adaptation model" in the field of visual adaptation. I am very grateful to Bev Murphey who persevered in the typing of this manuscript. This work was supported by U.S. Public Health Service Grant No. EY002930. Computer facilities were made available by Air Force Office of Scientific Research Grant AFOSR 75–2804. The author was supported by U.S. Public Health Service Training Grant GM00572.

References

Aquilar, M., Stiles, W.S.: Saturation of the rod mechanism at high levels of stimulation. Optica Acta **1**, 59–65 (1954)

Barlow, H.B.: Increment thresholds at low intensities considered as signal/noise discrimination. J. Physiol. **136**, 469–488 (1957)

Barlow, H.B.: Optic nerve impulses and Weber's law. Cold Spring Harbor Symposia on Quantitative Biology **30**, 539–546 (1965)

Barlow, H.B.: Dark and light adaptation: Psychophysics. In: Handbook of sensory physiology, Vol. VII/4, pp. 5–9. Jameson, D., Hurvich, L.M. ed. Berlin, Heidelberg, New York: Springer 1972

Boynton, R.M., Whitten, D.N.: Visual adaptation in monkey cones: recordings of late receptor potentials. Science **170**, 1423–1426 (1970)

Daw, N., Pearlman, A.L.: Rod saturation in the cat. Vision Res. **11**, 1361–1364 (1971)

Dodt, E., Echte, K.: Dark and light adaptation in pigmented and white rat as measured by electroretinogram threshold. J. Neurophysiol. **24**, 427–445 (1961)

Dowling, J.E.: The site of visual adaptation. Science **155**, 273–279 (1967)

Dowling, J.E.: Receptoral and network mechanisms of visual adaptation. Neurosci. Res. Progr. Bull. **15**, 397–407 (1977)

Dowling, J.E., Ripps, H.: Adaptation in skate photoreceptors. J. Gen. Physiol. **60**, 698–719 (1972)

Frank, R.N.: Properties of "neural" adaptation in components of the frog electroretinogram. Vision Res. **11**, 1113–1123 (1971)

Grabowski, S.R., Pinto, L.H., Pak, W.L.: Adaptation in retinal rods of Axolotl: intracellular recordings. Science **176**, 1240–1243 (1972)

Green, D.G.: Scotopic and photopic components of the rat electroretinogram. J. of Physiol. **228**, 781–797 (1973)

Kleinschmidt, J., Dowling, J.E.: Intracellular recordings from Gecko photoreceptors during light and dark adaptation. J. of Gen. Physiol. **66**, 617–648 (1975)

MacLeod, D.I.A.: Visual sensitivity. Ann. Rev. Psychol. **29**, 613–645 (1978)

Naka, K.I., Rushton, W.A.H.: S-potentials from colour units in the retina of fish (*Cyprinidae*). J. of Physiol. **185**, 536–555 (1966)

Normann, R.A., Werblin, F.S.: Control of retinal sensitivity. I. Light and dark adaptation of vertebrate rods and cones. J. Gen. Physiol. **63**, 37–61 (1974)

Norren, D. van, Baron, W.S.: Increment spectral sensitivities of the primate late receptor potential and b-wave. Vision Res. **17**, 807–810 (1977)

Riggs, L.A.: Vision. In: Woodworth and Schloberg's experimental psychology, Third Edt., Chap. 9, pp. 286–289. Kling, J.W., Riggs, L.A., ed. New York: Holt, Rinehart, and Winston, 1971

Rushton, W.A.H.: The Ferrier lecture, 1962. Visual adaptation. Proc. R. Soc. B**162**, 20–46 (1965)

Rushton, W.A.H., Henry, G.H.: Bleaching and regeneration of cone pigments in man. Vision Res. **8**, 617–631 (1968)

Tong, L., Green, D.G.: Adaptation pools and excitation receptive fields of rat retinal ganglion cells. Vision Res. **17**, 1233–1236 (1977)

Weinstein, G.W., Hobson, R.R., Dowling, J.E.: Light and dark adaptation in the isolated rat retina. Nature **215**, 134–138 (1967)

Williams, T.P., Gale, J.G.: A critique of an incremental threshold function. Vision Res. **17**, 881–882 (1977)

Witkovsky, P., Nelson, J., Ripps, H.: Action spectra and adaptation properties of carp photoreceptors. J. Gen. Physiol. **61**, 401–423 (1973)

Received: April 11, 1978

S. M. Dawis
Dept. of Physiology
424 Millard Hall
Univ. of Minnesota
Minneapolis, MN 55455, USA

Biol. Cybernetics 30, 195—208 (1978)

Biological Cybernetics

On the Interaction between the Central Nervous System and the Peripheral Motor System

J. H. M. van Dijk

Department of Medical and Physiological Physics, University of Utrecht, The Netherlands

Abstract. The problem of the control of voluntary human movements is considered from a cybernetic point of view. The human motor system is considered to be divided into a central part and a peripheral part. The peripheral part is relatively well known and may be regarded as a set of subsystems with well known input-output relations. The interaction between the peripheral part and the central part is related to the mechanisms of the peripheral motor part. With regard to the central part two different types of control mechanisms are possible, a) an intricate functioning of the central part which generates the control signals with regard to internal and external dynamical factors, b) the central part has some degree of independence with respect to the dynamics of the peripheral motor part. In the latter case the central part prescribes the desired movement exactly, but the final performance of the movement is also brought about by peripheral feedback mechanisms. As a functional form of the interaction between the central part and the peripheral part it might be that the control signals are encoded in a way that is related to the muscle lengths.

1. Introduction

In research on the control of voluntary human movements five types of approach may be distinguished:

Neuro-anatomical research, i.e. the investigation of parts of the brain that are involved in the control of voluntary movements. This investigation deals with the afferent and efferent pathways of the brain parts, the neuronal organization of the brain parts, the organization of the afferent and efferent pathways (for instance a somatotopical organization), the information that is conveyed by the signals, and the operations effected on the input signals by the brain parts. The operations on the input signals are related to the organization of the afferent pathways and the type of information that is mediated. General operations can include: convergence or divergence of signals, summation, occlusion, spatial contrasting, differentiation of signals, transformation of signals to another code, etc. (Brooks et al., 1971).

Psychological research. Typical for this approach is the focusing on higher functions of the brain. The consequence of this approach is that less attention is paid to the physiological mechanisms and neuro-anatomical networks underlying the motor control and the details of the execution of the movements. The execution of a motor task is generally described as a stimulus-response scheme: desired movement→selection of the appropriate motor programme→motor performance. Aspects considered in this approach include various reaction, decision and calculation times, (learned) motor or action programmes for the execution of particular movements, anticipation of future events, the cues that are relevant for the execution of different types of movement, the evaluation of movements, the working of the memory that is involved in the evaluation of the movements and in the learning and execution of the stimulus-response schemes, etc. (Welford, 1968; Rabbitt, 1974; Stelmach, 1976).

Physiological approach. This approach focuses attention on elementary subsystems that can be distinguished in the performance of the peripheral motor system, i.e. detailed research on the properties of muscles (Huxley, 1974), muscle spindles (Cheney et al., 1976), tendon organs (Houk et al., 1967), the spinal reflex organization (Gottlieb et al., 1972), etc. By linking these separate physiological elements one obtains a model of the peripheral part of the motor system. From a cybernetic point of view special attention should be given to the input-output relations of the physiological elements.

Movement analysis. The primary aim of this research is the registration of the parameters of different types of movement. One can distinguish between var-

0340-1200/78/0030/0195/$02.80

ious aspects of movement, for example arbitrary, routine or following movements, the tempo of the movement, the limb that performs the movement, the relevant parameters that describe the movement, the type of the motor task, etc. In essence this research is pure phenomenological, i.e. it is concerned with establishing the relationship between the type and the parameters of the movement (Bernstein, 1967). Because muscle activity can also be measured, there is a direct link with physiological research and especially with the coordination of the elementary subsystems in the performance of a movement. There is also a link with the psychological approach, because movement analysis and psychological research both focus attention on overall aspects of motor control: input→output, motor task→movement, stimulus→response.

Cybernetic analysis. In this approach the experimental data and information on the first four fields of interest are interpreted from a cybernetic angle (Gawronski, 1971). The whole motor system may be considered as a set of (sub-) systems with certain input-output relations and with certain relations (interactions) between the systems. A distinction may be made between a central part and a peripheral part. The peripheral part consist of the limbs and the lower control centres (i.e. the motorneuronal pools). Higher control centres belong to the central part (i.e. the central nervous system, CNS). The peripheral motor system is relatively well known and therefore the coordination or control of the peripheral part, involving the integrated action of many subsystems, may be a link with the control activity of the CNS that is needed in the performance of movements. In the interaction between the central and the peripheral part the coding of the afferent and efferent signals, i.e. what information is mediated and in what form it is encoded, is of primary interest. This interaction between the central and the peripheral part is the subject of the present study.

2. Cybernetic Analysis of the Control of Voluntary Movements

2.1. The Link between the Central and the Peripheral Part

The peripheral part of the motor system is relatively well known as far as input-output relations of the composing elements are concerned. The coordination or control of the peripheral motor system by the CNS depends on the possibilities and constraints of the peripheral part, i.e. the control has to match the properties of the peripheral motor part. As a consequence, the physiological properties of the elementary subsystems and of the peripheral mechanisms and their possible functions in the performance of a movement give some indication of the way in which the CNS controls the peripheral motor system. So, a first step in a cybernetic approach to the control of voluntary movements may be a sort of inventory of possibilities, elementary subsystems and mechanisms in the peripheral motor system. The cybernetic aspects of the peripheral motor system can be divided as follows:

Control variables, i.e. variables in the peripheral motor system that can be affected by the CNS; for instance the excitation or inhibition of the motorneurons, the adjustment or regulation of the reflex organization (i.e. the regulation of the static, dynamic and force feedback gains, the regulation of the neuronal couplings between the motorneurons of different muscles), the regulation of the lengths of the muscle spindles, etc.

Feedback of movement information. Information about the course of the movement is sent back to preceding control centres via internal (proprioceptive signals) and external (visual image) channels. Feedback signals carry information about the muscle lengths and their derivatives, the muscle tension, and about the position of the joints, etc.

The limb mechanics, i.e. the mechanical properties of the limb and the link between the limb and the muscles and receptors; for instance the link between the muscle forces and the moment contributions acting upon the limb segments; the link between the length l_i and force F_i of a muscle i and the position $\boldsymbol{p}$ of the limb:

$$(l_i, F_i) \leftrightarrow \boldsymbol{p} . \tag{1}$$

The properties of the subsystems, i.e. the properties of the physiological elements such as the force-velocity relationship, the force-length relationship and the properties of the excitation-contraction coupling of the muscles regarded as input-output relations. The muscles may be considered to be systems with at least three inputs, namely the neuronal excitation, the muscle length and the contraction velocity and one output, namely the muscle force (Fig. 3).

Peripheral mechanisms, i.e. properties that may arise from the coordinated cooperation of elementary subsystems. This cooperation may be for instance various types of feedback of intermediary information, i.e. aspects of the results of the effector systems such as are schematically shown in Figure 1; the regulation of damping in the limb by simultaneous activation of antagonistic muscles; the (possible) derivative components in the control signals that may arise in the motorneuronal pools as a result of the Renshaw inhibition; the recruitment of different types of muscle fibres on account of Renshaw inhibition (Scheibel et

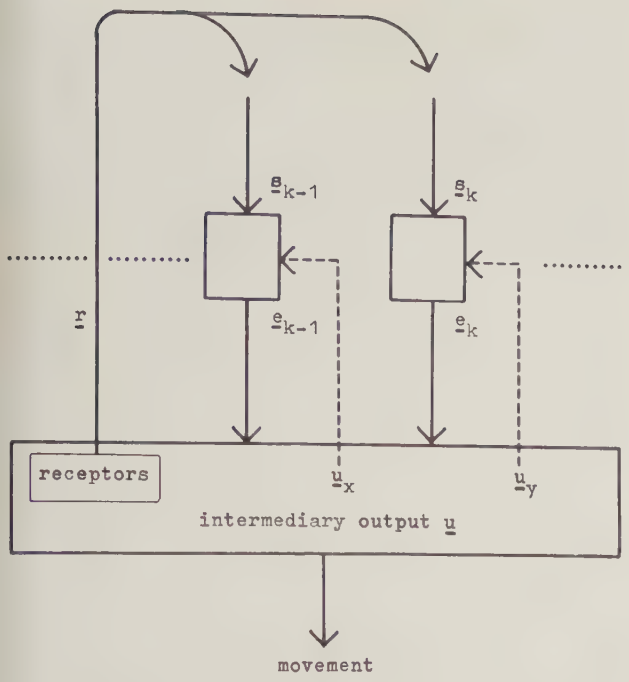

Fig. 1. Schematic representation of the various types of feedback. s_k and e_k are the respective input and output signals of the k^{th} effector subsystem. Some components u_x, u_y of the intermediary output (for instance muscle lengths, contraction velocities, and muscle tensions) influence the effector subsystems and may be interpreted as intrinsic feedback mechanisms. Some components r of the intermediary output are detected by internal (proprioceptive) receptors and are relayed to control centres at various levels, where this feedback information may be compared with control signals from the CNS, whereupon the resulting signals may be fed back to the effector system

al., 1971) or on account of supraspinal and muscle spindle signals (Burke et al., 1976).

2.2. *The Control of Voluntary Movements by the* CNS

The number of different movements that may be performed is practically unlimited and, furthermore, the dynamics of the peripheral motor system are very intricate and vary with different external influences. This complexity of the peripheral motor system can be matched by the CNS in two different ways:

1) The CNS generates the activation patterns of the muscles for each movement, taking account of the internal dynamics of the limb and the external load. This calculation of the control signals or the selection of the appropriate motor programme from the memory will be very intricate, because of the complex dynamics of the peripheral motor system. A control theory in this case would imply that mechanisms for the calculation or selection of the pattern of control signals for each particular movement in given external circumstances would have to be constructed. This would lead to a very cumbersome control theory, which would be difficult to assay because of the intricate relation with the neuro-anatomical substratum underlying these mechanisms.

2) The control by the CNS may have some degree of independence with respect to the dynamics of the peripheral motor system. In this case the patterns of control signals that are generated by the CNS have a broad form, implying that the peripheral part has to complete the motor performance. This method of cooperation would mean that the CNS prescribes the desired movement exactly, but that the final realization of the movement is also brought about by peripheral feedback mechanisms, which are closely related to the dynamics of the peripheral motor system. The more or less independent functioning of the CNS implies that various types of feedback are necessary, for instance a servo feedback, in which case the CNS has only to determine the norm (i.e. the parameters of the desired movement) without any need for detailed knowledge of the dynamics of the peripheral motor system and interaction with the environment (Fig. 2). Because the CNS has only to determine the desired movement and not the precise pattern of activation signals in order to cope with the internal dynamics and the external load, this type of central control is relatively simple. In the first instance it is necessary to investigate what kind of independent functioning the CNS may have in relation to the performance of the peripheral motor system (see Section 3). It is possible to distinguish two types of feedback that may be involved in this type of central control:

a) A servo type of control in which information about the movement is sent back to higher control levels in order to compare aspects of the real movement with aspects of the desired movement. Feedback information and control signals have to be comparable, i.e. the coding of the control signals has to match the coding of the feedback information and the organization of the effectors in the peripheral motor system. In Figure 2 the detailed activation signals $f(t)$ to the muscles are produced by the control signals $s(t)$ and the feedback signals $r(t)$. The control signals $s(t)$ only have to determine the desired movement and may have some degree of independence with respect to the dynamics of the peripheral motor system. $s(t)$ may contain components that can cope with some dynamical properties of the peripheral motor system and the influence of the environment, but whenever the response deviates from the desired movement the feedback signals will elicit a reaction to this deviation.

b) Intrinsic feedback mechanisms. These mechanisms may consist of intrinsic feedback loops of information; for instance the degree of activation of a muscle depends on the contraction velocity, so that the

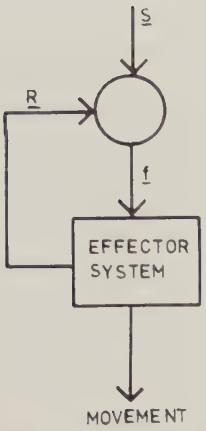

Fig. 2. Servo control diagram, showing that the control signals $f(t)$ to the effector system result from the central control signals $s(t)$ and the feedback signals $r(t)$ coming from the effector system. The control signals $s(t)$ may only contain information about the parameters of the desired movement, and may be rather independent with respect to the dynamics of the effector system. Deviation from the desired movement will cause a mismatch of the control signals $s(t)$ with the feedback signals $r(t)$. The less the control signals account for the dynamics of the effector system, the more is required from the capacity of the feedback mechanisms, in order to react to the deviations from the desired movement

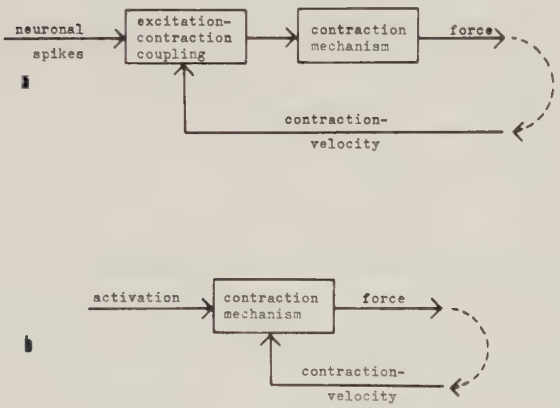

Fig. 3.a The excitation-contraction coupling is driven by neuronal input, but depends on the contraction velocity as well. As a consequence, the excitation-contraction coupling may be considered as an intrinsic feedback loop. **b** The contraction mechanism is driven by the activation from the excitation-contraction mechanism, but depends on the contraction velocity as well. As a consequence of this intrinsic property, the contraction mechanism may be regarded as an intrinsic feedback property

final activation is a result of the neuronal excitation (norm signal) and the contraction velocity (feedback signal); or the feedback may consist of intrinsic properties of the effectors, i.e. the working of the effectors depends on certain aspects of the overall or intermediary output. Both the intrinsic feedback loop and the output-dependent intrinsic properties may have the quality of a negative feedback (see Fig. 3 and Section 4.5), and as a consequence they are mechanisms that can cope with external and internal disturbances and dynamics (for instance the external load and the limb mechanics). The CNS may function in an independent or abstract way owing to the various feedback mechanisms. The feedback mechanisms may act at all levels of control and performance.

3. Control of Movement based on the Length Principle

3.1. The Hypothetical Relation between Movement, Muscle Length, and Muscle Activity (the Length Principle)

The independence of the CNS with respect to the dynamics of the peripheral motor system may have the form of general abstract codings of spatial positions and trajectories, such as Cartesian coordinates $(x, y, z)_t$ or joint angles $(\theta, \varphi, \ldots)_t$. The form of the abstraction is closely related to the interaction between the central and the peripheral part, i.e. it is related to the control variables, the feedback information from the periphery, and the organization of the peripheral motor system. The form of the abstraction determines in which form the control signals from the CNS to the peripheral motor system are encoded. A functional abstraction has to match the organization of the peripheral motor system. In order to come to a functional abstraction of the CNS the following aspects are considered: 1) The limb position and the limb movement are related to the lengths of the limb muscles. 2) A part of the afferent (feedback) signals is related to the muscles, for instance signals from the muscle spindles and tendon organs, and another part is linked to the muscle lengths (joint receptors). 3) The efferent control signals are directly related to the muscles, in fact the muscles are controlled in order to produce the limb movement. Further considerations are dealt with in Section 4. These considerations lead to the suggestion that a functional abstraction of the CNS with respect to the performance of the peripheral motor system may be the lengths of the limb muscles, i.e. the coding of the control signals is related to the muscle lenghts. The abstraction is in this form a "static principle", i.e. it is a transformation or recoding of a spatial pattern in muscle lengths. The length principle acquires a dynamical aspect by taking higher derivatives of the muscle lengths into consideration as well.

3.2. Generation of Control Signals on the Basis of the Length Principle

The control of voluntary movements can generally be considered as a transformation of a desired spatio-

temporal movement pattern $\boldsymbol{p}'(t)$ in a spatio-temporal pattern of activation signals $\boldsymbol{a}(t)$:

$$T0: \boldsymbol{p}'(t) \rightarrow \boldsymbol{a}(t), \tag{2}$$

in which the vector $\boldsymbol{p}'$ is an abstract coding of a spatial position (K dimensions for a limb with K degrees of freedom) and $\boldsymbol{a}$ is the vector of muscle activation signals (N dimensions for a limb with N muscles). This general transformation $T0$ is in this form a very intricate transformation. With the help of the length principle the general transformation $T0$ may be analysed in a number of steps:

$$\boldsymbol{p}'(t) \xrightarrow{T1} \boldsymbol{L}(t) \xrightarrow{T2} [\boldsymbol{L}(t), \dot{\boldsymbol{L}}(t), \ddot{\boldsymbol{L}}(t)] \xrightarrow{T3} \boldsymbol{a}(t), \tag{3}$$

in which $\boldsymbol{L}$ is the vector of the muscle lengths corresponding to the spatial position $\boldsymbol{p}'$. The vector $\boldsymbol{L}$ has N dimensions for a limb with N muscles. The three steps in this series of transformations are quite simple in comparison with the general transformation:

Transformation T1. This is a spatio to spatio transformation or a recoding of a spatial pattern, in which the time aspect is absent. For the execution of this transformation the CNS must have at its disposal a static internal model, i.e. the CNS must be able to recode the position of the limbs from a receptor-bound coding (peripheral $\boldsymbol{L}$-coding) to an abstract coding ($\boldsymbol{p}'$) of spatial positions and vice versa. Such a static internal model seems to be a minimum condition for every control theory; namely the visual and various proprioceptive spatial patterns have to be transformed in the same abstract coding, so that the positions of the limbs can be compared with each other and with respect to the environment. The suppositions by transformation $T1$ are related to the separation of the spatial and temporal aspects in the general transformation $T0$ and, furthermore, it is assumed that the peripheral coding ($\boldsymbol{L}$) is related to the corresponding muscle lengths.

Transformation T2. This transformation adds to the overall transformation a dynamical aspect and fits the control in some way to the dynamics of the peripheral motor systems (Sections 2.2 and 4). This transformation is simple to realize by using neurons or neuronal networks with a differential working.

Transformation T3. This transformation generates the control signals. A distinction may be made between the direct and the servo control, respectively the control signals $\boldsymbol{a}_d(t)$ to the alpha-motorneurons and the control signals $\boldsymbol{a}_s(t)$ to the gamma-motorneurons:

$$\begin{aligned} \boldsymbol{a}_d(t) &= c_1 \cdot \boldsymbol{L}(t) + c_2 \cdot \dot{\boldsymbol{L}}(t) + c_3 \cdot \ddot{\boldsymbol{L}}(t) \\ \boldsymbol{a}_s(t) &= c'_1 \cdot \boldsymbol{L}(t) + c'_2 \cdot \dot{\boldsymbol{L}}(t) + c'_3 \cdot \ddot{\boldsymbol{L}}(t). \end{aligned} \tag{4}$$

The direct control signals $\boldsymbol{a}_d(t)$ and the servo control signals $\boldsymbol{a}_s(t)$ have a common base according to (4), which may find express in an "alpha-gamma linkage" (Granit, 1970). Only the various weighting factors $c_{1,2,3}$ and $c'_{1,2,3}$ are different: the term $c_1 \cdot \boldsymbol{L}$ is relatively less relevant with respect to the other components, while the term $c'_1 \cdot \boldsymbol{L}$ is very relevant with respect to the other components. Differences may occur between the direct and the servo control signals as a consequence of these differences in the weighting factors. The transformations $T1$, $T2$, and $T3$ are represented in (4) in their basic form. In fact the peripheral $\boldsymbol{L}$-coding may reflect more than only the muscle lengths, so the transformations $T1$, $T2$, and $T3$ are somewhat more intricate in the general case (Section 4).

4. Evidence and Consequences of the Length Principle

4.1. Evidence for the Length Principle as an Abstraction of the CNS *with Respect to the Performance of the Peripheral Motor System*

The abstraction on the basis of the muscle lengths is one of the many possible forms of abstraction that the CNS could have with respect to the performance of the peripheral motor system. Evidence for this form of abstraction is: 1) The linkage between the muscle lengths and the limb position. 2) The linkage between the afferent signals and the muscle lengths. 3) The projection of the supraspinal signals to the motorneuronal pools of the muscles. 4) The parallel manipulation of the signals, which has certain advantages with respect to a scalar, composed signal. In the parallel manipulation of the signals the relation with the source of the signal and the relative simplicity of the coding of the information can be maintained. In combining the various signals to a scalar, composed variable (i.e. a potential), there arises an intricate relation with both the source and the destination of the signals. The CNS manipulates the signals for the most part in a parallel manner, as is shown from the somatotopical organization of the afferent and efferent pathways, in which the source and the destination of the information-flow is encoded in the specific organization of the parallel channels. 5) In a case when the afferent and efferent signals have a similar coding, for instance the peripheral $\boldsymbol{L}$-coding according to the length principle, the comparison of afferent and efferent signals is possible at many levels without any need for prior transformation of these signals. The result of the comparisons may be fed back to the peripheral motor system or may be sent to higher control levels as an indication about the difference between the intended and the real movement.

The parallel information-flow and the transformations are schematically shown in Figure 4. Besides

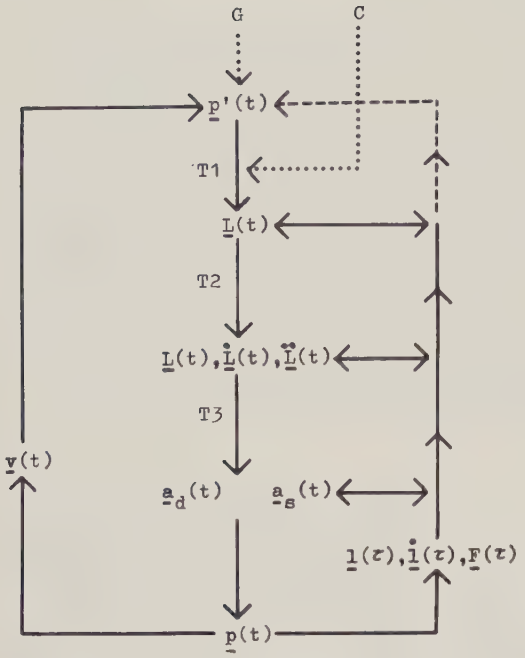

Fig. 4. Diagram of the parallel (multi-channel) information-flow and the transformations $T1$, $T2$, and $T3$, according to the length-principle hypothesis. $p'(t)$ is the desired (imaginary) movement in an abstract coding; $L(t)$ is the desired movement in the peripheral L-coding; $\dot{L}(t)$ and $\ddot{L}(t)$ are dynamical components generated by transformation $T2$; $a_d(t)$ and $a_s(t)$ are the respective direct and servo control signals; $p(t)$ is the real performed movement by the peripheral motor system; $v(t)$ is the visual feedback; $l(\tau)$, $\dot{l}(\tau)$ and $F(\tau)$ are the internal feedback of intermediary output, which may interact with the central control signals at various levels; G is the generation of the imaginary movement; C is the choice of the limb which has to perform the movement

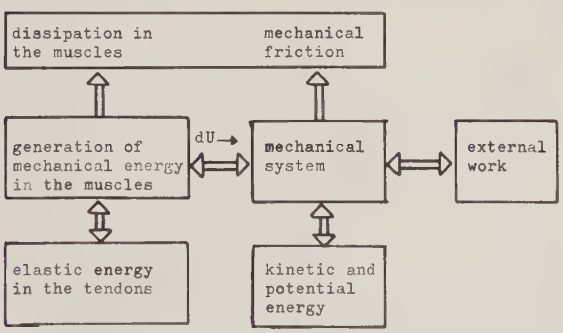

Fig. 5. Diagram of the energy-flow between the muscles, the mechanical system of the limb, and the environment

the indicated feedback loops the intrinsic feedback mechanisms are also relevant, for instance the direct control via the alpha-motorneurons makes use of intrinsic feedback mechanisms in the peripheral motor system (Section 4.5). Transformation $T1$ depends on the choice of the limb that has to perform the desired movement, since the peripheral L-coding is linked to the muscles of this limb. In the following sections some aspects of motor control will be considered in relation to the length principle.

4.2. Minimum Energy Criterion

The following energy consideration relates to a particular movement, considered during a small time interval dt and, for that movement it derives an energetically efficient way of muscle activation. The course of the movement is supposed to be completely determined by the higher control centres; only the relative contributions of the muscles remain to be determined. The length of a muscle i changes a fraction dl_i that corresponds to the course of the movement in the time interval t, $t+dt$. The work performed by the muscles on the mechanical system (Fig. 5) in that time interval is the sum of the effective work done by the separate muscles:

$$dU = \Sigma_i F_i \cdot (-dl_i) \tag{5}$$

The non-effective energy dissipated per sarcomere will be approximated by:

$$dW_{\text{sarc}} = \sigma \cdot (q + k \cdot \mu) \cdot dt\,. \tag{6}$$

This approximation would mean that the non-effective energy per sarcomere is proportional to the generated tension σ and increases whenever the contraction velocity μ per sarcomere is larger[1]. The stiffness of the muscle tendon is rather large (Hill, 1971), which supports the assumption that the contraction velocity $-\dot{l}_i$ of the whole muscle (inclusive the tendon) is practically equal to the contraction velocity of the muscle fibres and as a consequence the average contraction velocity μ per sarcomere is proportional to $-\dot{l}_i/l_{i0}$. For a muscle i with cross-section A_i and rest length l_{i0} the non-effective energy dW_i according to these approximations will be:

$$\begin{aligned} dW_i &= (A_i \cdot l_{i0}) \cdot \sigma \cdot (q + k \cdot \mu) \cdot dt \\ &= (A_i \cdot \sigma) \cdot (q - k' \cdot \dot{l}_i / l_{i0}) \cdot l_{i0} \cdot dt \\ &= F_i \cdot (q \cdot l_{i0} \cdot dt - k' \cdot dl_i)\,. \end{aligned} \tag{7}$$

The total amount of energy delivered by the muscles in the time interval $t, t+dt$ is:

$$\begin{aligned} dE &= \Sigma_i dW_i + dU \\ dE &= q \cdot dt \cdot \Sigma_i F_i \cdot l_{i0} \\ &\quad + k' \cdot \Sigma_i F_i \cdot (-dl_i) + \Sigma_i F_i \cdot (-dl_i)\,. \end{aligned} \tag{8}$$

The movement in that time interval is prescribed, and can be formulated by the condition:

$$\left.\begin{aligned} &dl_i \text{ is fixed for all muscles } i \\ &dU = \sum_i F_i \cdot (-dl_i) = C \text{ (fixed} \end{aligned}\right\}. \tag{9}$$

[1] The breakdown of ATP for a single isotonic tetanus of 400 ms duration is roughly proportional to: $q + k \cdot \Delta l / l_0 + \Delta l$ (Mommearts et al., 1962). The breakdown of ATP depends on the developed tension (Davies et al., 1967). Forced stretch of the activated muscle reduces the breakdown of ATP (Infante et al., 1964)

Optimum activation of the muscles would mean that (8) has to be minimised under condition (9). From (8) and (9) some broad rules with respect to an efficient coordination of the muscles can be derived. By assuming the same rest length l_0 for all the muscles, we can simplify (8) to:

$$dE=(q\cdot l_0\cdot dt)\cdot\Sigma_i F_i+(1+k')\cdot C\,. \tag{10}$$

Expression (10) is minimal if $\Sigma_i F_i$ is minimal under condition (9) and $F_i\geqq 0$ for all muscles. This rough energy consideration indicates that an efficient distribution of the muscle forces F_i is that distribution in which the muscles with the largest shortening $(-dl_i)$ in that interval of time have the largest forces. It can be concluded that the change of the muscle lengths in a particular movement is related to an energeticly efficient method of muscle coordination with respect to the relative contributions of the various muscles. It should be noted that this energy consideration is only a rough approximation and that specific muscle properties (rest length, fibre type) should be regarded if efficient cooperation of the various muscles is required. The generation of the control signals according to (4) may be expressed in a more general way in order to cope with these specific geometrical and fibre properties of the muscles [Section 4.3, Formula (22)].

4.3.Minimum Side Effects Criterion

There is a close relation between the moment arm of a muscle corresponding to a particular direction of rotation of a joint and the change in length of that muscle. In the general case of an arbitrary muscle trajectory and a non-ideal centre of rotation the following energy consideration can be derived with respect to the work performed by that muscle and the work needed to rotate the limb:

$$F\cdot(-dl)=M\cdot d\alpha+\text{“translation energy”}, \tag{11}$$

in which F is the muscle force, $-dl$ the contraction of the muscle, M the moment acting upon the limb and $d\alpha$ the rotation of the limb. The translation of the limbs relative to each other is very small and as a sequence the translation energy in (11) can be neglected:

$$F\cdot(-dl)=M\cdot d\alpha\,. \tag{12}$$

For a joint with three directions of rotation the energy equation will be:

$$F\cdot\left[\left(-\frac{\partial l}{\partial\alpha}\right)\cdot d\alpha+\left(-\frac{\partial l}{\partial\beta}\right)\cdot d\beta+\left(-\frac{\partial l}{\partial\gamma}\right)\cdot d\gamma\right] =M_\alpha\cdot d\alpha+M_\beta\cdot d\beta+M_\gamma\cdot d\gamma\,. \tag{13}$$

Muscle i causes in a particular joint a moment in a particular spatial direction. This moment may be analysed in a desired direction α and other directions (side effects):

$$F_i\cdot\left(-\frac{\partial l_i}{\partial\alpha}\right)=M_{\alpha i}$$

$$F_i\cdot\left(-\frac{\partial l_i}{\partial\beta}\right)=M_{\beta i} \tag{14}$$

$$F_i\cdot\left(-\frac{\partial l_i}{\partial\gamma}\right)=M_{\gamma i}\,.$$

Side effects generally cause compensatory activity in other muscles. Side effects do increase the total sum of muscle forces and also the total amount of non-effective energy, which would mean that side effects should be small from an energetically point of view.

Apart from the minimum-energy criterion the choice of the active muscles may be considered as a choice of those muscles that have minimum side effects and generate forces or moments as near as possible to the desired directions. This criterion would mean that muscles which cause moment contributions in directions as near as possible to the desired directions are preferred. From (14) it appears that the side effects can be translated in terms of changes of muscle lengths that occur when the limb rotates in the various directions. Smaller side effects would mean for a particular muscle that (Fig. 6):

$|\boldsymbol{M}_\alpha|$ is relevant with respect to $|\boldsymbol{M}_\alpha+\boldsymbol{M}_\beta+\boldsymbol{M}_\gamma|=|\boldsymbol{M}|=r\cdot F$

and as consequence:

$\left(-\frac{\partial l}{\partial\alpha}\right)^2$ is relevant with respect to $\left(-\frac{\partial l}{\partial\alpha}\right)^2+\left(-\frac{\partial l}{\partial\beta}\right)^2+\left(-\frac{\partial l}{\partial\gamma}\right)^2=r^2\,.$ (15)

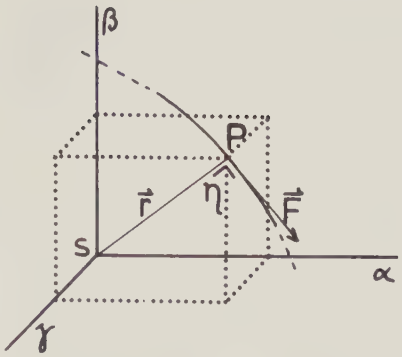

Fig. 6. α, β, and γ are directions of rotation for a joint. P is the point in which the radius $\boldsymbol{r}$ is perpendicular to the force $\boldsymbol{F}$ in the indicated muscle. The moment contribution of the muscle with respect to the centre of rotation S is $\boldsymbol{M}=\boldsymbol{r}*\boldsymbol{F}$, where $|\boldsymbol{M}|=|\boldsymbol{r}*\boldsymbol{F}|=|\boldsymbol{r}|\cdot|\boldsymbol{F}|\cdot\sin\eta=r\cdot F$

This would mean that muscles with a large potential change of length in the desired direction of rotation have relatively slight side effects. A muscle i has slight side effects with respect to other muscles, if for a particular rotation $d\alpha$:

$$\frac{1}{r_i}\cdot\left(-\frac{\partial l_i}{\partial \alpha}\right)\cdot d\alpha \tag{16}$$

is relatively large. Apart from a factor $1/r_i$, which depends on the muscle i and the position of the joint, muscles with a big change in length are preferable since they minimise side effects. This consideration is related to muscles over only one single joint. In the general case the limb is composed of K joints and to each of these joints there corresponds a rotation $d\alpha_k$ in the desired direction of rotation for a particular movement. Muscle i is suitable because of its minimal side effects if the amount of work delivered in the desired directions of rotation of the limb segments is large with respect to the sum of the absolute moment contributions of muscle i in the K joints. The measure for the suitability of a muscle i depends on the desired rotations $d\alpha_k$:

$$G_i(d\alpha_k, k=1,2,\ldots,K)=\frac{dU_i}{\Sigma_k|M_{i,k}|}. \tag{17}$$

dU_i is the amount of work delivered by muscle i for particular, small rotations $d\alpha_k$ of the limb segments:

$$dU_i=F_i\cdot\Sigma_k\left(-\frac{\partial l_i}{\partial \alpha_k}\right)\cdot d\alpha_k \tag{18}$$

The sum of the moment contributions of muscle i in the K joints is:

$$\Sigma_k|M_{i,k}|=F_i\cdot\Sigma_k r_{i,k}. \tag{19}$$

The suitability G_i of muscle i is for the desired rotations $d\alpha_k$:

$$G_i(d\alpha_k, k=1,2,\ldots,K)=\frac{\Sigma_k\left(-\frac{\partial l_i}{\partial \alpha_k}\right)\cdot d\alpha_k}{\Sigma_k r_{i,k}}, \tag{20}$$

which would mean that muscle i is suitable, i.e. has only slight side effects for the desired rotations $d\alpha_k$ of the K joints if:

$$d_i\cdot\Sigma_k\left(-\frac{\partial l_i}{\partial \alpha_k}\right)\cdot d\alpha_k \tag{21}$$

is large with respect to the corresponding fraction of the other muscles. d_i is a proportional factor, which depends on the position of the limb and which may be different for each muscle. $d_i=1/\Sigma_k r_{i,k}$, in which $r_{i,k}$ represents the moment arm of muscle i corresponding to joint k. In order to account for this factor d_i and other muscle dependent factors (for instance geometrical and fibre properties), the transformation $T3$ that generates the control signals according to (4) has to be expressed in a more general form. The scalar factors $c_{1,2,3}$ and $c'_{1,2,3}$ in (4) should be replaced by the vectors of coefficients $\boldsymbol{c}_{1,2,3}$ and $\boldsymbol{c}'_{1,2,3}$ in order to account for special differences between the muscles:

$$\begin{aligned}\boldsymbol{a}_d(t)&=\boldsymbol{c}_1\cdot\boldsymbol{L}(t)+\boldsymbol{c}_2\cdot\dot{\boldsymbol{L}}(t)+\boldsymbol{c}_3\cdot\ddot{\boldsymbol{L}}(t)\\ \boldsymbol{a}_s(t)&=\boldsymbol{c}'_1\cdot\boldsymbol{L}(t)+\boldsymbol{c}'_2\cdot\dot{\boldsymbol{L}}(t)+\boldsymbol{c}'_3\cdot\ddot{\boldsymbol{L}}(t).\end{aligned} \tag{22}$$

For instance, a muscle i that is particularly involved in fast movements would require relatively large c_{2i} and c_{3i} coefficients in (22). The coefficient vectors in (22) depend on the limb position and therefore on the muscle lengths, i.e. $\boldsymbol{c}=\boldsymbol{c}(\boldsymbol{L})$, which would mean that the generation of the control signals according to (22) is generally a non-linear transformation.

4.4. Movement Strategy

If for a particular movement, which is desired by the CNS, not every a priori degree of freedom of the limb is prescribed, there remains some possibility of using the remaining degrees of freedom in an optimal way, i.e. of performing the movement in such a way that certain criteria (for instance minimum energy, minimum side effect or minimum duration of the movement) are involved as much as possible in the performance of the movement. For instance, the CNS may prescribe the course of the wrist in the time-space without prescribing the position of the lower arm (semiprone, etc.) or the vertical position of the elbow. A generation of the control signals, which is based on the length principle, will lead automatically to a performance of the movement, which is optimal in some respects. From Section 4.3 it follows that the generation of the control signals based on the changes in muscle lengths which have to occur prefers those muscles that have slight side effects. If there were no compensatory muscle activity, then the non-prescribed degrees of freedom and the resulting limb movement would be determined by the active muscles, i.e. they would take a form that corresponds to the most effective directions of the muscle forces. This would mean that the non-prescribed degrees of freedom will tend to values at which the most suitable muscles (with slight side effects) will be more suitable. It is to be expected that in a control based on the length principle, the non-prescribed degrees of freedom will be automatically adjusted in such a way that the most active muscles will on the whole become more effective. This would mean a suitable adjustment with respect to the minimum energy criterion (less non-effective muscle forces and less energy dissipation in the muscles),

but the movement is speeded up as well (if the muscles are suitable with respect to the desired, effective forces the movement can be performed rapidly).

If gravitation is considered, the general conclusion remains the same, but the resulting values of the non-prescribed degrees of freedom will be different from the case without gravitation. A compromise will have to be reached with respect to the directions of the forces of the active muscles and the external (gravitation) force acting upon the limb. It is possible that muscles which oppose gravitation will benefit in respect to other muscles; for instance, the coefficients c_1 in (22) may be more relevant for anti-gravitation muscles than for other muscles. This strategy of muscle coordination is the subject of the following rough energy consideration. Suppose that for a particular movement of a limb with n degrees of freedom the a priori degrees of freedom ϱ_v with $v \in V$ are prescribed by the CNS and the degrees of freedom ϱ_u with $u \in U$ are left undetermined. The forces F_i of the various muscles and the external moments M_s (in the degree of freedom ϱ_s, $s \in (V \cup U)$) perform a total amount of work dU_{tot} in a time interval dt on the limb:

$$\begin{aligned} dU_{\text{tot}} &= \Sigma_i F_i \cdot (-dl_i) + \Sigma_s M_s \cdot d\varrho_s \\ &= \Sigma_i F_i \cdot \sum_{v \in V} \left(-\frac{\partial l_i}{\partial \varrho_v} \right) \cdot d\varrho_v + \Sigma_i F_i \\ &\quad \cdot \sum_{u \in U} \left(-\frac{\partial l_i}{\partial \varrho_u} \right) \cdot d\varrho_u + \Sigma_s M_s \cdot d\varrho_s . \end{aligned} \tag{23}$$

The energy delivered by the muscles in the non-prescribed directions is dU_{non}:

$$dU_{\text{non}} = \Sigma_i F_i \cdot \sum_{u \in U} \left(-\frac{\partial l_i}{\partial \varrho_u} \right) \cdot d\varrho_u . \tag{24}$$

The resulting movement in the non-prescribed directions is determined by the work done in these directions. From an energy point of view it is desirable that the work done by the muscles per non-prescribed direction ϱ_u should be as little as possible:

$$dU_{\text{non},u} = \Sigma_i F_i \cdot \left(-\frac{\partial l_i}{\partial \varrho_u} \right) \cdot d\varrho_u . \tag{25}$$

A suitable strategy will minimise:

$$\Sigma_i F_i \cdot \left(-\frac{\partial l_i}{\partial \varrho_u} \right) \tag{26}$$

for each non-prescribed degree of freedom ϱ_u. Minimising of (26) for each $u \in U$ is possible in two ways: 1) Minimising of the side effects for each muscle. This will lead to a problem similar to the one in Section 4.3. 2) By activating the muscles in such a way that the side effects of different muscles will compensate each other. In general this implies that the muscle forces in this case are larger than strictly necessary. The sum of all muscle forces must be as small as possible in order to minimise the energy dissipation, which would mean that this method of minimising (26) is less efficient than the method first mentioned. Because of the similarity with the problem of Section 4.3 it can be concluded that for a particular limb movement, in which the degrees of freedom ϱ_v, $v \in V$ are prescribed and ϱ_u, $u \in U$ are non-prescribed, a muscle i is more suitable if:

$$d_i \cdot \sum_{v \in V} \left(-\frac{\partial l_i}{\partial \varrho_v} \right) \cdot d\varrho_v \tag{27}$$

is larger (d_i is a proportional factor). This would mean that muscles, which shorten most in the prescribed directions, are suitable with respect to a combination of various criteria such as the minimum energy criterion and the minimising of the side effects. This consideration of the strategy of muscle coordination relates only to the choice of the active muscles and the efficiency of these muscles in a particular movement. Higher forms of strategy, for instance the determination of the course of the movement in the time-space and the way in which some a priori degrees of freedom will be prescribed and others will remain undetermined, are beyond the scope of this study.

4.5. Feedback Mechanisms

As was pointed out in Section 2.2 the length principle is an abstraction of the CNS with respect to the performance of the peripheral motor system, i.e. the control based on the muscle lengths of the limb does not account for the dynamics of the periphery in a direct way. In order to perform the right movements in spite of the abstraction of the control, various types of feedback mechanisms, such as feedback loops and intrinsic feedback must operate. Possible feedback mechanisms are: long loop and spinal reflexes, the force-velocity dependence for positive contraction velocities (shortening), the dependence of the excitation-contraction coupling on the muscle length and the contraction velocity (Taylor et al., 1970), etc. All these feedback mechanisms act as a negative feedback, i.e. they react to internal and external disturbances. For instance, in a particular movement at a point of time t the muscles i and j should shorten with the respective velocities $-\dot{L}_i$ and $-\dot{L}_j$. Suppose that, as a consequence of the dynamics of the limb or other internal and external factors, muscle i shortens rather too slowly and muscle j faster than was intended. As a result of the force-velocity relationship of the muscles an effect arises that has certain features of a negative feedback with respect to the ideal situation, in which

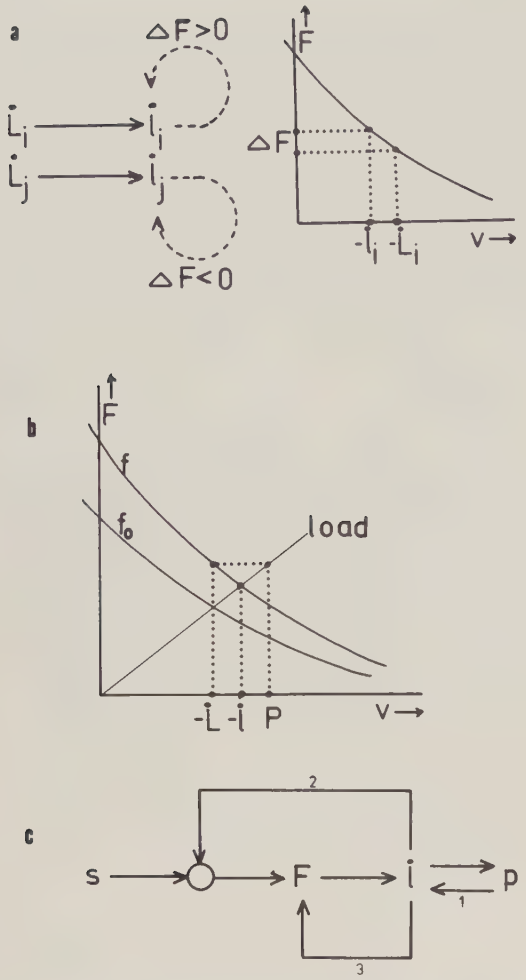

Fig. 7a. The difference between the intended contraction velocity $-\dot{L}$ and the real velocity $-\dot{l}$ in the performance of a particular movement causes a force ΔF in that muscle, which opposes this difference. **b** f_0 is the activation of a muscle, which would correspond to a desired contraction velocity $-\dot{L}$. If the control signals deviate from the precise pattern of control signals needed for the exact execution of the desired movement, deviation from the desired movement will occur. The activation $f > f_0$ will cause a contraction velocity $(-\dot{l}) > (-\dot{L})$. As a consequence of the force velocity relationship this deviation will be smaller than if the muscle force did not depend on the contraction velocity (indicated by point *P*). **c** Diagram of the velocity feedback. *1* is the interaction between the limb and the environment (load, disturbances), *2* is the proprioceptive feedback, *3* is the intrinsic velocity feedback in the muscles. $s(t)$ is the central control signal, F is the muscle force, $-\dot{l}$ is the contraction velocity and p is the real performance of the movement

muscle i and muscle j had shortened as was intended. In muscle i the force F_i will increase somewhat and in muscle j the force will decrease somewhat with respect to the ideal performance of the movement (Fig. 7a). Internal disturbances, such as an improper control signal, will elicit effects that may be understood as negative feedback, which oppose the disturbances (legend Fig. 7b and c).

4.6. Psychological Aspects of the Length Principle

Experimental data (Russell, 1974) suggest that location cues are important for good motor performance. This finding fits well with the control scheme as presented in Section 3, since the length principle is in the first instance a static principle i.e. it is based on spatial and limb positions in an abstract or a peripheral coding (transformation $T1$). In the second instance "dynamic" components are generated (transformation $T2$) in order to add to the control signals certain components that will be able to cope with the inertia of the peripheral motor system.

By far the most intricate transformation in the series $T1$, $T2$, $T3$ is transformation $T1$. This transformation needs a static internal model in which the positions of the limbs can be compared with each other and with external positions. From the hypothesis stated in Section 3 it may be derived that whenever in the mental development this static internal model is established, the performance of good movements will generally present no problem, unless not enough time is left for the execution of transformation $T1$ on account of the velocity of the movement or of many unknown external factors.

If transformation $T1$ is the limiting factor that restricts the fast performance of a movement, then independently of the form or size of the movement the inequality holds:

$$H(\text{movement}) \leqq C(T1) \cdot \Delta T, \tag{28}$$

i.e. the information contents or variety H of a movement are limited by the capacity C of transformation $T1$ and the interval of time ΔT in which the movement has to take place. If the movement is performed as fast as possible and transformation $T1$ is the limiting factor, the following equation holds:

$$H(\text{movement}) = C(T1) \cdot \Delta T. \tag{29}$$

This would mean that, whenever the desired movement is more varied, the time to perform that movement is proportional to the degree of variety, but does not depend on the form or size of the movement. For fast limbs such as the fingers and the hand, it may be that the capacity of transformation $T1$ or higher psychological functions are indeed the limiting factor in the speed of performance of the movements. The transformation $T1$ and the capacity $C(T1)$ depend on the limb that performs the movement.

Movements can be performed in the imagination without real execution. Transformation $T1$ may be involved in this "in imago" performance of a particular movement. The execution of transformation $T1$ may be evaluated with the help of feedback loops (Fig. 4). If this is true then it may be expected that a future real

performance of that movement will show a positive transfer (Trumbo et al., 1965).

In the hypothetical control scheme based on the length principle there is a close link between the movement and the sequence of positions that will be passed through. This will facilitate the evaluation of a performed movement: the desired sequence of positions and the sequence of positions that are passed through by the limb in the real performance of the movement may be compared with each other at many levels and even during the movement (see Fig. 4, in which the feedback and comparison of signals are indicated). Various types of feedback information may be used in this evaluation, both visual and proprioceptive information, encoded in an abstract or a receptor-bound coding. Anticipation of the final position, corresponding to a pattern of control signals for a goal directed movement, is a triviality in this control scheme, because the pattern of control signals is derived from the positions that have to be passed through and also from the final position. The evaluation of the movements performed and the link between the pattern of control signals and the anticipation of the goal of the movement, are rather intricate in other control theories (Schmidt, 1975; Pew, 1966).

The imaginary movement $\boldsymbol{p}'(t)$ (Fig. 4) is encoded in an abstract coding, i.e. the coding is not directly linked to the limb that has to perform the movement. A learned movement that has been performed by a particular limb will show some positive transfer of the movement characteristics, if that movement has suddenly to be performed by another limb. It should be noted that many peripheral feedback mechanisms are involved in the performance of the movement, so this transfer of movement characteristics cannot be expected to be very large.

4.7. Experimental Assay

The control signals according to the length principle are given by (22). The control signals depend on the direction in which the limb is instructed to move or is to be moved by an external force. If an external force acts upon the limb in a particular direction, then the muscle lengths will change in a particular way if the limb moves slightly in that direction. This motion may be very slight or even imaginary. From (22) it can be derived that the control signal for a muscle i is proportional to the slight shortening $-\Delta l_i$ of that muscle in a static situation:

$$a_i :: -\Delta l_i, \tag{30}$$

in a rough approximation, since various feedback mechanisms may act and change the control signal in order to fit the control to the mechanics of the limb or other internal factors. In order to check this expectation (30) about the relationship between the muscle activation and the change in muscle length in a static situation, the following pilot experiment was done.

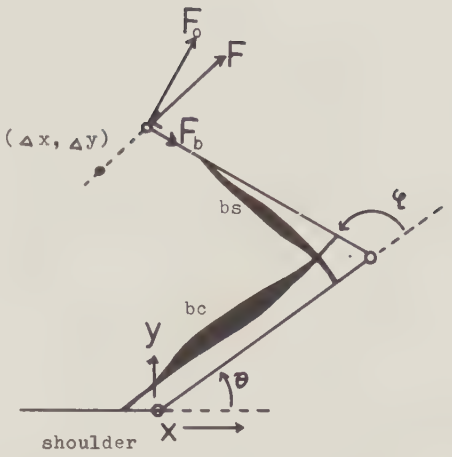

Fig. 8. Representation of the upper and lower arm. F_0 and F_b are external forces. $(\Delta x, \Delta y)$ is a small displacement of the wrist along the resulting force F in an opposite direction. *bc* and *bs* represent the m.biceps and the m.brachioradialis. The lengths of the upper and lower arm in the calculation (32) and in the simulations are equal (0.32 m)

The right arm of a subject is in a horizontal plane through the shoulders. The lower arm is in a semiprone position and the elbow is supported in such a way that the arm is able to move freely in the horizontal plane. Two external forces F_0 and F_b are attached to the wrist (Fig. 8). The subject is asked to hold his arm in a fixed position. The activities (i.EMG) of the m.biceps and the m.brachioradialis are measured, while the subject balances his arm with the resulting force F. By changing the direction of the external force F, the small displacements of the limb and the corresponding changes of the muscle lengths $-\Delta l_{bc}$ and $-\Delta l_{bs}$ of respectively the m.biceps and the m.brachioradialis also change. These changes in muscle lengths may be caused by the difference between the prescribed muscle lengths $\boldsymbol{L}$ and the real muscle lengths $\boldsymbol{l}$ (as a consequence of the external force F), or it may be that the CNS opposes directly the external force by transforming the direction of the external force in the corresponding changes of the muscle lengths $\Delta \boldsymbol{L}$. The changes Δl_{bc} and Δl_{bs} are calculated as follows (Fig. 8). The imaginary or small real displacement $(\Delta x, \Delta y)$ of the wrist along the resulting force F in the opposite direction is:

$$\begin{aligned} \Delta x &= -\varepsilon \cdot [F_b \cdot \cos(\theta+\varphi) - F_0 \cdot \sin(\theta+\varphi)] \\ \Delta y &= -\varepsilon \cdot [F_b \cdot \sin(\theta+\varphi) + F_0 \cdot \cos(\theta+\varphi)]. \end{aligned} \tag{31}$$

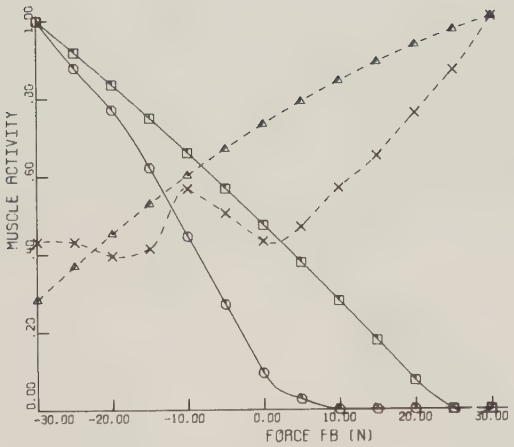

Fig. 9. The arm of the subject balances with an external force F (which is attached to the wrist) of which the component F_b along the lower arm is varied (shown in Fig. 8). □ and ○ are the respective calculated and measured activities of the m.brachioradialis; △ and × are the respective calculated and measured activities of the m.biceps. The muscle activities are in arbitrary units

ε is a small, positive constant. The angles θ and φ can be derived from (32) and the position (x, y) of the wrist.

$$\varphi = \text{a}\cos((x^2+y^2)/(2\cdot s^2)-1)$$
$$\theta = \pi/2 - \text{a}\tan(x/y) - \varphi/2\,. \tag{32}$$

s is the length of the upper and lower arm. For every combination of F_b and F_0 from (31), (32) and the rest position $x_0=0$, $y_0=0.32\,m$, the changes $\Delta\theta$ and $\Delta\varphi$ corresponding to the displacement $(\Delta x, \Delta y)$ can be derived. The corresponding length changes of the m.biceps and the m.brachioradialis are approximated by:

$$\Delta l_{bc} = -0.019\cdot\Delta\theta - 0.026\cdot\Delta\varphi\, m$$
$$\Delta l_{bs} = -0.048\cdot\Delta\varphi\, m\,. \tag{33}$$

In this experiment F_0 was fixed (15 N) and F_b varied from -30 up to $+30$ N. In Figure 9 the calculated and measured activities of the m.biceps and the m.brachioradialis are shown.

4.8. Simulations

A circle movement of the human arm is simulated, in which the control signals are derived from the muscle lengths corresponding to the prescribed imaginary movement. The imaginary movement is given by (34); see also Figure 8 and 10

$$x(t) = 0.15\cdot\cos(8t) \quad m$$
$$y(t) = 0.32 + 0.15\cdot\sin(8t) \quad m \tag{34}$$
$$0 \leqq t \leqq 1.5 \quad s\,.$$

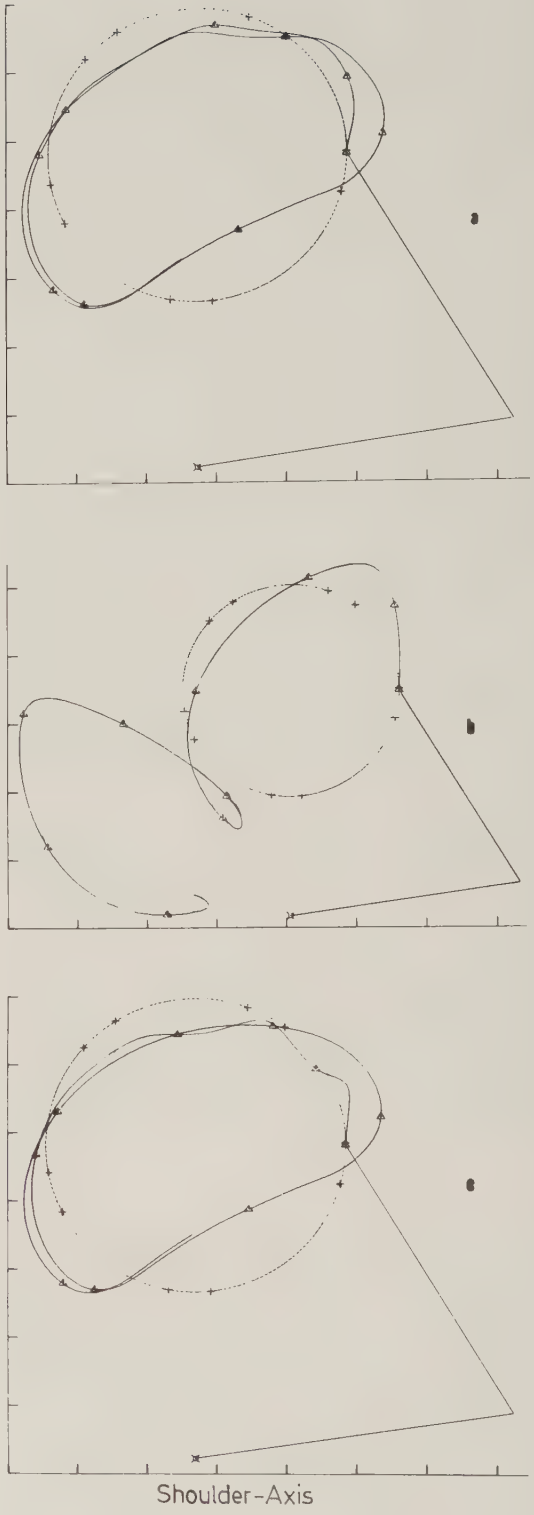

Fig. 10a—c. Simulation of a circle movement of the human arm. **a** Servo control. **b** Direct control. The CNS receives no information about the performed movement; this explains the drift in the simulated movement in this case. **c** A combination of servo and direct control. The markers on the simulated movements indicate time intervals of 150 ms duration

Three types of control are simulated and the control signals in these cases are:

1) Servo control. The control signals $a_{s,i}(t)$ for each muscle i are simulated by:

$$a_{s,i}(t) = c'_1 \cdot L_i(t). \tag{35}$$

Comparison with the feedback signals results in the activation signals f_i:

$$f_i = 8\dot{l}_i + 100 \cdot (l_i - L_i) - 0.03 \cdot F_i. \tag{36}$$

The activation signals $f_i(s^{-1})$ and the feedback signals l_i, $\dot{l}_i(s^{-1})$ and F_i(N) are delayed in order to account for the various conduction delays. For details of the model of the human arm and other simulations the reader is referred to van Dijk (1978).

2) Direct control. The control signals $a_{d,i}(t)$ and the activation signals $f_i(t')$ are simulated by ($\dot{L}$ in s^{-1}, $\ddot{L}$ in s^{-2}):

$$a_{d,i}(t) = -1.0 \cdot \dot{L}_i(t) - 0.05 \cdot \ddot{L}_i(t) = f_i(t'). \tag{37}$$

3) A combination of direct and servo control. The activation signals f_i in this case are:

$$f_i = -1.0 \cdot \dot{L}_i - 0.05 \cdot \ddot{L}_i + 6 \cdot \dot{l}_i + 75 \cdot (l_i - L_i) - 0.02 \cdot F_i. \tag{38}$$

The generation of the activation signals according to (35)–(38) is a simplification of the general form of transformation $T3$ (22). In (35)–(38) no distinction has been made between specific muscle properties (fast or slow type fibres) or different geometrical factors (rest length; factors which depend on the limb position: moment arms, force-length relationship). The coefficient vectors $\boldsymbol{c}_{1,2,3}$ and $\boldsymbol{c}'_{1,2,3}$ could account for these specific properties.

The control signal to the motorneuronal pool of a particular muscle i is mediated by a number of parallel nerve fibres in which the components $L_i(t)$, $\dot{L}_i(t)$ and $\ddot{L}_i(t)$ may be weighted with different factors. Different types of muscle fibres in the same muscle i may be distinguished in this way. On account of the different strength of the components $L_i(t)$, $\dot{L}_i(t)$, and $\ddot{L}_i(t)$ in the various parallel nerve fibres, these components can be distinguished and compared with the corresponding components $l_i(t')$, $\dot{l}_i(t')$ and $\ddot{l}_i(t')$ in the feedback signals.

5. Discussion

In Sections 2 and 4.5 it is pointed out that the peripheral motor system contains many feedback mechanisms, in which the feedback may consist of a loop of information following an explicit pathway, or may exist as an intrinsic feedback of intermediary output variables, or as an intrinsic property of the subsystems. The feedback mechanisms may simplify the control of the peripheral motor system. The feedback mechanisms oppose disturbances. These disturbances may be considered in a broad sense, i.e. they may be external disturbances (for instance the load), or internal disturbances which may be caused by a non-optimal matching of the control signals with the dynamics of the periphery. Feedback mechanisms make the control by the CNS less dependent on the dynamics of the peripheral motor system and the interaction with the environment (Legends Figs. 2 and 7). This fact offers the CNS the possibility of having a certain degree of independence with respect to the performance of the peripheral motor system. The CNS has to prescribe the desired movement exactly, but the control signals do not need to cope with all the dynamical aspects in the periphery. The degree of independence or abstraction of the CNS with respect to the performance of the peripheral motor system depends on the degree to which the feedback mechanisms are able to compensate the internal and external disturbances; this degree is related to the complexity of the peripheral motor system and the capacity of the feedback mechanisms.

A functional abstraction of the CNS has to match the control variables, the type and form of the feedback information and in general the properties of the peripheral motor system. As a functional abstraction the length principle may be considered. The length principle states that the muscle lengths and their derivatives for a particular, desired movement are related to the control signals conveyed to the alpha and gamma motorneurons; moreover the CNS generates the control signals on the basis of muscle lengths and this generation of control signals is rather independent on the peripheral dynamics. This abstraction implies that the communication between the CNS and the peripheral motor system exists in a form that is related to the muscle lengths. The source and destination of the signals are known by the specific somatotopical organization of the afferent and efferent parallel channels. The communication between the CNS and the peripheral motor system may be considered as a peripheral ***L***-coding, i.e. a somatotopical organization of parallel channels that relay information which is closely related to the muscle lengths. The generation of the control signals for a particular movement may be represented by a series of three transformations, $T1$, $T2$, and $T3$. $T1$ transforms a desired position $\boldsymbol{p}'$ from an abstract coding to a peripheral coding $\boldsymbol{L}$. In the case of a movement this would mean a successive (or continuous) transformation of positions $\boldsymbol{p}'(t)$. $T2$ generates dynamical components which may be used in the generation of the control signals in order to cope with some aspects of the peripheral motor system (for

instance the inertia of the limb, the preference of fast or slow type muscles, etc.). *T*3 generates the alpha and gamma control signals from the position sequence (encoded in the ***L***-coding).

The hypothetical control scheme, based on the length principle, has many attractive aspects. In Section 4.2 it is concluded that a muscle is more suitable as an active muscle in a particular part of a movement, whenever the muscle has to shorten more. Furthermore, the ***L***-coding of the control signals is well adapted to the organization of the peripheral motor system (Sections 2 and 4.1). So, this abstraction of the CNS and the encoding of the afferent and efferent signals fits well with neuro-anatomical and physiological data, with the criteria of minimum energy and minimum side effects and with other data (Section 4).

In this study of the control of voluntary movements only the relative coordination of the muscles is considered, i.e. their relative suitability as active muscles in a particular part of a movement. The overall or absolute intensity of the control signals is not considered. The length principle deals with the interaction between the CNS and the peripheral motor system. Further attention will be paid to the extension of the control scheme based on the length principle and to the functioning of the CNS in the execution of voluntary movements.

I would like to thank G. van Meer for his assistance with the experiment, and Prof. J. J. Denier van der Gon and Dr. A. Crowe for critisism of the manuscript. The author is indebted to the Netherlands Organization for the Advancement of Pure Research (Z.W.O.) for financial support of this investigation.

References

Bernstein, N.: The co-ordination and regulation of movements. Oxford: Pergamon Press 1967

Brooks, V.B., Stoney, Jr., S.D.: Motor mechanisms: the role of the pyramidal system in motor control. Ann. Rev. Physiol. **33**, 337—392 (1971)

Burke, R.E., Rymer, W.Z., Walsh, Jr., J.V.: Relative strength of synaptic input from short-latency pathways to motor units of defined type in cat medial gastrocnemius. J. Neurophysiol. **39**, 447—458 (1976)

Cheney, P.D., Preston, J.B.: Effects of fusimotor stimulation on dynamic and position sensitivities of spindle afferents in the primate. J. Neurophysiol. **39**, 20—30 (1976)

Davies, R.E., Kushmerick, M.J., Larson, R.E.: ATP, activation, and the heat of shortening of muscle. Nature **214**, 148—151 (1967)

van Dijk, J.H.M.: Simulation of human arm movements controlled by peripheral feedback. Biol. Cybernetics **29**, 175—186 (1978)

Gawronski, R., ed.: Bionics. Amsterdam: Elsevier 1971

Gottlieb, G.L., Agarwal, G.C.: The role of the myotatic reflex in the voluntary control of movements. Brain Res. **40**, 139—143 (1972)

Granit, R.: The basis of motor control. London: Academic Press 1970

Hill, A.V.: First and last experiments in muscle mechanics. Cambridge: University Press 1970

Houk, J., Henneman, E.: Responses of Golgi tendon organs to active contractions of the soleus muscle of the cat. J. Neurophysiol. **30**, 466—481 (1967)

Huxley, A.F.: Review lecture "Muscular contraction". J. Physiol. **243**, 1—43 (1974)

Infante, A.A., Klaupiks, D., Davies, R.E.: ATP changes in muscles doing negative work. Science **144**, 1577—1579 (1964)

Mommearts, W.F.H.M., Seraydarian, K., Marechal, G.: Work and chemical change in isotonic muscular contractions. Biochim. Biophys. Acta **57**, 1—12 (1962)

Pew, R.W.: Acquisition of hierarchical control over the temporal organization of a skill. J. Exp. Psych. **71**, 764—771 (1966)

Rabbitt, P.M.A., ed.: Attention and performance. Vol. 5. New York: Academic Press (1974)

Russell, D.G.: Spatial location cues and movement production. In: Motor control; issues and trends. pp. 67—85. Stelmach, G.E., ed. New York: Academic Press 1976

Scheibel, M.E., Scheibel, A.B.: Inhibition and the Renshaw cell: a structural critique. Brain Behav. Evol. **4**, 53—93 (1971)

Schmidt, R.A.: The schema as a solution to some persistent problems in motor learning theory. In: Motor control; issues and trends. pp. 41—65. Stelmach, G.E. ed. New York: Academic Press 1976

Taylor, S.R., Rüdel, R.: Striated muscle fibres: inactivation of contraction induced by shortening. Science **167**, 882—884 (1970)

Trumbo, D., Ulrich, L., Noble, M.E.: Verbal coding and display coding in the aquisition and retention of tracking skill. J. Applied Psych. **49**, 368—375 (1965)

Welford, A.T.: Fundamentals of skill. London: Methuen 1968

Received: May 4, 1978

Dr. J. H. M. van Dijk
Physics Laboratory
Dept. of Med. and Physiol. Physics
Sorbonnelaan 4
Utrecht, The Netherlands

Biol. Cybernetics 30, 209—220 (1978)

Biological Cybernetics

Is the Adaptation Model a Valid Description of the Vestibulo-Ocular Reflex?*

A. W. Sills[1]**, V. Honrubia[1], and R. W. Baloh[2]

[1] Department of Surgery Division of Head and Neck (Otolaryngology), and
[2] Department of Neurology, UCLA School of Medicine, Los Angeles, California

Abstract. The pendulum model of the vestibulo-ocular reflex, including the effects of adaptation, has been evaluated using the responses of 36 normal subjects to impulsive stimuli of 128 and 256°/s. Estimates of the model parameters such as the time constants, the slow velocity threshold, and the minimum stimulus required to produce an after-nystagmus have been obtained using a new analytical technique. Although some of the data support the validity of the adaptation model, evidence is presented to demonstrate that the overall applicability of the model is limited.

Introduction

The vestibulo-ocular reflex is a complex mechanism that acts to maintain gaze direction by compensating for disturbances in eye position due to head movements. Although it is often approximated by a three-neuron arc connecting the labyrinthine organs to the eye muscles, there are important collateral connections, both afferent and efferent, involving many parts of the central nervous system. The need to improve the detection and diagnosis of vestibular dysfunction in humans has been a strong motivating force for the development of mathematical models of the vestibulo-ocular reflex arc. These models ignore complicating features of the anatomy to varying degrees. Thus, one may ask, how well do the models describe the experimental data?

Based on the experimental work of Steinhausen (1931), who made direct observations of the cupular motion in the pike, and Lowenstein and Sand (1940), who studied the response of vestibular nerve fibers in the ray fish, van Egmond et al. (1949) proposed that the kinematic behavior of the cupula could be adequately described by the same second-order linear differential equation used to describe an aperiodic pendulum. Work by Lorente de Nó (1931) and others (Steinhausen, 1933; Dohlman, 1938; van Egmond et al., 1952; Hallpike and Hood, 1953) suggested that the slow phase eye velocity (SV) during nystagmus is correlated with changes in the neural firing rate of the vestibular nerve, presumably due to deviations of the cupula, so that the same equation could be used to describe SV as a function of time.

This pendulum model has two parameters that describe the time course of the response: a long time constant T_1, approximately 10–20 s in magnitude, and a short time constant T_2, on the order of 0.001–0.1 s. Since the pioneering work of van Egmond et al., attempts have been made to use estimates of these time constants obtained from measurements of the nystagmus response induced by different labyrinthine stimuli as diagnostic tools for detecting vestibular dysfunction (Groen, 1957a; Aschan et al., 1952; Melvill-Jones et al., 1964; Hallpike and Hood, 1953). These efforts have been largely unsuccessful, in part because this model is incomplete. The model predicts that the SV resulting from an impulsive stimulus will rise quickly to a peak at a rate governed by T_2, then decay back to zero at a rate governed by T_1. At low stimulus magnitudes, the pendulum model appears to predict the correct behavior of SV. However, at higher stimulus magnitudes the response does not merely decay back to zero, but after a brief latency period, appears again beating in the opposite direction. This latter phase is referred to as after-nystagmus in contrast to the initial or post-rotatory phase. Such behavior cannot be explained solely on the basis of the linear pendulum model.

A second failure of the model concerns the predicted duration of the post-rotatory phase (TDUR). According to van Egmond et al. (1949), the time at

* This work was supported by NIH Grant NS 09823
** To whom reprint requests should be addressed

0340-1200/78/0030/0209/$02.40

which the post-rotatory phase ends should increase as the logarithm of the stimulus magnitude. But experimentally (Konrad et al., 1975) TDUR stops increasing for impulses above 64°/s.

Finally, there is a more subtle effect that discredits the pendulum model, even where it seems to work correctly. At low stimulus magnitudes where only the post-rotatory phase is experimentally observed, the model predicts that the descending part of the SV response is exponential. A plot of the logarithm of SV versus time should result in a straight line with a negative slope of $-1/T_1$. However, experimental responses often deviate from a straight line by displaying a sharper rate of descent. This curvature, as well as the existence of after-nystagmus, has been attributed to adaptation effects (Groen, 1956, 1957; Groen, 1957b; Malcolm and Melvill-Jones, 1970).

These adaptation effects are of a complex nature, but several authors have created models that attempt to take them into account (Young and Oman, 1969; Schmid, 1970; Schmid et al., 1971a). Malcolm and Melvill-Jones (1970) tried to determine the time constants using experimental data from a few normal subjects and fitting the data to the theoretical model by trial and error. Guedry et al. (1971) and Stockwell et al. (1973) used triangular velocity profiles in normal subjects to evaluate t_r, a theoretical parameter derived from the adaptation model. There have also been attempts (Meiry, 1965; Young, 1968; McClure et al., 1976; Schmid et al., 1971b) to determine the adaptation model parameters by fitting the cumulative slow phase eye position data using nonlinear regression analysis. One problem with these efforts is a lack of uniqueness among the determined values of the parameters. Several authors (Julius, 1972; van Mastrigt, 1977) have demonstrated the non-uniqueness of solutions obtained by fitting sums of exponentials to experimental data. Another obstacle to the practical application of time constant analysis to large groups of subjects has been the tedium of arbitrarily adjusting multiple time constants to fit the data of individual subjects. Even when done with a minicomputer, nonlinear regression can take as long as twenty minutes to find a solution that fits the subject's response.

Probably for these reasons no one to our knowledge has evaluated the time constants for a large group of normal subjects and then applied the normative results to patient groups. In order for us to do this it was first necessary to develop a fast, accurate and convenient method for determining unique values of the model parameters from the experimental data. This methodology provided more rigorous estimates of the model parameters than were possible earlier. These values were determined for a large population of normal subjects to answer the following questions:

a) Are the model parameters independent of stimulus magnitude?

b) Are they time invariant?

c) What correlations exist between the different parameters?

d) In the light of these data, is the adaptation model adequate?

e) How well do time constants derived from the adaptation model discriminate between normal subjects and patients?

Methods

Experimental. The subjects were strapped in a chair mounted on a Goerz-Inland model 811 rotating table inside a light-tight, electrically shielded room. The Inland 811 controller delivers maximum torque of 10 ft lb which provides a weight-dependent constant accelaration of approximately 140°/s^2. The stimuli consisted of changes in velocity of 16, 32, 64, 128, and 256°/s, done at maximum table acceleration. The changes were done in clockwise (CW) and counterclockwise (CCW) directions presented in random order, except that the 256°/s impulses were never given first and were always paired to avoid excessive angular velocities. The tests were conducted in the dark with the subject's eyes open. It is critical in rotatory testing that the subject's head be firmly fixed so that the actual stimulus is close to the presented one. This is particularly true of the impulsive test because the duration of the impulse is so brief that any large countermovement of the head markedly reduces the effective stimulus to the labyrinth. We used a holder that pressed the subject's head back firmly against a head rest tilted downward thirty degrees to align the horizontal canals with the plane of the floor. Long extensions along the temples prevented lateral movement of the head. Above this piece a separate *U*-shape section projected forward to fit around the subject's forehead. It could be adjusted backwards to give a tight but comfortable fit. With this device very little lateral head motion was possible. The adequacy of the head holder was confirmed by mounting an accelerometer on the heads of several subjects and measuring the actual head movement as the two largest impulses were presented. The accuracy of the accelerometer was sufficient to detect changes in head velocity of less than 10°/s. The head velocity reproduced the table velocity profile to within this experimental error for subjects weighing between 100 and 185 lb (Gilman et al., submitted for publication). Criteria for establishing the normality of the subjects, details of the EOG apparatus, and methods of data analysis are given in Sills et al. (1977).

Nystagmus data were obtained using binocular electroocular recordings. Calibrations were done in the light at the beginning and end of the test sequence using targets 15° apart. Changes in the calibration were compensated for using a linear interpolation

weighted by the elapsed time for each test response. Subjects were kept alert by subtracting sevens aloud from a large number.

An example of the SV responses to the five impulsive stimuli is shown in Fig. 1. The appearance of after-nystagmus is seen at the higher magnitudes of stimulation. Fig. 2 shows the experimental response to a change in velocity (DV) of 256°/s. The figure shows the six experimental quantities that are required to fix the values of the time constants and several other model parameters of interest. The two points corresponding to the maximum responses for the post-rotatory and after-nystagmus phases are described by the maximum values, SVMX1 and SVMX2 respectively, and their times of occurrence, TMX1 and TMX2. These four quantities are sufficient to uniquely determine the three time constants, T_1, T_2, and T_A as well as a constant of proportionality, K_i, that relates the hypothetical cupular deflection to the observed slow phase eye velocity.

The end of the post-rotatory phase is called TDUR. Between TDUR and the beginning of after-nystagmus lies a space of time called ΔT. This time is related to the threshold value of SV (SV_{th}) below which nystagmus does not occur. TDUR and ΔT can be used to determine the value of SV_{th}, and also the minimum impulse (DV_{min}) required to produce after-nystagmus.

Thirty-six normal subjects were tested and the values of the six experimental quantities described were obtained for responses to stimuli of 128°/s and 256°/s. CW and CCW responses were considered separately so the maximum number of responses for each stimulus would be 72. However, at 128°/s not all subjects had an after-nystagmus reaction. In addition, the data were screened for outliers by computing the mean and standard deviation over all the subjects for each stimulus (e.g., DV = 256°/s) and then eliminating those points that fell outside the range of two standard deviations on either side of the mean value. The mean and standard deviation were then recomputed without the discarded points. The values for the six quantities are listed in Table 1.

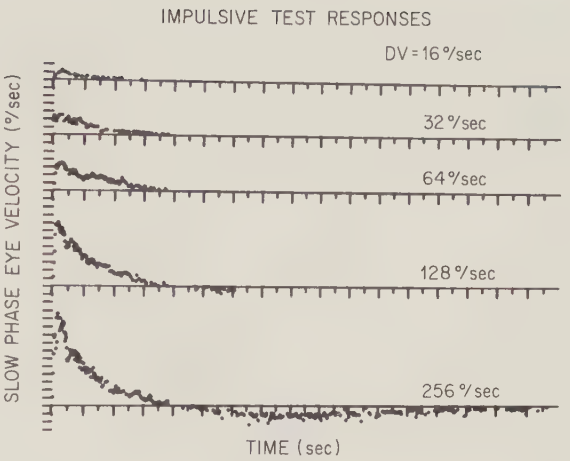

Fig. 1. The slow phase eye velocity of nystagmus responses are shown for progressively greater impulsive stimuli (DV). The horizontal axis is marked every 5 s and the vertical axis is marked every 10 °/s. The axes are the same for all the responses. Up to an impulse of 64 °/s, the response resembles that predicted by the damped pendulum model, but above this magnitude an after-nystagmus appears, beating in the direction opposite to the post-rotatory nystagmus

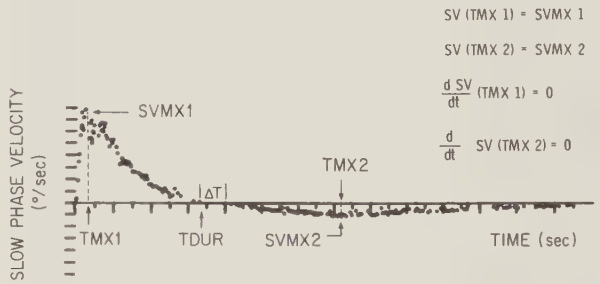

Fig. 2. The six experimental quantities measured in all subjects are shown on this response to an impulse of 256°/s. SVMX1 and TMX1 mark the amplitude and time of occurrence of the peak response during the post-rotatory phase. SVMX2 and TMX2 mark the corresponding quantities during the after-nystagmus phase. TDUR is the time of the last beat of the post-rotatory phase, and ΔT is a measure of time between the last beat of the post-rotatory phase and the first beat of the after-nystagmus. The horizontal and vertical axes are the same as Fig. 1. The four equations on the right allow one to obtain unique values for the time constants. The equation for SV is given in the text

Table 1. The mean values, standard deviations, and number of responses on which these measures are based are given for the six experimental quantities defined in Figure 2. The coefficient of variation (C.V.) is also listed for each measure. The left side gives the appropriate values for a stimulus of 128 °/s, and the right side for 256 °/s. A *t*-test of stimulus magnitude dependence for each quantity is shown at the right. The only quantity independent of stimulus magnitude is TMX2

	μ 128 °/s	σ	N	C.V.	μ 256 °/s	σ	N	C.V.	*t*-test
SVMX1	82.2	20.4	40	25%	130.6	35.0	63	27%	***
SVMX2	10.0	4.2	40	42%	16.2	6.9	66	42%	***
TMX1	2.2	0.6	40	27%	2.7	0.7	65	26%	***
TMX2	69.7	11.6	39	17%	69.6	9.8	65	14%	ns
TDUR	37.1	7.0	40	19%	34.2	5.6	65	16%	*
ΔT	15.0	5.2	39	35%	9.0	3.8	64	42%	***

* = $P < 0.05$
*** = $P < 0.001$
"ns" = not significantly correlated

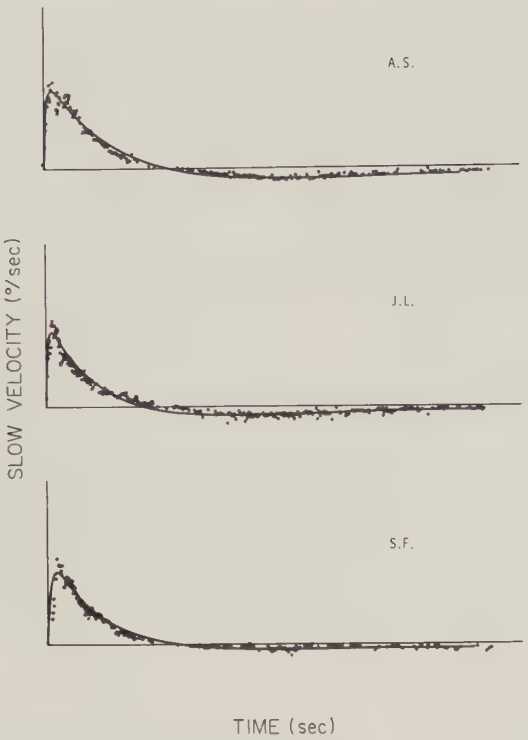

Fig. 3. The excellent fit of the theoretical model (continuous line) to the experimental data (dots) is demonstrated for the three normal subjects. The vertical axis is the slow phase eye velocity in degrees/second, the horizontal axis is the time in seconds. The parameters of the theoretical model were determined separately for each response from the values of the maximum points as in Fig. 2

Theoretical. The basic assumption of the adaptation model is that the subject's response is modified by the recent history of that response. There is imagined to be a central reference level with which changes in the neural signal from the labyrinths are compared (Malcolm and Melvill-Jones, 1970). The slow phase eye velocity is now assumed to be proportional to the *difference* between the neural signal and the reference level. The direction of the slow phase is determined by the sign of this difference. For sufficiently long or powerful stimuli, the reference level is assumed to move away from its steady state position in such a way as to minimize the difference between itself and the neural signal from the labyrinth. The rate at which this reference level moves is described by the adaptation time constant T_A. The decay of the response to the original position of the reference level undershoots the current position of the level, and this change in the sign of the difference between the two manifests itself as after-nystagmus.

The approximate solution to the differential equation that describes this behavior for impulsive stimuli is

$$SV(t)=\frac{DV\cdot T_1\cdot T_2}{K_i(T-T_2)}\left[\left(\frac{T_A}{T_A-T_2}-\frac{T_A}{T_A-T_1}\right)\cdot\exp(-t/T_A)+\frac{T_A}{T_A-T_1}\exp(-t/T_1)-\frac{T_A}{T_A-T_2}\exp(-t/T_2)\right]. \quad (1)$$

"t" is time, and all other quantities have been defined. All the factors in front of the bracket are assumed to be constant. At $t=0$ the terms inside the bracket cancel out as they must since $SV(t=0)$ is zero. The exponentials are arranged in progressively faster acting order. As the last exponential dies away, the curve described by this equation rises to a peak. As the second term begins to diminish, the curve drops back to zero. Eventually the first term dominates and the curve drops below zero, returning at last to zero for the final time. Equation (1) can be evaluated directly at the two maxima by substituting the value of SVMX1 or SVMX2 on the left side and setting t equal to the corresponding value of TMX1 or TMX2 on the right side. This yields two equations. Since there are four unknowns (T_1, T_2, T_A, and K_i), two more equations are necessary. These are obtained by noting that at the times of the maxima, the first derivative of (1) is zero.

Table 2. The mean values, standard deviations, and number of responses on which these measures are based are given for six parameters of the adaptation model. See text for explanation of the parameters. The values for a stimulus of 128 °/s are given on the left side, those for a stimulus of 256 °/s on the right. All the parameters are stimulus dependent except T_1 and T_A

	μ 128 °/s	σ	N	C.V.	μ 256 °/s	σ	N	C.V.	t-test
T_1	18.2	6.6	34	36%	16.6	5.5	52	33%	ns
T_A	95.5	44.0	34	46%	99.5	52.3	51	52%	ns
T_2	0.69	0.26	33	38%	0.97	0.38	50	39%	***
K_i	0.89	0.38	33	43%	1.45	0.52	50	36%	***
SV_{th}	4.2	2.2	31	52%	5.7	2.9	49	51%	*
DV_{min}	70.5	42.2	33	60%	100.9	49.1	47	49%	**

* = $P<0.05$
** = $P<0.01$
*** = $P<0.001$
"ns" = not significantly correlated

Evaluating the first derivative at these two times yields the remaining equations. The system of equations is shown schematically on the right side of Fig. 2. The system is in principle well determined (i.e. four equations and four unknowns), but the resulting set of equations is too complex to solve in this form. They can be simplified by making two assumptions.

The rising part of the post-rotatory phase is governed by the value of T_2. Since the rising part is completed within 2 s after the impulse, T_2 is less than 0.4 s. Theoretical hydrodynamical analyses of the semicircular canal (Steer, 1967; Oman and Young, 1972; Melvill-Jones, 1972), electrophysiologic data from vestibular neurons (Fernandez and Goldberg, 1971), and visual inspection of nystagmus responses indicate a range from 0.001–0.01 s for T_2. Similar studies (van Egmond et al., 1949; Groen, 1956, 1957) have placed the value of T_1 at 10–20 s. T_A varies from 30 s to infinity (Malcolm and Melvill-Jones, 1970; Young and Oman, 1972; Stockwell et al., 1973). These estimates are conservatively broad, but because of the differences in their magnitudes it is possible to drop some terms in the system of equations to achieve the necessary simplification.

A) Since T_A is so much larger than T_2, the ratio $T_A/(T_A - T_2)$ is approximately equal to one.

B) Since TMX2 is so much larger than T_2 (see Table 1), the exponential term in (1) containing T_2 is approximately zero when evaluated at $t = \text{TMX2}$.

These two assumptions are the key to the uniqueness of the set of values that simultaneously solve the four equations. The derivation of the equations is shown in Appendix A.

The procedure for finding the simultaneous solution of these equations can be simply programmed on a minicomputer. The search for the correct triad of time constants usually takes less than five seconds. Once T_1, T_2, and T_A have been determined, the value of K_i can be obtained directly from (1). Several examples of the excellent fit obtained by this technique are shown in Fig. 3.

It is also possible to derive quantitative expressions from the adaptation model for the slow phase velocity threshold (SV_{th}) and the minimum impulse required to produce after-nystagmus. The appropriate equations are given in Appendix A.

Results

The data were screened for outliers in the previously explained manner. CW and CCW responses were considered separately. Table 2 lists the means and standard deviations of the parameters for both stimulus magnitudes, the number of cases remaining after screening, and whether the difference between the mean values for the two magnitudes was statistically significant. The latter was computed using a t-test. One of the basic assumptions of a linear model is that the parameters are independent of the stimulus magnitude. The long time constant (T_1) and the adaptation time constant (T_A) are independent of the stimulus, but the short time constant (T_2) is not. This can be attributed to the torque limitations of the rotatory table. Since the peak acceleration possible with a normal adult subject in the chair is approximately 140°/s², an impulse of 256°/s takes approximately 2 s to accomplish, and this is much too long to allow an accurate measurement of T_2. Note that the average value for TMX1 (Table 1) is also strongly stimulus-dependent, lending support to the idea that the 256°/s impulse influences T_2 estimates because it lasts too long. However, the values of K_i, SV_{th}, and DV_{min} are also stimulus-dependent, although they are expected to be independent of the stimulus magnitude.

Table 3. The correlations between supposedly independent parameters of the adaptation model are given for a stimulus of 128°/s in the lower left triangle, and for a stimulus of 256°/s in the upper right triangle. The correlations are between parameter values obtained for the same response (i.e. either clockwise or counterclockwise), and then tested for correlation over the entire subject population

	T_1	T_A	K_i	SV_{th}
T_1		***	ns	ns
T_A	***		*	*
K_i	*	ns		ns
SV_{th}	ns	ns	ns	

256°/s
128°/s

* = $P < 0.05$
*** = $P < 0.001$
"ns" = not significantly correlated

A second basic assumption of a linear model is that the parameters are independent of each other. This implies that the correlation between any two of the parameters determined from the same subject's response is zero when averaged over the entire normal population. Table 3 shows the various interparameter correlations. Where significant, the level of significance is shown. "ns" means the correlation was not significantly different from zero. Although the correlation between most of the parameters determined for the same conditions of stimulus magnitude and direction is weak or absent as expected, T_1 and T_A are a notable exception. Their strong correlation results from the

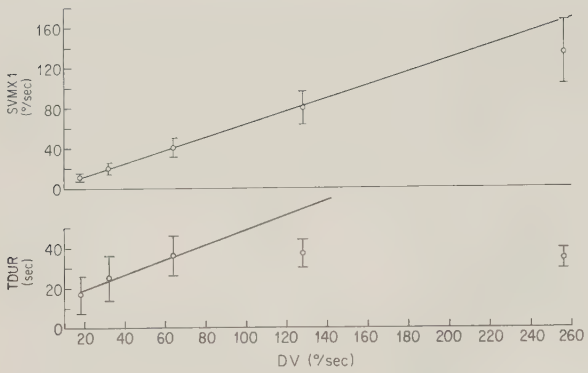

Fig. 4. The upper part of the figure shows the increase in SVMX1 as the impulsive stimulus (DV) is increased. Although there is some evidence of saturation at the highest stimulus magnitude, the increase is largely linear. The lower part of the figure shows the clear saturation in the value of TDUR that occurs above DV = 64°/s. The saturation occurs as a result of adaptation (see text and Fig. 5)

constancy of TMX2 (Table 1) and the form of (2). Why TMX2 should be a constant is uncertain, but the fact that it is results in the breakdown of the assumption that T_1 and T_A are independent parameters in the model.

If the model is to be clinically useful, values of the parameters estimated from CW and CCW responses in the same normal subject should be highly correlated with each other since no directional dependence in the responses of normal subjects is expected. In comparing the CW and CCW responses, this expected high correlation was not found. Neither T_1 nor T_A were significantly correlated between the two directions for either stimulus magnitude. SV_{th} was marginally correlated ($P<0.05$) between CW and CCW directions for DV = 256 °/s, and K_i was marginally correlated ($P<0.05$) between the two directions for both stimulus magnitudes. None of the parameters exhibited a degree of significance stronger than 0.05 between the two directions.

The determination of T_1, T_A, and K_i depends on the values of the experimental SVMX and their respective times of occurrence. It might be argued that some of these results impugning the adaptation model are invalid because they are based on experimental data that are hard to estimate. One test of the precision with which the experimental quantities are determined is the coefficient of variation. The coefficients given in Table 1 indicate that the experimental quantities are determined with reasonable accuracy (i.e., the inter-subject variation is less than 50%).

Another test of the effects of error in determining the experimental quantities can be made by seeing how drastically the values of the model parameters change as the experimental quantities are varied over a range around their estimated true values. A more complete description of this analysis is given in Appendix B, but the results imply that the parameters are not critically sensitive to the experimental quantities.

The correlation analysis of the parameters casts some doubt on the validity of the adaptation model. However, the next result is a strong piece of evidence in favor of the model. Figure 4 demonstrates the saturation of TDUR as the magnitude of the impulsive stimulus is increased. This behavior is in contrast to the prediction of the simple pendulum model. However, it will be shown that this saturation effect is predicted by the adaptation model. Equation (5a) can be rewritten in the form

$$\exp(-t_z)-a\cdot\exp(-t_z\cdot a)=K_i(b-1)(1-a)SV_{th}\cdot T_1^{-1}/DV, \tag{10}$$

where, as before, $a=T_1/T_A$, $t_z=TDUR/T_1$, and $b=T_1/T_2$. K_i, SV_{th}, and DV have their usual definitions. The right side numerator consists of constants, and since "b" is greater than one and "a" is less than one, it is always positive. The saturation effect occurs because the left side of the equation can become negative in the presence of adaptation. The differences between the simple pendulum model and the adaptation model are shown in Fig. 5. The left part of the figure shows the plot of the left and right sides of (10) for "a" = 0 (no adaptation) plotted against $\exp^{-t_z}$. The various horizontal lines represent the right side, which is constant for a given DV and which decreases as DV is increased. The line at 45° is the left side which in the absence of adaptation is equal to the abscissa. The intersection of the two curves defines the predicted TDUR. Smaller values of TDUR lie further to the right than larger values because the abscissa is plotted as $\exp^{-t_z}$. As DV increases, the constant line drops closer to abscissa and the intersection moves continuously toward zero (i.e., larger values of TDUR). Since the left side of (10) in the absence of adaptation is positive all the way to zero, the simple pendulum model predicts that as DV increases, TDUR increases without limit, which contradicts the experimental evidence that TDUR saturates at about 35 s.

The right part of Fig. 5 shows the same plot, but now "a" is not zero. This corresponds to a finite value of T_A. The figure has been plotted for $a=0.19$ which follows from the values of T_1 and T_A for DV = 128 °/s in Table 2. The presence of adaptation causes the left side of (10) to become negative to the right of zero. Since the right side of (10) can never be negative, TDUR can never be greater than the value defined by the point where the left side of the equation crosses the horizontal axis and becomes negative, if the adaptation model is appropriate. For the values derived from the 128 °/s data in Table 2, the maximum TDUR is 37.3 s. For the parameters associated with a DV = 256 °/s, the maxi-

mum TDUR is 35.7 s. Both of these are close to the actual values shown in Fig. 4.

Discussion

The adaptation model was first proposed to account for several effects that were not predicted by the simple pendulum model. The appearance of an after-nystagmus following sufficiently large magnitudes of stimulation is a striking example of such an effect. By introducing a third time constant, the pendulum model could be modified to explain the existence of after-nystagmus. Several other, more quantitative predictions are also possible using the adaptation model. The prediction of a maximum TDUR value, in contrast to the theoretical behavior of TDUR found using the simple pendulum model, and the correct estimate of its value from parametric values obtained using the adaptation model to fit experimental data are also impressive evidence of the model's validity. The values of the parameters derived by the methods described are shown in Table 2. The estimates of T_1 and T_A for the two magnitudes are comparable to the results of others (Malcolm, 1968; Schmid et al., 1971a; Guedry et al., 1971; Stockwell et al., 1973). The magnitude of SV_{th} is comparable to that given by Aschan et al. (1952).

Further support for the model is found in Table 1 where TMX2 appears to be almost identical for the two stimulus magnitudes when averaged over all the subjects. In fact, TMX2 displays an astonishingly constant value from subject to subject, compared to all the other experimental quantities in Table 1, which are strongly dependent on the stimulus magnitude. By comparing the coefficients of variation among the various experimental and parametric quantities listed in Tables 1 and 2, it can be seen that the spread of values for TMX2 is much smaller than any of the other quantities. Furthermore, the constancy of the TMX2 value appears in the data of others as well. Ranacher (1977) has examined impulsive data and found a similar value for TMX2. Guedry (1974) shows a result due to Parsons (1970), obtained from magnitude estimates of subjective angular velocity during and after a 9 °/s² angular acceleration that also reproduces the value of TMX2. This latter result is all the more remarkable since the sensation of rotation requires more complex pathways than does the production of nystagmus. This difference in pathways is confirmed by the frequent observation that time constants derived from measuring subjective sensation and nystagmus responses differ significantly (van Egmond et al., 1949; Aschan et al., 1952). Yet the value of TMX2 is almost identical in the data from these two techniques. Equation (2) predicts that TMX2 will be a constant if T_1 and T_A are independent of the stimulus. The success

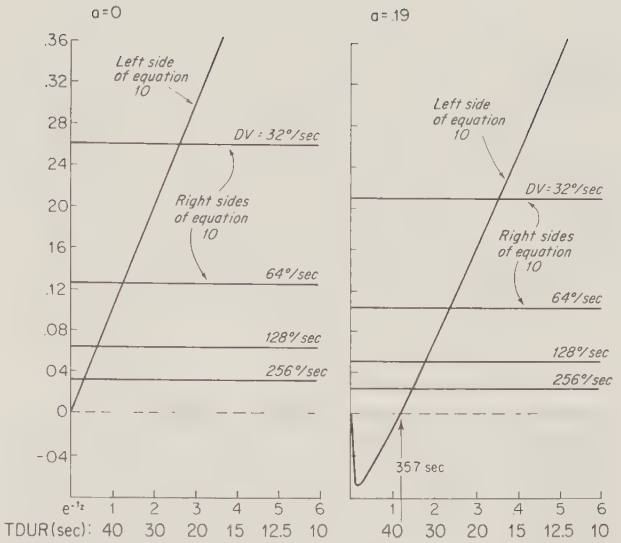

Fig. 5. The theoretical value of TDUR predicted by the simple pendulum model ("a" = 0; no adaptation) is shown graphically in the left part of the figure. The left and right sides of Eq. (10) have been plotted separately. The intersection of the two curves defines the predicted TDUR value. The left side of Eq. (10) for the case of no adaptation is the same as the abscissa, which gives the ascending straight line. The right side of Eq. (10) is a function of the impulsive stimulus (DV). Larger values of DV lie closer to the abscissa. The abscissa is plotted as e^{-t_2}, but the corresponding value of TDUR is given in the scale below that. Note that TDUR theoretically increases without limit as DV increases. The right part of the figure gives the corresponding plot for the adaptation model ("a" = 0.19; adaptation). The scales are the same as above. Note that for "a" non-zero, there is a region where the left side of Eq. (10) becomes negative. Since the right side of Eq. (10) is always positive, there are no solutions for TDUR greater than the value defined by the crossing of the abscissa by the curve representing the left side of Eq. (10). TDUR increases up to this value and then remains a constant no matter how DV is increased. Compare to the values of TDUR shown in Fig. 4

of this prediction is further support for the adaptation model.

Finally, it is an experimentally confirmed fact that after-nystagmus is seldom seen for DV = 64 °/s, appears commonly for DV = 128 °/s, and is almost inevitable for DV = 256 °/s. This implies that the minimum DV needed to produce after-nystagmus lies somewhere between 64 and 128 °/s. Using the constants in Tables 1 and 2 and (9), the minimum stimulus (DV_{min}) required is seen to be between 70 and 100 °/s (Table 2). This range was obtained by computing DV_{min} for each subject, and averaging DV_{min} over all the subjects for either stimulus magnitude. Another estimate of DV_{min} is obtained by using the average values of the required parameters from the tables to evaluate (9). This yields a narrower range for DV_{min}, lying between 78 and 86 °/s. Both estimates support the validity of the adaptation model.

On the other hand, there are several indications that this model does not satisfactorily describe the responses of certain subjects. Among those data using

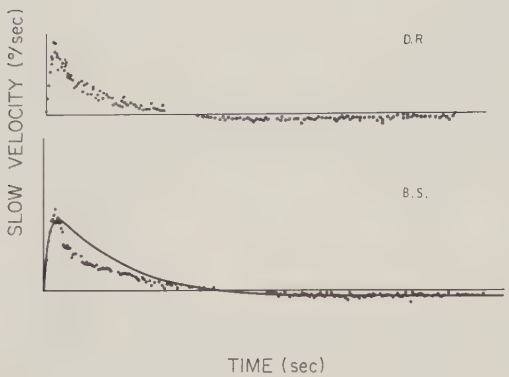

Fig. 6. The upper figure shows a response that had no simultaneous solution for the time constants. It appears to be similar in most respects to those responses shown in Fig. 3. The same subject had a solution for the response in the opposite direction. All experimental quantities are similar except the value of SVMX1, which is 79°/s for the response with a solution and 54°/s for the response pictured here. The ratio of SVMX1/SVMX2 for the response with a solution was 8, the ratio for the response without was 6. The lower figure had a solution for the response, but as can be seen from the superimposed line of the solution, it is not a good fit. There appear to be multiple relaxation processes occurring during the post-rotatory phase, so a model using a single exponential term to describe the descending part of the response would be inappropriate

DV = 128 °/s, 15% of the subjects failed to have any solution to the set of coupled equations. For the data corresponding to DV = 256 °/s, 19% failed to have solutions. No solution means that either the only solution found was for $T_1 = T_A$, which violates the assumption that T_1 is less than T_A, or that there was no mutual intersection of the four equations, so that no triad of time constants satisfied all the equations simultaneously. These failure rates agree closely with Aschan's findings that 20% of the observed normal subjects failed to behave as predicted theoretically (Aschan, 1955, 1956). The failure of certain subjects to have any simultaneous solution to the equations is the more surprising since a study of parameter sensitivity (see Appendix B) can demonstrate that wide variations are allowable in the parametric values as the experimental values are changed. Such flexibility suggests that these failures are fundamentally different from the rest of the responses.

By varying the experimental values used in the equations the causes of the failures could sometimes be identified. In all cases where no solution could be found, the ratio of SVMX1/SVMX2 was low, typically 3–6, relative to data that did have solutions. Such data typically had ratios of 12–18, but there were solutions for ratios as low as 6. These suggest that the failures are due to excessive deviations from the average response in either the post-rotatory phase (SVMX1 too low) or the after-nystagmus phase (SVMX2 too high). Note that had the data been fitted by estimating the goodness of fit visually, these failures would not have been seen. Figure 6A shows an impulsive response that failed to have a simultaneous solution in the sense discussed above. There is nothing overtly odd about the appearance of this response.

There is a second type of failure that is completely different in its nature. Figure 6B shows a response that did have a solution, but which is inappropriate. The response clearly shows a discontinuity in the post-rotatory decline, perhaps indicative of multiple processes (two T_1 time constants), or some nonlinearity in the subject's response. The correct parameters are not fixed by the SVMX and TMX values.

These types of failures indicate that the adaptation model has a limited validity in describing the responses of some normal subjects. More complex models have been suggested. A model proposed by Schmid et al. (1970, 1971a, b, 1976) performs very well in predicting the vestibular response to a complex series of rotatory stimuli. The model is essentially similar to the one presented here, but with the addition of a nearly-perfect integrator representing a saccadic pathway in parallel with the vestibulo-ocular pathway. In its entirety, the model contains six adjustable parameters. There are several reasons for admiring this model, but the experimental data is simply not adequate for accurately evaluating six experimental parameters and one must resort to the arbitrary assignment of certain parameters in order to fit the others.

Furthermore, there are some indications that even when the adaptation model fits the data well and has well-defined solutions, some of the assumptions inherent in the model's derivation break down. The linear model assumes that the parameters of the model are constant over time, independent of magnitude, and linearly independent of each other. But Melvill-Jones and Milsum (1970) found a relationship between the neural gain and the stimulus amplitude (angular velocity) at a given sinusoidal frequency, such that the gain was proportional to the angular velocity raised to the -0.28 power. Since the gain is directly proportional to $T_1 \cdot T_2$ [see (1)], this implies that these time constants change their values with the magnitude of the stimulus. We have found that all the parameters vary with the magnitude except T_1 and T_A, and these two are strongly correlated with each other (Tables 2 and 3).

Groen (1957b) has pointed out that the values for the long time constant can be altered according to the experimental conditions (e.g. visual fixation, staring in the dark, closed eyes, mental arithmetic, etc.) in much the same way that the gain of the VOR can be altered (Barr et al., 1976; Robinson, 1976; Gonshor and Melvill-Jones, 1976a, b). This variation in T_1 is not surprising since the gain of a second-order linear system such as the damped pendulum model is related

to the time constants of the system (Milsum, 1966). The sensitivity of T_1 to experimental conditions could presumably be stabilized by the proper control of these conditions, were it not for its coupling to T_A. The variations in T_A do not seem to bear any relationship to the experimental conditions. The constancy of TMX2 suggests that all subjects peak at the same time during after-nystagmus, regardless of differences in the post-rotatory phase. This behavior is consistent with a central, stimulus-independent process that is also independent of such physical factors as endolymphatic viscosity or canal size. Young and Oman (1969) agree that "it has become increasingly evident that behavioral dynamics deviate significantly from the second-order characteristics of the fluid dynamics because of intervening neurological processes". They include a central nervous system delay time of 0.3 s in their model. They also include a threshold velocity of 3.8 °/s which agrees with the SV_{th} estimate of 4–6 °/s derived from our data. The large coefficient of variation for T_A (46–52 %) is probably responsible for the variations in T_1 (33–36 %), as a result of the coupling between the two time constants.

Finally, there is the question of whether the parameters of the adaptation model are clinically useful. That is, if the parametric values are determined for a group of normal subjects and applied to patients suffering from well-determined vestibular dysfunctions, do these parameters distinguish between the normal subjects and the patients in any meaningful way? Although McClure et al. (1976) reported some success in detecting patients by abnormalities in their rotatory responses, we applied this technique to groups of patients with peripheral and central lesions and failed to consistently separate their responses from those of the normal subjects. This work is more fully reported in Sills and Honrubia (1977) and Baloh et al. (1978).

In conclusion, the technique presented in this paper has been applied to the impulsive data from normal subjects to deduce the numerical values of a number of parameters in an attempt to evaluate the validity of the adaptation model. Although the original pendulum model was based on considerations of cupular dynamics, the data used here derive from movements of the eyes. Thus, the parameters actually relate to the VOR itself. The values of the parameters, their dependence on stimulus direction and magnitude, and their correlations with each other have been assessed. The adaptation model has several striking successes, but equally true, there are a number of times when the assumptions of the model are invalid. There are a substantial (20 %) number of responses from normal subjects that cannot be adequately described by the adaptation model. The agreement between the numerical estimates of the parameters made using this technique and estimates made by others using different techniques seems to rule out the possibility that these disagreements between theory and experiment are due to the technique itself. Instead, we conclude that the erratic appearance of various central effects probably precludes the consistent use of any single model.

Appendix A

In deriving the four equations needed to determine values of T_1, T_2, T_A, and K_i, it is convenient to introduce the following dimensionless quantities: $a = T_1/T_A$; $b = T_1/T_2$; $\mathrm{TM1} = \mathrm{TMX1}/T_1$; $\mathrm{TM2} = \mathrm{TMX2}/T_1$; and $R = \mathrm{SVMX1/SVMX2}$. With these, using assumptions A and B (see text), and evaluating (1) and its first derivative at the two maxima, one can derive three equations:

$$\mathrm{TM2} = [2n(a)]/(a-1), \tag{2}$$

$$b \cdot \exp(-\mathrm{TM1} \cdot b) = [a^2 \cdot \exp(-\mathrm{TM1} \cdot a) - \exp(-\mathrm{TM1})]/(a-1), \tag{3}$$

$$R = \frac{a \cdot \exp(-TM1 \cdot a) - \exp(-\mathrm{TM1}) + (1-a)\exp(-\mathrm{TM1} \cdot b)}{a \cdot \exp(-\mathrm{TM2} \cdot a) - \exp(-\mathrm{TM2})}. \tag{4}$$

The first two equations are the first derivative of (1) evaluated at TMX2 and TMX1 respectively. The last is the ratio of (1) evaluated at TMX1 and TMX2. These three equations contain three unknowns; a, b, and T_1, or equivalently T_A, T_2, and T_1. To find the simultaneous solution of these equations, T_1 is assigned an arbitrary value. Equation (2) is solved for the unique value of "a" corresponding to the T_1 value. The value of "a" is unique because the right side of (2) is a monotonically descending function of "a". This value of "a" is used to fix the right side of (3) which determines the value for "b". The left side of (3) is actually multivalued (i.e., the same value on the right side determines two values of "b"), but "b" can still be determined uniquely as follows. The function $b \cdot \exp(-\mathrm{TM1} \cdot b)$ has its maximum value at $b = 1/\mathrm{TM1}$. From the definitions of "b" and TM1, this is where $T_2 = \mathrm{TMX1}$. Since T_2 is actually much less than TMX1, the correct values of "b" will always lie to the right of the maximum value of $b \cdot \exp(-\mathrm{TM1} \cdot b)$ and in this region the left side of (3) is a monotonically descending function of "b".

Finally, the arbitrary value of T_1 and the resultant values of "a" and "b" are used to evaluate the right side of (4), which is compared to the experimentally determined value of R on the left side. If the two sides do

not agree, the value of T_1 is changed in a direction consistent with the observed difference between the two sides of (4) and the procedure is repeated starting with (2).

To derive the expressions for SV_{th} and DV_{min}, it is convenient to introduce two more dimensionless quantities, $t_z = TDUR/T_1$ and $\Delta t = \Delta T/T_1$, and evaluate (1) at the two times, $t = TDUR$ and $t = TDUR + \Delta T$. Since TDUR is much greater than T_2, the last exponential in (1) can be dropped so that

$$SV_{th} = \frac{DV \cdot T_1}{K_i(b-1)(a-1)} [a \cdot \exp(-t_z \cdot a) - \exp(-t_z)], \quad (5a)$$

$$-SV_{th} = \frac{DV \cdot T_1}{K_i(b-1)(a-1)} [a \cdot \exp -(t_z + \Delta t)a - \exp -(t_z + \Delta t)]. \quad (5b)$$

Note that if $\Delta t = 0$, $SV_{th} = -SV_{th}$ which implies that $SV_{th} = 0$, or that there is no threshold. This underlines the connection between the existence of a threshold velocity and the appearance of a pause (ΔT) between the post-rotatory and after-nystagmus phases.

Using a first-order expansion for SV_{th} around $t = t_z$

$$SV(t_z + \Delta t) \cong SV(t_z) + dSV/dt]_{t=t_z} \Delta t . \quad (6)$$

But $SV(t_z + \Delta t) - SV(t_z) = -2SV_{th}$ by definition so that

$$2SV_{th} = -dSV/dt]_{t=t_z} \Delta t . \quad (7)$$

This expression assumes that the transition through the threshold region is linear so that the difference between SV evaluated at $t = t_z + \Delta t$ and $t = t_z$ can be approximated by $2SV_{th}/\Delta t$. Figures 1 and 2 show several examples of the apparent linearity of SV through the threshold region. dSV/dt can be obtained by taking the first derivative of (5a) so that

$$SV_{th} = -\frac{DV \cdot \Delta T}{2 \cdot K_i(b-1)(a-1)} [-a^2 \cdot \exp(-t_z \cdot a) + \exp(-t_z)]. \quad (8)$$

DV, K_i, a, b, ΔT, and t_z are all known and SV_{th} may be evaluated. In principle, SV_{th} could have been determined directly from either (5a) or (5b), but near the threshold the duration of the nystagmus beats lengthens, and an error of 1–2 s in estimating TDUR is possible if only one measurement is used. By using the last beat of the post-rotatory phase and the first beat of the after-nystagmus phase, this error is minimized as much as possible.

By definition SVMX2 must be greater than SV_{th} or there would be no after-nystagmus. Since an expression for SVMX2 can be explicitly written using (1), it is possible to express the minimum impulse required to produce after-nystagmus by requiring that $SVMX2 = SV_{th}$. This leads to

$$DV_{min} = \frac{DV \cdot \Delta t_{DV}}{2} \left[\frac{a^2 \cdot \exp(-t_z \cdot a) - \exp(-t_z)}{a \cdot \exp(-TM2 \cdot a) - \exp(-TM2)}\right], \quad (9)$$

where Δt_{DV} is the experimentally observed dimensionless time corresponding to the actual impulse (DV) used to induce the experimental response.

Appendix B

The determination of K_i and the time constants T_1 and T_A depends on the values of the two SVMX and their respective times of occurrence. A random sample of nine subjects was selected and the variation in the values of these three parameters as a function of the four experimental quantities was studied. The sensitivity to changes was estimated from the slope of the line that resulted from holding all but one of the experimental quantities constant and varying the remaining one over a range of values. Where the parameters showed a significant curvature (e.g. T_A as a function of SVMX2), a linear approximation to the slope over a small range of values to either side of the true value was used.

The mean slopes and standard deviations for T_1, T_A, and K_i as functions of the four experimental

Table 4. Mean slopes of parameters versus experimental data

	TMX1	TMX2	SVMX1	SVMX2	N
T_A	24.9±15.1	0.8 ±0.8	1.3 ±0.4	−17.9 ±13.2	9
T_1	−1.6±0.9	0.3 ±0.1	−0.13 ±0.14	1.0 ± 0.7	9
K_i	0.7± 0.2	−0.006±0.003	−0.011±0.006	− 0.012± 0.004	9

Equivalence table

T_A (s)	±25	±4	±26	±54
T_1 (s)	± 1.6	±1.5	± 2.6	± 3
K_i (s)	± 0.7	±0.03	± 0.2	± 0.03

Values from Table 2

μ=99.5 ±52.3 s
μ=16.6 ± 5.5 s
μ= 1.45± 0.52 s

quantities are given in Table 4, based on nine subjects. To give more meaning to the slope values, an equivalence table is presented in the lower part of Table 4. The values are derived by multiplying the slopes from the upper part of Table 4 by the following estimates of uncertainty in the measurement of the experimental quantities: TMX1 = ±1 s, TMX2 = ±5 s; SVMX1 = ±20 °/s; SVMX2 = ±3 °/s. These estimates of uncertainty are based on the analysis of data such as in Figs. 1 and 2. From this equivalence table one can see that uncertainties in the estimates of SVMX2 can lead to large variations in T_A and T_1, comparable to the inter-subject variance. The estimated value of K_i depends strongly on the value of TMX1. For these reasonable estimates of the uncertainty in the experimental quantities, the uncertainty in the values of the parameters is less than the inter-subject variability. This implies that the parameters are not critically sensitive to the estimates of the experimental quantities.

References

Aschan, G.: The mechanism of the cupula ampullaris in man. Acta Soc. Medicorum Upsaliensis **60**, 77–88 (1955, 1956)

Aschan, G., Nylen, C.O., Stahle, J., Wersall, R.: The rotation test. Cupulometric data from 320 normals. Acta Otolaryngol. (Stockh.) **42**, 451–459 (1952)

Baloh, R.W., Sills, A.W., Honrubia, V.: Impulsive and sinusoidal rotatory testing. A comparison with results of caloric testing Laryngoscope (in press)

Barr, C.C., Schultheis, L.W., Robinson, D.A.: Voluntary, non-visual control of the human vestibulo-ocular reflex. Acta Otolaryngol. (Stockh.) **81**, 365–375 (1976)

Dohlman, G.: On the mechanism of transformation into nystagmus on stimulation of the semicircular canals. Acta Otolaryngol. (Stockh.) **26**, 425–442 (1938)

van Egmond, A.A.J., Groen, J.J., Jongkees, L.B.W.: The mechanics of the semicircular canal. J. Physiol. **110**, 1–17 (1949)

van Egmond, A.A.J., Groen, J.J., Jongkees, L.B.W.: The function of the vestibular organ. Pract. ORL **14**, Suppl. 2 (1952)

Fernandez, D., Goldberg, J.M.: Physiology of peripheral neurons innervating the semicircular canals of the squirrel monkey. II. The response to sinusoidal stimulation and the dynamics of the peripheral vestibular system. J. Neurophysiol. **34**, 661–675 (1971)

Gilman, S., Dirks, D.D., Hunt, S.: Measurement of head movement during auditory localization (submitted for publication)

Gonshor, A., Melvill-Jones, G.: Short-term adaptive changes in the human vestibulo-ocular reflex arc. J. Physiol. (Lond.) **256**, 361–379 (1976a)

Gonshor, A., Melvill-Jones, G.: Extreme vestibulo-ocular adaptation induced by prolonged optical reversal of vision. J. Physiol. (Lond.) **256**, 381–414 (1976b)

Groen, J.J.: The semicircular canal system of the organs of equilibrium. II. Central inhibition. Phys. Med. Biol. **1**, 225–242 (1956, 1957)

Groen, J.J.: Cupulometry. Laryngoscope **67**, 894–905 (1957a)

Groen, J.J.: Adaptation. Pract. ORL **19**, 524–530 (1957b)

Guedry, F.E.: Psychophysics of vestibular sensation. In: Handbook of sensory physiology. Vol. VI. Part 2, p. 40. Kornhuber, H.H., ed. Berlin, Heidelberg, New York: Springer 1974

Guedry, F.E., Stockwell, C.W., Norman, J.W., Owens, G.G.: Use of triangular waveforms of angular velocity in the study of vestibular function. Acta Otolaryngol. (Stockh.) **71**, 439–448 (1971)

Hallpike, C.S., Hood, J.D.: The speed of the slow component of ocular nystagmus induced by angular acceleration of the head: Its experimental determination and application to the physical theory of the cupular mechanism. Proc. Roy. Soc. (Biol.) **141**, 216–230 (1953)

Julius, R.S.: The sensitivity of exponentials and other curves to their parameters. Computers Biomed. Res. **5**, 473–478 (1972)

Konrad, H.R., Sills, A.W., Baloh, R.W., Honrubia, V.: The impulsive test in man. Trans. AAOO **82**, ORL-232-238 (1976)

Lorente de Nó, R.: Ausgewählte Kapitel aus der vergleichenden Physiologie des Labyrinths. Die Augenmuskelreflexe beim Kaninchen und ihre Grundlagen. Ergeb. Physiol. **32**, 73–242 (1931)

Lowenstein, O., Sand, A.: The individual and integrated activity of the semicircular canals of the elasmobranch labyrinth. J. Physiol. (Lond.) **99**, 89–101 (1940)

Malcolm, R.: Progress in vestibular modeling. Part III. A quantitative study of vestibular adaptation in humans. Fourth symposium on the role of the vestibular organs in space exploration. NASA SP-**187**, 369–380 (1968)

Malcolm, R., Melvill-Jones, G.: A quantitative study of vestibular adaptation in humans. Acta Otolaryngol. (Stockh.) **70**, 126–135 (1970)

van Mastrigt, R.: Constant-step approximation of multi-exponential signals using a least-squares criterion. Comput. Biol. Med. **7**, 231–247 (1977)

McClure, J.A., Lycett, P., Bicker, G.R.: A quantitative rotational test of vestibular function. J. Otolaryngol. **5**, 279–288 (1976)

Meiry, J.L.: The vestibular system and human dynamic space orientation. ScD thesis, MIT Department of Aeronautics and Astronautics, Man-Vehicle Laboratory (1965)

Melvill-Jones, G.: Transfer function of labyrinthine volleys through the vestibular nuclei. Prog. Brain Res. **37**, 139–156 (1972)

Melvill-Jones, G., Barry, W., Kowalsky, N.: Dynamics of the semicircular canals compared in yaw, pitch, and roll. Aerospace Med. **35**, 984–989 (1964)

Melvill-Jones, G., Milsum, J.H.: Characteristics of neural transmission from the semicircular canals to the vestibular nuclei of cats. J. Physiol. (Lond.) **209**, 295–316 (1970)

Milsum, J.H.: Biological control systems analysis. New York: McGraw-Hill 1966

Oman, C.M., Young, L.R.: The physiological range of pressure difference and cupula deflections in the human semicircular canal. Theoretical considerations. Acta Otolaryngol. (Stockh.) **74**, 324–331 (1972)

Parsons, R.D.: Magnitude estimates of the oculogyral illusion during and following angular acceleration. J. Exp. Psychol. **84**, 230–238 (1970)

Ranacher, G.: Die Nystagmusdauer nach Drehbeschleunigung in Theorie und Experiment. Arch. ORL **214**, 313–318 (1977)

Robinson, D.A.: Adaptive gain control of vestibuloocular reflex by the cerebellum. J. Neurophysiol. **39**, 954–969 (1976)

Schmid, R.M.: Systems analysis of the vestibulo-ocular system. Fifth symposium on the role of the vestibular organs in space exploration. NASA SP-**314**, 237–249 (1970)

Schmid, R., Lardini, F.: On the predominance of anti-compensatory eye movements in vestibular nystagmus. Biol. Cybernetics **23**, 135–148 (1976)

Schmid, R., Stefanelli, M., Mira, E.: Mathematical modeling. A contribution to clinical vestibular analysis. Acta Otolaryngol. (Stockh.) **72**, 292–302 (1971a)

Schmid, R.M., Stefanelli, M.V., Anzaldi, E.G.: A mathematical model accounting for nonzero final value of the cumulative eye position in post-rotational nystagmus. Digist 9th Int. Conf. Med. Biol. Engineering, 37 (1971b)

Sills, A.W., Baloh, R.W., Honrubia, V.: Caloric testing. II. Results in normal subjects. Ann. ORL Suppl. **43** (1977)

Sills, A.W., Honrubia, V.: A new method for determining impulsive time constants and their application to clinical data. Otolaryngology **86**, ORL 81-89 (1978)

Steer, R.W.: The influence of angular and linear acceleration and thermal stimulation on the human semicircular canal. ScD Thesis, MIT Department of Aeronautics and Astronautics, Man-Vehicle Laboratory (1967)

Steinhausen, W.: Über den Nachweis der Bewegung der Cupula in der intakten Bogengangsampulle des Labyrinths bei der natürlichen rotatorischen und calorischen Reizung. Pflügers Arch. ges. Physiol. **228**, 322–328 (1931)

Steinhausen, W.: Über die Beobachtung der Cupula in den Bogengangsampullen des Labyrinths des lebenden Hechts. Pflügers Arch. ges. Physiol. **232**, 500–512 (1933)

Stockwell, C.W., Gilson, R.D., Guedry, F.E.: Adaptation of horizontal semicircular canal responses. Acta Otolaryngol. (Stockh.) **75**, 471–476 (1973)

Young, L.R.: A control model of the vestibular system. In: Technical and biological problems of control – A cybernetic view, p. 543. Iberall, A.S., Reswick, J.B. (eds.) Distributed by Instrument Soc. America (1968)

Young, L.R., Oman, C.M.: Model for vestibular adaptation to horizontal rotation. Aerospace Med. **40**, 1076–1080 (1969)

Received: May 8, 1978

Dr. A. W. Sills
Div. of Head and Neck Surgery
UCLA School of Medicine
Los Angeles, CA90024, USA

Biol. Cybernetics 30, 221—230 (1978)

Biological Cybernetics

A Study of the Motor Unit Action Potential by Means of Computer Simulation

P. A. M. Griep, K. L. Boon, and D. F. Stegeman

Bio-Information Group, Department of Electrical Engineering, Twente University of Technology, Enschede, The Netherlands

Abstract. In order to study the motor unit action potential a computer simulation model was developed. It is based on the superposition of single muscle fibre potentials of the fibres belonging to the motor unit. The parameters which characterize each fibre (spatial position, diameter, and a dispersion of arrival time of the potential at the electrode) are chosen from statistical distributions which can be derived from anatomical and physiological data. The electrode type, position and dimensions can be specified. Simulated motor unit action potentials are presented in the time and frequency domain. The simulation results refer to (1) the influence of the electrode position and dimensions with respect to the motor unit territory, (2) the meaning of this model for the study of pathological phenomena, (3) the variability of some parameters characterizing the motor unit, (4) the selectivity of uni- and bipolar electrodes and finally (5) the influence of the geometrical situation of the motor end-plates within the muscle, on the shape of motor unit action potentials.

1. Introduction

The motor unit action potential (muap; plural muaps) of a certain motor unit can be considered as a summation of fibre potentials of all muscle fibres belonging to that motor unit. In Figure 1a a schematic drawing of one motor unit in a muscle is given. The shape of a muap and its power spectrum depend on a number of factors. The most important are: scattering and diameter of the muscle fibres of one motor unit, position and dimension of the electrode, conduction velocity, and the time dispersion meaning that there is a spread in the arrival time of the fibre potentials at the electrode. This dispersion is mainly caused by the different conduction velocities and lengths of the nerve twigs and the different positions of the motor end-plates. All these factors together form an obstacle for analytic calculation of possible muap shapes. However by means of simplified models some mathematical investigations have been carried out by George (1970) and very extensively by Lindström et al. (1976).

These models are interesting, especially in order to see how certain measurable parameters may reveal information about the motor unit, but their meaning with respect to a number of specific muap shapes is limited. The present study was undertaken to explore possible muap shapes and their relation to important anatomical and physiological parameters by means of computer simulation. This simulation was based on the fact that with a computer it is feasible to calculate the potential contribution of each muscle fibre of a motor unit separately. The potential contribution of a muscle fibre as used in the model, is described by Lorente de Nó (1947) and Rosenfalck (1969). Ekstedt and Stålberg (1973) described the influence of the size of the leading-off surface.

2. Methods

2.1. General Principle

The muap is calculated basically as a summation of separate muscle fibre potentials according to the superposition principle. The spatial position, the diameter of each fibre and the dispersion in arrival time of each fibre potential at the electrode can be chosen according to probability density functions (pdf's) or if preferred in a deterministic way. These pdf's depend of course on the muscle being studied and can be estimated from anatomical and physiological data. The proposed model is very well suited for general use because it can be adjusted in a simple way to anatomical and physiological data. In this study unipolar or bipolar "registrations", based on data of the *m.* soleus of the rat have been calculated. The potential of the indifferent electrode, when unipolar "registrations" were made, was taken as zero: in reality this means

0340-1200/78/0030/0221/$02.00

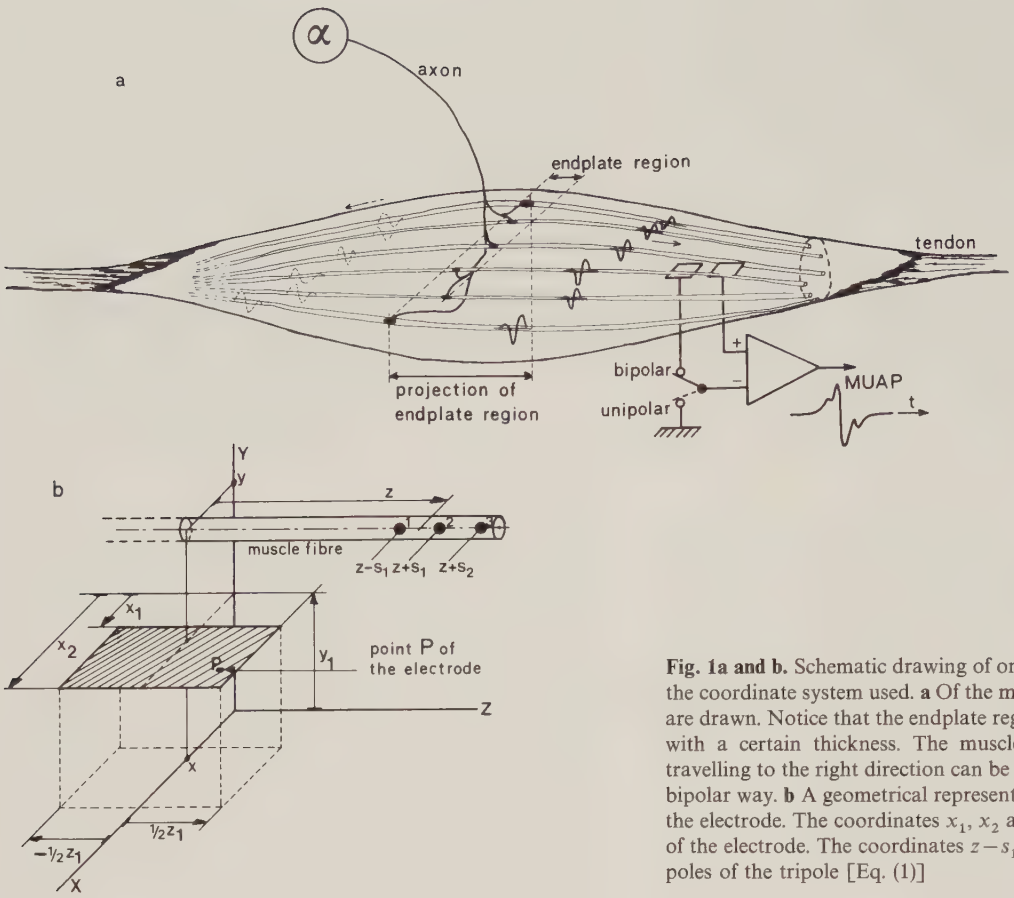

Fig. 1a and b. Schematic drawing of one motor unit in a muscle and the coordinate system used. **a** Of the motor unit only 6 muscle fibres are drawn. Notice that the endplate region resembles to a skew zone with a certain thickness. The muscle fibre potentials which are travelling to the right direction can be "measured" in an unipolar or bipolar way. **b** A geometrical representation of one muscle fibre and the electrode. The coordinates x_1, x_2 and z_1 refer to the dimensions of the electrode. The coordinates $z-s_1$, $z+s_1$ and $z+s_2$ refer to the poles of the tripole [Eq. (1)]

that such an electrode was thought to be placed in an electrical passive region.

The model was programmed in Fortran and can be executed on a minicomputer[1]. In Fig. 2 a typical example of a simulation result is shown. In the upper right quadrant the fibre positions and the electrode position and dimensions are drawn. The electrode position is intraterritorial in this example. In the upper left quadrant the dispersion, the distance to the electrode-centre and the fibre diameter for each fibre are drawn. The calculated muap is plotted in the lower left quadrant. The power spectrum is shown in the lower right quadrant since it is an important characteristic of the motor unit (Lindström, 1974; Leifer and Pinelli, 1976). In this study it will be shown how the coordinates of the low frequency maximum are related to specific anatomical and physiological data of the motor unit.

1 A detailed description and program listing is given in an internal technical report MUAP8 which can be obtained on request

2.2. *Calculation of a Fibre Potential*

The muscle tissue was considered as an homogeneous, anisotropic, unlimited volume conductor. The potential field in a volume conductor is subjected to the Laplace equation (Rosenfalck, 1969). Using the assumption that the potential field of one fibre is caused by three moving point-shaped current sources, this field can easily be described by the tripole model of Rosenfalck. It was assumed that the measuring electrode was placed far enough from the end-plate region so that the end-plate potentials and fibre potentials originating from parts of the fibre on the other side of the end-plate can be neglected. The potential contribution φ'_p for an isotropic medium of each muscle fibre with coordinates (x, y) to a point p (see Fig. 1b) with coordinates (x_i, y_i, z_i) of the electrode was calculated according to the tripole model mentioned above. It was assumed that the amplitude increases proportionally to the square of the fibre diameter (Lorente de Nó, 1947; Håkansson, 1957).

$$\phi'_p(r,z)=\frac{a^2\sigma_i K}{4\sigma_e}\left\{\frac{\overbrace{\alpha}^{\text{Pole}}}{\sqrt{(z-s_1-z_i)^2+r^2}}-\frac{1}{\sqrt{(z+s_1-z_i)^2+r^2}}+\frac{1-\alpha}{\sqrt{(z+s_2-z_i)^2+r^2}}\right\}, \tag{1}$$

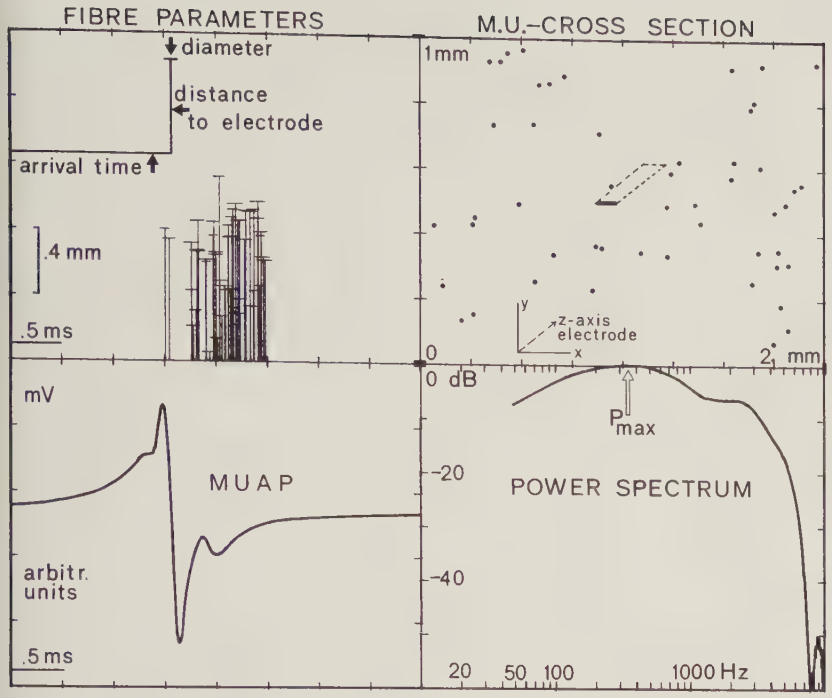

Fig. 2. Example of a computer plot of a muap simulation. Besides the muap and its power spectrum the parameters used are also drawn. In the upper left quadrant for each fibre; the dispersion of arrival time, distance to the electrode-centre and fibre diameter are presented as indicated. In order to visualize the main parameters, in the upper right quadrant a cross section is given showing the muscle fibres of the motor unit and the electrode. The horizontal line of the electrode is an exact representation of the X and Y coordinates of this electrode (see also Fig. 1b). The electrode was directed along the Z-direction which is perpendicular to the drawing. This is indicated in perspective by means of a dotted oblique line. Using these data the muap is calculated. It is presented in the lower left quadrant. Finally a normalized power spectrum of this muap is plotted in the lower right quadrant (to log/dB scales). The arrow at the power spectrum indicates the low frequency maximum

In (1): a represents the fibre diameter; a is chosen from pdf; r represents the distance from point P to a muscle fibre $r^2=(x-x_i)^2+(y-y_i)^2$; x and y are chosen from a pdf; z represents the position of the tripole on the muscle fibre $z=z_0+d-v\cdot t$; the dispersion d is chosen from a pdf, and v represents the velocity of the action potential and t is time; σ_i and σ_e are the conductivities in the intracellular and extracellular medium; K is a constant.

The function ϕ_p for a cylindrically *anisotropic* medium can be easily obtained by taking:

$$\phi_p(r,z)=3\cdot\phi'_p(r\sqrt{5},z)\left(\text{Rosenfalck '69}:\frac{\sigma_r}{\sigma_z}=1/5\right). \quad (2)$$

The tripole parameters (s_1,s_2,α) were assumed to be equal for each fibre. A possible relation between diameter and conduction velocity can be taken into account by adjusting the pdf of arrival times.

2.3. *Influence of Electrode Dimensions*

The potential of a not point-shaped electrode was approximated by spatial integration of the function ϕ_p with respect to the electrode surface A

$$\phi_{el}=\frac{1}{A}\int\int_A \phi_p(r,z)dA. \quad (3)$$

For the calculation of this integral it was convenient to take a rectangular lead-off area parallel to the muscle fibres. In Figure 1b the position of the electrode and one muscle fibre is shown to indicate the coordinates used. The solution of the integral (3) was obtained in two steps:

Step 1: integration in the X-direction in an analytic way (Boon and Griep, 1976). In order to evaluate the contribution of the different poles [see Eq. (1)] the contribution of Pole 1 to the electrode potential ϕ_{el} is considered,

$$\begin{aligned}\phi_{el;\,Pole\,1}&=\frac{1}{A}\int\int\phi_{p;\,Pole\,1}(r,z)dA\\&=\frac{1}{A}\int\int 3\cdot\phi'_{p;\,Pole\,1}(r\sqrt{5},z)dA\\&=\frac{a^2\sigma_i K}{4\sigma_e}\cdot\frac{3}{A}\cdot\int_{-1/2z_1}^{+1/2z_1}\underbrace{\int_{x_1}^{x_2}\frac{\alpha}{\sqrt{(z-s_1-z_i)^2+r^2\cdot 5}}dx_i}_{I_x}\cdot dz_i.\end{aligned} \quad (4)$$

For the solution of integral I_x see the Appendix.

Step 2: Integration in the Z-direction by means of a numerical convolution (Griep and Boon, 1977).

Equation (4) can be written as:

$$\phi_{el;\,Pole\,1}(z)=\frac{a^2\sigma_i K}{4\sigma_e}\cdot\frac{3}{A}\int_{-1/2z_1}^{+1/2z_1} I_x(z-z_i)\cdot dz_i, \quad (5)$$

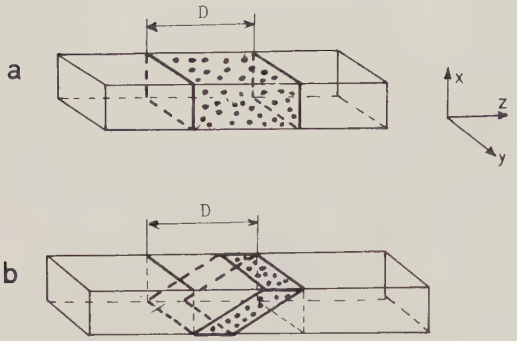

Fig. 3a and b. Distribution of arrival times. In these two figures the muscle is represented very schematically as a bar, the position of each dot represents an arrival time of a single fibre potential to the electrode. Two situations are drawn: **a** the arrival times are randomly spread in a rectangular zone; **b** the arrival times are randomly spread in a relatively small oblique zone (mainly due to the possible positions of the end-plates)

or in the time domain:

$$\phi_{\text{el;Pole 1}}(t) = \frac{a^2\sigma_i K}{4\sigma_e}\cdot\frac{3}{A}\cdot -v \int_{+1/2\frac{z_1}{v}}^{-1/2\frac{z_1}{v}} I_x(vt-vt_i)dt_i$$

$$= \frac{a^2\sigma_i K}{4\sigma_e}\cdot\frac{3}{A}\cdot v \int_{-\infty}^{+\infty} I_x(vt-vt_i)h(t_i)dt_i \quad \text{with} \quad h(t_i)$$

$$= 1 \quad \text{for} \quad -\frac{z_1}{2v} < t_i < \frac{z_1}{2v} = 0 \quad \text{else}$$

$$= \frac{a^2\sigma_i K}{4\sigma_e}\cdot\frac{3}{A}\cdot v\cdot I_x(yt) * h(t) \qquad (6)$$

Where $*$ means a convolution.

For all fibres of the motor unit with the same conduction velocity a convolution with the same impulse response has to be carried out. Therefore it is more convenient to summate the contributions of all fibres and calculate the convolution only once afterwards.

2.4. Choice of Parameters

The parameters used in this study all refer to the *m.* soleus of the rat. This muscle contains about 2500 muscle fibres (Close, 1967; Frank et al., 1975). It was assumed that there are approximately 30–50 motor units (own measurements). A motor unit containing 50 fibres was simulated. Considering the results of Kugelberg (1973) it seems reasonable to accept a random (rectangular) distribution for the pdf of the fibre positions. For the cross section of a motor unit territory an area 1×2 mm was chosen. It was not possible to obtain a precise pdf of the fibre diameters. Several authors investigated the fibre diameter (Yonemura, 1967; Edström and Kugelberg, 1968). Finally a random (rectangular) distribution between 20–40 μm was chosen. Although a Gaussian distribution might be more appropriate, it can be argued that the effects of this difference are subordinate.

The pdf, describing the dispersion in arrival time can be chosen in different ways. These ways are illustrated in Fig. 3.

In Figure 3a the pdf is chosen randomly (within a certain interval) so the arrival times are independent of the fibre positions within the muscle. In Figure 3b however the pdf is chosen randomly in an oblique zone. In this way the arrival times depend also on the fibre position within the muscle. Such an pdf is in accordance with a skew situation of the end-plate zone in a muscle (see Fig. 1a). The width D of the distribution function of arrival times was taken 1.25 ms. If the dispersion is only caused by the distribution of end-plate positions, this interval relates to an upper limit of the motor end-plate region of $1.25\times 4 = 5$ mm (conduction velocity: 4 m/s). This region is comparable with the results of Frank et al. (1975) and Ypey (1976). Besides these authors indicate a skew end-plate region within the muscle.

At situation 3b the shape of the "parallelogram function" can be indicated by a factor DCOR between 0 and 1. In this case the arrival time for the i^{th} fibre D_i is:

$$D_i = D\left\{\text{DCOR}\frac{x_i}{x_{\text{terr}}} + (1\text{-DCOR})\cdot R_i\right\}, \qquad (7)$$

where: x_i is the x-coordinate of the i^{th} fibre; x_{terr} is the territory dimension in the x-direction; R_i is a random variable between 0 and 1.

All the simulation results refer to a value of DCOR = 0.5, unless otherwise indicated.

The tripole parameters were adjusted on the basis of literature (Rosenfalck, 1969) and our own measurements. With a conduction velocity of 4m/s the value of $s_1 = 250$ μm. The other parameters were stated on: $\alpha = 0.8$ and $s_2 = 1250$ μm. The dimensions of the electrode can be chosen arbitrary.

All the simulation results refer to unipolar electrodes, unless otherwise indicated.

3. Results

3.1. Influence of Dimensions and Position of the Electrode on the Muap

Figure 4 shows simulated muaps, "registrated" both by a point-shaped electrode and a macro electrode (dotted curves: lead-off area: 200×760 μm) each for two different positions with respect to the motor unit territory.

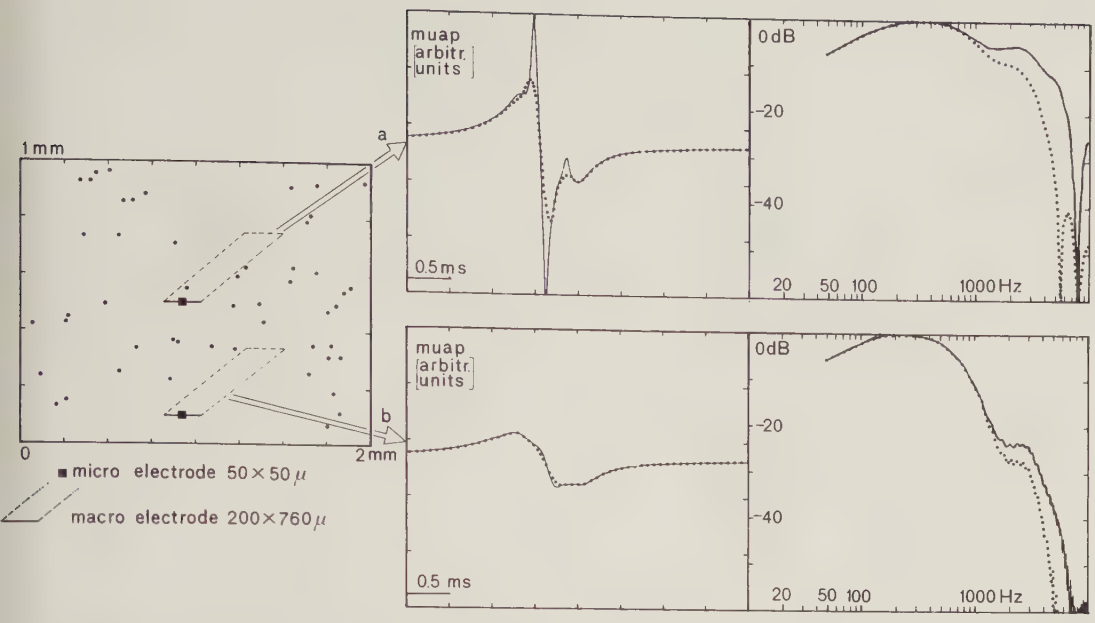

Fig. 4. Influence of the electrode dimensions and position with respect to the motor unit territory on the muap and its power spectrum. The dotted lines refer to an electrode with a lead-off area of 200 × 760 μm, while the continuous lines refer to a micro electrode. In the left part the positions of the muscle fibres with respect to the electrode and the dimensions of the electrode are indicated. The change to smooth biphasic patterns, for both electrode types from intra to extraterritorial registrations is clearly shown

Within the motor unit territory (Fig. 4a) there was a noticeable difference in amplitude and slightly in shape. The actual shape of an intraterritorial observed muap is of course strongly determined by the fibres lying right next to the electrode. This dependence was most obvious for a point shaped electrode: rapid changes caused by the nearby fibres were observed. These rapid changes correspond to high frequency components in the power spectrum. A low-pass filtering effect of the macro electrode, caused by the surface integration, was obvious in intraterritorial observation (Fig. 4a). If the electrode was at the "border" of the motor unit territory there was hardly any difference between the "registrations" of the two electrodes.

Low and high frequency contents could be distinguished in the power spectra. The low frequency part was characterized by the first maximum in the spectrum (see Fig. 2).

As will be shown the amplitude of this first maximum contained general information about the motor unit. Extraterritorial simulated muaps were dominated by the low frequency content. In the time domain only a biphasic slow wave appeared.

In the intraterritorial simulated muaps a similar slow biphasic wave could also be recognized as a part of the total muap. The low frequency maximum in the spectrum showed a slight shift to lower frequencies with increasing distance of the electrode (Fig. 4a and b) as can be expected considering the volume conduction theory.

3.2. Influence of the Distribution of Motor End-plates on the Muap

It is interesting to know if the shape of a muap depends on the arrival times of the individual single fibre potentials at the electrode.

Anatomical studies of the situation of the motor end-plates (Coërs and Woolf, 1957; Ypey, 1975; Frank et al., 1975) suggest a small oblique zone as shown in Figure 3b. Such a geometrical situation implies a correlation between the position of a fibre with respect to the cross section of the muscle and the arrival time. Therefore the influence of some arrival time functions, which depend on the distribution of motor end-plates on the muap was simulated.

In Figure 5a the end-plates were randomly distributed in the whole end-plate region, a polyphasic muap was obtained. If the end-plates were exactly positioned in a skew plane a biphasic muap with a higher amplitude was obtained. In this Figure 5c it is assumed that differences in arrival times are only caused by different end-plate positions with respect to the cross-section of the muscle (DCOR = 1).

In Fig. 5b an "intermediate" structure (DCOR = 0.3) was simulated: the muap is almost biphasic.

It is obvious that the structure of the motor end-plane zone influences the muap strongly.

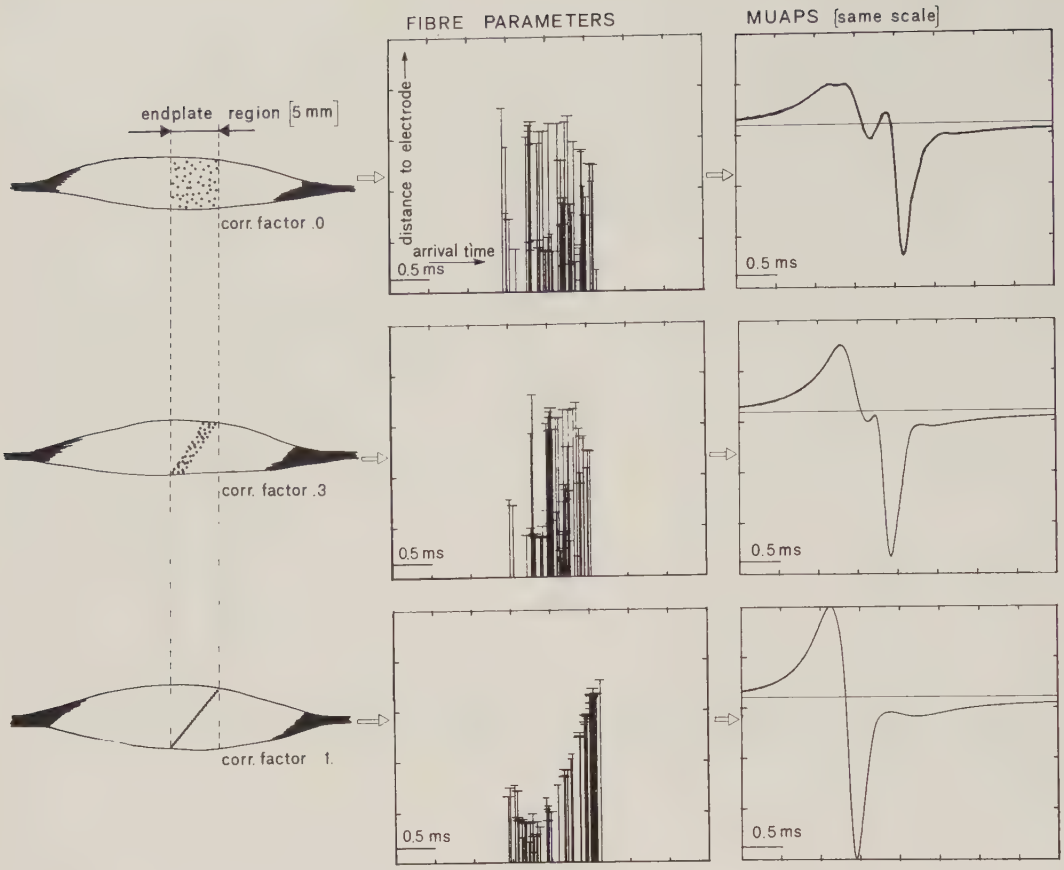

Fig. 5. The influence of the distribution of motor end-plates on the muap. From the simulation result in the top of the figure, it can be seen that with end-plates, that are randomly distributed in the whole end-plate region, a more polyphasic muap is obtained. If the end-plates are exactly positioned in a skew plane a biphasic muap with a higher amplitude is obtained. With these simulations it is assumed that differences in arrival times are only caused by different end-plate positions with respect to the cross-section of the muscle. In the middle of the figure an "intermediate" structure (DCOR = 0.3) was simulated: the muap is almost biphasic. It is obvious that the structure of the motor end-plate zone influences the muap strongly

3.3. Changing Dispersion of Arrival Times

An enlarged dispersion is an important pathological symptom which can be observed e.g. in neuropathies (Hopf and Struppler, 1974).

Figure 6 shows an example of the influence of an enlarged dispersion on the form of the muap and its spectrum.

The electrode ($100 \times 380\,\mu m$) was situated in the centre of the motor unit territory. In Fig. 6a the maximum difference in arrival time is 1.25 ms. A muap which is almost biphasic is apparent. An enlargement to 2.5 ms (Fig. 6b) generated a more polyphasic muap. The change to a more polyphasic pattern is a well-known phenomenon. Usually a polyphasic muap resulted in more power in the high frequency part of the power spectrum.

From Fig. 6 it can be concluded that the maximum of the low frequency part shifts to a lower frequency. In general this frequency was related to the duration of the muap and therefore also to the dispersion of the arrival times.

3.4. Fibre Density and Motor Unit Territory

Differences in the number of fibres of one motor unit can be caused by variations in the fibre density and/or variations in the size of the motor unit territory. Of course an increasing number of fibres will result in a higher amplitude of the muap as long as compensation, due to superposition of negative and positive phases of different fibre potentials is not complete. It can be argued that such a compensation will hardly affect the low frequency part of the spectrum.

Therefore the amplitude of the low frequency maximum is an interesting parameter. In Fig. 7 this parameter is plotted as a function of the number of fibres in an intraterritorial simulated muap. Besides this parameter also the top-top amplitude of the time registration is indicated in Fig. 7 in order to compare this with the low frequency maximum parameters.

In Fig. 7a the number of fibres is proportional to the density by taking a constant cross-section area of the motor unit, whereas in Fig. 7b this number is proportional to the area, implicating a constant

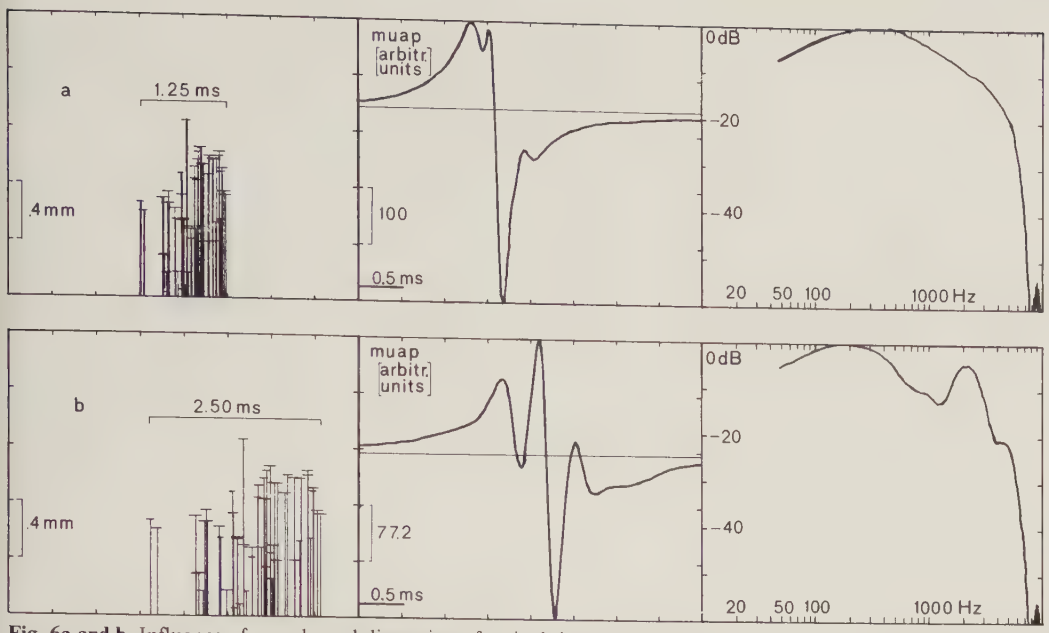

Fig. 6a and b. Influence of an enlarged dispersion of arrival times on the muap and its power spectrum. In the upper part a situation is shown, with a certain dispersion of arrival times, while in the lower part the dispersion is (proportionally) enlarged. This causes a more polyphasic pattern and usually more power in the high frequency part of the power spectrum. The maximum of the low frequency part is shifted to a lower frequency

density. In the first case the parameter $\sqrt{P_{max}}$ shows an almost linear dependence with respect to the fibre density while the muap amplitude is rather variable.

In the lower part the fibre density is taken constant. There is a clear relation between $\sqrt{P_{max}}$ and territory diameters.

The muap amplitude is almost independent of the territory diameter. All results are obtained by assuming the lead-off electrode to be positioned in the geometrical centre of the territory.

3.5. Variability of Some Muap-parameters with Respect to the Electrode Position

In order to illustrate the influence of the actual position of the electrode within a certain area of the motor unit territory we simulated 42 muaps. The electrode positions were chosen arbitrary in the following areas:

1. approximately in the geometrical centre of the motor unit, 2. at the boundary, 3. between the centre and the boundary, and 4. outside the motor unit territory (three different regions).

From the simulation results means and standard deviations were calculated based on the muaps led off from one area, characterized by its mean distance to the geometrical centre of the motor unit.

In Fig. 8 the ratio standard deviation/mean is plotted as a function of the distance to the centre, both for the top-top amplitude as well as for the square root of the amplitude of the low frequency maximum. Apparently the latter is a less fluctuating parameter. Therefore it can be used in order to obtain information

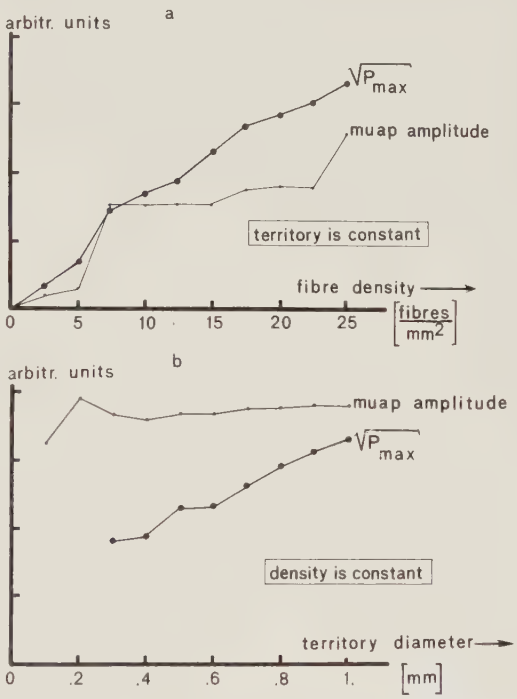

Fig. 7a and b. Relation between two muap parameters and fibre density or territory diameter. In this figure it is presented how the muap amplitude and $\sqrt{P_{max}}$ are related to the fibre density as well as the territory diameter. In the upper part the m.u. territory is taken constant and the fibre density is taken variable. The parameter $\sqrt{P_{max}}$ shows an almost linear dependence with respect to the fibre density while the muap amplitude is rather variable. In the lower part the fibre density is taken constant. There is a clear relation between $\sqrt{P_{max}}$ and territory diameter. The muap amplitude is almost independent of the territory diameter. All results are obtained by assuming the lead-off electrode to be positioned in the centre of the territory

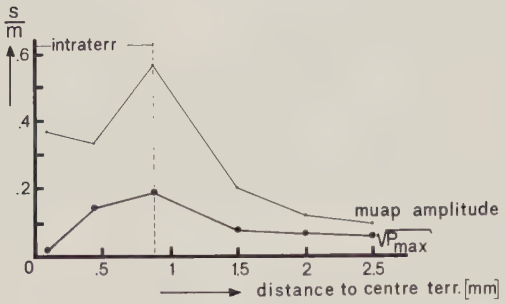

Fig. 8. The ratio standard deviation/mean $\left(\frac{s}{m}\right)$ of simulated muaps, plotted as a function of the distance from the electrode to the centre of the motor unit, indicating the variability of the parameters top-top amplitude and $\sqrt{P_{max}}$ (electrode $100 \times 380\,\mu m$). The upper curve refers to the top-top amplitude, the lower curve refers to the square root of the low frequency maximum, which is clearly a less fluctuating parameter and therefore it is an interesting parameter for further investigations

about fibre density and/or size of the motor unit (see also Fig. 7).

3.6. Selectivity of an Unipolar and a Bipolar Electrode

The usefulness of the model for studying different kind of electrodes will be illustrated by a case study concerning the selectivity-characteristics of unipolar and bipolar electrodes. Two partly overlapping motor units (Nos. 1 and 2) were considered, which were supposed to be active at the same time (see Fig. 9).

The recording electrode was positioned in the centre of motor unit 1 and just outside motor unit 2.

In the upper part of Fig. 9 it is shown how the muap of motor unit 1 simulated with an unipolar electrode, is strongly interfered with the muap originating of the second motor unit.

However if a bipolar registration (the two electrodes are positioned parallel to the muscle fibres) is simulated, with an interelectrode distance of 0.3 mm, the electrode seems to be very selective for the activity of motor unit 1 (lower part of Fig. 9).

The powerspectra are also indicated. It can be seen that a bipolar electrode acts as a high-pass filter.

On increasing the interelectrode distance the power spectrum showed regular "dips" from which the conduction velocity of the muscle fibres can be derived (Linström, 1974).

4. Discussion

In the previous part results of the stochastic simulation model have been presented. It was a first attempt to study possible muaps by "imitating" the motor unit both from the anatomical as the physiological point of view. The model is strongly based on the description of the single fibre potential. If only one fibre is considered, the model is equivalent to the model of Ekstedt and Stålberg (1973) however in our model also anisotropy is taken into account. The description of the single fibre potential has its limitations especially when the muscle fibre is too close to the electrode. In this case a more correct description can be given as is shown by Rosenfalck (1969) and Dimitrov and Dimitrova (1974).

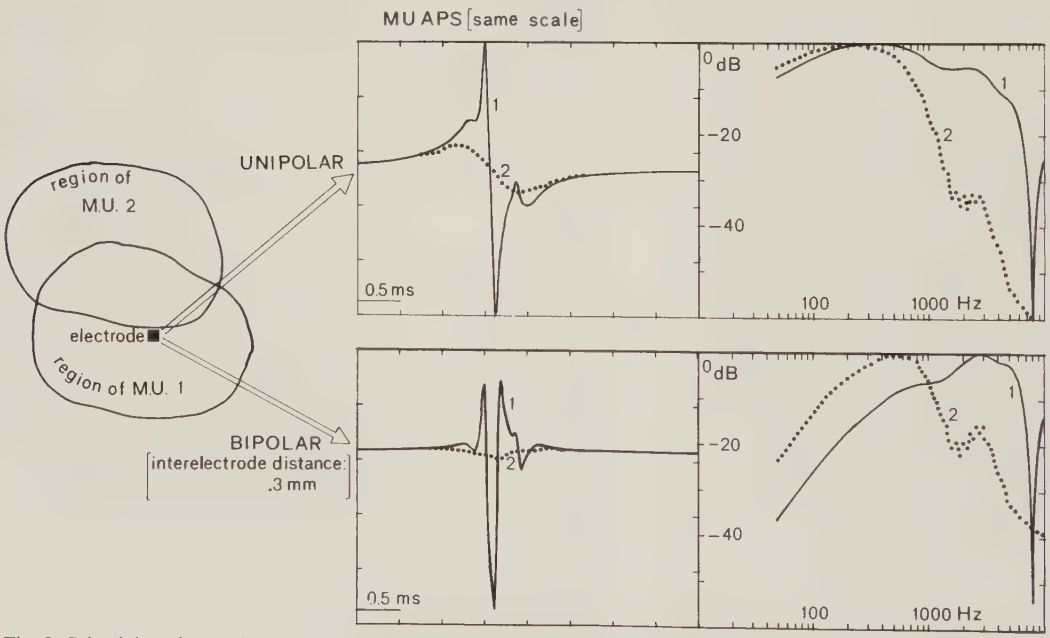

Fig. 9. Selectivity of an unipolar and bipolar electrode. Two overlapping motor units are active at the same time. If an unipolar recording is simulated, the activity of motor unit 1 is strongly interfered with the activity of motor unit 2. However if a bipolar recording is simulated with a small interelectrode distance (0.3 mm) this electrode seems to be very selective for the activity of motor unit 1; a distorted muap is obtained. The power spectra are also indicated. It can be seen that a bipolar electrode acts as a high-pass filter

It is of course possible to build in a more general applicable algorithm. This was not done here because then the program would be very time-consuming while improvements were expected to be marginal.

In literature a number of EMG-models concerning single muaps are mentioned (George, 1970; Lindström et al., 1976). These models have an analytical approach and the results refer to so called expected values (statistical first and second order moments). Especially the work of Lindström et al. on this subject is very extensive. They calculated expected values of the Fourier and power spectrum using the pdfs of the stochastic variables already mentioned. With the stochastic simulation model these pdfs are basically *not* used as a weighting function. In our model the contribution of each fibre to the muap is individually simulated using parameter-values that were chosen for each fibre separately, according to the same pdf. Only if the number of fibres in a motor unit is large enough, do the two methods lead to comparable results. For extraterritorial recorded muaps the number of fibres is usually large enough so that the expected values form a good approximation of a particular muap. For intraterritorial recorded muaps however it is the fraction of fibres in close vicinity to the electrode that is not large enough to make statistical weighting useful. Single fibre potentials that can be observed separately in reality and also which our model (see Fig. 3) are "smoothed away" by using the pdf as a weighting function.

This effect can be easily demonstrated by assuming only one fibre potential $f(t-\tau)$. Only the arrival time τ is taken as a stochastic variable. On this way the expected value is:

$$E[f(t-\tau)] = \int_{-\infty}^{\infty} f(t-\tau)p(\tau)d\tau = \bar{f}(t), \tag{7}$$

where $p(\tau)$ represents the probability density function of the arrival time τ. The shape of $\bar{f}(t)$ differs of course of the shape of $f(t)$ and it must be emphasized that *always* (at every experiment) the shape $f(t)$ will be measured and never $\bar{f}(t)$!

The results of the simulation model will depend strongly on the parameters used, just as in reality. This implies that when using the model to study a specific muscle, these parameters must be investigated as well as possible. Now an extensively investigation is in progress to test the simulation model thoroughly by measuring all the parameters used in the model by electrophysiological and histochemical techniques.

From our own measurements it appeared that the variability in registrated muap-shapes is comparable with our simulation results.

In the model the dispersion of arrival times is taken statistically dependent on the position of the muscle fibre within the motor unit. From the simulation results (Fig. 5) it can be concluded that on increasing the correlation between position and arrival time, less polyphasic patterns will appear.

Physiological studies of Stålberg and Thiele (1975) by means of single fibre electromyography indicate a relatively small time interval between two fibres that are lying close to each other; this also reveals a possible correlation. Another parameter of interest is the pdf which determines the scattering of muscle fibres of a motor unit within the muscle. In the simulation model the fibres were randomly (homogeneous) spread. Although in literature these pdf's are not investigated quantitatively, data about motor unit cross sections (also of different muscles) suggest such a random distribution (Kugelberg, 1973; Burke et al., 1974; Brandstater and Lambert, 1973; Stålberg et al., 1976). In pathological cases however this distribution can change significantly.

On observing several simulated muaps it is apparent how variable the top-top amplitude of the muap can be if an intraterritorial registration is made. The variations of the distance between the closest fibres and the electrode are very dominant.

This is the reason why Gydikov (1976) only used extraterritorial muaps for studying the motor units (see also Buchthal and Rosenfalck, 1973). Stålberg (1976) argued that the intraterritorially recorded muap is only determined by the "closest fibres" already mentioned.

With respect to these findings the simulation results indicate that the low-frequency maximum of the power spectrum is an interesting parameter to describe features concerning the *whole* motor unit, both with extraterritorial as with intraterritorial measurements.

As could be expected it is evident from the simulation results with micro- and macro-electrodes (Fig. 4) that the micro electrode gives a signal that is mainly related to the closest single fibre(s) while the macro electrode signal (especially for larger macroelectrodes) is more related to all the fibres of the motor unit. This suggests an electrode construction for practical investigations in which a macro and micro electrode are combined. For example the micro electrode signal can be used as a trigger signal to average the macro electrode signal, in order to obtain isolated "macro-muaps" from interference patterns (Leifer and Pinelli, 1976).

Finally it is concluded that the presented simulation model is a powerful tool for investigating the basic problems: behaviour of existing and new types of electrodes and the relations between muaps and motor unit parameters.

Appendix

Solution of integral I_x:

$$I_x = \int_{x_1}^{x_2} \frac{\alpha}{\sqrt{(z-s_1-z_i)^2 + r^2 \cdot 5}}\, dx_i$$

$$= \alpha \int_{x_1}^{x_2} \frac{1}{\sqrt{(z-s_1-z_i)^2 + 5(x-x_i)^2 + 5(y-y_i)^2}}\, dx_i$$

$$= \frac{\alpha}{\sqrt{5}} \cdot \ln \frac{(x-x_1) + \sqrt{(x-x_1)^2 + b^2}}{(x-x_2) + \sqrt{(x-x_2)^2 + b^2}} \quad \text{for:} \quad x > x_2$$

$$= \frac{\alpha}{\sqrt{5}} \cdot \ln \frac{(x_2-x) + \sqrt{(x_2-x)^2 + b^2}}{(x_1-x) + \sqrt{(x_1-x)^2 + b^2}} \quad \text{for:} \quad x < x_1,$$

where: $b^2 = (y-y_i)^2 + \frac{(z-s_1-z_i)^2}{5}$.

For $x_1 < x < x_2$ a combination of these formulas is obtained.

Acknowledgements. The authors wish to express their appreciations to Willemien Wallinga-de Jonge and Herman Boom for their suggestions and comments, Dr. F. H. Lopes da Silva for reading and Truus Steijlen for typing the manuscript.

References

Boon, K.L., Griep, P.A.M.: The motor unit action potential: study with a stochastic simulation model and development of a new mechano-electrical needle transducer. In: Proc. Isek, Arrigo, A. ed. Pavia, pp. 22–26, 1976

Buchthal, F., Rosenfalck, P.: On the structure of motor units. In: Desmedt, J.E. ed., New Developm. Electromyogr. clin. Neurophysiol., Vol. 1, pp. 71–85. Basel: Karger 1973

Burke, R.E., Levine, D.N., Salcman, M., Tsairis, P.: Motor units in cat soleus muscle: physiological, histochemical, and morphological characteristics. J. Physiol. **238**, 503–514 (1974)

Brandstater, M.E., Lambert, E.H.: Motor unit anatomy. In: New Developm. Electromyogr. clin. Neurophysiol., Vol. 1, pp. 14–22, Desmedt, J.E., ed. Basel: Karger 1973

Close, R.: Properties of motor units in fast and slow skeletal muscles of the rat. J. Physiol. **193**, 45–55 (1967)

Coërs, C., Woolf, A.L.: The innervation of muscle, pp. 14–20. Oxford: Blackwell 1957

Dimitrov, G., Dimitrova, N.: Extracellular potential field of a single striated muscle fibre immersed in anisotropic volume conductor. Electromyogr. clin. Neurophysiol. **14**, 423–436 (1974)

Edström, L., Kugelberg, E.: Histochemical composition, distribution of fibres, and fatiguability of single motor units. J. Neurol. Neurosurg. Psychiat. **31**, 424–433 (1968)

Ekstedt, J., Stålberg, E.: How the size of the needle-electrode leading-off surface influences the shape of the single muscle fibre action potential in electromyography. Comp. Progr. Biomed. **3**, 204–212 (1973)

Frank, E., Jansen, J.K.S., Lømo, T., Westgaard, R.: The interaction between foreign and original motor nerves innervating the soleus muscle of rats. J. Physiol. **247**, 725–743 (1975)

George, R.E.: The summation of muscle fibre action potentials. Med. Biol. Eng. **8**, 357–365 (1970)

Griep, P.A.M., Boon, K.L.: Study of the motor unit action potential (muap) by means of a simulation model. Electroenceph. clin. Neurophysiol. **43–4**, 610 (1977)

Gydikov, A.: The potential of the motor unit. In: The motor system: neurophysiology and muscle mechanisms. Shahani, M. ed., pp. 49–71. Amsterdam: Elsevier 1976

Håkansson, C.H.: Action potentials recorded intra- and extra cellularly from the isolated frog muscle fibre in ringer's solution and in air. Acta physiol. scand. **39**, 291–318 (1957)

Hopf, H.C., Struppler, A.: Elektromyographie. Stuttgart: Thieme 1974

Kugelberg, E.: Properties of the rat hind-limb motor units. In: New Developm. Electromyogr. clin. Neurophysiol., Vol. 1, pp. 2–13. Desmedt, J.E., ed. Basel: Karger 1973

Leifer, L., Pinelli, P.: Analysis of motor units by computer aided electromyography. In: Proceedings Isek, Arrigo, A., ed. Pavia, 1976

Lindström, L.: Contributions to the interpretation of myoelectric power spectra. Doctoral Dissertation, Chalmers University of Technology, Göteborg, 1974

Lindström, L., Broman, H., Magnusson, R.: Determinants of myoelectric power spectra. In: Proceedings Isek, pp. 92–95. Arrigo, A., ed.: Pavia, 1976

Lorente de Nó, R.: A study of nerve physiology. Studies from the Rockefeller Institute for Medical Research, 1947

Rosenfalck, P.: Intra- and extracellular potential fields of active nerve and muscle fibres. Doctoral Dissertation, København: Akademisk Forlag 1969

Stålberg, E., Thiele, B.: Motor unit fibre density in the extensor digitorum communis muscles. J. Neurol. Neurosurg. Psychiat. **38**, 874–880 (1975)

Stålberg, E.: Single fibre electromygraphy for motor unit study in man. In: The motor system: neurophysiology and muscle mechanisms, pp. 79–92. Shahani, M., ed. Amsterdam: Elsevier 1976

Stålberg, E., Schwartz, M.S., Thiele, B., Schiller, H.H.: The normal motor unit in man. J. Neurol. Sci. **27**, 291–301 (1976)

Yonemura, K.: Resting and action potentials in red and white muscles of the rat. Jap. J. Physiol. **17**, 708–719 (1967)

Ypey, D.L.: Effects of muscle length and stimulus frequency on neuromuscular transmission in the frog. Doctoral Dissertation, Leiden University, 1975

Ypey, D.L.: Personal communications 1976

Received: May 13, 1978

P. A. M. Griep
Department of Electrical Engineering
Twente University of Technology
P.O. Box 217
7500 AE Enschede, The Netherlands

Biol. Cybernetics 30, 231—240 (1978)

Biological Cybernetics

Anisotropic Connectivity and Cooperative Phenomena as a Basis for Orientation Sensitivity in the Visual Cortex

S. Finette[1], E. Harth[1], and T. J. Csermely[2]
[1] Department of Physics
[2] Department of Electrical and Computer Engineering, Syracuse University, Syracuse, NY, USA

Abstract. A computer simulation model of the neural circuitry underlying orientation sensitivity in cortical neurons is examined. The model consists of a network of 3000 neurons divided into two functionally distinct cell types: excitatory (*E*-cells) and inhibitory (*I*-cells). We demonstrate that both orientation sensitivity and shape selectivity can be accounted for by making the following assumptions: 1) thalamic afferents to a sheet of cortical neurons are retinotopically organized; 2) thalamic afferents come from a single neuron, or at most a few neurons, in the lateral geniculate nucleus; 3) cortical activity is cooperative, i.e. largely dependent on intracortical connections, some of which have anisotropies along directions parallel to the pial surface. Anisotropies are specified only by the distribution of cells which are postsynaptic to a particular neuron, without specifying the axonal or dendritic contributions. In this paper, orientation sensitivity arises through cooperative interactions among neurons having anisotropic excitatory, and isotropic inhibitory connections.

Introduction

Information processing in the mammalian visual system is often described as a succession of mappings between layers or sheets of neurons (Szentágothai and Arbib, 1974). The mappings from the retina to lateral geniculate nucleus (LGN) and from the LGN to striate cortex are retinotopically organized: neighboring areas on one sheet correspond to neighboring areas on the successive sheet. Single cell receptive fields are generally circular with center-surround structure in the retina and LGN and – in the case of primates – also in lamina IVc of striate cortex. However, in other layers of the visual cortex, the dominant field configuration shows distinct elongation with rather precise orientation for the stimulus eliciting the highest response. Neurons having similar orientation preferences are located within structures termed orientation *columns* or *slabs* (Hubel and Wiesel, 1977). Anatomical confirmation of the orientation slabs has recently been found by use of 2-deoxyglucose autoradiography (Hubel et al., 1978; Skeen et al., 1978).

The neural circuity underlying the transition from neurons with more or less circular receptive fields to those with elongated, orientation sensitive fields is unknown. The earliest theory of orientation sensitivity (Hubel and Wiesel, 1962) suggested that a simple cortical receptive field may be the result of specific geniculate afferents selected by the dendritic tree of the cortical cell. In this model, a simple cortical field elongated in the horizontal direction is formed by inputs from geniculate cells whose receptive fields are displaced along a horizontal direction. Similarly, complex and hypercomplex cells would be constructed from simple cell receptive fields, thus forming a hierarchy of cortical receptive field structures.

An alternative view interprets orientation sensitivity as arising from intracortical connections rather than from a selection of specific afferent fibers. Braitenberg (1974) suggests that axon collaterals of pyramidal cells, often seen traveling laterally over considerable distances, might form strong recurrent interactions with neighboring pyramidal cells. In this scheme, axon collateral distributions may play an important role in determining the shape of cortical receptive fields. Colonnier (1964) and McIlwain (1976) believe that oriented receptive fields in the visual cortex may reflect anisotropies in the dendritic fields of cortical cells. Creutzfeldt et al. (1974) have considered anisotropic inhibitory connections as a contributing factor in orientation selectivity, while Schiller et al. (1976) suggest a number of schemes based on anisotropic cortical connectivities in order to explain both orientation selectivity and direction specificity.

0340-1200/78/0030/0231/$02.00

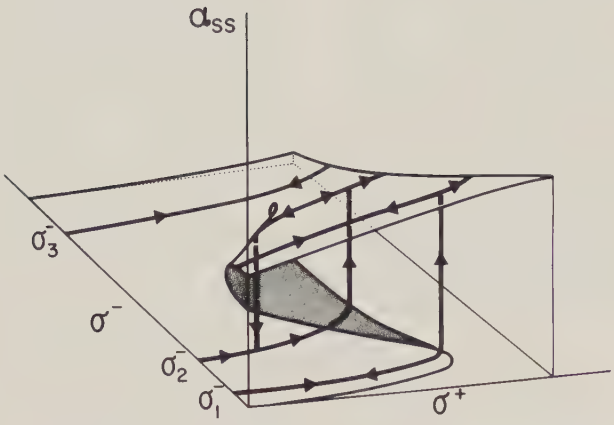

Fig. 1. Schematic diagram illustrating a netlet's stationary activity states, α_{ss}, due to slowly varying afferent stimulation. Parameters σ^- and σ^+ represent the strengths of afferent inhibitory and excitatory stimulation, respectively. Three qualitatively different steady state trajectories, corresponding to three constant levels of inhibitory inputs, are illustrated

Our present study is part of a program in which we apply both experimental and theoretical methods to the problem of orientation sensitivity. Quantitative features of cortical structure are determined by computer-aided measuring techniques using tissue sections from striate cortex, stained by the rapid Golgi method. This work is now carried out with a new measuring system (Gentile and Harth, to be published), incorporating a novel technique for alignment of serial sections (Gentile and Harth, 1978). The classic anatomical methods, using Golgi and similar staining techniques, have in the past failed to exhibit any of the striking features related to either ocular dominance or orientation selectivity in the cortex. This has been a major puzzle in neurophysiology. Lack of anatomical confirmation of physiological facts has been largely overcome by the new radiographic techniques which have vastly extended our knowledge of the spatial arrangement of ocular dominance and orientation columns. However, the fundamental question of the neuronal *wiring* responsible for orientation sensitivity still remains unanswered. It is our belief that a new quantitative look at the details of cortical fiber structure will be required to reveal the relevant circuitry.

To this end we will continue our program of quantitative neuroanatomy, using computer-aided three dimensional measurements of neuronal fiber structure. The rapid Golgi stain will be the chief method for revealing structural details at the light microscope level.

In the theoretical phase of this program, the dynamics of neural structures is studied by mathematical analysis and computer simulation. By incorporating results of morphological studies such as our own, we hope that new insights will be gained into the relationship between cortical structure and function.

We approach the problem of the neuronal origin of orientation sensitivity by building upon our background of large-scale computer simulations of diverse neuronal networks. The aim here will be simply to demonstrate that orientation sensitivity can be accounted for by making the following assumptions: I) cortical neurons are retinotopically organized (Daniel and Whitteridge, 1961; Tusa et al., 1978); II) the thalamic afferents to a cortical neuron come from a single cell, or at most a few cells in the LGN (Creutzfeldt and Ito, 1968; Creutzfeldt et al., 1974), and III) cortical activity is cooperative, i.e. largely dependent on intracortical connections, some of which have anisotropies along directions parallel to the pial surface (Paldino and Harth, 1977; Marin-Padilla and Stibitz, 1974; Szentágothai and Arbib, 1974). Supporting evidence is strongest for the first, weakest for the last of these assumptions. No attempt will be made here to describe either the topology of the mapping of visual space onto the visual cortex, or the detailed structure of cortical receptive fields at the single cell level.

Theoretical Considerations

We have been studying cooperative phenomena within neural populations for some time (Harth et al., 1970; Anninos et al., 1970; Wong and Harth, 1973; Harth et al., 1975). Several researchers have found similar effects under a rather wide range of assumptions (Amari, 1971; Wilson and Cowen, 1972; 1973; Grossberg, 1973; Amari and Arbib, 1977). We have introduced the concept of *netlets*, defined as functionally distinct populations of neurons randomly connected without distance bias. The significant dynamic variable in a netlet was taken to be its *activity*, i.e. the fraction of neurons firing at a given time. It was found that small changes in the strength of afferent excitatory (inhibitory) stimulation can produce sudden phase transitions between two activity levels and hysteresis effects in a single netlet or between coupled netlets. In Fig. 1 we depict a typical surface corresponding to the set of stationary states for a single netlet. In this graph, the vertical axis, α_{ss}, denotes the activities corresponding to stationary states of the net. The horizontal axes, σ^+ and σ^-, refer to the strengths of excitatory and inhibitory inputs distributed over the net. Three vertical cuts through the surface for different levels of inhibitory input indicate three qualitatively different curves. These describe the trajectories of the system for slowly varying excitatory inputs and constant levels of inhibition. It is seen that phase transitions and hysteresis effects occur for particular values of excitatory

and inhibitory stimulation where two stable stationary states exist.

These concepts have been extended to *locally connected* two dimensional homogeneous, isotropic networks (unpublished data). Such structures can be made *locally addressable*, meaning that activity generated by local input remains within or near the region of stimulation. In such a case, this local region approximates the dynamics of a netlet. For the isotropic network the natural stimulus shape is circular – local activity is elicited most readily by a compact stimulus. The property of local addressability remains when the interconnections within the network are *anisotropic*. Such networks can be made to respond selectively to elongated stimuli of particular orientations. Once again, sharp phase transitions and hysteresis are obtained in response to changes in stimulus intensity. These cooperative effects are quite sensitive to stimulus shape and orientation.

Details of the Model

We consider the dynamics of spatially localized cell populations within a single orientation column, i.e. a network tuned to a specific stimulus orientation. Ocular dominance effects are not considered. In this model, orientation tuning of neuronal populations will be shown to arise through cooperative interactions among neurons with anisotropic excitatory interconnections.

In the computer simulations presented here we have considered a network consisting of two functionally distinct cell types: excitatory (E-cells) and inhibitory (I-cells). Estimates are that approximately 70–85% of all cells in the visual cortex of cat are pyramidal cells, though for primates this fraction is smaller (Sholl, 1955; Blinkov and Glezer, 1968). Such neurons are thought to be excitatory in their postsynaptic effects (Szentágothai, 1975). Other cell types, collectively termed stellate cells, may be either excitatory or inhibitory (Szentágothai, 1975). Because of the presumably small number of I-cells we have assumed that intracortical connections are of the types $E-E$, $E-I$, and $I-E$, but not $I-I$. A number of studies support the view that geniculo-cortical synapses are excitatory (Creutzfeldt and Ito, 1968; Colonnier and Rossignol, 1969; Toyama et al., 1974; Garey, 1976). Anatomical observations on the terminal site of geniculate fibers indicate contacts may occur with both pyramidal and stellate cells (Globus and Scheibel, 1967; Lund, 1973). Intracellular recordings (Creutzfeldt and Ito, 1968; Creutzfeldt et al., 1974) suggest that, on the average, only one or two geniculate fibers project to a single cortical cell. However, the relative proportion of cortical excitatory and inhibitory neurons contacted by the primary visual afferents is unknown. Consequently, we choose the afferents to our networks as excitatory, and randomly distributed to both groups of cells within a specified region. On the average, only one afferent fiber will synapse with a given excitatory or inhibitory cell.

Anisotropies within the cortex may arise from either dendritic or axonal asymmetries, or both. Marin-Padilla (1974), Marin-Padilla and Stibitz (1974) have found strong anisotropies in both the axonal and dendritic distribution of basket cells in both motor and visual cortices. Basket cells (type 1 cell of Jones) have been seen in primate somatosensory cortex with their axons organized in antero-posteriorly oriented planes (Jones, 1975). Szentágothai (1973) tentatively concludes that basket cells in the visual cortex of cat and monkey have axons oriented in narrow planes. He also reports on another large stellate interneuron whose axon arborizes within a vertical plane perpendicular to that of the basket cell arborization of the same region (Szentágothai, 1975). Paldino and Harth (1977) have found anisotropies in collaterals of axons from pyramidal cells in rat visual cortex. Degeneration studies in monkey and cat striate cortex (Fisken et al., 1975; Creutzfeldt et al., 1977) indicate strong asymmetries in degeneration of intrinsic fiber connections. Physiological studies have suggested orientation selective intracortical inhibition in cat cortex (Creutzfeldt et al., 1974; Nelson and Frost, 1978). Anisotropies of connectivity are specified in our model only by the distribution of cells that are postsynaptic to a particular neuron. The separate contributions of axonal and dendritic fields to this effect will be considered in future work. This paper specifically deals with networks containing anisotropic excitatory connections and isotropic inhibitory connections.

The pattern of orientation columns and their relation to the system of ocular dominance columns was studied in monkey by Hubel et al. (1978). The minimum length of an orientation column, corresponding to one eye, is obtained by assuming that it crosses the ocular dominance column at right angle. The average density of neurons in monkey visual cortex is 110 per $0.001\,\text{mm}^3$ (Cragg, 1967) and the average thickness 1.5–2.0 mm. We therefore estimate that about 4–5000 neurons are found in a single column. Our network consists of 3000 neurons, 85% E-cells and 15% I-cells. At present, the dimension perpendicular to the pial surface is neglected for the sake of simplicity. All neurons and their interconnections lie in a plane. The positions of the somas are distributed randomly over the plane in order to conform to the general appearance of Nissl stained tissue and to avoid any orientation bias inherent in a lattice structure.

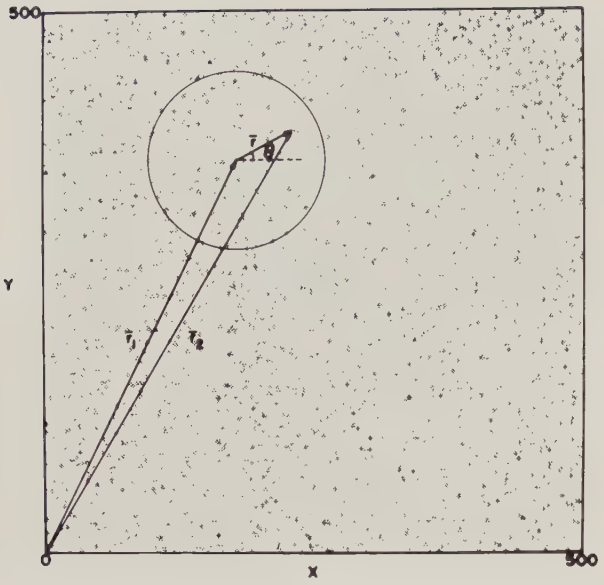

Fig. 2. Coordinate system used to determine the spatial distribution of connections from an arbitrary neuron, located at $\bar{r}_1$, to postsynaptic cells within a given circular region of influence. Symbols denote neurons which are distributed randomly on a 500×500 grid. In general, the probability of a connection depends on $|\bar{r}|$ and θ

The spatial distribution of cells that are postsynaptic to a particular neuron is assumed to be a function of distances and orientation as follows. Let $\bar{r}_1$ and $\bar{r}_2$ be the locations of pre- and postsynaptic cells, respectively (Fig. 2). We now choose a set of polar coordinates r, θ centered on the soma of the presynaptic cell. For an isotropic homogeneous distribution, the probability that a cell located at $\bar{r}_2$ receives a synapse from a neuron located at $\bar{r}_1$ depends on $|\bar{r}| = |\bar{r}_2 - \bar{r}_1|$ and is independent of θ. We assume further that our probability density function is constant for $|\bar{r}| \leqq \varrho$ and zero for $|\bar{r}| > \varrho$. This choice of a uniform distribution receives partial support from the characteristic arborization pattern of pyramidal axon collaterals and its consequent effect on pyramidal-pyramidal cell connections (Szentágothai, 1975). Thus, for the homogeneous isotropic case, the distribution of postsynaptic neurons is uniform over a circular region of radius ϱ. For the anisotropic connectivity distribution, an elliptical region of influence is assumed, centered on the presynaptic cell body. Such a field is characterized by the lengths of the major and minor axes. For a given anisotropic connectivity distribution, neurons have their regions of influence parallel to each other (i.e., the major axes are parallel).

Single Cell Properties and Parameters

Every cell in each functionally defined group is characterized by a set of distinct parameters: threshold, summation time, number of efferent fibers, absolute and relative refractory periods, and rate of spontaneous activity. Each cell linearly sums the postsynaptic potentials impinging on it. Conduction times for postsynaptic and action potentials over the relatively short intracortical fiber distances are neglected. Synapses are characterized by a synaptic delay, and the strengths of excitatory (EPSP) or inhibitory (IPSP) postsynaptic potentials. These synaptic strengths refer to the effects produced at the axon hillock of a postsynaptic cell by the firing of a presynaptic neuron. Intracellular recordings from neurons in cat visual cortex (Toyama et al., 1974) often indicate that EPSPs have a short time course while IPSPs result in long lasting hyperpolarizations. The convergence of a large number of presumably inhibitory synaptic terminals surrounding pyramidal cell bodies and initial segments of the apical dendrites (Szentágothai, 1973; Marin-Padilla, 1974; Garey, 1976) suggest a possible anatomical explanation for this long lasting inhibition if the IPSPs arrive with different latencies.

Accordingly, mean EPSP magnitudes are chosen to be smaller than mean IPSP magnitudes, and the summation time for EPSPs is chosen to be shorter than the summation time for IPSPs. The *E*-cell thresholds are fixed so that, on the average, two to three EPSPs are needed to fire an *E*-cell in its resting state. We have chosen the average magnitude of the EPSPs generated by *E*-cells within the network to be approximately half that of EPSPs generated by afferent stimulation. These choices were to some extent guided by experimental data (Creutzfeldt and Ito, 1968).

Spontaneous activity of neurons in the visual system decreases considerably as one goes from the retina to LGN to striate cortex (Jacobs, 1972). We have chosen the maintained activity rate to be zero in these runs.

Procedure and Simulation Results

All results in this section are from a single, representative network. That is, connections among neurons remain fixed, though parameters describing single cell and stimulus properties may vary as indicated. A network of 3000 neurons is simulated by storing the parameters describing the individual *E* and *I*-cells and the details of the connectivity in core memory of an IBM 370/155 computer. There are approximately thirty functional connections per cell; each connection may correspond to a number of synapses. Every neuron is characterized by the location of its cell body on a plane. Cell positions on this surface are selected by a random number generator (flat distribution) with an additional algorithm used to reduce excessive cell density fluctuations.

The simulations were carried out in batch mode using FORTRAN IV. We employed a sequential simulation technique, whereby the network is searched at fixed time intervals for neural events and the state of each cell in the network is systematically updated. Time is quantized in units of synaptic delay. Though restriction of firing rates to discrete time intervals may introduce artifacts in the dynamics, the relative importance of such artifacts is model dependent. Wong and Harth (1973) have shown that, in the netlet approach, effective desynchronization is achieved by allowing the time intervals to become small relative to synaptic delays and other neural time constants. Their analysis of netlet dynamics demonstrated that the qualitative nature of the dynamics remains unchanged under these conditions.

A stimulus consists of both spatial and temporal components. Fixed regions of the network which may receive excitatory afferent stimulation are chosen as either circular or elliptical with different orientations. Neurons lying within those regions may receive "afferent EPSPs" at certain rates and for a prescribed length of time, the stimulus duration. Stimuli are presented near the center of the plane to minimize edge effects.

Stimulus intensity is determined in the following way. At each time interval (iteration) a certain fraction of cells within the chosen region receive EPSPs of constant size. The number of EPSPs generated by afferent stimulation is termed *stimulus intensity*. These neurons are selected by a random number generator using a uniform distribution. The time course of the stimulus is either rectangular or a linearly changing function of time. Throughout this paper, stimuli have been presented to fixed regions of the network. No motion dependence has been considered. Though cortical cells respond more vigorously to moving stimuli than to stationary stimuli, these cells exhibit the same optimal orientation and similar tuning curves for both stimulus types (Henry et al., 1974). The effects of moving stimuli are presently under study and will be considered elsewhere.

A flexible group of programs (Finette, to be published) allows us to monitor most features of the resulting network dynamics. For example, the computer furnishes graphic representations (via a Calcomp plotter) of the distribution of cells in the stimulus region, and the spatial and temporal distribution of depolarization, hyperpolarization and spike activity in all neurons.

For our study we have selected network parameters such that the response to localized stimulation remains localized. Within this locally addressable region, the response characteristics fall into four distinct categories (Fig. 3): *unresponsive*, *transient*, *sustained*, and *reverberatory*. In the first case, the activity remains low throughout, and is mostly due to afferent rather than intracortical connections. Transient activity shows a distinct initial burst in the excitatory fraction of the neuron population, followed by a relatively silent period, analogous to what is called a *phasic* response in sensory physiology. Sustained activity is more or less constant for the duration of the stimulus but ceases shortly after stimulus cessation, whereas reverberatory activity remains locked in, once initiated. It will be noted in our data that *E* and *I*-cell activities, in general, follow a distinctly different time course. We will consider the *E*-cell activity as representing the output of the cortex.

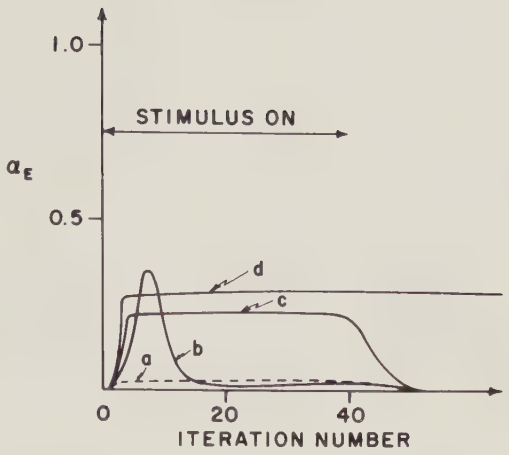

Fig. 3. General characteristics of the time course of activity, α_E, in a population of *E*-cells. Stimulus duration is indicated by double arrow. Four qualitatively different responses are shown: *a* unresponsive, *b* transient, *c* sustained, and *d* reverberatory

A variety of networks were tested for selectivity in regard to orientation and shape. Orientation sensitivity was examined by prescribing elliptical stimuli of constant intensity, differing only in orientation. Shape selectivity was tested by comparing responses to stimuli of circular and elliptical shapes. The areas stimulated were equal and contained approximately 200 neurons. For the elongated stimuli the ratio of major to minor axis was chosen to be 5:1 in all simulations.

The network is homogeneous in its connectivity, but anisotropic in the distribution of excitatory connections. We have chosen for the field of influence of an inhibitory neuron a circular region surrounding the neuron containing about 130 *E*-cells of which 25% are randomly chosen as postsynaptic to the given *I*-cell. The excitatory field was elliptical, with a ratio of 20:1 for the major and minor axes. An *E*-cell's field of influence covers an area consisting of approximately 35 neurons, 85% of which were chosen to be postsynaptic to a given cell.

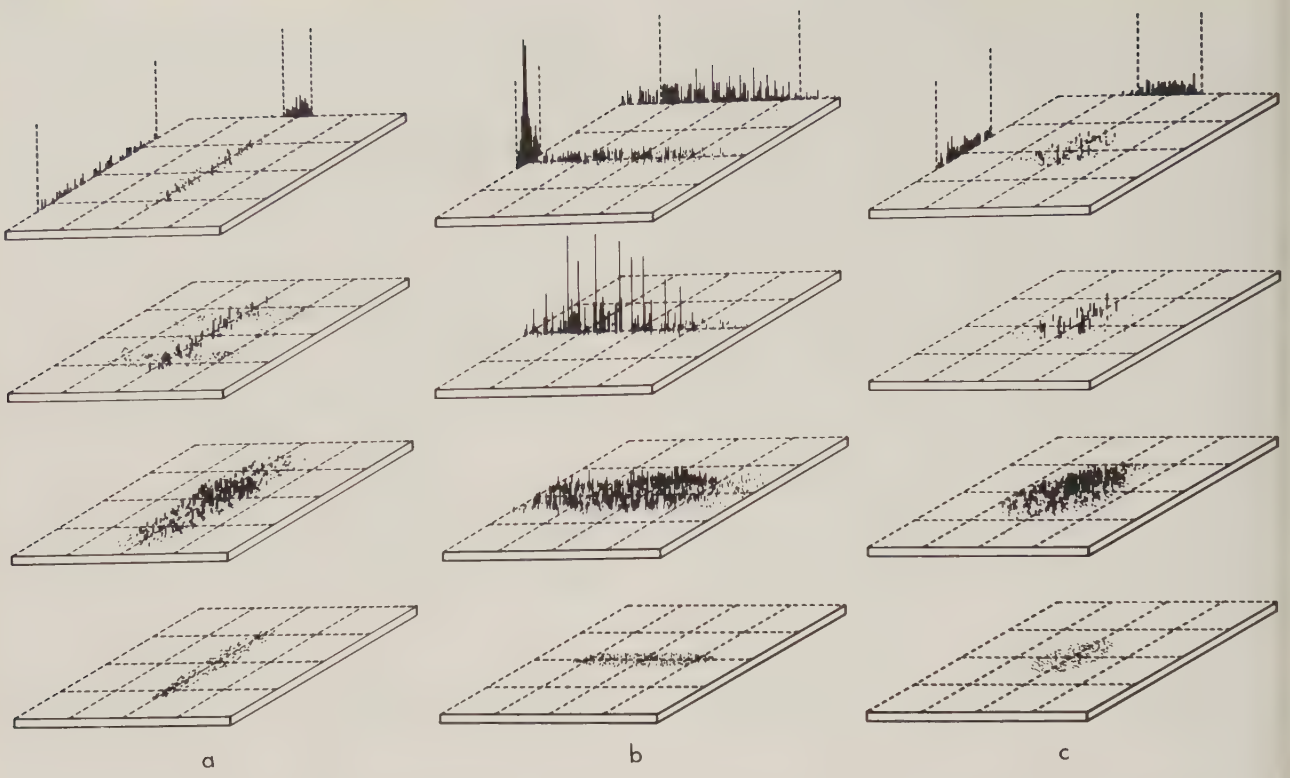

Fig. 4a–c. Stimulus pattern and spatial distribution of cell bodies (lower plane), hyperpolarized cells (2nd plane), depolarized cells (3rd plane) and cell firings (upper plane). The summed spike responses are shown against the projections of the boundaries of the stimulated region (vertical lines, upper plane). A set of planes for each stimulus configuration **a–c** indicate the state of the network at a given instant chosen arbitrarily from the middle of the runs. Distortions of the stimulus shape from elliptical or circular results from using an oblique coordinate system

In Fig. 4 we show the characteristic response of a network to two elongated stimuli of different orientations (Fig. 4a and b) and the responses to a circular stimulus (Fig. 4c).

Each set of planes in Fig. 4a–c represents the status of the network at a given instant and for a given stimulus shape and/or orientation. The iteration presented was chosen arbitrarily from the middle of the run. The only changes in parameters in these separate runs were those associated with stimulus shape and/or orientation. The spatial distribution of the stimulus on the network is depicted in the lower planes where each set of points correspond to the position of some of the neuron cell bodies which lie within the stimulated region. A randomly chosen subset of these cells (in this case, a constant 40%) receive afferent EPSPs for the first 40 iterations. The next plane indicates the spatial distribution of those neurons which are in a hyperpolarized state, while the third plane from the bottom shows the distribution of depolarized cells. The magnitude of depolarization or hyperpolarization for a given neuron is proportional to the height of the lines in these two planes. Finally, the upper plane shows a cumulative histogram of the number of firings of each cell in the network in response to the stimulus depicted in the lower plane. Here, the height of each line is proportional to the number of firings of a given neuron at the specified position. Also shown in the upper plane are the summed responses (total number of spikes) in two projection planes. The dashed vertical lines represent the projections of the boundaries of the stimulus on the two projection planes. Distortions of the stimulus shape from elliptical or circular are caused by the use of an oblique coordinate system to present the data.

A comparison of Fig. 4a and b demonstrates orientation sensitivity in this network. The stimulus is oriented such that its major axis is normal to that of the field of influence of *E*-cells in Fig. 4a but parallel in Fig. 4b. Shape selectivity is revealed in Fig. 4c, where the same stimulus strength is applied to a circular region whose area is equal to the area of the elliptical regions in Fig. 4a and b. The activities of both cell populations are displayed in Fig. 5 where the insets in the lower, middle, and upper graphs correspond to the stimulus shapes and orientations shown in the lower planes in Fig. 4a–c, respectively. Here, $\alpha_{E,I}$ refers to the respective fractions of *E* and *I*-cell populations

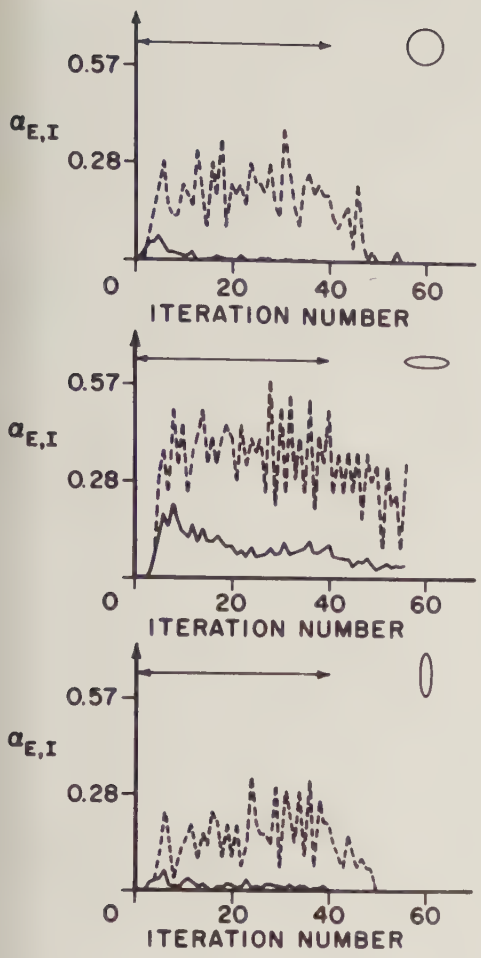

Fig. 5. Activities of *E* and *I*-cell populations within the stimulus region. $\alpha_{E,I}$ denotes the fraction of the *E*-cell (solid lines) and *I*-cell (dashed lines) populations firing at a given instant. Stimulus shapes and orientations are shown schematically as insets and the double-headed arrow indicates stimulus duration

(solid and dashed lines, respectively) within the stimulated region which are firing in a given iteration. The rather large fluctuations in *I*-cell activity are related to the small number of *I*-cells present (about 30). The "properly" oriented stimulus produces a reverberatory response in the *E*-cell population. For the "wrong" orientation, the *E*-cell activity can be described as unresponsive. This is also the case for the response to the circular stimulus.

Orientation tuning curves were determined for both subpopulation responses under a number of different parameter conditions. This was accomplished by choosing elliptical stimuli (same area as in Fig. 4) of equal intensity, and presenting them for forty iterations in several orientations in separate runs. Examples of tuning curves are given in Fig. 6 which show the total number of firings (E_T, I_T) within the network for both *E*-cell (solid line) and *I*-cell (dotted line) populations as a function of stimulus orientation.

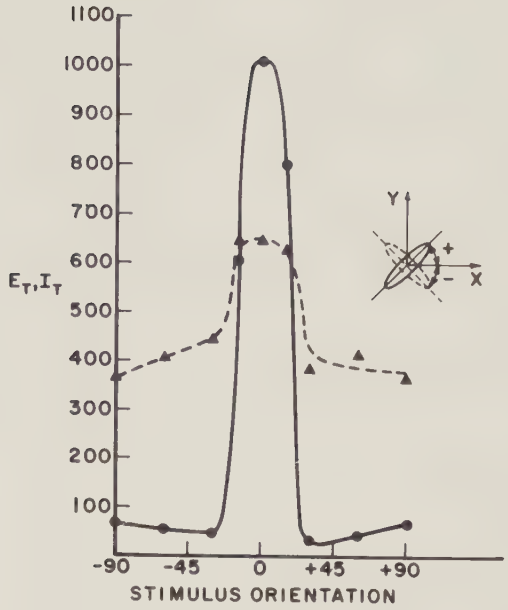

Fig. 6. Orientation tuning curves for *E*-cell (solid line) and *I*-cell (dotted line) populations. Each point corresponds to the total number of spikes in the population of *E* or *I*-cells within the stimulated region for the entire run. Stimuli were elliptical and were presented in different orientations in separate runs. The sense of the angle θ is shown in the inset

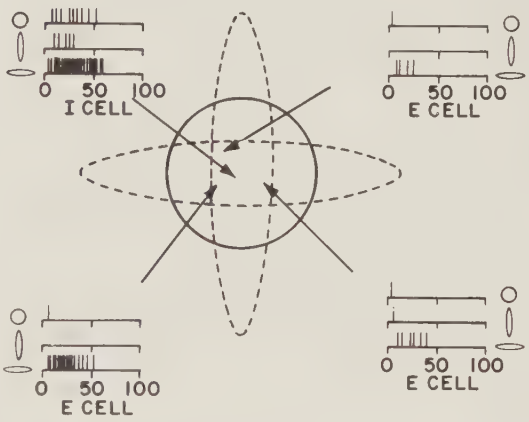

Fig. 7. Time course of action potentials for three *E*-cells and one *I*-cell, randomly selected from the central region common to all three stimuli. Two elliptical and one circular stimulus, covering equal areas were presented at constant intensity for 40 iterations. Each vertical line on the time scale represents a single action potential

In this example, the time course of activity in the *E*-cell population for different orientations ranged from sustained to unresponsive. *I*-cell activities were sustained in all orientations.

The corresponding time course of action potentials for single neurons, in response to circular and elliptical stimuli, is depicted in Fig. 7. These neurons were randomly selected from the central region common to

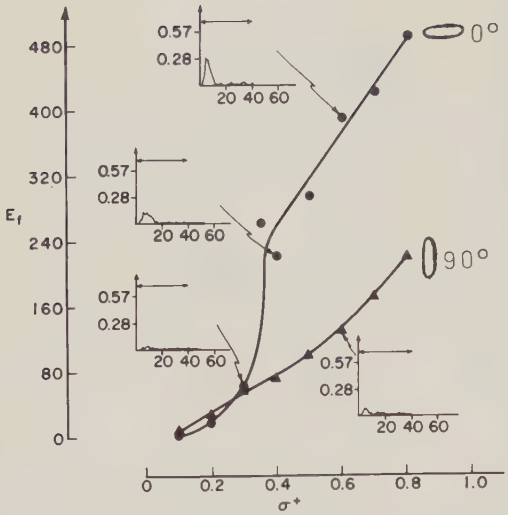

Fig. 8. Stimulus-response curves for rectangular pulses of different intensities, proportional to σ^+, presented in separate runs. Each point corresponds to the total number of *E*-cell firings, E_f, within the stimulated region over a period of 45 iterations. Two stimulus orientations are shown (corresponding to the orientations $\theta=0°$ and $\theta=90°$) and the insets depict the time course of *E*-cell activity for representative points within the locally addressable region

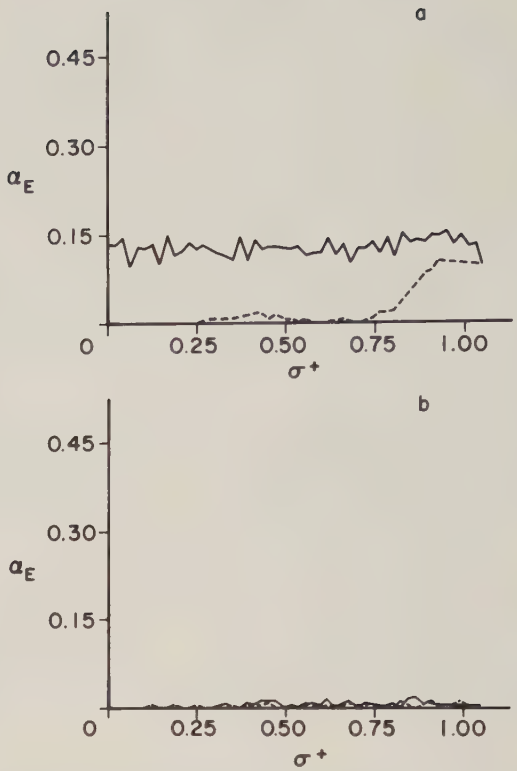

Fig. 9a and b. *E*-cell activity, α_E, in response to a slowly increasing and decreasing stimulus presented in a single run (dotted and solid lines, respectively). In **a**, an elliptical stimulus in the "correct" orientation evokes hysteresis in the *E*-cell population, while in **b** the same stimulus in the "wrong" orientation causes unresponsive activity

all three stimuli. The stimulus orientations and area were identical to those in Fig. 4. The three *E*-cells show markedly higher response to the "correct" orientation. This was true for practically all *E*-cells examined. Single *I*-cells were not as predictable in their response characteristics.

A prominent dynamic response in these networks is the appearance of phase transitions and hysteresis effects. In Fig. 8 we have chosen parameters which yield a transient response in the *E*-cell population when the stimulus intensity exceeds a critical value. Here, the curves labeled 90° and 0° correspond to the stimulus-response curves for the orientations shown in Fig. 4a and b, respectively. Each point on the stimulus-response curves corresponds to the total number of cell firings within the stimulated region over a period of 45 iterations. The points are determined in separate runs, in which the same net is tested for the response to constant stimulation. Insets show the time course of activity in the *E*-cell population for representative points. In the "wrong" orientation (90°) the *E*-cell population is unresponsive, except for very high stimulus intensities where a transient response begins to form. The response increases approximately linearly with stimulus intensity. The curve for "optimal" orientation (0°) follows essentially the same course up to $\sigma^+=0.3$, where σ^+ is proportional to stimulus intensity. However, between $\sigma^+=0.3$ and $\sigma^+=0.35$, the *E*-cell population undergoes a phase transition to a high *transient* response.

In Fig. 9 we show responses of *E*-cells to stimuli which are initially increasing (dotted lines), and then decreasing (solid lines). Two orientations are shown. The stimulus time dependence has a small slope to allow *E*-cell activity to approximate the steady state for any level of stimulation. The responses for each orientation were determined in a single run. Hysteresis is present (Fig. 9a) for the "optimal" orientation (0°), while the "wrong" orientation (90°) yields activity which is unresponsive (Fig. 9b). This hysteresis reflects the cooperative character of the *E*-cell activity. Increasing the level of inhibition within the network causes the effect to disappear as the response changes from reverberatory to unresponsive. This illustrates the separate effects of excitation and inhibition which can also be derived from the schematic surface shown in Fig. 1. Here, successive cuts through the surface for three different levels of inhibition show the same transition from the high reverberatory to the low unresponsive character.

Discussion and Conclusions

In the primary visual system of cat and monkey, the transformation from circular to elongated receptive

fields is abrupt. In cat, the transition occurs in the projection of LGN to striate cortex, while in monkey the transformation takes place between layer IVc and layers above and below IVc in the visual cortex. Hence, the circuitry underlying orientation sensitivity in monkey has a cortical origin, while in cat this may or may not be the case.

A fundamental question concerning the nature of the neuronal circuitry responsible for this transition remains unanswered. There appear to be two alternatives. The orientation sensitive neurons may receive multiple inputs from specific orientation insensitive cells, and thus be individually *wired* to respond to elongated visual stimuli of specific location and orientation. Since many cells in a given cortical region exhibit the same orientation selectivity (the orientation columns), this field-specific wiring would have to be repeated a large number of times. An alternative possibility is that orientation sensitivity is accomplished by locally anisotropic connectivity and cooperative interactions among these cells.

We have investigated a particular version of this alternative. Here, visual afferents to a sheet of cortical neurons are assumed to be retinotopically organized in such a manner that stimulation of a certain region in the retina would produce excitatory inputs into a geometrically similar region on the cortical sheet. Orientation sensitivity of individual neurons is accomplished by cooperative, anisotropic excitatory interactions within a local population constituting an orientation column.

To examine this idea we have carried out computer simulations for a network consisting of 3000 interacting elements representing excitatory and inhibitory neurons. Local interconnections were generated in which the cells postsynaptic to excitatory neurons were distributed evenly over elliptical regions, while connections originating from inhibitory neurons were chosen to be isotropic. Regions of different sizes and shapes were stimulated by distributing EPSPs among cells within the selected region. The stimuli were maintained for a fixed time span, and then turned off. Parameters of the network were chosen in such a way that neural activity remained localized within or near the region of stimulation. Activity resulting from stimulation was classified as unresponsive, transient, sustained, or reverberatory.

Results demonstrate that such networks can be tuned quite sharply to respond preferentially to stimuli of a given orientation. The nets were also selective for elongated shapes (of appropriate orientation), being relatively unresponsive to circular stimuli. Population responses often exhibit characteristic phase transitions and hysteresis effects observed by us previously in isotropic and randomly connected neural networks. The model is relatively independent of any one of the network parameters in the following sense: qualitative changes in the network's response due to variations in a given parameter can be compensated by other parameter changes. The important variable is the level of excitability in the network.

The origin of the anisotropy of excitatory neurons is not specifically attributed to either the axonal or dendritic fields. We have also studied connectivities incorporating both excitatory and inhibitory anisotropies, and anisotropic inhibitory connections with an isotropic excitatory distribution. In both cases, orientation tuning was observed for certain ranges of parameters.

The results point up a simplicity of design with which the observed transition from circular fields to elongated, orientation sensitive fields may be accomplished in the mammalian brain. The data therefore suggest that an intensified search be undertaken for an anatomical confirmation of such a scheme. Quantitative studies in which dendritic and axonal fields of cortical neurons in area 17 are carefully examined for anisotropies are now carried out in this laboratory. It is hoped that the accumulation of such structural data will make possible more realistic models of the visual cortex.

Acknowledgement. This work was in part sponsored by Grants NS10917 and EY01215 from the National Institutes of Health.

References

Amari, S.: Characteristics of randomly connected threshold element networks and network systems. Proc. IEEE **59**, 35–47 (1971)

Amari, S., Arbib, M.: Competition and cooperation in neural nets. In: Systems neuroscience. Metzler, J., ed. New York: Academic Press 1977

Anninos, P.A., Beek, B., Csermely, T.J., Harth, E., Pertile, G.: Dynamics of neural structures. J. Theor. Biol. **26**, 121–148 (1970)

Blinkov, G.M., Glezer, I.I.: The human brain in figures and tables. New York: Plenum Press 1968

Braitenberg, V.: Thoughts on the cerebral cortex. J. Theor. Biol. **46**, 421–447 (1974)

Colonnier, M.: The tangential organization of the visual cortex. J. Anat. **98**, 327–344 (1964)

Colonnier, M., Rossignol, S.: Heterogenity of the cerebral cortex. In: Basic mechanisms of the epilepsies. Jasper, H.H., Ward, A.A., eds. Boston: Little, Brown 1969

Cragg, B.G.: The density of synapses and neurons in the motor and visual areas of the cerebral cortex. J. Anat. **101**, 639–654 (1967)

Creutzfeldt, O.D., Ito, M.: Functional synaptic organization of primary visual cortex neurons in the cat. Exp. Brain Res. **6**, 324–352 (1968)

Creutzfeldt, O.D., Kuhnt, U., Benevento, L.A.: An intracellular analysis of visual cortical neurons to moving stimuli: responses in a cooperative neuronal network. Exp. Brain Res. **21**, 251–274 (1974)

Creutzfeldt, O.D., Garey, L.J., Kuroda, R., Wolff, J.: The distribution of degenerating axons after small lesions in the intact and isolated visual cortex of the cat. Exp. Brain Res. **27**, 419–440 (1977)

Daniel, P.M., Whitteridge, D.: The representation of the visual field on the cerebral cortex in monkeys. J. Physiol. **159,** 203–221 (1961)

Fisken, R.A., Garey, L.J., Powell, T.P.S.: The intrinsic, association and comissural connections of area 17 of the visual cortex. Phil. Trans. B **272**, 487–536 (1975)

Garey, L.J.: Synaptic organization of afferent fibers and intrinsic circuits in the neocortex. Electroenceph. Clin. Neurophysiol. **2**, pt. A 57–85 (1976)

Gentile, A.N., Harth, E.: The alignment of serial sections by spatial filtering. Comput. Biomed. Res. **11**, (1978) (in press)

Globus, A., Scheibel, A.B.: Synaptic loci on visual cortical neurons of the rabbit: the specific afferent radiation. Exp. Neurol. **18**, 116–131 (1967)

Grossberg, S.: Contour enhancement, short term memory, and constancies in reverberating neural networks. Stud. Appl. Math. **52**, 213–257 (1973)

Harth, E., Csermely, T.J., Beek, B., Lindsay, R.D.: Brain functions and neural dynamics. J. Theor. Biol. **26**, 93–120 (1970)

Harth, E., Lewis, N.S., Csermely, T.J.: The escape of Tritonia: dynamics of a neuro-muscular control mechanism. J. Theor. Biol. **55**, 201–228 (1975)

Hubel, D.H., Wiesel, T.N.: Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex. J. Physiol. (Lond.) **160**, 106–154 (1962)

Hubel, D.H., Wiesel, T.N.: Functional architecture of macaque monkey visual cortex. Proc. R. Soc. Lond. B **198**, 1–59 (1977)

Hubel, D.H., Wiesel, T.N., Stryker, M.P.: Anatomical demonstration of orientation columns in macaque monkey. J. Comp. Neur. **177**, 361–380 (1978)

Jacobs, G.H.: Spontaneous activity in visual systems. Am. J. Optom. Physiol. Opt. **49**, 905–921 (1972)

Jones, E.G.: Varieties and distribution of non-pyramical cells in the somatic sensory cortex of the squirrel monkey. J. Comp. Neurol. **160**, 205–268 (1975)

Lund, J.: Organization of neurons in the visual cortex, area 17, of the monkey *(Maccaca mulatta)*. J. Comp. Neur. **147**, 455–496 (1973)

Marin-Padilla, M.: Three-dimensional reconstruction of the pericellular nests (baskets) of the motor (area 4) and visual (area 17) of the human cerebral cortex. A Golgi study. Z. Anat. Entwicklungsgesch. **144**, 123–135 (1974)

Marin-Padilla, M., Stibitz, G.R.: Three-dimensional reconstruction of the basket cell of the human motor cortex. Brain Res. **70**, 511–514 (1974)

McIlwain, J.T.: Large receptive fields and spatial transformations in the visual system. In: International review of physiology, neurophysiology. II, 10. Porter, R., ed. Baltimore, Md.: University Park Press 1976

Nelson, J.I., Frost, B.J.: Orientation-selective inhibition from beyond the classic visual receptive field. Brain Res. **139**, 359–365 (1978)

Paldino, A.M., Harth, E.: Some quantitative results on Golgi impregnated axons in rat visual cortex using a computer assisted video digitizer. J. Comp. Neur. **176**, 247–262 (1977)

Schiller, P.H., Finlay, B.L., Volman, S.F.: Quantitative studies of single-cell properties in monkey striate cortex. V. Multivariate statistical analyses and models. J. Neurophysiol. **39**, 1362–1374 (1976)

Sholl, D.A.: The organization of the visual cortex in the cat. J. Anat. **89**, 33–46 (1955)

Skeen, L.C., Humphrey, A.L., Norton, T.T., Hall, W.C.: Deoxyglucose mapping of the orientation column system in the striate cortex of the tree shrew, *Tupaia glis*. Brain Res. **142**, 538–545 (1978)

Szentágothai, J.: Synaptology of the visual cortex. In: Handbook of sensory physiology, Vol. VII/3. Part B, pp. 269–324. Central processing of visual information. Jung, R., ed. Berlin, Heidelberg, New York: Springer 1973

Szentágothai, J., Arbib, M.A.: Conceptual models of neural organization. Neurosciences Research Program Bulletin **12**, No. 3. Cambridge: MIT Press 1974

Szentágothai, J.: The module concept in cerebral cortex architecture. Brain Res. **95**, 475–496 (1975)

Toyama, K., Matsunomi, K., Ohno, T., Tokashiki, S.: An intracellular study of neuronal organization in the visual cortex. Exp. Brain Res. **21**, 45–66 (1974)

Tusa, R.J., Palmer, L.A., Rosenquist, A.C.: The retinotopic organization of area 17 (striate cortex) in the cat. J. Comp. Neur. **177**, 213–236 (1978)

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

Wilson, H.R., Cowen, J.S.: A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. Kybernetik **13**, 55–80 (1973)

Wong, R., Harth, E.: Stationary states and transients in neural populations. J. Theor. Biol. **40**, 77–106 (1973)

Received: June 9, 1978

Dr. Steven Finette
Department of Physics
Syracuse University
Syracuse, NY 13210, USA

Biol. Cybernetics 30, 241—248 (1978)

Biological Cybernetics

Biological Oscillators Can Be Stopped—Topological Study of a Phase Response Curve

M. Kawato and R. Suzuki

Faculty of Engineering Science, Osaka University, Toyonaka-shi, Osaka, Japan

Abstract. Many biological oscillators are stable against noise and perturbation (e.g. circadian rhythms, biochemical oscillators, pacemaker neurons, bursting neurons and neural networks with periodic outputs). The experiment of phase shifts resulting from discrete perturbation of stable biological rhythms was developed by Perkel and coworkers (Perkel et al., 1964). By these methods, they could get important insights into the entrainment behaviors of biological rhythms. Phase response curves, which are measured in these experiments, can be classified into two types. The one is the curve with one mapping degree (Type 1), and the other is that with zero mapping degree (Type 0) (Winfree, 1970). We define the phase response curve mathematically, and explain the difference between these two types by the homotopy theory. Moreover, we prove that, if a Type 0 curve is obtained at a certain magnitude of perturbation, there exists at least one lower magnitude for which the phase response curve cannot be measured. Some applications of these theoretical results are presented.

1. Introduction

Perkel and his coworkers investigated the consequences of inhibitory or excitatory synaptic input between pacemaker neurons in abdominal ganglia of Aplysia and in strech receptors of Procambarus (Perkel et al., 1964). In this work they exploited the experiment of phase shifts resulting from discrete perturbation of rhythms of pacemaker neurons. By this method, they achieved success in getting important insights into the entrainment behaviors of pacemaker neurons. Generally, the phase resetting experiments can be used to investigate the behaviors of stable biological oscillators (e.g. circadian rhythms, biochemical oscillators, pacemaker neurons, bursting neurons and neural networks with periodic outputs). Their intrinsic properties as oscillators, their entrainment behavior under periodic forces and entrainment behavior of stable oscillators interacting each other can be investigated by this method. Winfree pointed out that two different types of phase response curves were measured in the phase resetting experiment of circadian rhythms (Winfree, 1970). The one is the curve with average slope of unity (one mapping degree, i.e. the Type 1 curve), and the other is that with zero average slope (zero mapping degree, i.e. the Type 0 curve). These two types of phase response curves have been found also in neural pacemakers (Winfree, 1977a). Why these two types of the phase response curve appear? To study this, we redefine the phase response curves mathematically and explain the difference between two types by homotopy theory. Applications of these mathematical results are developed.

2. Definition of the Phase Response Curve

In this chapter we explain the method of phase resetting experiments and define the phase response curve after Winfree (Winfree, 1973). Objects of these experiments are only stable oscillators. Consider a physical quantity x into which the biological oscillator projects its internal state. Let the period of oscillation be τ. We define phase ϕ as follows. At an arbitrary reference event (e.g. a maximum of x), $\phi=0$, otherwise $\phi=t/\tau(\mathrm{mod}\ 1)$ (see Fig. 1). Perturbation lasting for T is applied to the free running oscillator at phase $\phi-T/\tau$. Here, we call ϕ as the old phase. A thermal pulse and a light pulse are examples of perturbations for circadian rhythms. Injection of some chemical substance is that for biochemical oscillators. An electric current across the membrane, orthodromic and antidromic stimuli are that for neural oscillators. The elasped time of the i-th reference event after perturbation from the end of

0340-1200/78/0030/0241/$01.60

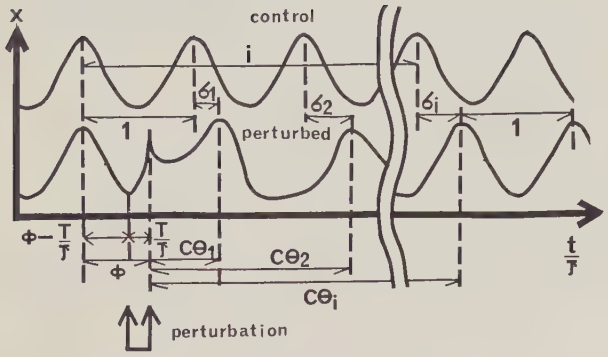

Fig. 1. Measuring method of a phase response curve

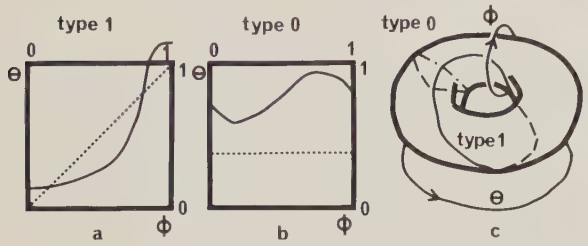

Fig. 2a—c. Two types of PRC with different mapping degree (the Type 1 PRC and the Type 0 PRC)

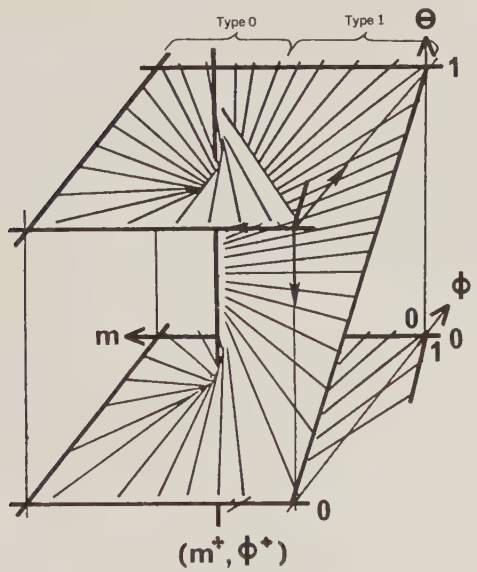

Fig. 3. The phase response surface (PRS) of an oscillation of glycolysis in yeast by injecting oxygen dissolved in buffer (from Winfree, 1975)

perturbation is normalized by τ and is denoted by the i-th cophase $C\theta_i$.

The i-th new phase θ_i is defined as

$$\theta_i = i - C\theta_i (\text{mod } 1) .$$

The i-th delay σ_i is defined as

$$\sigma_i = \phi + C\theta_i - i = \phi - \theta_i (\text{mod } 1)$$

(see Fig. 1). Because the oscillator is stable, its period returns to τ after a long period. So, θ_i and σ_i have their limits $\theta = \lim_{i \to \infty} \theta_i$ and $\sigma = \lim_{i \to \infty} \sigma_i$ respectively. We call θ the new phase, and σ the delay. Both θ and σ are functions of ϕ. Generally both $\theta(\phi)$ and $\sigma(\phi)$ are called as the phase response curve (PRC). Obviously, $\sigma(\phi) = \phi - \theta(\phi)$ holds. In this paper we call $\theta(\phi)$ the PRC, unless we say the contrary. In some cases $C\theta_1(\phi)$, $\theta_1(\phi)$, and $\sigma_1(\phi)$ are called PRC. In these cases PRCs are generally discontinuous and applications of PRCs are hard, so we don't adopt this definition. We discuss this problem fully in Chapter 6.

3. The Shape of a Phase Response Curve

PRCs are measured in various fields of biology. Though objects of the experiments are different, very fundamental, common properties of PRCs exist. First, because both θ and ϕ are mod 1, the PRC $\theta(\phi)$ is biperiodic. Second, as we shall prove in Chapter 5, $\theta(\phi)$ is a continuous function of ϕ. So, while ϕ changes 0 to 1, θ can change only by an integral number. PRCs can be classified into separate types according to these integral numbers (i.e. average slopes of PRCs). This integer number is called as mapping degree in topology. We may expect that PRCs of all integer mapping degree have been found. In biology, however, only the PRCs of one mapping degree (Type 1 PRC) and the PRCs of zero mapping degree (Type 0 PRC) were found. We show PRCs of Types 1 and 0 in Figure 2a and b respectively. Because of biperiodicity of the PRC, it is more natural to represent the PRC on the torus T^2 on which θ is scaled in the direction of the equator and ϕ is scaled in the direction of the meridian as in Figure 2c. PRCs drawn with dotted line in Figure 2a and b are the typical PRCs of Types 1 and 0 respectively. In Figure 2a $\theta(\phi) = \phi$ which means that eventually perturbation doesn't affect the rhythm at all. In Figure 2b $\theta(\phi) = \text{const}$, so the new phase is independent of the old phase.

Winfree pointed out that both the Types 1 and 0 PRC were measured by varying the magnitude of the perturbation for an identical circadian rhythm (Winfree, 1970). To examine this fully, he measured the PRC of an oscillation of glycolysis in yeast by injecting oxygen dissolved in buffer (Winfree, 1975). Varying the disturbance dosage of oxygen, he got a phase response surface (PRS) in Figure 3. The term "surface" means that the magnitude of perturbation is also varied. Here, an antero-posterior axis scales the old phase ϕ, and a right to left axis scales the disturbance dosage of oxygen m, and vertical axis scales the new phase θ. The perturbation of the magnitude m^* delivered at the old phase ϕ^* switched off the oscillation. Winfree called the line (m^*, ϕ^*) in 3-dimensional space as a singular

axis. PRCs with less dosage than m^* were Type 1 and PRCs with more dosage than m^* were Type 0. He predicted that PRSs of biological oscillators generally resemble that of Figure 3 (Winfree, 1977b). In the following chapters we shall prove mathematically that this prediction is true under certain restrictions.

4. Mathematical Formulation of the PRC and the PRS

All terminology about the dynamical system in this chapter are explained in Appendix A.

We redefine the PRC and PRS mathematically for the proof of theorems in Chapter 5. If a stable oscillation in biology can be described by a set of ordinary differential equations, it has a stable limit cycle because of its stability against noise and perturbation. We assume that the oscillator in its free running state can be described by a smooth flow (A) $\Psi: M \times R \to M$ on a smooth manifold M, or by the following system of a differential equation with an equivalent vector field $X(x) = \delta\Psi/\delta t|(x, 0)$

$$dx/dt = X(x). \tag{1}$$

The flow Ψ [or the system (1)] is assumed to have an elementary (AAA) stable limit cycle γ with a period τ. Here "elementary" is an extra assumption. But it is reasonable if we endow (1) with structural stability. We also assume that the oscillator can be described by the following Equation (2), or its equivalent flow $\Psi(m): M \times R \to M$, when perturbation of the magnitude m is applied on it.

$$dx/dt = X(x) + m \cdot Z(x, m). \tag{2}$$

When $m = 0$, (2) accords with (1) and $\Psi(0) = \Psi$ holds.

The following theorem has been known regarding the dynamical system with an elementary stable limit cycle (Guckenheimer, 1975).

Theorem (Guckenheimer). *Let $\Psi: M \times R \to M$ be a smooth flow with an elementary stable limit cycle γ. The stable set $W^s(x)$ (AA) of each $x \in \gamma$ is a) a cross section (AAAA) of γ, b) a manifold diffeomorphic to euclidean space. Moreover, the union of the stable manifolds $W^s(x)$, $x \in \gamma$, is an open neighborhood of γ and the stable manifold of γ.*

This theorem enables us to define eventual phase ϕ_e of (1) as follows.

Definition 1. Let a phase map $\phi: \gamma \to S^1$ be a homeomorphism such that for each $x \in \gamma$, $\phi(\Psi(x, t)) = \phi(x) + t/\tau (\mathrm{mod}\ 1)$. Where S^1 is a unit circle.

Definition 2. The eventual phase map $\phi_e: W^s(\gamma) \to S^1$ can now be defined as follows. For arbitrary $y \in W^s(\gamma)$, $\phi_e(y) = \phi(x)\left\{x \in \gamma ; \lim_{t \to \infty} \underset{M}{\mathrm{dis}}\,(\Psi(y, t), \Psi(x, t)) = 0\right\}$.

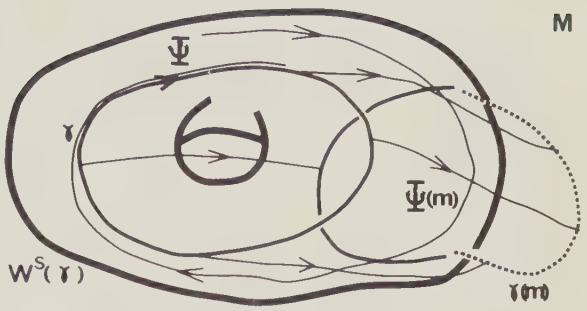

Fig. 4. An illustration of simple closed path $\gamma(m)$

From definition, for $y \in W^s(x)$, $\phi_e(y) = \phi(x)$. We assume that the state point of the oscillator is on γ at the beginning of perturbation, and duration of perturbation is T regardless of the magnitude m. If $z(x, m)$ has an appropriate regularity with respect to x, then for fixed m,

$$\Psi(m)(\gamma, T) = \{\Psi(m)(x, T); x \in \gamma\}$$
$$= \{\Psi(m)(\phi^{-1}(c), T); c \in S^1\}$$

is a simple closed arc in M because of continuity of the solution of (2) with respect to initial conditions and because of the uniqueness of the time-inversed system of (2). We call this closed arc $\gamma(m)$. $\gamma(0) = \gamma$ holds. $\gamma(m)$ is a set of state points at the ending of perturbation (see Fig. 4).

Definition 3. The Phase resetting map (PRS)

$$\theta(m, c): R^+ \times S^1 \to S^1$$

is defined for (m, c) such that

$$\Psi(m)(\phi^{-1}(c - T/\tau), T) \in W^s(\gamma),$$

as follows

$$\theta(m, c) = \phi_e(\Psi(m)(\phi^{-1}(c - T/\tau), T)).$$

For the fixed magnitude of perturbation m, $\theta(m, \cdot)$ accords with the PRC, $\theta(\phi)$ defined in Chapter 2. When the magnitude m is also variable, $\theta(\cdot, \cdot)$ accords with the PRS in Chapter 3.

5. The Proof of the Existence of a Singular Phase Response Curve

All terminology about homotopy group in this chapter are explained in Appendix B.

In this chapter, we prove the existence of a singular phase response curve, and that Winfree's conjecture is true under certain conditions.

Lemma 1. *If $Z(x, m)$ satisfies appropriate regularity conditions with respect to (x, m) on $M \times D$ (where D is a section of R^+ containing zero), then $\theta(m, c)$ is continuous with respect to (m, c) on $D \times S^1$.*

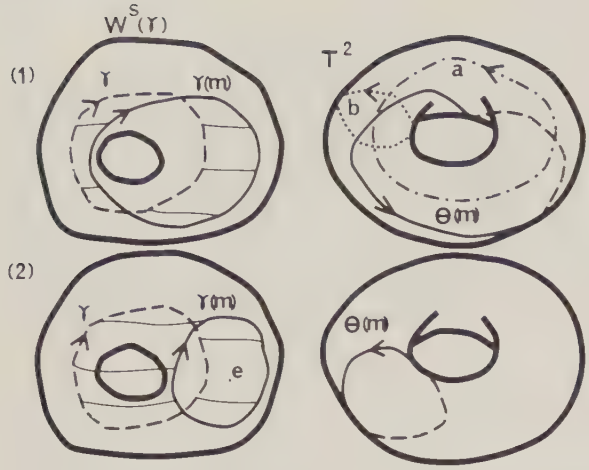

Fig. 5. An illustration of the Theorem 1

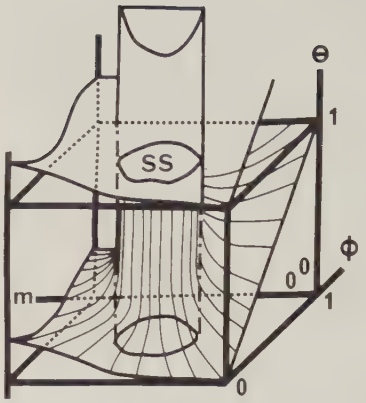

Fig. 6. An example of PRS with *SS* of finite extension

Proof. Step 1: Let us consider the following system.

$$\begin{cases} dx(t)/dt = X(x(t)) + m(t)Z(x(t), m(t)) \\ dm(t)/dt = 0. \end{cases} \tag{3}$$

The flow of the system (3) is denoted by

$$F : (M \times D) \times R \to M \times D.$$

Because of regularity of $Z(x, m)$ with respect to (x, m), the solution of (3) has continuity with respect to initial conditions. That is, for every $x_0 \in M$, $m_0 \in D$, and for every $\varepsilon > 0$, there exists $\delta > 0$ with the properties that for (x, m):

$$\operatorname{dis}_{M \times D} ((x, m), (x_0, m_0)) < \delta,$$

$$\operatorname{dis}_{M \times D} (F((x, m), T), F((x_0, m_0), T)) < \varepsilon$$

holds. The latter inequality can be written as

$$\operatorname{dis}_{M \times D} ((\Psi(m)(x, T), m), (\Psi(m_0)(x_0, T), m_0)) < \varepsilon.$$

This means that

$$\operatorname{dis}_{M} (\Psi(m)(x, T), \Psi(m_0)(x_0, T)) < \varepsilon.$$

Step 2: Because γ is asymptotically orbitally stable, for every $y_0 \in W^s(\gamma)$ and every small neighborhood $U(\gamma)$ of γ, there exists $s > 0$ such that $\Psi(y_0, s) \in U(\gamma)$. Because of continuity of the solution of (1) with respect to initial conditions, for arbitrary $\varepsilon > 0$, there exists $\delta > 0$ with the properties that for every y; $\operatorname{dis}(y, y_0) < \delta$, $\Psi(y, s) \in U(\gamma)$ and

$$\operatorname{dis}_{M} (\Psi(y, s), \Psi(y_0, s)) < \varepsilon$$

holds.

Step 3: Guckenheimers theorem asserts that for every $x \in \gamma$, $W^s(x)$ is cross section of γ. So if neighborhood $U(\gamma)$ of γ is sufficiently small, then for arbitrary $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\operatorname{dis}_{M} (y_1, y_2) < \delta (y_1, y_2 \in U(\gamma))$$

implies

$$\operatorname{dis}_{M} (x_1, x_2) < \varepsilon (y_1 \in W^s(x_1), y_2 \in W^s(x_2)).$$

Definition 3, Steps 1–3 show the proof. Continuity of the first derivative of $Z(x, m)$ with respect to (x, m) is sufficient.

Lemma 2. *γ is a deformation retract* (BBB) *of $W^s(\gamma)$.*

Proof. Guckenheimer's theorem guarantees the existence of invariant, smooth and sufficient small neighborhood $U(\gamma)$ of γ whose boundary integral curves of (1) cross transversally. Apparently γ is a deformation retract of $U(\gamma)$. The rest to be proved is that $U(\gamma)$ is a deformation retract of $W^s(\gamma)$. Because γ is asymptotically orbitally stable, for every $y \in W^2(\gamma)$, $\notin U(\gamma)$ there exists continuous $T(y)$ such that $\Psi(y, T(y))$ is on the boundary of $U(\gamma)$. We define homotopy (B)

$$r_t : W^s(\gamma) \times I \to W^s(\gamma)$$

for every $y \in W^s(\gamma)$ and every $t \in I$. If $y \in U(\gamma)$ then $r_t(y) = y$ and if $y \notin U(\gamma)$ then $r_t(y) = \Psi(y, tT(y))$. Because both $T(y)$ and Ψ is continuous and $r_0(y) = y$, $r_1 : W^s(\gamma) \to U(\gamma)$, r_t is a required homotopy.

It follows that $W^s(\gamma)$ is homotopy equivalent (BB) to γ from Lemma 2. So the fundamental group (BBBB) of $W^s(\gamma)$ is isomorphic to the integer group Z and its generator is a closed path $[\gamma = \gamma(0)]$. On the other hand generators of the fundamental group of the torus T^2 are the equator [a] and the meridian [b] (see Fig. 5).

Theorem 1. *1) If $\gamma(m)$ is homotopic* (B) *to $\gamma = \gamma(0)$ in $W^s(\gamma)$, then $\theta(m, \cdot)$ is homotopic to $a + b$ in the torus T^2.*

2) If $\gamma(m)$ is homotopic to a constant mapping e in $W^s(\gamma)$, then $\theta(m, \cdot)$ is homotopic to b in the torus T^2 (see Fig. 5).

Proof. 1) Both $\gamma(m) = \Psi(m)(\phi^{-1}(\cdot), T) : S^1 \to W^s(\gamma)$ and $\theta(m) = \theta(m, \cdot) : S^1 \to S^1$ are continuous from Lemma 1. $\theta(0, c) = c$, that is $\theta(0, \cdot) \simeq a + b$. Because $\gamma(m) \simeq \gamma$ in

$W^s(\gamma)$ there exists a homotopy $G(c,t)$ such that $G: S^1 \times I \to W^s(\gamma)$; $G(\cdot,0)=\gamma(0)$, $G(\cdot,1)=\gamma(m)$ and for fixed $t \in I=[0,1]$, $G(\cdot,t)$ is a closed path in $W^s(\gamma)$. For each $t \in I$, we define closed path $H(\cdot,t)$ on torus T^2 as we have defined a PRS in Definition 3. That is $H(\cdot,t) = \phi_e(G(\cdot - T/\tau, t))$. Because $G(c,t)$ and ϕ_e are continuous, $H(c,t)$ is continuous with respect to (c,t). Moreover, it is a homotopy $H: S^1 \times I \to S^1$; $H(\cdot,0) = \theta(0,\cdot)$, $H(\cdot,1)=\theta(m,\cdot)$, so $\theta(m,\cdot) \simeq \theta(0,\cdot) \simeq a+b$.

2) We can define a closed path θ_e on the torus T^2 from a constant mapping as Definition 3, then $\theta_e \simeq b$ in T^2. The rest of the proof is same as 1).

$\theta(m,\cdot) \simeq a+b$ implies that the PRC for the magnitude m is Type 1. $\theta(m,\cdot) \simeq b$ implies that the PRC is Type 0. If $\gamma(m) \simeq n\gamma$, then $\theta(m,\cdot) \simeq na+b$ and the PRC is Type n. However this has not been observed in biology. Now all $\gamma(m)$ is listed up from Lemma 2.

Theorem 2. *If there exists $m_0 \in D$ such that $\theta(m_0,\cdot)$ is homotopic to b, then there exists at least one $m^*(0<m^*<m_0)$ with the property that*

$$\gamma(m^*) \cap W^s(\gamma)^c \neq \{\phi\},$$

in other words the PRC for this magnitude m^* is singular. Where, $W^s(\gamma)^c$ is complement of $W^s(\gamma)$ in M.

Proof. If for all m; $0<m<m_0$, $\gamma(m) \cap W^s(\gamma)^c = \{\phi\}$, then there exists homotopy $F(c,t)$ such that

$$F(c,t)=\theta(m_0 t, c): S^1 \times I \to S^1,$$

$F(c,0)=\theta(0,c)$, $F(c,1)=\theta(m_0,c)$ from Lemma 1. It follows that

$$a+b \simeq \theta(0,c) \simeq \theta(m_0,c) \simeq b$$

in T^2. This is a contradiction.

Theorem 2 says that if Type 0 PRC is measured for a certain magnitude of perturbation, then there exists at least one smaller magnitude for which a PRC is not defined. That is, the perturbation of magnitude m^* delivered at an appropriate old phase stops oscillation.

In phase resetting experiments of circadian rhythm, PRCs are often measured for varied durations with fixed magnitude of perturbation. In this case we can prove the following as same as before. If a Type 0 PRC is measured for certain duration of perturbation, then there exists at least one shorter duration for which a PRC is not defined.

If the PRC for m is not defined, then

$$\gamma(m) \cap W^s(\gamma)^c \neq \{\phi\}.$$

Now we define a singular set $SS \in D \times S^1$ as follows.

$$SS = \{(m,c) \in D \times S^1 : \Psi(m)(\phi^{-1}(c - T/\tau), T) \in W^s(\gamma)^c\}.$$

SS may be a point (m^*, c^*), or may be a (un)connected domain(s). In the former case, Winfree called (m^*, c^*) as a singular axis. In the latter case, an example of a PRS is drawn in Figure 6. If $(m,c) \in SS$, then perturbation of magnitude m at phase c turns off the oscillation.

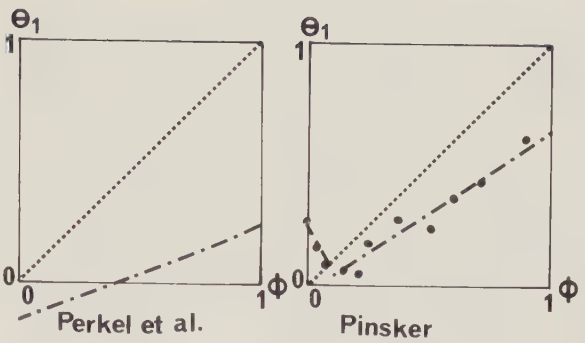

Fig. 7. Discontinuous PRCs measured by Perkel et al. (1964) and Pinsker (1977)

The PRS is very steep near the singular axis in Figure 3. It is also very steep near the frontier of SS in Figure 6. Guckenheimer proved that this is true for many (in some sense) dynamical systems with an elementary stable limit cycle (Guckenheimer, 1975). From this, we can say that $\theta(m,c)$ is continuous on SS^c but it is not uniformly continuous for many dynamical systems.

After all we can say that if a Type 0 PRC is measured for some magnitude of perturbation, the PRS seems like a full or hollow cork screw whether SS is a point or a domain(s) because of its periodicity as to ϕ and θ. In the following chapters applications of these results are developed.

6. A Discontinuous Phase Response Curve

As Winfree pointed out, discontinuous PRCs have been obtained in neurophysiology (Winfree, 1977a; Perkel et al., 1964; Pinsker, 1977). These discontinuous PRCs are different from singular PRCs in Chapter 5. Above stated two discontinuous PRCs are re-drawn in the form of the first new phase *vs* the old phase, $\theta_1(\phi)$, in Figure 7. The reference state ($\phi=0$) is the time of a spike for the PRC of Perkel et al., and it is the beginning of burst for the PRC of Pinsker. In both cases, the PRCs are discontinuous at $\phi=0$. As we have stated in Chapter 2, discontinuity of these PRCs comes from their definition, that is, these PRCs are of the first newphase θ_1, the first delay σ_1 or the first cophase $C\theta_1$.

For simplicity, we explain this discontinuity in a two dimensional dynamical system on a (X,Y)-plane. Let us consider the system in Figure 8a with a stable limit cycle γ, the phase mapping ϕ on γ, stable manifold of every point of limit cycle (this is a curve in the two dimensional case) and a simple closed path $\gamma(m)$ which is a set of state points perturbed for duration T. The reference state is set at the time of maximum of Y. Spike line (denoted SL in Figure 8a) is

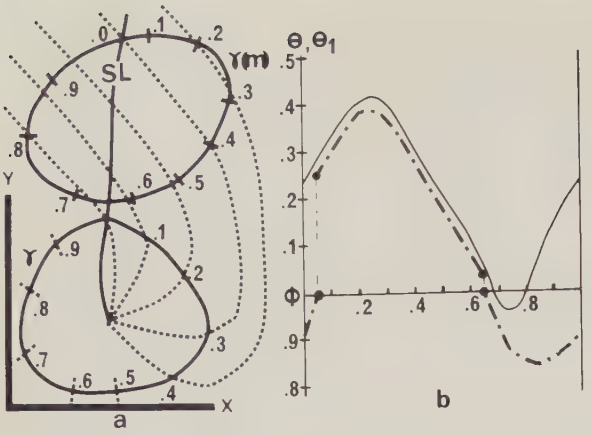

Fig. 8a and b. An illustration of discontinuous PRCs. $\theta(\phi)$ and $\theta_1(\phi)$ of a certain 2 dimensional dynamical system

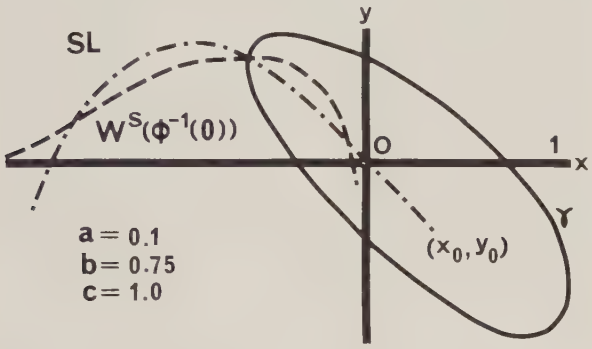

Fig. 9. Spike Line (SL) and stable manifolds of BVP model

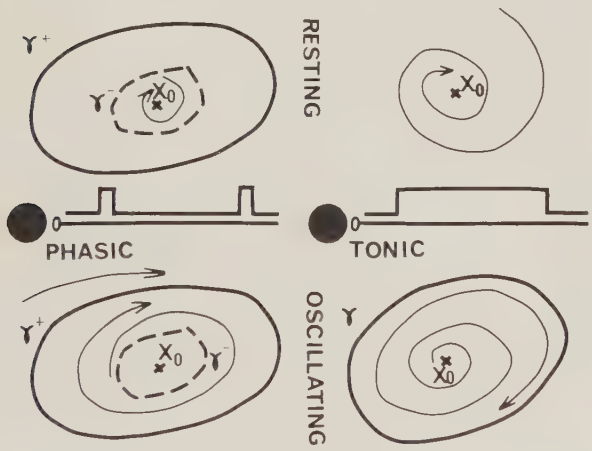

Fig. 10. Two modes of control of oscillators in neurobiology. (P) Phasic control. (T) Tonic control

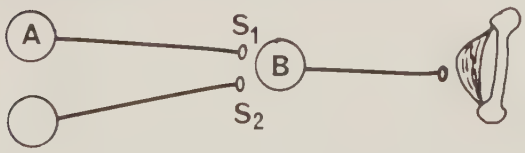

Fig. 11. Two neural oscillators one of which is entrained by the inputs from the other

defined as a set of points where Y gets maximum along the orbit of the system. The PRC of the new phase *vs* the old phase, $\theta(\phi)$, is drawn by a solid line in Figure 8b on the basis of stable manifolds. As we have proved in Chapter 5, this PRC is continuous. On the other hand, if we draw the PRC of the first new phase *vs* the old phase, $\theta_1(\phi)$, on the basis of the first reference state after perturbation, then it is discontinuous as a dotted line in Figure 8b. The reason of this is as follows. If perturbation is applied immediately after the maximum of Y, then the state point of the system comes fully close to a stable limit cycle at the time of the first reference state after the perturbation. So θ_1 is close to θ because SL and $W^s(\phi^{-1}(0))$ meet on the limit cycle. On the other hand, if Y reaches its maximum immediately after the end of perturbation, then θ_1 is very different from θ. So, $\theta_1(\phi)$ is discontinuous at the old phase where SL crosses $\gamma(m)$. Let us consider Fitzhugh's BVP-model to examine this problem concretely. BVP-model is as follows.

$$\begin{cases} dx/dt = c(y + x - x^3/3) \\ dy/dt = -(x - a + by)/c\,, \end{cases} \tag{4}$$

where, x is the minus quantity of a membrane potential, and y is the quantity of refractoriness. a, b, and c are constants satisfying $0 < b < 1$, $c > 0$, $b < c^2$. From the stability analysis of the equilibrium point and Poincaré-Bendixons theorem, we can say that the system has a stable limit cycle if

$$|a| < (1 - b/c^2)^{1/2}(1 - 2b/3 - b^2/(3c^2))\,.$$

In this parameter region we calculated stable manifolds by computer simulation. The reference state is a set at the minimum of x (i.e. maximum of the membrane potential). So, SL is the left part of a set of points $\{dx/dt = 0$ i.e. $y = x^3/3 - x\}$ to the equilibrium point (see Fig. 9).

Perkel et al. said that conditions of entrainment of a pacemaker neuron with a natural period τ to synaptic inputs with a period s is that the function $\sigma_1(\phi) + (\tau - s)/\tau$ has its zero point with a slope between 0 and 2 and the stable phase delay is the old phase at this zero point (Perkel et al., 1964). But, it does not matter when the oscillator is at its reference state after the perturbation and it do matter where the statepoint of the oscillator is at the beginning of a next perturbation. So $\sigma(\phi)$ must substitute for $\sigma_1(\phi)$.

Other reasons of discontinuities of PRCs $\theta_1(\phi)$ may be possible. To begin with, pacemaker neurons and bursting neurons may not be able to be described by a system of ordinary differential equations. Nevertheless, it seems like that $\theta(m,c)$ is continuous under certain appropriate conditions for general evolution equations such as partial differential equations or ordinary differ-

ential equations with time lag. Even if one would like to measure $\theta_1(\phi)$, it is desirable to measure $\theta(\phi)$ at the same time.

7. Applications to Neurobiology

In this chapter applications of the PRC and PRS to neurobiology are developed.

7.1. Two Modes of Control of the Neural Oscillators

Some neural oscillators must be switched on and off by the central nervous system (CNS). In these controls two modes are possible. The one mode is a phasic control (P) and the other is a tonic control of an oscillator (T). In case of the phasic control, the starting signal and the stopping signal from the CNS to the oscillator are phasic signals. In the tonic control the CNS must keep sending a pulse train to the oscillator to maintain its oscillation or to keep it in the resting state. We can decide whether a certain neural oscillator is controlled under the phasic mode or under the tonic mode by examining the PRS of that oscillator. In case of P, the *SS* of the PRS is a domain with finite extension, and in case of T it is a point.

For simplicity we explain these two modes of controls in the two dimensional dynamical system (see Fig. 10). In the case of the phasic control, the state point is on the stable equilibrium point X_0 in the resting state. It is driven out of the inside of an unstable limit cycle γ^- by the starting signals from the CNS, and it keeps rounding along a stable limit cycle γ^+ until it is driven into the inside of γ^- by the stopping signals from the CNS. CNS can switch on and off the oscillator quickly. The period of the oscillation is inherent in the oscillator, and stable against various perturbations. The CNS must be informed of phase of the oscillator to switch it off. Though the switching off mechanism is different from the phasic control, the neural network of the escape of Tritonia is the example of the phasic control (Harth et al., 1975). In engineering, these types of oscillators are called as hard oscillators.

In the case of the tonic control, the oscillator at rest is described by a dynamical system with a unique stable equilibrium point X_0. In oscillation, it is described by a dynamical system with an unstable equilibrium point X_0 and a stable limit cycle γ (see Fig. 10T). When the starting signal from the CNS arrives or the stopping signal ceases, X_0 becomes unstable and γ is formed around it. Then the state point becomes to round γ. In this control, the CNS takes a relatively long time to turn on regular oscillation or to turn off the oscillation at all. The period of the oscillation is a function of the frequency of a pulse train from the CNS and it is sensitive to perturbation. The oscillators of flight in a locust and the oscillator of a swimmeret of a crayfish are examples of oscillators under tonic control (Wilson, 1961; Stein, 1974). These oscillators are called as soft oscillators in engineering.

7.2. Hierarchy of Neural Oscillators

In some cases, each oscillator interacting each other can oscillate apart from the rest of the organ. In this case, can we decide which oscillator is the true pacemaker? Let us consider two oscillators. The one oscillator (A) acts on the other (B) through the synapse (S1) at its spike (see Fig. 11). If the oscillator B is synchronized to the output of A, in other words if B is governed by A, then the PRC $\theta(\phi)$ of B for perturbation such as a synaptic input through other synapse S2 than S1 or electric current across the membrane is a typical Type 1 PRC. Vice versa. The PRC $\theta_1(\phi)$ for these perturbation is not typical Type 1. If we dissect the axon from A to B and measure the PRC of B with orthodromic stimulation of this axon, then both $\theta(\phi)$ and $\theta_1(\phi)$ are not typical Type 1.

8. Conclusions

PRCs can be applied to investigate properties of oscillators, their entrainment behavior to periodic external inputs, controls of oscillators and interactions between oscillators. Here topological properties of PRCs are studied by the theory of a dynamical system and the homotopy theory. PRC is especially efficient to investigate human oscillatory neural networks which reject analytic investigations such as electrophysiology. We are studying psychologically human finger tapping neural networks by the PRC on the basis of theoretical results developed here (Yamanishi et al., 1978).

Theoretical questions are left about generalization of discussions in Chapters 4 and 5 to general evolution equations. Why in biology only the Types 1 and 0 PRCs have been measured? Is it because the technique of delivering perturbation is poor? If the dynamical system describing the biological oscillation is two dimensional, we can prove that the PRCs are only Types 1 or 0.

Appendix A

(A) flow: Flow $\Psi : M \times R \to M$ is defined as $\Psi(x,t) = y$ if the state x becomes the state y after t units of time. The map Ψ is to satisfy the usual flow properties $\Psi(x,0) = x$ and $\Psi(x, t_1 + t_2) = \Psi(\Psi(x,t_1), t_2)$.

(AA) stable set: If Ψ is a flow on M and S is a subset of M, then the stable set of S, denoted $W^s(S)$, is the set of points y for which $d(\Psi(y,t), \Psi(S,t)) \to 0$ as $t \to \infty$. If stable set is also a manifold, it is called a stable manifold.

(AAA) elementary stable: The Poincaré map of γ is denoted as Θ. One says that γ is an elementary (or hyperbolic) limit cycle if the matrix $D\Theta_x$ of first partial derivatives of Θ at $x \in \gamma$ has no eigenvalues of absolute value one.

(AAAA) cross section: Let $\Psi : M \times R \to M$ be a flow with a periodic orbit γ of period τ and let $x \in \gamma$. A cross section of γ at x is a submanifold $N \subset M$ with the following properties:

1) $x \in N$ and $\bar{N} \cap \gamma = \{x\}$. $\bar{N}$ is the closure of N in M.
2) $T_x N + T_x \gamma = T_x M$.

Appendix B

(B) homotopic: Let X and Y be topological spaces. If for two continuous mappings $f, g : X \to Y$, there exists continuous mapping $F : X \times I \to Y$ with the properties that $F|(X \times \{0\}) = f$, $F|(X \times \{1\}) = g$, then one says that f and g is homotopic and denote $f \simeq g$. F is called homotopy between f and g.

(BB) homotopy equivalent: Let X, Y be topological spaces, f, g be continuous mappings;

$$f : X \to Y, g : Y \to X \quad \text{and} \quad 1_x : X \to X, 1_y : Y \to Y$$

be an identity mapping. If $gf \simeq 1_x$, $fg \simeq 1_y$, then X is homotopy equivalent to Y.

(BBB) deformation retract: If a continuous mapping r from X to A which is a subset of X satisfies the following properties:

1) $r|A = 1_A$.

2) If we consider r as a continuous mapping from X to X, then $r \simeq 1_X$ and its homotopy $\{r_t\}$ ($r_0 = 1_X$, $r_1 = r$) satisfies $r_t|A = 1_A$,

then one says that A is a deformation retract of X and call $r : X \to A$ as a retraction. If A is a deformation retract of X, then X is homotopy equivalent to A.

(BBBB) fundamental group: Let X be an arcwise connected topological space. The set of all paths which have $x_0 \in X$ as a base point $\{w; w : I \to X, w(0) = w(1) = x_0\}$ classificated by the relation of homotopic equivalent ($\simeq$) is called as a homotopy class of closed paths having x_0 as a base point and is denoted $\pi_1(X, x_0)$. We introduce the structure of a group into $\pi_1(X, x_0)$ by defining product $w \cdot w'$ of two elements $w, w' \in \pi_1(X, x_0)$ as a closed paths having x_0 as a base point. This group is called as a fundamental group of X with x_0 as a base point (or 1-dimensional homotopy group or Poincaré group) and is denoted $\pi_1(X, x_0)$ again. Because X is arcwise connected, the fundamental group is decided independent of a base point. We call this as a fundamental group of X and denote $\pi_1(X)$. If X is homotopy equivalent to Y, then $\pi_1(X)$ is isomorphic to $\pi_1(Y)$.

Acknowledgement. One of the authors (M.K.) should like to express his gratitude to Professor S. Sato, Dr. K. Kobayashi, Dr. O. Sueda and our colleagues in the Department of Biophysical Engineering for many helpful comments.

References

Guckenheimer, J.: Isochrons and phaseless sets. J. Math. Biol. **1**, 259—273 (1975)

Harth, E., Lewis, N.S., Csermely, T.J.: The escape of Tritonia: dynamics of a neuromuscular control mechanism. J. Theor. Biol. **55**, 201—228 (1975)

Perkel, D.H., Schulman, J.H., Bullock, T.H., Moore, G.P., Segundo, J.P.: Pacemaker neurons: effects of regularly spaced synaptic input. Science **145**, 61—63 (1964)

Pinsker, H.M.: Aplysia bursting neurons as endogeneous oscillators. I. Phase-response curves for pulsed inhibitory synaptic input. J. Neurophysiol. **40**, 527—543 (1977)

Stein, P.S.G.: Neural control of interappendage phase during locomotion. Amer. Zool. **14**, 1003—1016 (1974)

Wilson, D.M.: The central nervous control of flight in a locust. J. Exp. Biol. **38**, 471—490 (1961)

Winfree, A.T.: Integrated view of resetting a circadian clock. J. Theor. Biol. **28**, 327—374 (1970)

Winfree, A.T.: The investigation of oscillatory processes by perturbation experiments. In: Biological and biochemical oscillators. pp. 461—501. Chance, B., Pye, E., Gosh, A., Hess, B., eds. New York: Academic Press 1973

Winfree, A.T.: Resetting biological clocks. Physics Today **28**, 34—39 (1975)

Winfree, A.T.: Phase control of neural pacemakers. Science **197**, 761—763 (1977a)

Winfree, A.T.: Some principles and paradoxes about the phase control of biological oscillators. J. Interdiscip. Cycle Res. **8**, 1—14 (1977b)

Yamanishi, J., Kawato, M., Suzuki, R.: Investigations of human finger tapping neural networks by PRC. The 17-th congress of Japan society of medical electronics and biological engineering 1978 (in Japanese)

Received: June 19, 1978

Dr. M. Kawato
Fac. of Eng. Science
Osaka Univ.
Toyonaka-shi
Osaka, 560, Japan

Biomathematics

Managing Editors: K. Krickeberg, S.A. Levin

Volume 1:
Mathematical Topics in Population Genetics
Edited by K. Kojima
55 figures. IX, 400 pages. 1970
Cloth DM 88,–; US $ 38.80
ISBN 3-540-05054-X

Contents: Random Drift and the Shifting Balance Theory of Evolution.–Changes in Mean Fitness under Natural Selection.–Models and Analyses of Dispersal Patterns.–Avoidance and Rate of Inbreeding.–Genetic Loads and the Cost of Natural Selection.–Stochastic Processes in Population Genetics, with Special Reference to Distribution of Gene Frequencies and Probability of Gene Fixation.–Theory of Limits to Selection with Line Crossing.–A Theory of Limits in Artificial Selection with Many Linked Loci.–The Evolution of Dominance.–Survival of Mutant Genes as a Branching Process.–The Incomplete Binomial Distribution.–Evolutionary Significance of Linkage and Epistasis.–Fitness and Optimization.

From the review:
"...It is far and away the most solid product I have ever seen labelled biomathematics."
American Scientist

Volume 2: E. Batschelet
Introduction to Mathematics for Life Scientists
2nd edition
227 figures. XV, 643 pages. 1975
Cloth DM 64,–; US $ 28.20
ISBN 3-540-07293-4
Also available as "Springer Study Edition"

Contents: Real Numbers.–Sets and Symbolic Logic.–Relations and Functions.–The Power Function and Related Functions.–Periodic Functions.–Exponential and Logarithmic Functions I.–Graphical Methods.–Limits.–Differential and Integral Calculus.–Exponential and Logarithmic Functions II.–Ordinary Differential Equations.–Functions of Two or More Independent Variables.–Probability.–Matrices and Vectors.–Complex Numbers.–Appendix (Tables A-K).–Solutions to Odd Numbered Problems.

From the reviews:
"A sincere attempt to relate basic mathematics to the needs of the student of life sciences."
Mathematics Teacher

"...Professor Batschelet is to be congratulated on the production of a book which will undoubtedly make mathematics more accessible to life scientists."
The Journal of Ecology

Volume 3*/Volume 4*:
M. Iosifescu, P. Tautu
Stochastic Processes and Applications in Biology and Medicine
*Distribution rights for the Socialist Countries: Centrală Cartii, Bucharest
Part 1: **Theory**
331 pages. 1973
Cloth DM 68,–; US $ 30.00
ISBN 3-540-06270-X

Contents: Discrete Parameter Stochastic Processes: Denumerable Markov Chains. Noteworthy Classes of Denumerable Markov Chains. Markov Chains with Arbitrary State Space. Continuous Parameter Stochastic Processes: Some General Problems. Processes with Independent Increments. Markov Processes.

Part 2: **Models**
337 pages. 1973
Cloth DM 58,–; US $ 25.60
ISBN 3-540-06271-8

Contents: Preliminary Considerations.–Population Growth Models.–Population Dynamics Processes.–Evolutionary Processes.– Models in Physiology and Pathology.

From the reviews:
"...the two-volume set, with its very extensive bibliography, is a survey of recent work as well as a textbook. It is highly recommended by the reviewer."
American Scientist

Volume 5: A. Jacquard
The Genetic Structure of Populations
Translated by D. Charlesworth, B. Charlesworth
92 figures. XVIII, 569 pages. 1974
Cloth DM 96,–; US $ 42.30
ISBN 3-540-06329-3

Contents: Basic Facts and Concepts.–A Reference Model: Absence of Evolutionary Factors.– The Causes of Evolutionary Changes in Populations.–The Study of Human Population Structure: Genetic Distance.–Appendices.

From the reviews:
"...should take its place as a major reference work..."
Science

Volume 6: D. Smith, N. Keyfitz
Mathematical Demography
Selected Papers
31 figures. XI, 515 pages. 1977
Cloth DM 78,–; US $ 34.40
ISBN 3-540-07899-1

Contents: The Life Table.–Stable Population Theory.–Attempts at Prediction and the Theory they Stimulated.–Parameterization and Curve Fitting.–Probability Models of Conception and Birth.–Branching Theory and Other Stochastic Processes.–Cohort and Period, Problem of the Sexes, Sampling.

This collection of readings brings together the major historical contributions that form the base of current population mathematics tracing the development of the field from the early explorations of Graunt and Halley in the seventeenth century to Lotka and his successors in the twentieth. The volume includes 55 articles and excerpts with introductory histories and mathematical notes by the editors.

Volume 7: E.R. Lewis
Network Models in Population Biology
187 figures. XII, 402 pages. 1977
Cloth DM 64,80; US $ 29.90
ISBN 3-540-08214-X

Contents: Foundations of Modeling Dynamic Systems.–General Concepts of Population Modeling.–A Network Approach to Population Modeling.–Analysis of Network Models.–Appendices: Probability Arrays, Array Manipulation. Bernoulli Trials and the Binomial Distribution.

Directed toward biologists who are looking for an introduction to biologically motivated systems theory, this book provides a simple, heuristic approach to quantitative and theoretical population biology.

Prices are subject to change without notice

Springer-Verlag
Berlin
Heidelberg
New York

5303/4/1

Springer books in

Pattern Recognition

NEW

T. PAVLIDIS

Structural Pattern Recognition

173 figures, 11 tables. Approx. 290 pages. 1977
Cloth DM 43,—; US $19.80
(Springer Series in Electrophysics, Volume 1)
ISBN 3-540-08463-0

Contents: Mathematical Techniques for Curve Fitting. — Graphs and Grids. — Fundamentals of Picture Segmentation. — Advanced Segmentation Techniques. — Scene Analysis. — Analytical Description of Region Boundaries. — Syntactic Analysis of Region Boundaries and Other Curves. — Shape Description by Region Analysis. — Classification, Description and Syntactic Analysis.

The book deals primarily with the encoding of pictures into mathematical structures which can be handled by classical pattern recognition techniques. It emphasizes methodology, showing the connection between various approaches, and represents the first book to provide a systematic description of the process of obtaining measurements from pictorial data.
This volume is intended as a text for a one-semester advanced course in pattern recognition and is based on class notes for a course on this subject given at Princeton University.

Springer-Verlag
Berlin Heidelberg New York

Further titles

Digital Pattern Recognition

Editor: K. S. FU
With contributions by T. M. Cover, E. Diday, K. S. Fu, A. Rosenfeld, J.-C. Simon, T. J. Wagner, J. S. Weszka, J. J. Wolf
54 figures, 4 tables. XI, 206 pages. 1976
Cloth DM 79,80; US $35.20
(Communication and Cybernetics, Volume 10)
ISBN 3-540-07511-9

Syntactic Pattern Recognition, Applications

Editor: K. S. FU
With contributions by J. E. Albus, R. H. Anderson, J. M. Brayer, R. DeMori, H.-Y. F. Feng, K. S. Fu, S. L. Horowitz, B. Moayor, T. Pavlidis, W. W. Stallings, P. H. Swain, T. Vamos
135 figures, 19 tables. XI, 270 pages. 1977
Cloth DM 85,—; US $37.40
(Communication and Cybernetics, Volume 14)
ISBN 3-540-07841-X

U. GRENANDER

Pattern Synthesis

Lectures in Pattern Theory, Volume 1
120 figures, 18 tables. VII, 509 pages. 1976
DM 36,20. US $16.00
(Applied Mathematical Sciences, Volume 18)
ISBN 3-540-90174-4

Picture Processing and Digital Filtering

Editor: T. S. HUANG
With contributions by H. C. Andrews, F. C. Billingsley, J. G. Fiasconaro, B. R. Frieden, T. S. Huang, R. R. Read, J. L. Shanks, S. Treitel
113 figures. XIII, 289 pages. 1975
Cloth DM 79,80; US $35.20
(Topics in Applied Physics, Volume 6)
ISBN 3-540-07202-0

Digital Picture Analysis

Editor: A. ROSENFELD
With contributions by S. J. Dwyer, R. M. Haralick, C. A. Harlow, G. Lodwick, R. L. McIlwain, K. Preston, A. Rosenfeld, J. R. Ullmann
114 figures, 47 tables. XIII, 351 pages. 1976
Cloth DM 72,—; US $31.70
(Topics in Applied Physics, Volume 11)
ISBN 3-540-07579-8

Prices are subject to change without notice